

BESZÉDKUTATÁS 2009

**HELYZETKÉP
A BESZÉDTUDOMÁNYRÓL**

BUDAPEST

BESZÉDKUTATÁS 2009

HELYZETKÉP A BESZÉDTUDOMÁNYRÓL

**Szerkesztette:
GÓSY MÁRIA**

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest 2009

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

A szerkesztőbizottság elnöke: Gósy Mária

A szerkesztőbizottság tagjai: Hoffman Ildikó

Mády Katalin

Markó Alexandra

Marton Klára

Olaszy Gábor

Siptár Péter

Vago, Robert

A korpusz szerkesztésében részt vettek:

Bóna Judit

Gósy Mária

Gráczy Tekla Etelka

Gyarmathy Dorottya

Horváth Viktória

Imre Angéla

Markó Alexandra

Neuberger Tilda

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztő:

Markó Alexandra

ISSN 1218-8727

© MTA Nyelvtudományi Intézet

Felelős kiadó: Kenesei István igazgató

Készült: Regiszter Kiadó és Nyomda Kft.

TARTALOM

Markó Alexandra – Grácz Tekla Etelka – Bóna Judit: Zöngésségi hasonulás a spontán beszédben és a felolvasásban (esettanulmányok)	5
Beke András – Horváth Viktória: A nazális koartikuláció variabilitása a spontán beszédben	28
Gósy Mária: Szóejtés és szóészlelés: változatosság és adaptálódás	46
Bóna Judit: Az idős életkor tükröződése a magánhangzók ejtésében	76
Markó Alexandra: Stigmatizált hanglejtésforma a spontán beszédben	88
Bata Sarolta: Beszélőváltások a beszédpartnerek személyes kapcsolatának függvényében	107
Grácz Tekla Etelka: Temporális jellemzők a beszédpartnerek ismeretségének függvényében	121
Olaszy Gábor – Rácz Zsuzsanna Zsófia – Bartalis Mátyás: Formánsmérések automatizálása, formánsadatbázisok létrehozása	134
Beke András – Szaszák György: A svávariációk automatikus felismerése magyar nyelvű spontán beszédben	148
Gósy Mária – Gyarmathy Dorottya – Horváth Viktória: A beszéd természetességéről alkalmazott fonetikai szempontból	170
Neuberger Tilda: A spontán beszéd lejegyzése – a BEA adatbázis tapasztalatai alapján	182
Gyarmathy Dorottya: A beszélő bizonytalanságának jelzései: ismétlések és újraindítások	196
Bakti Mária: Megakadások a szinkrontolmácsok spontán beszédében	217
Váradi Viola: Hallásalapú és vizuális alapú közlések	228
Horváth Viktória – Imre Angéla: A diszlexia tünetei a spontán beszédben	240
„Nyelvbtlás”-korpusz	257
1. Téves szó	259
2. Grammatikai hiba	260
3. Kontamináció	261
4. Téves kezdés	262

5. „Nyelvem hegyén van” jelenség	263
6. Perszeveráció	263
7. Anticipáció	264
8. Metatézis	265
9. Egyszerű nyelvbontás	265
10. Több típusba sorolható jelenségek	266
A pedagógusjelöltek Kazinczy-versenye	268
A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 37. országos döntője • <i>Fehér Éva</i>	268
Szabó Zoltán: Szerelmes földrajz (részlet)	270
A 2009-es döntő kötelező szövegének értő-értető előadása • <i>Bóna Judit</i>	271
A magyar és a nemzetközi tudományos élet eseményei	273
Beszéd kutatás 2009 konferencia • <i>Erdős Klaudia</i>	273
A kötet szerzői	278

ZÖNGÉSSÉGI HASONULÁS A SPONTÁN BESZÉDBEN ÉS A FELOLVASÁSBAN (ESETTANULMÁNYOK)

Markó Alexandra – Gráci Tekla Etelka – Bóna Judit

Bevezetés

Az elmúlt évtizedekben számos kutatás vizsgálta a fonológiai folyamatok fonetikai vetületét, megvalósulását (pl. Barry 1992, Recasens–Pallarès 2001, Jansen 2007). Egyre több tanulmány fókuszál a zöngésségi hasonulás okozta neutralizációra, amely során egy fonéma zöngésségi jegye kiterjed a másik fonéma realizációjára (Burton–Roblee 1997, Jansen–Toft 2002, Gow–Im 2004 stb.). A zöngésségi hasonulás fonetikai aspektusú vizsgálatának kérdései között a zöngésségi oppozíció neutralizálódása mellett a hasonulás természet, iránya és mértéke is fontos. A magyar nyelv gazdag vizsgálati területet kínál a zöngésségi hasonulás elemzéséhez.

A különböző beszédstílusok eltérő fonetikai jellegzetességeket mutatnak, mind szegmentális, mind szuprasegmentális szempontból. Számos koartikulációs modell tekintetbe veszi a beszédstílusokat mint a megvalósulástípusok kiváltó okát. Lindblom a H&H elméletben (1990) felveti, hogy az artikuláció (a feszesség-lazaság tekintetében) a túl- és az alulartikulált beszéd két végpontja között egy skálán írható le. Az elmélet alapja, hogy a hangokat (ebből következően a hangkapcsolatokat szintén) soha nem ejtjük azonos módon. Az ejtés több dimenzió mentén változik, így például a kontextus, a hangerő, a beszédtempó, a beszédstílus mentén. A beszélő folyamatosan monitorozza, ellenőrzi a hallgatóval megosztani kívánt információ szempontjából azt, hogy artikulációjának mennyire kell érthetőnek lennie. Feltételezésünk szerint a spontán beszéd és az olvasás ennek a H&H-skálának két különböző intervallumában helyezkedik el. A hangos olvasás általában a gondozott, túlartikulált beszédhez, míg a spontán beszéd típusai az alulartikulált beszédhez állnak közelebb, és emiatt eltérő beszédhangjellemzőkkel (pl. koartikuláció, redukció) írhatók le. A két beszédstílus tervezési sajátosságait tekintve is különböző. Olvasás során a beszélőnek nem kell magasabb szintű beszédtervezési folyamatokat végrehajtania, figyelmét a pontosabb artikulációra fordíthatja. A spontán beszéd ezzel szemben párhuzamosan működő kognitív folyamatokat feltételez, amiknek következtében az artikuláció lényegesen automatizáltabban működik. Feltételezzük, hogy ezek a beszédstílusok közötti eltérések a zöngésségi hasonulásban is eltérő sajátosságokhoz vezetnek.

A magyarban a szomszédos mássalhangzók zöngésségük tekintetében többnyire megegyeznek (pl. *bukta*, *kesztyű*, *labda*, *pezsgő* stb.). Ehhez alkalmazkodnak a különböző zöngésségű fonémakapcsolatot tartalmazó szóátvételek is (pl. *futball* [fudbəl], *abház* [ɔphaːz]; Siptár–Törkenczy 2000).

A tövégi zöngétlen mássalhangzó zöngésedik, amennyiben zöngés konzonánssal kezdődő toldalék járul hozzá, a zöngés pedig zöngétlenedik zöngétlen mássalhangzóval kezdődő toldalék előtt. A zöngésségi hasonulás szóés szerkezethatáron is átível, amennyiben nem szakítja meg a mássalhangzókapcsolatot (hosszabb) szünet (Siptár–Törkenczy 2000).

1. Zöngésedés

zsák [ʒaːk] → *zsákb**an* [ʒaːɡbɔn]

kis [kiʃ] → *kis béka* [kiʒ beːkɔ]

2. Zöngétlenedés

kéz [keːz] → *kéztől* [keːstøːl]

nagy [nɔʃ] → *nagy kuty**a* [nɔc kuɔ]

A zöngésségi hasonulás többtagú mássalhangzó-kapcsolatok esetén nem csak a közvetlenül szomszédos, hanem ismétlődően az összes konzonánsan végbemegy (pl.: *test-ben* [teʒdben]) (Kenesei–Vágó–Fenyvesi 1998). A magyar zöngésségi hasonulást kötelező posztlexikális szabályként tartják számon, amelynek a bekövetkezése nem függ a beszéd/artikulációs tempótól sem (Siptár–Törkenczy 2000).

A magyar nyelvben három mássalhangzó aszimmetrikus viselkedést mutat a zöngésségi hasonulás tekintetében. A /v/ zöngétlenedik (*szívtől* [siːftøːl]), de nem zöngésít (*hatvan* [hɔtvɔn]). [A v kétarcú fonológiai viselkedésének vizsgálatáról lásd Bárkányi és Kiss több munkáját (pl. 2006).] A v viselkedésének fordítottja jellemzi a /h/-t [és/vagy /x/-t – Siptár–Törkenczy (2000)]; vagyis ez a mássalhangzó zöngétleníti a megelőző zöngés obstruenseket (*adhat* [ɔthɔt]), de önmaga nem zöngésedik a követő zöngés konzonáns hatására. A szonoráns /j/ leggyakrabban előforduló megvalósulása, alaprealizációja (zöngés) palatális approximáns (*ajtó* [ɔjtoː]), azonban amennyiben mássalhangzó előzi meg, és szünet követi, spiránsként realizálódik, valamint felveszi a megelőző konzonáns zöngésségi jegyét (pl. *dobj* [dobj], *férj* [feːrj], *kapj* [kɔpɕ]). A j tehát nem vált ki zöngésségi hasonulást, azonban megelőző zöngétlen obstruens hatására zöngétlenedik (Blaho 2008).

A fentiekben hivatkozott írásokon kívül számos további fonológiai elemzés foglalkozik a magyar zöngésségi hasonulással. Vago (1980) a magyar fonológiai folyamatok teljes rendszerét írta le (egy SPE-típusú kétjegyű fonológiai modellben). Szigetvári (1998) a /h/, /j/ és a /v/ problematikájával foglalkozott. Ritter (2000) kormányzásfonológiai elméletben elemezte a magyar zöngésségi hasonulást. Jansen (2004) funkcionális megközelítésű laboratóriumi fonológiai vizsgálata főként a résztvevő mássalhangzók feszességét érintette. A magyar zöngésségi hasonulást akusztikai vizsgálatokkal elemző ta-

nulmányok megegyeznek abban, hogy a folyamat eredményét a tökéletesen zöngés és a teljesen zöngétlen végpontok közötti átmenetnek tekintik. Jansen és Toft (2002) vizsgálata azt az eredményt hozta, hogy a magyar zöngésségi hasonulás eredménye fokozati jellegű. Gow és Im 2004-es kutatásában azt találta, hogy a magyar zöngésségi hasonulás eredményeként megjelenő részhangok a zöngés és zöngétlen spiránsok közötti köztes realizációt mutatnak. Bárkányi és Kiss (2006) a magyar *v* képzésének aerodinamikáját és a fonológiai szembenállás különböző kontextusbeli észlelhetőségét elemezte.

A korábbi kutatások többségét olvasott anyagon (izolált szavak, szókapcsolatok) végezték. Spontán beszédben Gósy (1999, 2002) végzett elemzéseket. Tanulmányaiban a szóhatáron átívelő zöngésségi hasonulást elemezte a szavak között megjelen(het)ő szünet időtartamának függvényében, hogy a mentális lexikon aktivációs folyamatairól nyerhessen képet. Eredményei azt mutatták, hogy a szünetidőtartamok jól elkülöníthetően három szakaszra bonthatóak a zöngésségi hasonulás bekövetkezésének tekintetében. 55 ms-nál nem hosszabb szünet esetében minden alkalommal végbement a folyamat, 314 ms felett azonban egyszer sem. Az 55–314 ms közötti időtartamban megjelent szünetek esetében a hasonulás vagy bekövetkezett, vagy nem.

A magyar zöngésségi hasonulás fonológiai leírásai és akusztikai elemzései között ellentmondás fedezhető fel. Az elméleti munkák (kivéve Vago 1980) kizárólag regresszív, kötelező és kategorikus jelenségnek tekintik azt, amennyiben nem jelenik meg szünet a részt vevő mássalhangzók között. Az akusztikai elemzések azonban ezeknek a leírásoknak az egyes pontjaival ellenkező eredményekre jutottak (l. fentebb).

Saját korábbi kutatásunkban (Bóna–Grácz–Markó 2008) három beszédstílust elemeztünk a zöngésségi hasonulás szempontjából: a spontán és az irányított spontán vagy interpretatív/félspontán beszédet, illetve a felolvasást. 10 magyar anyanyelvű beszélő vett részt a kísérletben, a következő feladatokat kapták. Egy 13 mondatból álló ismeretterjesztő szöveget és 14 önálló mondatot kellett felolvasniuk. A másik feladatban két meghallgatott szöveg tartalmát kellett összefoglalniuk a résztvevőknek. A harmadik feladatban különböző témákról kellett beszélniük az adatközlőknek. A kísérletben azokat a szó belseji és szóhatáron megjelenő kétagú mássalhangzó-kapcsolatokat elemeztük, amelyekben a résztvevő konzonánsok zöngés és zöngétlen obstruensek voltak (bármely sorrendben), és nem jelent meg közöttük szünet. Az összesen 1190 kettős mássalhangzó-kapcsolat 88,5%-ában a szabályoknak megfelelően realizálódott a zöngésségi hasonulás. A további 11,5%-ot négy fő csoportba soroltuk:

1. A zöngésségi hasonulás nem ment végbe (3,1%).
2. A zöngésségi hasonulás progresszíven hatott (1,3%).
3. A realizálódott beszédhang(ok) legalább egyike félzöngés volt (4,5%).
4. Az első mássalhangzó nem jelent meg (így a hasonulás megtörténtéről nem tehető megállapítás) (2,6%).

A zöngésedés és a zöngétlenedés között nem volt eltérés abban a tekintetben, hogy milyen arányban ment végbe a fonológiai szabályoknak megfelelően a hasonulás (a zöngésedést 85,9%-ban, a zöngétlenedés 89,3%-ban következett be szabályosan). Jelentősebb különbségeket az olvasott és a spontán beszéd (spontán és félspontán) között találtunk.

A korábbi eredmények számos kérdést vetettek fel, amelyek megválaszolásához újabb elemzéseket végeztünk. A korábbi kutatás nem számolt a zöngés és zöngétlen obstruenseket is tartalmazó hármas mássalhangzó-kapcsolatokkal, illetve a szünetek által megszakított kéttagú konzonáncsoportokkal. Az itt következő esettanulmányok egyikében a hármas mássalhangzó-kapcsolatokat elemezzük a zöngésségi hasonulás szempontjából. Ugyancsak megvizsgáljuk azokat a kéttagú mássalhangzó-kapcsolatokat, amelyeket szünet szakított meg. A kapott eredményeket Gósy (1999) kutatásával vetjük egybe.

A korábbi kutatásunkban az irreguláris realizációk között a félzöngés megvalósulás volt a leggyakoribb. Az eredmények azt mutatták, hogy ebben a realizációs típusban jellemzőek az egyéni eltérések, vagyis megjelenésük beszélspecifikus artikulációs jellemzőket mutat. A harmadik esettanulmány ezt a realizációs típust elemzi nagyobb elemszámon. Az adatok alátámasztják, hogy a magyar zöngésségi hasonulás is fokozatosnak tekinthető.

Anyag és módszer

Kutatásunkhoz a BEA adatbázisból választottuk ki 10 beszélő (5 nő, 5 férfi; 29–64 évesek) hangfelvételét. Mindannyian magyar anyanyelvűek, korrigálatlan, problémát okozó hallási és látási nehézségük nem volt. A hanganyagból kétféle beszédtypust, spontán beszédet és felolvasást elemeztünk.

Az olvasott korpusz mindegyik beszélőtől egy 13 mondatos szöveg, illetve 25 önálló mondat (összesen 433 szó) felolvasását tartalmazta. A felolvasások időtartama 209,6–300,0 másodperc volt, így összesen mintegy 43,3 percnyi hangos olvasást elemeztünk.

A másik beszédtypus kétféle hanganyagból állt: félspontán-, illetve spontánbeszéd-felvételekből. Az első beszédhelyzetben a résztvevőknek hallás után össze kellett foglalniuk két rövid szöveg tartalmát, míg a másodikban személyes témákról (általában a munkájukról vagy a hobbijukról) beszéltek. A két beszédtypusból nyert adatokat az elemzéskor együttesen kezeltük, mivel a hasonulások szempontjából a korábbi kutatásunkban nem találtunk közöttük különbséget (vö. Bóna–Gráczki–Markó 2008). A beszédminták hossza 6,7–11,3 perc volt, és 638–1337 szóból álltak: összesen 84,3 percnyi spontánbeszéd-felvételt (10 717 szó) elemeztünk.

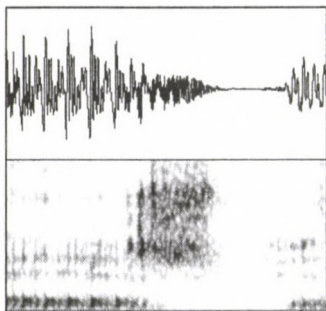
Az adatok elemzését három szempontból végeztük el:

(i) Vizsgáltuk azokat a hármas mássalhangzó-kapcsolatokat, amelyekben egymást követő helyzetben szerepelt két különböző zöngésségű obstruens. Elemeztük a CCC-kapcsolat mássalhangzóinak sorrendjét, a hangkapcsolat fonetikai pozícióját (szó belseji vagy szóhatáron átvivő), illetve a harmadik

mássalhangzó típusát. A zöngésségi hasonulás realizációit elemeztük, tekintetbe véve a harmadik mássalhangzó viselkedését is. Ebből a vizsgálatból kizártuk azokat a mássalhangzó-kapcsolatokat, amelyekben a hangok között szünetet tartott a beszélő.

(ii) Elemeztük azokat a C_1C_2 kapcsolatokat, amelyekben a két beszédhang között szünet szerepelt (akár a szó belsejében, akár szóhatárok között). Meghatároztuk a szünetek típusát (néma és/vagy jellel kitöltött) és időtartamát; majd elemeztük a zöngésségi hasonulás realizációt: szabályosan vagy szabálytalanul valósult meg; milyen fonetikai pozícióban fordult elő a hangkapcsolat; illetve mi okozhatta a szünetek megjelenését. A lehetséges magyarázat(ok)hoz megvizsgáltuk a mássalhangzó-kapcsolat szűkebb kontextusát (beleértve a megakadásjelenségeket is), illetve a beszélő artikulációs sajátosságait is. Ha a C_2 zöngétlen explozíva vagy affrikáta volt, csak akkor tekintettük szünetnek a C_1 és a C_2 közötti néma szakaszt, ha az legalább 100 ms időtartamban valósult meg. Így elkerültük azt, hogy a mássalhangzó zárszakaszát szünetnek tekintsük. Ezekben az esetekben a „szünet” utolsó 50 ezredmásodpercenyi tartamát nem számítottuk hozzá a szünettartamhoz.

(iii) A harmadik vizsgált típusba a félzöngés realizációk tartoztak. Azokat a $C_1(\#)C_2$ kapcsolatokat elemeztük, amelyekben nem tartott szünetet a beszélő a két hang között, és amelyekben a mássalhangzók zöngéssége nem volt egyértelmű. Meghatároztuk a zöngés rész arányát, majd osztályoztuk a félzöngés realizációkat. A zöngésséget szubjektív és objektív szempontok alapján határoztuk meg, oszcillogram és spektrogram segítségével: tekintetbe vettük a spektrogramon látható zöngés rész arányát, illetve a Praat által mért periodicitást. Zöngésnek tekintettük a hangot, ha az időtartamának legalább 80%-ában tartalmazott zöngét. Zöngétlennek tekintettük a hangot, ha az időtartamának kevesebb, mint 20%-ában tartalmazott zöngét. A 20–80% közötti értékek esetén a hangot félzöngésnek tekintettük (1. ábra). Nem vizsgáltuk a hangkapcsolatok helyét, illetve a hármas mássalhangzó-kapcsolatokat.



1. ábra

Az /ε3bi/ szekvencia oszcillogramja és spektrogramja (a teljes bizonyossággal szókapcsolatból)

Mindhárom esetben kizártuk a vizsgálatból azokat a hangkapcsolatokat, amelyeket érdes zöngével képzett a beszélő, vagy amelyek zajos volt (nevetés, háttérzaj hallatszott). Az adatokat a beszédstílusok függvényében elemeztük: megvizsgáltuk, hogy vannak-e, illetve milyen különbségek vannak a felolvasás és a spontán beszéd között. Az akusztikai elemzéshez a Praat 5.0 szoftvert használtuk (Boersma–Weenink 2008), a statisztikai elemzést (független párosított t -próba és egytényezős ANOVA) az SPSS 16.0 verziójával végeztük.

Eredmények

(i) Hármassálhangzó-kapcsolatok

A felolvasandó szöveg öt hármassálhangzó-kapcsolatot tartalmazott, így a tíz adatközlő beszédprodukciójában összesen 50 realizációt adatoltunk. A hangkapcsolatokban a két obstruens mellett egy szonor mássálhangzó állt (n, r, j) vagy a zöngésségi hasonulás szempontjából irreguláris v . A harmadik mássálhangzó a $C_{obstr1}C_{obstr2}$ kapcsolatot két esetben megelőzte (*szervezetünkbe, hazánkban*), három esetben pedig követte (*súlyosabb problémát, visszamaradt veგი, azt jelenti*). Az elvárt zöngésedések és zöngétlenedések, illetve a szó belseji és a szóhatáron átívelő realizációk aránya 40 : 60% volt.

A realizációk közül 13 esetben (26%) a C_{obstr1} törlődött: *szervezetünbe, hazánban*. Ezekben az esetekben nem dönthető el, hogy a zöngésségi hasonulás megtörtént-e, mivel a potenciálisan hasonló hang kiesett a hangkapcsolatból. Megjegyezzük, hogy ezek a példák az ún. mássálhangzótörlések azon alcsoportjába tartoznak, amelyet „egyértelműen a nem-gondozott beszéd jelensége”-ként tart számon a szakirodalom (vö. Ács–Siptár 1994: 566).

37 esetben (74%) a zöngésségi hasonulás szabályosan végbement. Ebből két szóban a harmadik mássálhangzó törlődött: [hoza:gbon]. További három esetben a zöngésségi hasonulás megvalósult a fonológiai szinten, de a C_{obstr2} kiesett: [ɔsɟjenti]. Egy esetben a hasonulás végbement (az eredmény a szabály szerinti zöngétlenedés lett), de nem az elvárt módon: a beszélő ejtésében a C_2 affrikálódott [ɔstʃs jelenti] az eredeti explozív helyett [ɔst jelenti].

A spontán beszédben 78 hármassálhangzó-kapcsolatot adatoltunk. Zöngésedést 12 esetben (15,4%) vártunk, míg zöngétlenedést 66 esetben (84,6%). Szóhatáron 57 hangkapcsolat (73,1%), szó belsejében 21 darab (26,9%) állt.

A spontán beszédben adatolt hármassálhangzó-kapcsolatokat két fő csoportba oszthatjuk. Az első típusba tartoznak azok, amelyekben a harmadik mássálhangzó nem befolyásolja a zöngésségi hasonulás létrejöttét, mivel szonoráns vagy v . A második típusban mindhárom mássálhangzó obstruens (nem v), itt ismételt zöngésségi hasonulást vártunk el.

Szonor mássálhangzó (m, n, j, r) vagy v volt a harmadik hang 60 esetben (a spontán beszédben adatolt CCC-kapcsolatok 76,9%-ában). Ennek a 60 esetnek az 56,7%-ában (34 eset) a harmadik mássálhangzó megelőzte a

$C_{obstr1}C_{obstr2}$ kapcsolatot: például *majd pedig, harcban*. 43,3%-ban (26 esetben) a harmadik mássalhangzó a $C_{obstr1}C_{obstr2}$ kapcsolat után állt, például: *azt jelentette, megpróbál*.

Három esetben (5%) a C_{obstr1} törlődött, így nem lehetséges annak meghatározása, hogy a zöngésségi hasonulás megtörtént-e: például a *bankban* [bambɒn]-ként realizálódott, a *volt gyerekkoromban* [vol jɛrɛk:orombɒn]-ként. Ezek a realizációk a nem gondozott beszédre jellemzőek.

49 esetben (81,7%) a zöngésségi hasonulás szabályosan megvalósult; ezek közül 38-ban (a 49 eset 77,6%-ában) mindhárom mássalhangzó az elvártnak megfelelően realizálódott. A zöngésségi hasonulás szabályosan megtörtént, de a C_{szon} törlődött a *pénztár* [pe:stár] szóból. 9 esetben (18,4%) a C_{obstr2} törlődött, míg a C_{obstr1} zöngétlenedett: például *ezt nagyon* [es nojon], *azt megfizetni* [ɔs mekfizetni], *azt látod* [ɔs la:tod]. Még egy esetet adatoltunk, amelyben a hangkapcsolat nem az elvártnak megfelelően realizálódott, de a zöngésedés megtörtént: a *harcban* elvárt ejtése [hɔrdzɒn], de [hɔrzbɒn] ejtéssel realizálódott.

8 irreguláris realizációt (13,3%) adatoltunk ebben a típusban: például *hangszer* [hɔngsɛr], *mint dolgozó* [mint dolgozo:]. Egy esetben a C_{obstr2} törlődött anélkül, hogy a zöngésségi hasonulás megtörtént volna: *az francia* [ɔz rɒntsiʃɔ] – talán megakadási jelenség következményeként.

A hármas mássalhangzó-kapcsolatok másik fő típusába, amelyben mindhárom mássalhangzó obstruens (de nem *v*), 18 eset tartozott (ami a spontán beszédben adatolt hármas hangkapcsolatok 23,1%-a). Az ebbe a típusba tartozó hangkapcsolatok nagy része szabályosan realizálódott (16 eset, 88,9%). 7 alkalommal $C_{obstr1}C_{obstr2}C_{obstr3}$ -ként realizálódott a kapcsolat (az ún. nem gondozott beszéd ellenére): például *azt hogy* [ɔst hoj], *azt hitte* [ɔst hit:ɛ].

8 esetben a C_{obstr2} nem jelent meg a felszínen, de a zöngésségi hasonulás létrejött, például *azt fogják* [ɔs fogja:k], *azt hogy* [ɔs hoj]. Két okot feltételezhetünk a jelenség hátterében. Az egyik lehetőség az, hogy ismétlődő zöngésségi hasonulás ment végbe (a C_{obstr3} -tól kiindulva), és egy későbbi szakaszban törlődött a C_{obstr2} : /zth/ > /sth/ > [sh]. A másik lehetőség az, hogy a kiesés miatt előbb törlődött a C_{obstr2} , majd a C_{obstr3} hasonította a C_{obstr1} -t: /zth/ > /zh/ > [sh]. Egy esetben ebben a típusban még egy lépés történt: az *azt hiszem* [ɔsɪsem] formában, a *h* törlésével realizálódott, ami a kifejezés gyakoriságával magyarázható. Még egy hangkapcsolatot adatoltunk, amelyben a C_{obstr3} törlődött: *azt hogy* [ɔst oj].

A maradék két esetben is irregulárisan realizálódott a zöngésségi hasonulás. A szabály szerint a C_{obstr3} is részt vett volna a hasonulásban, de a hasonulás csak a C_{obstr1} és a C_{obstr2} között jött létre, majd a C_{obstr2} törlődött: *azt gondolom* [ɔs gondolom], *azt befejeztem* [ɔs befejestem].

(ii) Szünettel megszakított mássalhangzó-kapcsolatok

Mind az olvasott, mind a spontán beszédprodukcióban adatoltunk olyan C_1C_2 kapcsolatokat, amelyek tagjai eltérő zöngésségű obstruensek, és közöt-

tük szünet jelent meg. Az olvasott szövegben 25, a spontán beszédben 64 ilyen eset fordult elő. A jelenség hátterében nyilvánvalóan más okokat kell feltételeznünk a két beszédmódban, mivel az olvasás és a spontán beszéd eltérő beszédtervezési műveleteket igényelnek. Ezért a kétféle beszédmódban előforduló szünetes C_1C_2 -kapcsolatokat egymástól függetlenül elemezzük.

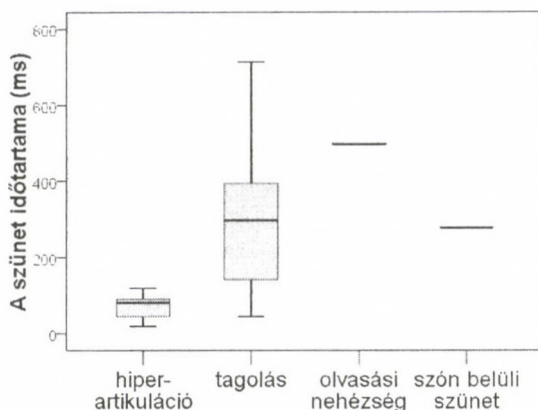
Az olvasott hanganyagban 22 alkalommal vártunk zöngésedést, és mindössze 3 alkalommal zöngétlenedést. A kettős mássalhangzó-kapcsolatokat megtörő szünetek hosszúsága 19 és 714 ms közötti, mindegyik néma szünet volt. A zöngésségi hasonulás egy kivétellel minden esetben elmaradt, az egyetlen esetben, amelyben érvényesült, a szünet 277 ms volt. Ez az eset volt az egyetlen, amikor a szünet szó belsejében jelent meg a szótő és a toldalék között (*szerekből*: [sereg | bœ:l]). A szünettartást nyelvbottlás vagy olvasási bizonytalanság okozhatta, ahol a zöngésedés a vizuális információnak köszönhetően jött létre – a felolvasó dekódolta a *b* zöngésségét, de nem volt biztos a szöveg folytatásában.

A 24-ből 15 esetben, amelyben a hasonulás nem jött létre, a szünet olyan szavak között jelent meg, amelyeket az írott szövegben vessző választott el egymástól (felsorolás tagjai voltak): *zavarokat*, *daganatos [megbetegedéseket]*; *zölkségek*, *gyümölcsök*. Ezekben a példákban a tagolás felülírta a fonológiai folyamatokat. A szünetek hossza 45 és 714 ms közötti, az átlagos szünettartam 286,3 ms, az átlagos eltérés 148,5 ms volt.

A maradék 8 esetet egyértelműen a hiperartikuláció okozta: *amelyeket gyanútlanul, gyerekek bukfencezni, kalauz szigorúan*. A hiperartikuláció legpregnansabb megjelenése az a realizáció volt, amikor a nő adatközlő svát ejtett a C_1 és a C_2 hang között: *kalauz* [szünet: 47 ms] *szigorúan*. Ebben a típusban a szünetek időtartama 19 és 120 ms között szóródott. Az átlag 71,6 ms, az átlagos eltérés 26,5 ms volt. A tagolás és a hiperartikuláció következtében megjelenő szünetek közötti különbség a 2. ábrán látható (az időtartamértékek terjedelmében van átfedés).

Egy alkalommal, amikor a *tulipánágvasok gyomlálásával* szókapcsolatban 497 ms hosszúságú szünet jelent meg, a szünettartást a 64 éves adatközlő átmeneti olvasási nehézsége okozhatta.

Az eredeti írott szövegben 48 olyan kettős mássalhangzó-kapcsolat található, amelyben zöngésségi hasonulást várnánk, 15 realizációt szóhatáron (8-at az izolált mondatokban, 7-et a szövegben). Az adatok azt mutatják, hogy a beszélők általában működtetik a fonológiai szabályt ezekben az esetekben – szünet megjelenése nélkül. Ha összevetjük a 24 elemzett hangkapcsolatot a lehetséges 150 hellyel (a tíz beszélőre vetítve), akkor azt találjuk, hogy az esetek mindössze 16%-ában realizálódott a mássalhangzó-kapcsolat szünettel. Mindazonáltal minden esetben, amikor a szünet szóhatáron állt, a zöngésségi hasonulás nem jött létre, függetlenül a szünet hosszától. Az egyes beszélőket tekintve legalább egyszer mindegyikük ejtésében előfordult ilyen hangkapcsolat, a legtöbb négy eset volt ugyanazon beszélő beszédprodukcójában.



2. ábra

A C_1C_2 kapcsolatokban megjelenő szünettartamok eloszlása a felolvasásban: szavak között (hiperartikuláció, tagolás, olvasási nehézség) és szó belsejében

A spontán beszédben adatolt 64 db C_1 [szünet] C_2 -kapcsolat mind szóhatáron állt; percenként 1,3-szer fordult elő a vizsgált hanganyagban. Ez azt jelenti, hogy minden ötödik $C_1\#C_2$ -kapcsolat esetén szünet jelent meg a szóhatáron.

A C_1 [szünet] C_2 -kapcsolatok közül 11-ben (17,2%) jött létre szabályos zöngésségi hasonulás, a többi 53 esetben (82,8%) elmaradt. A lehetséges zöngétlenítések száma jelentősen meghaladta a lehetséges zöngésítések számát. 47 hangkapcsolatban vártunk zöngétlenítést, ebből 11-ben (23,4%) valósult meg szabályosan. 17 esetben vártunk zöngésítést, de a szünet miatt egyik esetben sem jött létre a hasonulás.

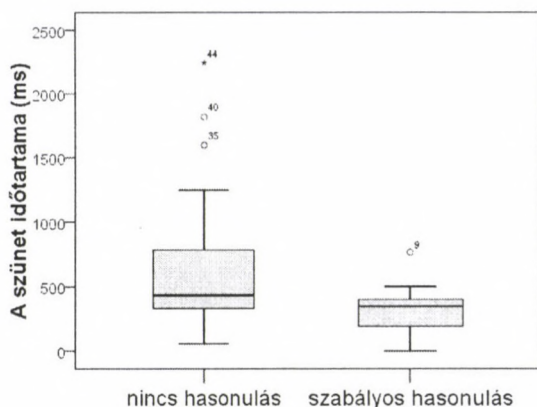
A szabályos hasonulások esetében a szünetidőtartam 105 és 767 ms között szóródott, az átlagos szünettartam 328,8 ms, az átlagos eltérés 146,9 ms volt. Több realizációval kapcsolatban felmerül a kérdés, hogy vajon a zöngétlenedés valóban a zöngésségi hasonulás szabályos megvalósulásának következménye-e. A leghosszabb szünettartamot a következő példában mértük: a beszélő a *hogy* kötőszót ismétli (*hogy ... hogy* [hoc | hoj]), miközben az ismétlések között szünetet tart, ezzel is időt nyerve a beszédtervezéshez. Az ismétlés és a szünettartás megakadásjelenségek, amelyek időt adnak a beszélőnek a megnyilatkozás folytatásának megtervezésére. Ugyanez a stratégia figyelhető meg a második leghosszabb szünetet (501 ms) tartalmazó példában is. A *hogy ... felismerik* [hoc | felismerik] szavak között kombinált szünet áll: a 285 ms tartamú néma szünetet egy 216 ms-os svászerű hezitáció követi. Amíg a néma szünetnek a beszédben számos különböző funkciója lehet (tagolás, lélegzetvételtre biztosított idő; segíti a beszédtervezési folyamatokat, az önmonitorozást stb.), addig a hezitáció a beszédtervezési folyamat diszharmonijára

utal (vö. Gósy 2005). Felmerül a kérdés, hogy ha a beszélő nem tudja, hogy hogyan folytassa a közlést, akkor miért valósítja meg szabályosan a zöngésségi hasonulást. Feltételezhető, hogy valójában nem zöngésségi hasonulás, hanem szakaszvégi/szóvégi zöngétlenítés történt az említett két (és több másik) esetben.

A szóvégi zöngétlenedést (az a folyamat, amelynek eredményeképpen a szóvégi helyzetben a zöngés mássalhangzó zöngétlenné válik) nem tekintik a magyar fonológiai szabályrendszer részének (vö. Szigetvári 1998); mindemellett egyéni artikulációs sajátosság lehet. Az adataink meglehetősen különbözőek az egyes beszélők között abban a tekintetben, hogy egyénenként változó a szünettel megszakított C_1C_2 kapcsolatok aránya, illetve ezek realizációs típusai is egyénenként eltérő eloszlásokat mutatnak. Ezért sem a zöngésségi hasonulás, sem a szóvégi zöngétlenítés nem igazolható egyértelműen minden esetben (ehhez nagyobb adatszámra volna szükség személyenként). Ugyanakkor három megállapítás fogalmazható meg. Az első az, hogy szabályos zöngésítést egyáltalán nem adatoltunk a C_1 [szünet] C_2 -kapcsolatok között. Másodszor, a 11 szabályos zöngétlenítést öt adatközlő beszédprodukciónak adatoltuk, ötöt egyetlen beszélő beszédében. Végül, a korábbi kutatásunkban 1190 db C_1C_2 -kapcsolatban (a korpuszban nem szerepeltek a C_1 [szünet] C_2 -kapcsolatok) a zöngésítés és a zöngétlenítés arányában nem volt különbség: a zöngésítés 85,8%-ban, a zöngétlenítés 89,3%-ban valósult meg a szabály szerint (Bóna–Grácz–Markó 2008). Ennélfogva a zöngésségi hasonulás valószínűleg nem magyarázza önmagában az adatok aszimmetriáját, a szó végi zöngétlenítés feltehetően a beszélők egy részének artikulációs sajátossága.

Azokban az esetekben, ahol nem jött létre a zöngésségi hasonulás (függetlenül az elvárt zöngésítéstől vagy zöngétlenítéstől), a szünetek szóródása 57 és 2245 ms közötti volt. Az átlagos szünettartam 585,2 ms, az átlagos eltérés 321,7 ms volt-nak adódott. (A megvalósult zöngésségi hasonulások esetében az átlagos szünettartam 328,8 ms, az átlagos eltérés 146,9 ms, az adatok szóródása 105–767 ms volt!) A 3. ábra a szünetidőtartamok szóródását mutatja annak függvényében, hogy megvalósult-e a hasonulás. A független kétmintás t -próba szignifikáns különbséget mutat a két csoport között: $p = 0,051$ [$t(62) = -1,991$].

Az 53-ból 34 esetben (64,2%) a beszédtervezésre volt visszavezethető a zöngésségi hasonulás elmaradása, amit a különféle megakadásjelenségek is jeleztek. 25 esetben hezitációt (21-szer néma szünettel kombinálva), 5 esetben töltelékszót, kétszer magánhangzónyújtást, egyszer újraindítást és egyszer pedig befejezetlen mondatot adatoltunk. (Természetesen a többi eset háttérében is feltételezhetünk beszédtervezési folyamatokat, de azokban nem erősíti meg ezt a feltételezést megakadásjelenség megjelenése.) Ezekben az esetekben a szünetidőtartam 146–2245 ms között szóródott, az átlag 714,7 ms, az átlagos eltérés 359,9 ms volt.



3. ábra

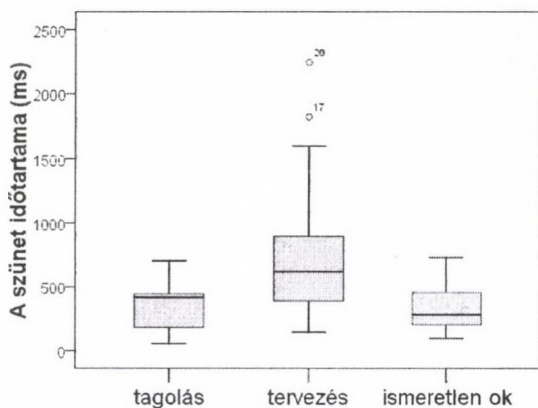
A C_1 és C_2 közötti szünettartamok szórása a spontán beszédben: szabályosan megvalósult és elmaradt zöngésségi hasonulás

További 10 alkalommal (18,9%) a tagolás okozta szünet miatt nem jött létre a zöngésségi hasonulás. Ezekben a példákban a szünetek rövidebbek voltak: a szünettartamok 57–703 ms között realizálódtak, az átlag 364,9 ms, az átlagos eltérés 156,4 ms volt.

A további 9 (17,0%) C_1 [szünet] C_2 realizáció esetében a felszíni szerkezet és a kontextus alapján nem dönthető el, milyen folyamat írhatta felül a zöngésségi hasonulást. Természetesen feltételezhető tervezési bizonytalanság ezeknek a realizációknak is a hátterében, még ha a felszínen nem is figyelhető meg a nyoma. Ugyanakkor a szünetek rövidebbek, mint azokban az esetekben, amelyek egyértelműen a tervezésre vezethetők vissza. A szünetidőtartamok ebben a típusban 98 és 730 ms közé estek, az átlagos szünettartam 341 ms, az átlagos eltérés 160 ms volt. A korábbi kutatásunkban az 1190 elemzett obstruenskapcsolatban 37 olyan realizációt találtunk (3,1%), amelyben a zöngésségi hasonulás egyáltalán nem jött létre, pedig a mássalhangzók között nem állt szünet. Ez azt jelenti, hogy a zöngésségi hasonulás elmaradását nem szükségszerűen szünettartás okozza. Azt is meg kell jegyeznünk, hogy a 9 esetből öt ugyanannak a beszélőnek a beszédprodukciónak jelent meg, aki viszonylag gyakran nem valósította meg a nem szünetes C_1C_2 kapcsolatokban sem a zöngésségi hasonulást.

A 4. ábráról leolvasható, hogy a tagolás miatt létrejött szünetek jelentősen rövidebbek, mint a beszédtervezési nehézségre visszavezethető szünetek. A dobozdiagramok szerint azok a szünetek, amelyeknek a hátterében bizonytalan okok húzódnak meg, a tagolásnál mért adatokhoz hasonlóak. Az egyté-

nyezős ANOVA szerint a csoportok között szignifikáns különbség van: $p = 0,013$ [$F(2, 50) = 4,785$].



4. ábra

A C_1 és C_2 közötti szünettartamok szórása a spontán beszédben az elmaradt zöngésségi hasonulások esetén: tervezés, tagolás, meghatározhatatlan ok

(iii) Félzöngés realizációk

328 olyan realizációt adatoltunk, amelyben legalább egy mássalhangzó félzöngéssé vált. Nagy eltérést találtunk azok között az esetek között, amikor a szabály szerint zöngésítés, illetve zöngétlenítés az elvárható. (A hármas hangkapcsolatokban és a szünetes realizációkban nem találtunk különbséget ebből a szempontból.)

Az olvasott hanganyagban 89 félzöngés realizációt adatoltunk, ennek 55,1%-ában (49 db) zöngétlenítést, 44,9%-ában (40 db) zöngésítést vártunk. Az elemzett hangkapcsolatok körülbelül azonos arányban jelentek meg szóhatáron és szó belsejében. A zöngétlenítések 55,1%-a (27 eset) és a zöngésítések 57,5%-a (23 eset) szó belsejében szerepelt; míg a maradék 44,9%-nyi (23db) zöngétlenítés és a 42,5%-nyi (17 db) zöngésítés szóhatáron állt. Az olvasásban adatolt félzöngés realizációk három főbb kategóriába, illetve további hat altípusba sorolhatók.

Az első főbb kategóriában a C_1 vált félzöngéssé; ez 37 olyan esetben (75,5%) fordult elő, amikor a szabály szerint zöngétlenítést vártunk. Ezek az esetek két altípusba sorolhatók. A C_2 zöngéssége nem változott, míg a zöngésségi hasonulás részlegesen végbement a C_1 -ben; például *ügyfeleknek* [y_hfeleknek]. Ebbe az altípusba 36 eset tartozott (97,3%-a a főcsoport elemeinek). A maradék egy realizáció (2,7%) oly módon realizálódott, hogy a C_2 zöngéssége teljesen megváltozott, miközben a C_1 félzöngéssé vált: *kalauz szigorúan* [kolozu_h zigoru:ɒn] – feltehetően az adatközlő hiperartikulációja

okozta a nem várt formát (a homorgán mássalhangzók között egy svát ejtett ahelyett, hogy hosszú hangként ejtette volna ki a kapcsolatot).

A második fő kategóriában a C_2 vált félzöngéssé; ez 8 esetben fordult elő, amikor zöngétlenítést vártunk (a 49 zöngétlenítés 16,3%-a). A realizációkat két altípusba sorolhatjuk ebben a csoportban is: 7 esetben (87,5%) a C_1 hangon létrejött a zöngésségi hasonulás, például *okozhatnak* [okosʰɔtnɔk]. A másik altípusban (egyetlen esetben, ez a főcsoport elemeinek 12,5%-a) a C_1 törlődött: *legtöbb* [lɛgɔbɪ].

A zöngétlenítések harmadik nagy csoportjába sorolt példákat is két altípusra oszthatjuk. Ezekben az esetekben nem történik sem kiesés, sem hiperartikuláció, de a realizálódó mássalhangzó(k) félzöngéssé válik(/nak). Négy ilyen esetet adatoltunk (az összes elvárt zöngétlenítés 8,2%-a). Három esetben a C_1 és a C_2 legalább két disztinktív jegyben különbözött, így az eredmény két különböző félzöngés mássalhangzó-realizáció lett, például *több héten* [tɔʰ ʰɛtɛn]. A maradék egy esetben egyetlen hosszú félzöngés mássalhangzó realizálódott: *kalauz szigorúan* [kɔlɔuzːɟigoruːɔn]. Ez a jelenség csak akkor fordulhat elő, amikor a mássalhangzó-kapcsolat obstruensei csak a zöngésségben különböznek egymástól, és a zöngésségi hasonulás szabályos megvalósulása után az elvárt alak egyetlen hosszú beszédhang.

Az olvasott szövegben elvárt 40 zöngésedés a következőképpen realizálódott. Az első fő kategóriában, 12 esetben (30%) a C_1 vált félzöngéssé. Az egyik altípusban, 3 esetben (ez a főcsoport elemeinek 25%-a) a C_2 zöngéssége nem változott, például *gyermekek bukfencezni* [jɛrmeɛg bukfɛntʰɛzni]). 9 esetben (75%) pedig C_2 zöngéssége is megváltozott, például *túlzásba* [tuːlzaːʒpɔ].

A második főcsoportba, amelyben a C_2 vált félzöngéssé, 22 esetet soroltunk (a 40 eset 55%-a). Az egyik altípusban, amelyben a C_1 hangon létrejött a zöngésségi hasonulás, 7 realizációt (a 22 eset 31,8%-a) adatoltunk, például *teljes bizonyossággal* [tɛjːɛʒ ɓizɔnoʃːaːgːɔl]. A másik altípusba pedig, amelyben a C_1 törlődött, 15 eset tartozott (a 22 eset 68,2%-a), például *boltokban* [boltɔʰɔn] vagy *zöldségek, gyümölcsök* [zɔltʃɛːgɛ ɟymɔltʃɔk].

A harmadik főcsoportba 6 eset (15%) tartozott. Ebből 4-ben két különböző képzéshelyű és képzésmódú mássalhangzó realizálódott félzöngésként, például *háztartásban* [haːstɔrtaːʒɔn]. Két másik esetben pedig hosszú félzöngés mássalhangzó jött létre, például *zavarokat daganatos* [zɔvɔroko dːɔgɔnoʃɔʃ].

A spontán beszédben 239 félzöngés esetet adatoltunk. A legtöbb ilyen megvalósulásban (209 darab, 87,4%-a az eseteknek) zöngétlenítést vártunk, míg a maradék 30 esetben (12,6%) zöngésítést. A 209 elvárt zöngétlenítésből 129 (61,7%) szóhatáron, 80 eset (38,3%) pedig szó belsejében állt. Az elvárt zöngésedések fonetikai pozíciója ellentétesen alakult: 10 esetet (a 30 elvárt zöngésedés 33,3%-a) szóhatáron, míg 20 esetet (66,7%) szó belsejében adatoltunk. Az eseteket három főcsoportba oszthatjuk aszerint, hogy melyik mássalhangzó vált félzöngéssé.

A zöngétlenítések első fő kategóriájába 169 kettős mássalhangzó-kapcsolatot (a 209 elvárt zöngétlenítés 80,9%-a) sorolhatunk, amelyben a C_1 vált félzöngéssé. A C_2 függvényében három altípust különítettünk el. 152 alkalommal (a 169 eset 89,9%-a) a C_2 zöngéssége nem változott, például *hogyha* [hoj̥ho] vagy *egy filmet* [eɟ filmet]. 16 esetben (9,5%) a C_2 törlődött, például *biztos* [biɟos]. Egyetlen esetben (0,6%) pedig a C_2 zöngéssége is megváltozott: *úgyhogy* [uɟ̥hoɟ̥].

Azokban az esetekben, amikor csak a C_2 vált félzöngéssé (22 eset, az elvárt zöngétlenítések 10,5%-a), három altípust különítettünk el. Az első altípusban (10 eset, a 22 eset 45,5%-a) a C_1 zöngéssége megváltozott, például *úgyhogy* [u:ɟ̥hoɟ̥]. Négy alkalommal (28,6%) a C_1 törlődött, például *egy kicsit* [eɟit̪it̪]. A harmadik altípusba egyetlen példa (4,5%) tartozott, ekkor a C_1 nem változott, például *megcsinálni* [megd̥ʒina:lni].

A harmadik nagy csoportba 18 (8,6%) elvárt zöngétlenítés realizációja került. Ezeket a példákat négy altípusba soroltuk. Háromszor (16,7%) egy hosszú félzöngés mássalhangzó realizálódott, amelynek képzéshelye és/vagy módja különbözött az eredeti fonémák képzési sajátosságaitól, például *nagyszerű* [noɟ̥z̥ery]. Öt alkalommal (a 18 eset 27,8%-ában) mindkét mássalhangzó realizálódott, de nem volt külön-külön elemezhető, például *biztos* [biɟ̥tos]. A harmadik altípusba azok a realizációk tartoztak, amelyekben egy hosszú félzöngés mássalhangzó realizálódott, és amelyekben az eredeti fonémák zöngésségi párok voltak, például *tudtuk* [tuɟ̥ruk] vagy *meg katona* [me:ɟ̥ ɒtono]. Ezt a típust hat alkalommal (33,3%) adatoltuk. Az utolsó altípusba 4 realizációt (22,2%) soroltunk, ebben két félzöngés mássalhangzó realizálódott, például *hogyha* [hoj̥ho].

A spontán beszédben elvárt 30 zöngésedés típusai másképp oszlottak meg. 12 esetben (a 30 elvárt zöngésedés 40%-a) csak a C_1 vált félzöngéssé. Ebből háromszor (25%-ban) a C_2 zöngéssége nem változott, például *piac, de* [pi̯ɒɟ̥ de]; kilencszer (75%-ban) pedig a C_2 zöngétlenedett, például *és boldog* [e:ɟ̥ ɒldoɟ̥].

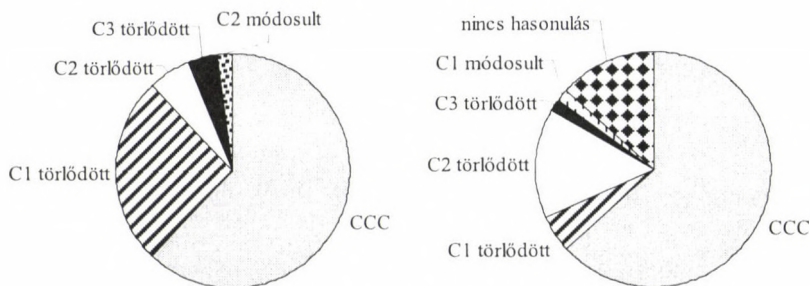
14 olyan esetet adatoltunk, amelyben a C_2 félzöngéssé vált (46,7%-a az elvárt zöngésedéseknek). Az első altípusban a C_1 zöngésedett, például *képzett* [keɟ̥zet̪]. Ebbe az altípusba 8 eset tartozott (a főcsoport elemeinek 57,1%-a). 4 alkalommal (28,6%) a C_1 törlődött, például *filmekben* [filmeɟ̥ben]. C_1 zöngéssége két esetben (14,3%) nem változott, például *török, de* [t̪or̥oɟ̥ de].

A harmadik főcsoportba, amelyben mindkét konszonáns félzöngésként realizálódott, 4 eset tartozott (az elvárt zöngésedések 13,3%-a). Három alkalommal két különböző félzöngés mássalhangzó jött létre, például *robotból* [roboɟ̥ɒ:l]. Egy alkalommal pedig egy hosszú mássalhangzó realizálódott, ebben a szomszédos fonémák zöngésségi párban álltak, például *szép bőr* [se:ɟ̥ɒr].

Az olvasott és a spontán beszéd eredményeinek összevetése

(i) A hármassalhangzó- kapcsolatok

Összesen 128 CCC-kapcsolatot adatoltunk a hanganyagban: 50 darabot a felolvasásban, 78-at a spontán beszédben. A harmadik mássalhangzó a leggyakrabban szonoráns volt vagy *v*: a felolvasásban mind az 50, a spontán beszédben pedig 60 esetben. Mindkét beszédtypusban a hármassalhangzó kapcsolatok szabályos realizációja volt a leggyakoribb (5. ábra), azaz a zöngésségi hasonulás szabályosan végbement, és mindhárom mássalhangzó megjelent a felszínen is. Mindkét beszédmódban voltak olyan realizációk, amelyekben az egyik fonéma törlődött, bár az altípusok aránya különböző. A legnagyobb különbség a két beszédstílus között az, hogy csak a spontán beszédben adatoltunk olyan esetet, amikor a zöngésségi hasonulás elmaradt. Ugyanakkor a C_1 törlődése gyakrabban fordult elő az olvasott szövegben.



5. ábra

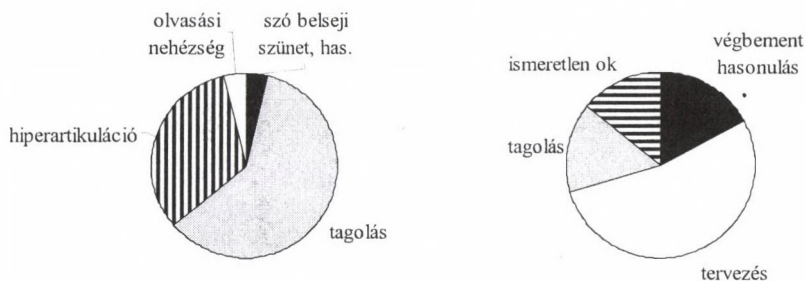
A szonoránst vagy *v*-t tartalmazó CCC-realizációk típusainak aránya felolvasásban (balra) és spontán beszédben (jobbra)

Több olyan realizációt adatoltunk az olvasott anyagban, amikor a C_{obstr1} törlődött; ezekben az esetekben az első mássalhangzó szonoráns volt. A szakirodalom szerint ez a jelenség a lezser vagy a gyors beszédre jellemző (vö. Ács–Siptár 1994), habár a beszélők szándéka alapján az olvasott anyag feltehetően gondozott beszéd volt. Azokban a beszédszakaszokban, amelyekben a mássalhangzó törlődése végbement, az artikulációs tempó értéke átlagosan 13,3 hang/s volt (a szórás 11,9–14,5 hang/s); ami megfelel az átlagos magyar köznyelvi artikulációs tempónak (Gósy 2004). Ezért ezeket az eseteket nem tekinthetjük a gyors beszéd következményeinek. Meg kell jegyeznünk azonban, hogy ezekben az esetekben a nazális kötelező érvényű képzéshely szerinti hasonulása nem a felszínen követő /b/-hez, hanem a törlődött veláris explozívához történt meg: *házánkban* [ho:za:ɳbɒn] szemben a lehetséges [ho:za:mbɒn]-nal – aminek kissé más jelentése van. A lenézelt forma megőrzi az eredeti [+ veláris] jegyet annak érdekében, hogy a percepció felismerje a szó fonoló-

giai szerkezetét. A hangos olvasást hagyományosan gondozott beszédstílusnak tartják. Mivel az adatközlők nem voltak képzett/gyakorlott beszélők/olvasók, felmerül, hogy a felolvasás standard normája közelebb kerül(t) a lezser hétköznapi beszédhez. Ez a kérdés későbbi kutatások tárgya lehet.

(ii) A szünettel megszakított mássalhangzó-kapcsolatok

A felolvasás és a spontán beszéd adatait összevetve azt figyelhetjük meg, hogy az eltérő beszédtervezési folyamatok okoztak hasonlóságokat és eltéréseket is a CC kapcsolatok ejtésében (6. és 7. ábra). A hiperartikulációra és a tagolásra visszavezethető szünetes hangkapcsolatok relatíve gyakori megjelenése az olvasásban meggyőzően alátámasztja azt, hogy ebben a beszédmódban a fonetikai tervezés és az artikulációs kivitelezés nagymértékben tudatos folyamatok. Ugyanakkor várhatóan a beszédtervezés a leggyakoribb oka a szünetes CC kapcsolatok megjelenésének a spontán beszédben. A tagolás itt is állhat a jelenség hátterében, de sokkal ritkábban, mint a felolvasásban. Fontos kiemelni, hogy a spontán beszédben sokkal gyakrabban jön létre zöngességi hasonulás a szünettartás esetén is, mint olvasáskor. Ez azt jelenti, hogy a zöngességi hasonulás meglehetősen automatikus beszédtervezési folyamat.



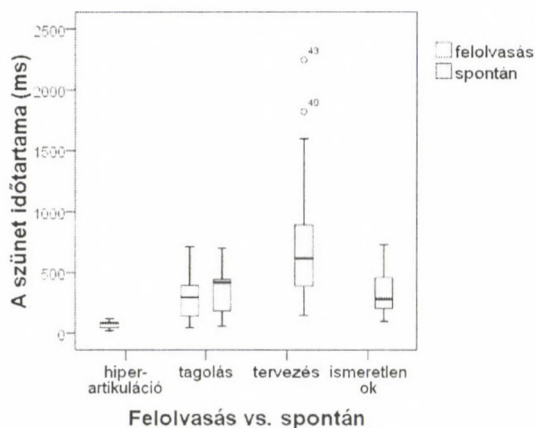
6. ábra

C_1 [szünet] C_2 -kapcsolatok realizációinak aránya az olvasásban (balra) és a spontán beszédben (jobbra)

Az elmaradt zöngességi hasonlások esetén mért szünettartamokról készült dobozdiagramok (7. ábra) azt mutatják, hogy a tagolás hasonlóan jelentkezett a felolvasásban és a spontán beszédben. A hiperartikuláció csak az olvasásban jelent meg. A beszédtervezésre visszavezethető jelenségek adatai szintén különböznek a kétféle beszédmódban. A spontán beszédben adatolt jelenségek utolsó csoportja, amelyeknek a hátterében bizonytalan okok húzódnak meg, hasonló eloszlást mutat, mint a tagolásra visszavezethető jelenségeké. Azonban amíg a tagolást a szintaktikai-szemantikai struktúra határozza meg, addig a bizonytalan háttérű realizációkat nem köthetjük grammatikai formákhoz (például *tovább* | *folyt*, *lesz* | *gépekre*). Két okot feltételezhetünk ezen

adatok háttérben: *a)* egyrészt lehetnek a felszínen nem megjelenő tervezési nehézségek; *b)* másrészt egyéni artikulációs sajátosságként jelentkezhetnek.

A spontán beszédben a szóhatáron megjelenő szünettartás 82,8%-ban okozta a zöngésségi hasonulás elmaradását. Ez az arány igen hasonlít Gósy (1999) eredményeire (77,2%). Az idézett kutatásban a szerző azt találta, hogy a szünetek időtartama befolyásolja a fonológiai kódolást: ha a szünet rövidebb, mint 55 ms, a zöngésségi hasonulás mindig létrejön, de ha a szünet hosszabb, mint 314 ms, sosem jön létre. A jelen kutatásban kapott eredmények szerint a zöngésségi hasonulás létrejöttékor mért szünetek tartama 105–767 ms között szóródott. A Gósy eredményei és a saját adataink között levő nagy különbségre a szövegi zöngétlenítés szolgálhat magyarázatul. Váratlan eredmény, hogy az olvasásban adatolt egyetlen létrejött zöngésségi hasonulás esetén 277 ms-os szünetet mértünk. Ez a realizáció azt mutatja, hogy egy nyelvbtlás okozhat relatíve hosszú szünetet a hangkapcsolatban, miközben a fonológiai kódolás az elvárt formát hívja elő a felszínen.



7. ábra

A C_1 és C_2 közötti szünettartamok szórása az elmaradt zöngésségi hasonulások esetén: az olvasás és a spontán beszéd összevetése

(iii) Félzöngés realizációk

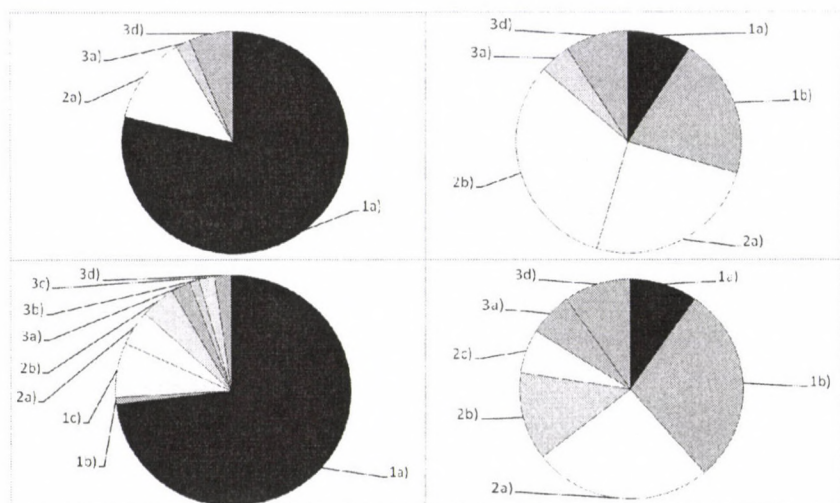
91 félzöngés realizáció jelent meg az olvasott szövegekben (26,3%), míg 255 a spontán beszédben (73,7%). Az adatok gyakoriságát a lehetséges zöngésségi hasonulási helyek arányában tekintve nem találtunk jellemző különbséget a két beszédstílus között. Az olvasások során mintegy minden 22., a spontán beszédben pedig körülbelül minden 25. vizsgált mássalhangzókapcsolatban jelent meg legalább az egyik hangzó félzöngésen. A félzöngés megvalósulások megjelenésében meghatározóbb tehát a beszédstílusnál a

zöngésségi hasonulás típusa, vagyis hogy zöngésítést vagy zöngétlenítést várhatunk a szabály alapján. Emiatt nem csak a beszédstílus, hanem a várható típus függvényében is elemezzük a félzöngés realizációkat. A félzöngés realizációk három fő csoportja hasonló megoszlást mutatott az olvasott és a spontán beszéd esetében (1. táblázat). A realizációk altípusai közül három csak a spontán beszédben jelent meg (8. ábra). Ezek azonban kevesebb mint ötször jelentek meg. A zöngésedés és zöngétlenedés között talált nagyobb eltérések az egyes beszédstílusokban a következők voltak. Azok a megvalósulások, amelyekben C_1 félzöngésen realizálódott, C_2 pedig megőrizte zöngésségét, mindkét beszédstílusban gyakoribbak voltak a szabályszerűen zöngétlenedést előidéző hasonulási helyeken (70% felett), mint ahol zöngésedést várunk (kevesebb mint 10%). Azon esetek száma, melyekben C_1 törődött, eltérő volt a két beszédstílusban. Az olvasott szövegekben gyakrabban fordult elő a zöngésedést feltételező mássalhangzó-kapcsolatokban (30% felett, míg a szabályszerűen zöngétlenítést előidéző kapcsolatokban nem jelent meg), míg a spontán beszédben gyakrabban fordult elő zöngétlenítési helyen (12,9%), mint zöngésítésin (4,9%). Azok a realizációk, melyekben C_1 az elvárt módon hasonult, azonban C_2 félzöngésen valósult meg, mindkét beszédstílusban azokban a zöngésedést feltételező mássalhangzó-kapcsolatokban voltak gyakoribbak (olvasás: 25%, illetve 12,8%; spontán: 25,8%, illetve 4,9%).

1. táblázat: A félzöngés realizációk kategóriáinak gyakorisága a beszédstílus és a zöngésségi hasonulás típusának függvényében

	Spontán beszéd		Olvasás	
	Zöngétlenedés	Zöngésedés	Zöngétlenedés	Zöngésedés
C_1 félzöngés	81,7%	38,7%	78,7%	29,5%
C_2 félzöngés	9,8%	45,2%	12,8%	56,8%
C_1C_2 félzöngés	8,5%	16,1%	8,5%	13,6%

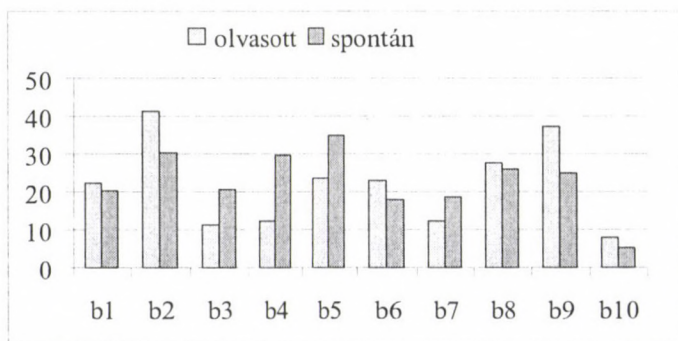
A félzöngés realizációk gyakorisága viszonylag nagymértékű beszélőn belüli és beszélők közötti eltéréseket mutatott (9. ábra). Az egyes beszélőknél megjelent félzöngés realizációk arányát (az összes lehetséges zöngésségi hasonulási helyhez viszonyítva, amelyeket két, szünet által nem megszakított obstruens alkotott) tekintve az olvasásban átlagosan 22,0%-os előfordulás volt jellemző 8,7%-os átlagos eltéréssel. A spontán beszédben hasonlóak az adatok: az átlagos gyakoriság 22,9%, az átlagos eltérés pedig 6,3%. A legnagyobb eltérést a két beszédstílusbeli realizációkat összevetve a 4-es beszélőnél tapasztaltuk. Olvasott beszédének 12,5%-ában, míg a spontán beszéd során mintegy kétszer gyakrabban, 29,7%-ban jelentek meg félzöngés realizációk. A beszélőknek mintegy felénél az olvasott beszédben, másik felénél pedig a spontán beszédben fordult elő gyakrabban félzöngés realizáció.



8. ábra

A félzöngés realizációk altípusainak előfordulási aránya a beszédstílus és a zöngésségi hasonulás típusának függvényében (fent olvasás, lent: spontán beszéd; bal oldalon: zöngétlenedés, jobb oldalon: zöngésedés)

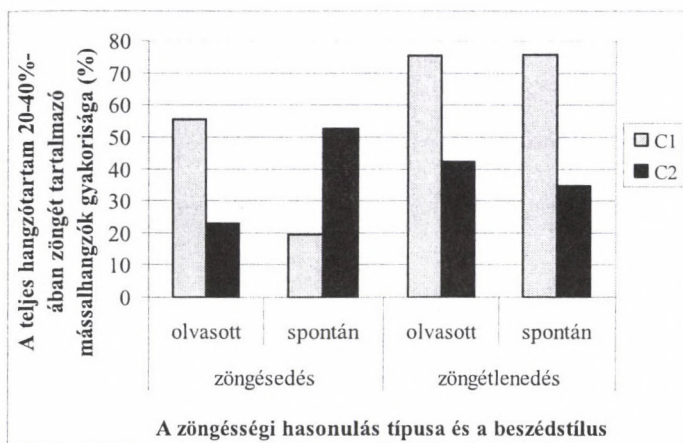
[1a) C_1 félzöngés, C_2 megőrizte a zöngésségét; 1b) C_1 félzöngés, C_2 hasonult; 1c) C_1 félzöngés, C_2 törlődött; 2a) C_1 hasonult, C_2 félzöngés; 2b) C_1 törlődött, C_2 félzöngés; 2c) C_1 megőrizte a zöngésségét, C_2 félzöngés; 3a) hosszú félzöngés C_1 ; 3b) C_3 félzöngés; 3c) C_{12} félzöngés; 3d) C_1 és C_2 is félzöngés)



9. ábra

A félzöngés realizációk gyakorisága a beszélők közötti és beszélőn belüli eltérések függvényében

A félzöngésen realizálódott hangok többsége, azaz 70,0%-a (238 a 340 előfordulásból) a teljes hangidőtartam 20–40%-ában volt zöngés. Gyakoriságuk miatt ezeket a realizációkat részletesebben elemezzük (10. ábra). Ezekben az esetekben artikulációs kényszer okozta zöngétlenedés, illetve a zöngé lecsengésének és megindulásának viszonylagos megnyúlása/késése okozhatja a részben bekövetkező hasonulás mellett. Amennyiben a C_1 zöngés részének aránya 20 és 40% között volt azokban a mássalhangzó-kapcsolatokban, amelyekben zöngésedést vártunk, feltehetően nem következett be a zöngésségi hasonulás. Olvasott beszédben ezek a realizációk mintegy 50%-ban, spontán beszédben pedig körülbelül 40%-ban jelentek meg. Amennyiben a második konszonáns zöngés részének aránya volt 20–40% között ezekben a mássalhangzó-kapcsolatokban, akkor feltételezzük, hogy a zöngésségi hasonulás progresszív módon (is) hatott. Ezek az esetek az olvasott szövegekben körülbelül 20%-ban, a spontán beszédben több mint 50%-ban fordultak elő. Azokban a mássalhangzó-kapcsolatokban, amelyekben zöngétlenedés a szabályos, feltehetően működött a zöngésségi hasonulás, ha a C_1 a tartamának 20–40%-ában zöngés. A spontán beszédben több mint 70%-ban, az olvasottban mintegy 50%-ban fordult elő ez a megvalósulás. Amennyiben zöngétlenedési hasonulást feltételez a mássalhangzó-kapcsolat, azonban a C_2 zöngés részének aránya 20–40%, feltehetően nem beszélhetünk progresszív hasonulásról.



10. ábra

Az időtartam 20–40%-ában zöngés realizációk gyakorisága a zöngésségi hasonulás típusa, a beszédstílus és a mássalhangzó kapcsolaton belüli helyének függvényében

Az eredmények összességében azt mutatják, hogy a félzöngés realizációk eltérően viselkednek számos tekintetben: mind a beszédstílus, mind a beszé-

lő, mind a hasonulás típusa, illetve egyéb fizikai tényezők is hatással vannak a zöngésségi hasonulás eredményének megvalósulására.

Következtetések

Feltételezéseink szerint a zöngésségi hasonulás realizációinak eltérése a két beszédstílusban visszavezethető a tervezési folyamatok eltérő működéseire. Olvasás során a beszélőnek nem szükséges a magasabb kognitív szintek működtetése, a beszéd során a makrotervezés, illetve a mikrotervezés néhány folyamata nem játszik szerepet. Mivel a kísérleti személyeknek volt lehetősége végigolvasni a szöveget a felolvasás előtt, feltételezhetjük, hogy az interpretáció megtörténhetett a meghangosítás előtt. Mindezek miatt a spontán beszédben gyakrabban jelentek meg nem szabályos megvalósulások. Olvasás során a tagolás írta felül több esetben a zöngésségi hasonulás szabályát, míg a spontán beszédben a beszédtervezés okozta a legtöbb szünetet a vizsgált más-salhangzók között.

Bizonyos szempontok tekintetben jelentős egyéni eltéréseket találtunk. A feltételezést, miszerint a magyarban létezik egyénfüggő szövegi zöngétlenedés, a mi eredményeink is megerősítik. A félzöngés realizációk megjelenése és altípusai változók voltak az egyes beszélők esetében. Az eredmények alátámasztják, hogy a zöngésségi hasonulás vizsgálata nemcsak fonetikai-fonológiai nézőpontból fontos, hanem hozzájárulhat a pszicholingvisztikai, szociolingvisztikai és kriminalisztikai fonetikai eredményekhez is.

A félzöngés realizációk tekintetében nagy eltéréseket találtunk a zöngésségi hasonulás zöngésítési és zöngétlenítési típusa között. Számos artikulációbeli hatás együttjárása (pl. zöngétlenedési kényszer) befolyásolhatja megjelenésük típusát, módját.

Kiinduló hipotézisünk, miszerint az olvasás a túl-, a spontán beszéd pedig az alulartikulált beszédhez áll közelebb a lindblomi H&H-elméletben, igazolódni látszik. Ugyanakkor az olvasás normája úgy tűnik, közeledik a lezser beszédéhez. Ezt támasztják alá az olvasott anyagban a kiesést mutató CCC realizációk is. (Hasonló eredményekre jutott Torstenson [2004], aki a svéd televíziós hírek felolvasásában adatolt redukciós jelenségeket. Azonban az olvasott beszéd általánosan a kanonikus formához való közelség benyomását kelti.) A spontán beszéd egyes típusai azonban mindenképpen az alulartikulált beszéd különböző fokú megjelenései.

Az adatok sokszínűsége alátámasztja, hogy a magyar zöngésségi hasonulás fokozatos, bizonyos esetekben részben progresszív folyamat, és még ha kötelező is, számos tényező felülírhatja. Ezen tényezők között kell számon tartanunk a beszédstílust is.

Irodalom

- Ács Péter – Siptár Péter 1994. Túl a gondozott beszédén. In: Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest, 550–581.
- Barry, Martin C. 1992. Palatalisation, assimilation, and gestural weakening in connected speech. *Speech Communication* 11(4–5). 393–400.
- Blaho, Sylvia 2008. *The syntax of phonology. A radically substance-free approach*. PhD dissertation. University of Tromsø.
- 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 – Gráczki, Tekla Etelka – Markó, Alexandra 2008. Coarticulation rules and speaking style dependency. In Sock, Rudolph – Fuchs, Susanne – Laprie, Yves (eds.): *International Seminar on Speech Production 2008*. Strasbourg, 245–248.
- Burton, Martha W., Robblee, Karen E. 1997. A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25. 97–114.
- Gósy, Mária 1999. Predictability of voicing assimilation in speech production. In: *Proceedings of the XIVth International Congress of Phonetic Sciences* (San Francisco) Vol. 3. 1905–1908.
- Gósy, Mária 2002. Temporal coding of voicing assimilation in speech production. *Acta Linguistica Hungarica* 49(3–4). 257–276.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gow, David W. Jr. – Im, Aaron M. 2004. A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language* 51. 279–296.
- Jansen, Wouter 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. Groningen Dissertations in Linguistics 47. Rijksuniversiteit Groningen.
- Jansen, Wouter 2007. Phonological ‘voicing’, phonetic voicing and assimilation in English. *Language Sciences* 29. 270–293.
- Jansen, Wouter – Toft, Zoe 2002. On sounds that like to be paired (after all): an acoustic investigation of Hungarian voicing assimilation. *SOAS Working Papers in Linguistics* 12. 19–52.
- Kenesei, István – Vago, Robert – Fenyvesi, Anna 1998. *Hungarian*. Routledge Descriptive Grammar. Routledge, London.
- Kiss, Zoltán – Bárkányi, Zsuzsanna 2006. A phonetically based approach to the phonology of [v] in Hungarian. *Acta Linguistica Hungarica* 53(2). 175–226.
- 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.
- Recasens, Daniel – Pallarès, Maria D. 2001. Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics* 29. 273–301.
- Ritter, Nancy A. 2000. Hungarian voicing assimilation revisited in Head-Driven Phonology. In Alberti, Gábor – Kenesei, István (eds.): *Papers from the Pécs Conference*. Approaches to Hungarian 7. JATE Press, Szeged.
- Siptár, Péter – Törkenczy, Miklós 2000. *Hungarian phonology*. Cambridge University Press, Cambridge.

- Szigetvári, Péter 1998. Voice assimilation in Hungarian: the hitches. In Varga, László (ed.): *The Even Yearbook* 3. ELTE SEAS Working Papers in Linguistics. Eötvös Loránd University, Budapest.
- Vago, Robert 1980. *The Sound Pattern of Hungarian*. Georgetown University Press, Washington, DC.

A kutatás a 78315 számú OTKA-pályázat támogatásával készült.

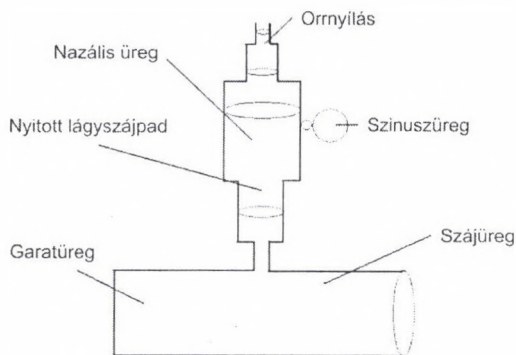
A NAZÁLIS KOARTIKULÁCIÓ VARIABILITÁSA A SPONTÁN BESZÉDBEN

Beke András – Horváth Viktória

Bevezetés

A koartikuláció a folyamatos beszéd természetes jelensége: a beszédképzés során az egyes beszédhangok nem önálló egységként realizálódnak, hanem különféle módon és mértékben módosítják egymás artikulációs konfigurációját. A hangképző szerveknek egyrészt bizonyos időre van szükségük ahhoz, hogy átálljanak az egyik hang artikulációs konfigurációjáról a másikéra; másrészt az artikulációs gazdaságosság miatt a beszédhangok hasonlóbbakká válnak a szomszédos hangokhoz. A koartikuláció tehát egyfelől fiziológiai okokra vezethető vissza, ezért univerzális jelenség, megvalósulása azonban nagymértékben nyelvspecifikus, bizonyos mértékben egyénfüggő (Gósy 2004).

A nazalizált magánhangzók realizációja a koartikulációnak egy sajátos területe, mivel képzésükkor egyszerre nyitott az orr- és a szájüreg: a lágy szájpad leereszkedik, de nem képződik zár a szájüregben, a levegő tehát folyamatosan áramlik az orr- és szájüregben keresztül (Rose 2002, Chen 1996, vö. 1. ábra).



1. ábra

Az artikulációs csatorna sematikus felépítése (Chen 1996 alapján)

A magánhangzók nazalizációjával kapcsolatban számos kutatás született különböző nyelvekre (vö. Beddor 2007, Chen et al. 2007, Delvaux et al. 2007), az eredmények azonban meglehetősen árnyalt képet mutatnak a jelenség meg-

valósulásáról. A vizsgálatok azt mindenesetre igazolták, hogy a nazalizáció hatása a magánhangzók akusztikai szerkezetére nagymértékben függ a hangkapcsolatban részt vevő beszédhangok minőségétől (Pickett 1999). A nazálisok hatásának átfogó leírását számos tényező nehezíti: a jelenség nyelvspecifikus realizációja, a magán- és mássalhangzó-minőségétől (Ladefoged 2005) és a szótag pozíciójától függő megvalósulás (Krakow 1993, Greenberg 1999, Beddor 1993), valamint a beszélők egyéni jellegzetességei.

A magyar nyelvre vonatkozó kutatások eredményei szerint a magánhangzó két nazális mássalhangzó között nazalizálódik (Földi 1989); a progresszív nazalizáció erősebb hatású, mint a regresszív (Horváth 2005, Horváth 2008); továbbá a jelenség nagymértékben függ a hangkapcsolatban részt vevő beszédhangok minőségétől. A VNC hangkapcsolatokban a nazális mássalhangzó időtartama a követő mássalhangzó képzésmódjától függ. Résmássalhangzó hatására a nazális időtartama rövidül, az előtte lévő magánhangzó időtartama – ellensúlyozva a nazális időtartamának csökkenését – nő (Gósy–Beke 2008).

A nazalizált magánhangzók akusztikai paramétereinek meghatározása a kutatási terület egyik legfontosabb kérdése. A szakirodalom azonban nem egységes a nazális mássalhangzók hatásának pontos akusztikai szerkezetével kapcsolatban, vagyis hogy milyen paraméterekkel írható le a jelenség; egy nazalizált magánhangzó milyen akusztikai jegyekben különbözik egy orális, ugyanolyan minőségű vokálistól.

Az akusztikai vizsgálatok egy része azt mutatta, hogy a nazálisok koartikulációs hatása a magánhangzó spektrumában az alsóbb tartományban a legerősebb; az alacsony frekvencián megjelenő póluson 200 Hz körül, az első formáns amplitúdójának csökkenésében, a sáv szélességében, extra csúcsok megjelenésében 1000–2500 Hz között, a spektrum kisímitásában 0–1300 Hz között mutatható ki (vö. Fant 1960, Maeda 1982, Stevens et al. 1987), és ezek a paraméterek a magánhangzó minőségétől függetlenül meghatározó akusztikai jegyei a nazalizációnak (Hawkins–Stevens 1985).

Az elmúlt évtizedek vizsgálatai alapján a magánhangzó-nazalizációt tíz jól mérhető akusztikai paraméterrel próbálták leírni, ezeket Pruthi (2007) foglalta össze. A nazalizált magánhangzók akusztikai struktúrájában a következő változások figyelhetők meg: a nazális hatására a magánhangzó első formánsának frekvenciaértéke és intenzitása csökken, a sáv szélesség szűkül, és extra pólus-zérus pár jelenik meg. Az akusztikai elemzések szerint a nazalizáció érinti a magánhangzó második és harmadik formánsának a frekvenciaértékét és amplitúdóját is. A nazalizált magánhangzók formánsszerkezetét a szájüregben kialakult formánsokon kívül a nazális formánsok, valamint a két csatorna interakciója miatt létrejövő antifórmánsok befolyásolják (Johnson 2003).

A nazalizált magánhangzók automatikus osztályozásával számos kutatás foglalkozott. Glass (1984, Glass–Zue 1985) az akusztikus paraméterek közül az 1 kHz-ig tartó tartomány tömegközéppontjával (center of gravity), annak átlagos eltéréssel és a pólus-zéró pár paramétereivel készítette el automati-

kus felismerőjét, amelyet egy 200 szavas adatbázison tanított, és azon is tesztelt. Az osztályozás eredménye 74% volt. Huffman (1990) a magánhangzó első formánsamplitúdó és a harmonikusok amplitúdóinak különbségének átlagát használta a nazalizált és nem nazalizált magánhangzók elkülönítésre. Eredményei hozzájárultak a nazalizáció percepciójának megismeréséhez. Chen (1995, 1996, 1997) három akusztikai paramétert mért ki a magánhangzók spektrumából: az első formáns amplitúdóértékét (A_1), a két formáns között megjelenő extra nazális csúcs amplitúdóértékét (P_1) és az első formáns és az alaphang között megjelenő nazális csúcs amplitúdóértékét (P_0). A mért paraméterek közül a P_0 -t a legnehezebb megállapítani, így a kísérletek többsége az A_1 -et és a P_1 -et használja. A két érték különbsége jól korrelál a nazalitás mértékével. Megállapította, hogy minél alacsonyabb az A_1 - P_1 értéke, annál nagyobb a nazális hatás.

A magánhangzók nazalizációjával kapcsolatos eddigi kutatási eredmények felolvasott hangsorok vizsgálatán alapultak mind a nemzetközi, mind a hazai szakirodalomban; ezért fontos kérdés, hogy miként valósulnak meg a nazalizált vokálisok a spontán beszédben.

A spontán beszéd során a beszédtervezés és a kivitelezés egy időben zajlik, a beszélő az adott pillanatban válogatja ki a közlésre szánt gondolatokat és tervezi meg azok nyelvi formáját; ezért a beszélők nagyobb figyelmet fordítanak a magasabb tervezési folyamatokra, mint az artikulációs kivitelezésre. Ennek következményeként a spontán beszédet kevésbé pontos artikuláció jellemzi, a beszédhangok minősége módosul (Beke 2009).

A nazalizáció megvalósulásának vizsgálata a spontán beszédben a koartikulációs folyamatok megismerésén túl számos gyakorlati jelentőséggel bír. Fontos szerepe van a mesterséges beszéd minél jobb minőségű előállításában, a beszédfelismerésben, és a beszélő személy azonosításában is, különösen az olyan nyelvekben, ahol a magánhangzó nazalizációja fonológiaiilag megkülönböztető jegy.

A jelen kutatás célja a bilabiális és az alveoláris nazális mássalhangzók hatásának vizsgálata a magánhangzók, spontán beszédben első ízben. A nazálisok hatását a spontán beszédben leggyakrabban előforduló magánhangzók akusztikai szerkezetében elemeztük.

A kutatás célja továbbá annak vizsgálata, hogy az automatikus osztályozás során milyen biztonsággal különíthetők el egymástól a nazalizált és az orális magánhangzók, illetőleg van-e különbség az egyes magánhangzók nazalizációjának mértékében, és ez hogyan jelenik meg az automatikus detektálásban.

Hipotéziseink szerint a nazálisok hatása eltérő mértékben valósul meg a magánhangzókra a koartikuláció során, vagyis a nazalizáció nagyrészt a vokális artikulációs konfigurációjának függvénye. Azt feltételeztük, hogy a nazális mássalhangzó minősége is befolyásolja a magánhangzó nazalizációját, a [m] és a [n] eltérő módon hat a vokális akusztikai szerkezetére. Korábbi nemzetközi (Hardcastle–Hewitt 1999) és hazai kutatások (Horváth 2008) ered-

ményei alapján azt feltételeztük, hogy a progresszív nazalizáció a spontán beszédben is nagyobb mértékben befolyásolja a magánhangzó módosulását, mint a regresszív.

Anyag és módszer

A nazalizált magánhangzók vizsgálatához a BEA spontánbeszéd-adatbázis felvételeit használtuk fel (Gósy 2008). Összesen 30 percnyi spontán beszédet címkéztünk fel hangkapcsolat-szinten a Praat programban. A kutatáshoz 19 beszélő (10 nő és 9 férfi) spontán beszédét használtuk fel. A női adatközlők átlagéletkora 33,4 év (20–54 év), a férfiaké 33 év (21–64 év). Mindannyian egynyelvűek, köznyelvet beszélők és ép hallók.

A jelen kutatásban a leggyakoribb magánhangzók (Gósy 2004) akusztikai szerkezetét elemeztük, ezek a következők: [ɔ], [a:], [ɛ], [i], [o].

Összesen 2236 hangkapcsolat magánhangzójának akusztikai paramétereit vizsgáltuk a következő szempontok szerint:

Van-e különbség a nazális mássalhangzót tartalmazó, illetve nem tartalmazó hangkapcsolatok magánhangzóiban?

Van-e különbség a magánhangzó spektrális szerkezetében, attól függően, hogy [m] vagy [n] mássalhangzóval együtt jelenik meg a hangkapcsolatban?

A [m] és a [n] hatása progresszív vagy regresszív irányban érvényesül erősebben?

A mássalhangzó-minőségtől függetlenül a progresszív vagy a regresszív nazalizáció eredményez nagyobb változást a vokális spektrális szerkezetében?

A statisztikai elemzéseket (ANOVA) az SPSS 13.0 programmal végeztük 95%-os szignifikanciaszinten. A csoportok közötti összehasonlításra a Tukey-féle post-hoc tesztet alkalmaztuk 95%-os szignifikanciaszinten.

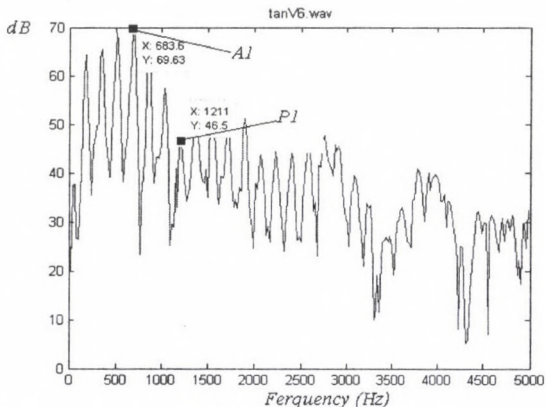
A jelen kutatásban az akusztikai paraméterek jellemzésére három különböző eljárást használtunk:

a) Formánsszerkezet. A CVC hangkapcsolatokban szereplő magánhangzók első és második formánsának értékét a Praat 5.1 szoftverben írt algoritmussal automatikusan határoztuk meg, és manuálisan ellenőriztük. Az algoritmus a magánhangzó tiszta fázisában 20 ms-os tartományban ad átlagos formánsértéket. A kiugró F_1 és F_2 értékű adatokat a vizsgálatból kihagytuk.

b) A_1 – P_1 . A magánhangzó F_1 amplitúdójának értéke és az első két formáns között megjelenő extra nazális csúcs amplitúdóértékének különbsége. Ennek mérésére MATLAB 5.0 verziójú szoftverben A_1 – P_1 mérő algoritmust hoztunk létre. Az algoritmus a magánhangzókra 512 pontos FFT analízist készít 20 ms-os Hamming-ablakolással, az A_1 -et és a P_1 -et pedig beépített csúcskereséssel találja meg, majd veszi ezek különbségét (2. ábra). Az algoritmus keresési paramétereit mindig be kell állítani a magánhangzó minőségétől függően.

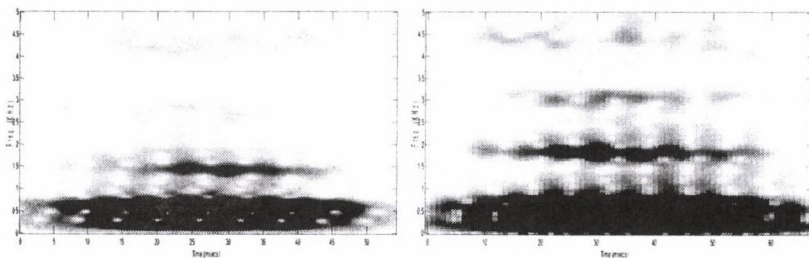
c) MFCC (Mel-frequency cepstral coefficients) és HMM (rejtett Markov-modell, HTK implementáció, vö. Vicsi-Szászák 2007). Az akusztikai fonetikai vizsgálatok általában a magánhangzó formánsstruktúrájával jellemzik a

magánhangzó akusztikai szerkezetét, noha a formánsok mérése sokszor nagyon nehéz, hosszadalmas, és mindenképpen kézi ellenőrzésre szorul. Az akusztikai paraméterek alternatívájaként MFCC-t használtunk a magánhangzók és alternánsaik automatikus osztályozásához (3. ábra), mivel ennek az akusztikai jegynek a meghatározása gyors és automatikus (Bunnell 2008).



2. ábra

Az [ɔ] nazalizált magánhangzó A_1 - P_1 értéke 512 pontos FFT analízissel készített spektrumon



3. ábra

Az [ɔ] magánhangzó spektruma nem nazális (balra) és nazális (jobbra) hangkörnyezetben

Az osztályozásra használatos tanuló algoritmusok közül (SVM, GMM, ANN) a HMM-et (rejtett Markov-modell) alkalmaztuk. A HMM tanításához a magánhangzók hullámformájának 13 MFCC jellemzőre és ezek első és második deriváltjára van szükségünk. A tanításhoz 3 állapotú monofon HMM-et használtunk, a kibocsátási valószínűséget pedig 4-8-16-32 Gauss függvényrel

leíró HMM-el vizsgáltuk. A spontán beszédben előforduló magánhangzók kétharmadán tanítottunk (1490 db), míg az egyharmadon teszteltünk (745 db).

Eredmények

Nazális hatás a magánhangzók formánsértékeiben

Az eredmények azt mutatták, hogy a nazális hatás mértékében – hipotézisünkkel ellentétben – nincs különbség a mássalhangzó-minőség függvényében: egyik magánhangzónál sem volt kimutatható, hogy a [m] vagy [n] nagyobb változást idézett elő a formánsértékekben. A spontán beszédben tehát a magánhangzók formánsszerkezetére gyakorolt nazális hatás nem függ a mássalhangzó képzési konfigurációjától.

Hipotézisünknek megfelelően a magánhangzók nazalizációja viszont meglehetősen árnyalt képet mutat a magánhangzó-minőség függvényében.

Az [a:] legszűkebb nyelvállású magánhangzó formánsstruktúrája spontán beszédben nem módosul a nazálisok környezetében: első és második formánsának értéke nem változik szignifikánsan a nazálisok hatására a nazálist nem tartozó kontextushoz képest sem a nők, sem a férfiak ejtésében (1. táblázat).

1. táblázat: Az [a:] formánsértékei nazális és nem nazális kontextusban

Kontextus	Férfiak				Nők			
	F ₁ (Hz)		F ₂ (Hz)		F ₁ (Hz)		F ₂ (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	632,16	59,67	1501,91	140,40	741,71	107,03	1841,73	165,09
Nazális	646,41	65,29	1524,96	141,06	716,05	104,11	1823,25	136,32

A férfiaknál mért formánsértékekben az a tendencia látszik, hogy az F₁ és F₂ is magasabb frekvenciaértéken realizálódik akkor, ha valamilyen nazális mássalhangzó szerepel a hangkapcsolatban, a nazálist nem tartalmazó kontextushoz képest. A nőknél fordított a tendencia: mindkét formáns értéke csökken nazális hatásra. Ezek a módosulások azonban statisztikailag nem szignifikáns mértékűek, tehát az [a:] formánsértékeiben nem mutatható ki jelentős változás a nazális mássalhangzók környezetében. Ez az eredmény eltér a felolvasásra kapott megállapításoktól: értelmetlen hangsorok meghangosításakor az [a:] F₂ és F₃ értékei a nők ejtésében nazális kontextusban eltérést mutattak a réshangkörnyezetben mért adatoktól (Horváth 2008). A különbség magyarázata valószínűsíthetően a spontán beszéd jellegzetességeiből adódik: míg felolvasáskor adott a szöveg, ezért a beszélő nagyobb figyelmet fordíthat a pontos artikulációra; spontán beszédben a magasabb tervezési folyamatok elvonják a figyelmet az artikulációs kivitelezéstől, ezért az pontatlanná, elnagyolttá válik. Ennek egyrészt az a következménye, hogy a magánhangzók, így az [a:] is redukálódik (Beke 2009), másrészt a koartikulációs folyamatok realizációja is módosul.

A veláris magánhangzók közül az [o] nazalizációja szintén korlátozott mértékben valósul meg, csak a nők ejtésében igazolható, és az első formáns értékeiben jelentkezik. (A 2. táblázatban *-gal jelöltük azokat az értékeket, amelyek között statisztikailag igazolható a különbség.)

2. táblázat: Az [o] formánsértékei nazális és nem nazális kontextusban

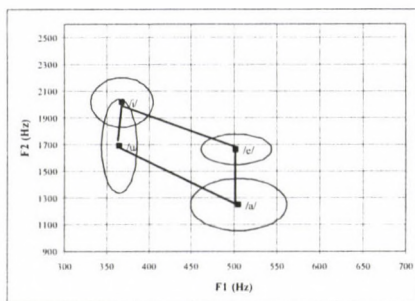
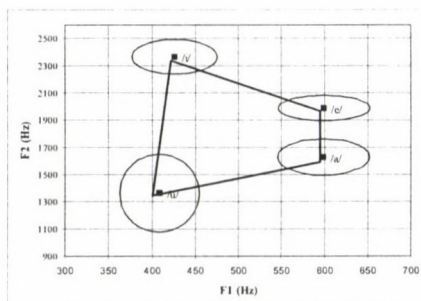
Kontextus	Férfiak				Nők			
	F ₁ (Hz)		F ₂ (Hz)		F ₁ (Hz)		F ₂ (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	467,26	50,94	1188,11	214,01	515,10*	67,94	1337,50	193,48
Nazális	461,63	46,95	1216,83	202,86	560,33*	93,88	1324,18	183,92

Az F₁ értéke a nőknél nazális kontextusban magasabb, mint nazálist nem tartozó hangkörnyezetben [egytényezős ANOVA: $F(1, 254) = 14,794$; $p < 0,001$]. Ez arra utal, hogy magánhangzó képzési konfigurációja a nyelvvállásokban módosul nazális hatásra: a nők nyitottabban ejtik a nazalizált [o]-t az orálishoz képest. Arra a kérdésre kerestük a választ, hogy ez a nazális hatás milyen irányból befolyásolja nagyobb mértékben a magánhangzó képzését. Az eredmények azt mutatták, hogy ha a mássalhangzó minőségétől függetlenül vizsgáljuk a koartikuláció irányfüggő hatását, akkor a progresszív és regresszív nazalizációban mért értékek között nincs szignifikáns különbség, ezért külön-külön is megvizsgáltuk a [m] és a [n] mássalhangzókat. A [n] környezetében nem változik nagymértékben az [o] F₁ értéke, ha a mássalhangzó megelőzi a vokálist, és akkor sem, ha követi. A bilabiális nazális esetében azonban szignifikáns különbséget találtunk a nőknél [egytényezős ANOVA: $F(4, 251) = 6,333$; $p < 0,001$]: a mássalhangzó akkor befolyásolja nagyobb mértékben az [o] F₁ értékét, ha követi a nazálist (Tukey-féle post-hoc teszt: $p = 0,031$). A hatás az F₁ értékének emelkedésében ragadható meg (vö. 3. táblázat), ami azt jelenti, hogy ha a magánhangzót [m] követi a hangkapcsolatban, akkor a női beszélők nyitottabban képzik az [o]-t. Az artikulációs szervek már a magánhangzó képzése közben felkészülnek a nazális mássalhangzóra, a légyszájpád ereszkedik és vele együtt a nyelv is, ami a vokális nyitottabb megvalósulását eredményezi. A nazálist nem tartalmazó kontextusban és a progresszív nazalizáció hatására mért F₁ átlagértékben egyáltalán nincs különbség, vagyis a nazális nem változtatja meg a követő [o] képzését – ennek megfelelően nemcsak a kétirányú nazalizáció között, hanem a nem nazális környezet és regresszív hatás között is szignifikáns a különbség ($p < 0,001$). A nők esetében a jelenség igazolhatóan szabályszerű, míg a férfiaknál nem. Ennek az lehet az oka, hogy a nők „artikulációs tere” nagyobb, mint a férfiaké (4. ábra), vagyis a magánhangzóminőségek jobban elkülönülnek egymástól, mint a férfiaknál. A 4. ábra azt mutatja, hogy a spontán beszédben mekkora a nők és a férfiak által használt artikulációs tér. Látható, hogy a férfiaké jelentősen kisebb, mint a nőké. A nőknél számottevően homogénebb

magánhangzó realizációk jelennek meg, így ezek elkülönítése is jelentősen könnyebb.

3. táblázat: Az [o] formánsértékei nem nazális és [m] környezetben

Kontextus	Férfiak				Nők			
	F ₁ (Hz)		F ₂ (Hz)		F ₁ (Hz)		F ₂ (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	467,26	50,94	1188,11	214,01	515,10*	67,94	1337,50	193,48
Progresszív nazalizáció	449,28	41,37	1129,96	73,65	515,47*	113,97	1235,30	250,10
Regresszív nazalizáció	465,64	59,28	1148,39	210,96	597,24*	78,42	1306,22	157,20



4. ábra

Az artikulációs tér az F₁ és F₂ értékek mentén egy nő (balra) és egy férfi (jobbra) spontán beszédében

A veláris magánhangzók közül csak az alsó nyelvállású [ɔ] nazalizációja igazolható mindkét nem beszélőinek ejtésében, a hatás az első formánsra jelentkezik. Az F₁ magasabb frekvenciaértéken realizálódik nazális mássalhangzó környezetében [nők: $F(1, 402) = 6,757$; $p = 0,010$, férfiak: $F(1, 212) = 4,811$; $p = 0,029$]. Ez azt jelenti, hogy a beszélők a nazalizált [ɔ]-t az eredetileg orális képzési konfigurációban alsó nyelválláshoz képest még nyitottabban képzik, vagyis a nyelv lejjebb ereszkedik a függőleges mozgást tekintve. Az [ɔ] második formánsa nem mutat nagymértékű frekvenciaérték-változást a kontextustól függően, de a tendencia az, hogy az F₂ értéke csökken a nazális környezetében (4. táblázat).

A mássalhangzó minőségétől független nazális hatás iránytól függő (progresszív vagy regresszív) mértékében itt sem volt különbség. A két nazalist külön is vizsgálva az eredmények azt mutatták, hogy a nők ejtésében a [n] eltérően befolyásolja a vokális F₂ értékét attól függően, hogy előtte vagy mögötte szerepel a hangkapcsolatban [$F(4, 399) = 2,790$; $p = 0,26$]. A progressz-

szív nazalizáció hatására nő az F_2 értéke, de csökken, ha a [n] követi a magánhangzót, és ez a csökkenés nagyobb mértékű változás (post hoc: $p = 0,031$). Ugyanez a tendencia kimutatható a férfiaknál is, de náluk a módosulás statisztikailag nem szignifikáns (5. táblázat).

4. táblázat: Az [ɔ] formánsértékei nazális és nem nazális kontextusban

Kontextus	Férfiak				Nők			
	F_1 (Hz)		F_2 (Hz)		F_1 (Hz)		F_2 (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	544,64*	62,47	1305,55	173,42	632,07**	73,34	1531,24	163,43
Nazális	564,56*	65,03	1272,04	134,73	653,06**	88,48	1515,95	168,29

5. táblázat: Az [ɔ] formánsértékei nem nazális és [n] környezetben

Kontextus	Férfiak				Nők			
	F_1 (Hz)		F_2 (Hz)		F_1 (Hz)		F_2 (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	544,64	62,47	1305,55	173,42	632,07	73,34	1531,24*	163,43
Progresszív nazalizáció	566,37	71,15	1343,75	111,55	646,18	88,24	1588,19	162,44
Regresszív nazalizáció	571,61	63,42	1286,83	164,56	655,77	78,76	1488,60*	78,42

A palatális [i] első formánsának értékében szintén szignifikáns változás következik be nazális hatásra. A nőknél és a férfiaknál egyaránt a frekvenciaérték emelkedik nazális kontextusban [egytényezős ANOVA: nők: $F(1, 166) = 21,579$; $p < 0,001$, férfiak: $F(1, 68) = 9,863$; $p = 0,003$], vagyis a nazalizált [i] képzésére alacsonyabb nyelvéllásfok jellemző. Az [i] nazalizációja is úgy valósul meg artikulációs szempontból, hogy a nyelv a lágyszájpaddal együtt mozogva szintén leereszkedik, és az orális képzésű [i]-hez képest nyitottabban valósul meg a nyelv függőleges mozgása szempontjából (6. táblázat).

6. táblázat: Az [i] formánsértékei nazális és nem nazális kontextusban

Kontextus	Férfiak				Nők			
	F_1 (Hz)		F_2 (Hz)		F_1 (Hz)		F_2 (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	336,48*	36,45	2138,45	117,44	405,83**	45,70	2333,44**	172,87
Nazális	367,93*	45,94	2121,68	117,75	445,52**	64,87	2424,10**	194,19

A nők ejtésében a nazalizált [i] második formánsa is magasabb frekvenciaértéken realizálódik az orális magánhangzóhoz képest [$F(1, 166) = 9,910$; $p = 0,002$], ez a képzési konfigurációban arra utal, hogy a nazális hatásra a nyelv előrébb tolódik a magánhangzó képzése során. Az [i] esetében a progresszív és regresszív irányú nazális hatás nem mutatott különbséget.

Az elemzett magánhangzók közül az [ε] érintett a nazalizációban a legnagyobb mértékben (7. táblázat). A nők ejtésében az első és a második formáns is szignifikáns mértékben változik a nazális hatására [egytényezős ANOVA F_1 -re: $F(1, 512) = 27,339$; $p < 0,001$; F_2 -re: $F(1, 512) = 27,991$; $p < 0,001$], a férfiaknál ez a különbség csak az első formáns értékében igazolható [egytényezős ANOVA: $F(1, 251) = 6,533$; $p = 0,011$]. A női ejtésben megvalósuló nazalizáció az [ε] F_2 értékének átlagosan 100 Hz-es csökkenését eredményezi, amely a nemzetközi szakirodalomban is fontos paramétere a koartikulációs jelenségnek (Pruthi 2007).

7. táblázat: Az [ε] formánsértékei nazális és nem nazális kontextusban

Kontextus	Férfiak				Nők			
	F_1 (Hz)		F_2 (Hz)		F_1 (Hz)		F_2 (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	526,48*	55,41	1654,05	137,42	611,13**	78,73	1954,35**	157,26
Nazális	546,88*	67,90	1639,54	134,57	647,82**	77,28	1876,58**	172,66

Az [ε] magánhangzónál igazolódott egyedül a progresszív és regresszív irányú nazális hatás különbsége is mindkét nem beszédében; a szignifikáns eltérés a magánhangzó második formánsán igazolható (8. táblázat).

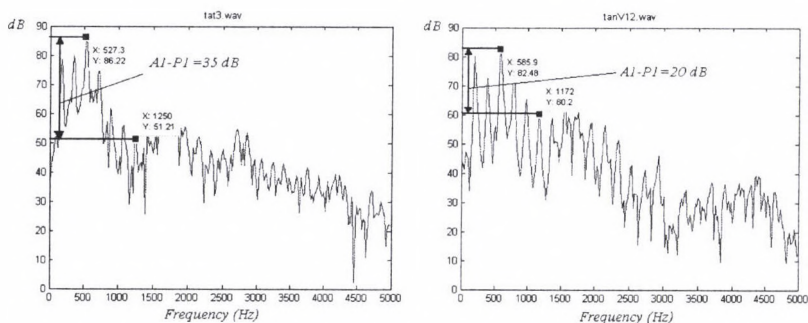
8. táblázat: Az [ε] formánsértékei nem nazális környezetben és az irányfüggő nazalizációban

Kontextus	Férfiak				Nők			
	F_1 (Hz)		F_2 (Hz)		F_1 (Hz)		F_2 (Hz)	
	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.	Átlag	Átl. elt.
Nem nazális	526,48	55,41	1654,05*	137,42	611,13	78,73	1954,35**	157,26
Progresszív nazalizáció	542,13	56,98	1688,55	126,88	654,55	81,74	1939,00	179,54
Regresszív nazalizáció	539,80	74,69	1600,10*	124,91	645,79	79,47	1847,35**	170,32

A férfiaknál a regresszív nazalizáció eredményez nagyobb változást az F_2 értékében [$F(3, 249) = 3,054$; $p = 0,029$]: ha az [ε]-t nazális követi, akkor szignifikánsan alacsonyabb frekvenciaértéken realizálódik a második formáns a progresszív irányú változáshoz képest ($p = 0,030$). A nőknél ugyanez a tendencia érvényesült, a nazális koartikuláció iránya meghatározza az [ε] második formánsának értékét [$F(3, 510) = 13,495$; $p < 0,001$]: regresszív nazalizáció hatására az F_2 szignifikánsan alacsonyabb frekvenciaértéket mutat a progresszív koartikulációhoz képest (post hoc: $p = 0,004$). A spontán beszédben a regresszív nazális hatás a magánhangzók formánsértékeiben erősebb, mint a progresszív, a tervezési folyamatok folytonossága miatt.

Nazális hatás a magánhangzók A_1 – P_1 különbségében

A nazalizált magánhangzó sajátos artikulációs konfigurációval realizálódik, mivel ez az egyetlen vokális, amelynek képzése alatt a tüdőből kiáramló levegő két úton keresztül áramlik, a száj- és orrüregben egyaránt. A nazalizált magánhangzó artikulációja közben az orrüreg irányában nyitott a vokális traktus, ezért a nazális üreg saját rezonanciája is megjelenik a spektrumban a pólus-zéró párként az átviteli függvény szerint. A számos akusztikai paraméter közül a nazalizáció a magánhangzó F_1 amplitúdójának (A_1) csökkenésében és a nazális csúcs (P_1) 1–1,5 kHz közötti megjelenésében ragadható meg. A két amplitúdóérték közötti különbség (A_1 – P_1) változása a nazalizáció egyik fontos paramétere a nemzetközi szakirodalom szerint. Az A_1 – P_1 érték a nazális környezetben kisebb lesz, mint nem nazális kontextusban, hiszen az A_1 értéke a nazalizálás során csökken, miközben az extra nazális csúcs intenzitása nő (Hawkins–Stevens 1985, Stevens et al. 1987, Pruthi 2007, vö. 5. ábra). A jelen kutatásban magyar nyelven elsőként kerestünk választ arra a kérdésre, hogy a nazális hatás hogyan jelentkezik ebben az akusztikai paraméterben.



5. ábra

A [o] magánhangzó nem nazális környezetben (balra: *tat*) és nazális környezetben (jobbra: *tan*) FFT spektruma és A_1 – P_1 értékei

Az A_1 – P_1 elemzésekor nem vettük figyelembe a női és férfi csoportokat, hiszen a formáns intenzitásában nincs különbség a két nem között. Az elemzést az [a:] és az [o] magánhangzókra nem végeztük el, mert ezeknél az automatikus detektálás során a formánsamplitúdót nem lehetett elkülöníteni a nazális csúcs amplitúdójától.

A veláris magánhangzók közül az [o] A_1 – P_1 értéke különbözik nazális és nem nazális kontextusban [ANOVA: $F(1, 722) = 4,227$; $p = 0,040$]. A nazális nem tartalmazó hangkapcsolathoz (átl.: 14,81, átl. elt.: 12,11) képest a nazális mássalhangzó környezetében az érték csökken (átl.: 12,97, átl. elt.: 10,68). Ez a tendencia egyezik a nemzetközi szakirodalomban megállapítottakkal: a magánhangzó F_1 amplitúdója csökken, ami az egyik legkonzekven-

sebb akusztikai jegy a nazális üreg jelenlétének (Huffman 1990), míg az extra nazális csúcs amplitúdója nő a légyszájpad emelkedése miatt (Pruthi 2007).

Eredményeink szerint az [ɔ] formánsértékeiben csak korlátozott mértékben lehetett igazolni a progresszív és regresszív nazalizáció különbségét. Az A_1 – P_1 értékét azonban eltérően befolyásolja a hatás az iránytól függően [$F(2, 721) = 5,117$; $p = 0,006$]. A progresszív nazalizáció szignifikánsan csökkenti a paraméter értékét ($p = 0,006$), a regresszív nazalizáció azonban nem változtatja meg; ezáltal különbség van a koartikuláció két irányának hatása között ($p = 0,039$). Ez a tendencia a mássalhangzó minőségének függvénye [$F(4, 719) = 3,876$; $p = 0,004$]. A bilabiális nazális akkor változtatja meg nagyobb mértékben az [ɔ] A_1 – P_1 értékét, ha megelőzi a vokálist ($p = 0,004$); ha követi, akkor az érték nem változik. A [m] progresszív koartikulációs hatása a változó értékének csökkenésében jelentkezik a nem nazális kontextushoz képest (nem nazális: átl.: 14,81, átl. elt.: 12,11; nazális: átl.: 8,69, átl. elt.: 10,88).

A palatális [i] első két formánsának frekvenciaértéke emelkedett nazális környezetben, de mindkét nem ejtésében csak az F_1 változása volt szignifikáns. Az A_1 – P_1 is jelentős különbséget mutat a kontextustól függően [$F(1, 420) = 21,781$; $p < 0,001$]. A nazálist nem tartalmazó hangkapcsolatokban az átlagérték 31,22 (átl. elt.: 9,57), nazális mássalhangzó hatására nagymértékben csökken (átl.: 26,90, átl. elt.: 9,01). A koartikulációs hatás előreható: a nazális akkor befolyásolja a értéket, ha megelőzi a vokálist [$F(2, 429) = 18,563$; $p < 0,001$], ha követi, akkor nem. A progresszív nazalizáció hatása ebben az esetben is függ a mássalhangzó minőségétől [$F(4, 717) = 9,275$; $p < 0,001$]; de az [i] A_1 – P_1 paraméterét nem a bilabiális nazális – mint az [ɔ] esetében, hanem az alveoláris nazális befolyásolja. A különbség az artikulációval magyarázható.

A nem nazális kontextusban mértékhez képest a progresszív nazalizációban csökken az érték (átl.: 25,34, átl. elt.: 8,95), a különbség statisztikailag szignifikáns ($p < 0,001$). Az [i] A_1 – P_1 jellemzőjét tehát eltérően befolyásolja a [n] attól függően, hogy megelőzi vagy követi a magánhangzót ($p = 0,046$).

Elemzéseink azt mutatták, hogy az [ɛ] formánsainak frekvenciaértéke változik a nazalizált ejtésben. Az A_1 – P_1 paraméter is változik nazális hatásra [$F(1, 1071) = 65,541$; $p < 0,001$], a koartikuláció az érték csökkenését eredményezi. Nem nazális környezetben az átlag 22,00 (átl. elt.: 11,97), nazális kontextusban pedig 15,17 (átl. elt.: 12,82).

Csak ennél a magánhangzónál igazolódott – és csak ebben az akusztikai paraméterben, a frekvenciaértékekben nem – az a hipotézisünk, hogy a mássalhangzó minősége is meghatározó a nazális koartikulációban: a statisztikai elemzés különbséget igazolt [m] és [n] között [$F(2, 1070) = 39,143$; $p < 0,001$]. Mindkettő környezetében csökken az A_1 és P_1 különbsége a nem nazális kontextushoz képest ($p < 0,001$), de a [m] hatására nagyobb mértékben (vö. 9. táblázat). A két nazális közül tehát a bilabiális változtatja meg jobban az A_1 – P_1 értéket ($p = 0,002$).

9. táblázat: Az $[\epsilon]$ A_1 – P_1 értéke a kontextus függvényében

Kontextus	Átlag	Átl. elt.
Nem nazális	22,00	11,97
[n]	17,68	11,86
[m]	13,40	13,53

Az $[\epsilon]$ magánhangzónál a progresszív és regresszív nazalizáció különbsége a második formáns értékében statisztikailag igazolható, az A_1 – P_1 paraméterben azonban nem volt különbség a hatás irányától függően (Stevens 2007).

A nazalizált magánhangzók automatikus osztályozása

A nazális és nem nazális magánhangzók elkülönítéséhez HMM-et alkalmaztunk. Az orális és nazalizált magánhangzók elkülönítésére két HMM modellt építettünk: OV és NV. Az OV az orális magánhangzók modellje volt, míg az NV bármilyen nazális hangkörnyezetben lévő magánhangzó modellje. Az ALL minden modellben az összes magánhangzót jelenti. A legjobb eredményt a 16 Gauss függvényvel leíró rejtett Markov-modell adta (10. táblázat).

10. táblázat: Az orális és nazalizált magánhangzók osztályozási arányai

Magánhangzók	Összes	Helyes találat (%)	Helyes találat törlés nélkül (%)	Törlés
Összes	875	75,82	82,00	203
OV	588	77,36	90,74	112
NV	287	72,84	94,71	91

Az eredmények szerint az algoritmus a magánhangzók 75,82%-át osztályozta helyesen. Ez az eredmény jobb Raja (2006) felolvasott beszédben elért eredményénél, ami 52,92%. A nazalizált magánhangzók 72,84%-át osztályozta helyesen az algoritmus, amely a spontán beszédben jó eredménynek számít a magánhangzók számos variációja miatt – az angol nyelvű spontán beszédben a nazalizált magánhangzók 48,61%-át tudták osztályoztatni MFCC 39 jellemzővel és SVM alkalmazásával (Pruthi 2007). Ha nem vesszük figyelembe a töröltek számát, akkor a 94,71%-os a helyes osztályozási aránya. Az eredmények azt mutatják, hogy a magánhangzók homogén csoportként nazalizálódnak a spontán beszédben.

A második tanítás során a nazalizált magánhangzókra két modellt építettünk a nazális mássalhangzó függvényében. A helyes találatok aránya gyengébb volt, mint az előző modellekben, ezért a továbbiakban nem vettük figyelembe a nazális mássalhangzó minőségét.

A nazális környezetben a magánhangzók a nemzetközi és hazai szakirodalom szerint nem azonos mértékben nazalizálódnak. Az alsóbb nyelvállású magánhangzók nagyobb mértékben módosulnak a koartikuláció hatására, mert a

lágyszájpadot ebben a konfigurációban könnyebb leereszteni, mint a felsőbb nyelvállásúak esetében (Johnson 2003) – ezért érdemes az egyes magánhangzó-minőségeket nyelvállásfok szerint külön HMM-ekkel modellezni. A második tanítás során a nazalizált és nem nazalizált magánhangzókra két-két modellt építettünk. A felsőbb nyelvállású magánhangzókat VF-fel jelöltünk, míg az alsóbbakat VA-val. A nazalizált magánhangzók esetében a felsőbb nyelvállásúakat NF-vel, az alsóbb nyelvállásfokúakat NA-val. Az osztályozásban a legjobb eredményt a 16 Gauss függvénnel leíró HMM adta (11. táblázat).

11. táblázat: Az orális és nazalizált magánhangzók osztályozási arányai a magánhangzók nyelvállásfoka szerint

Magánhangzók	Összes	Helyes találat (%)	Helyes találat törlés nélkül (%)	Törlés
Összes	1215	73	83,5	123
VF	153	83	90,0	13
VA	646	71	79,0	65
NF	54	48	59,0	10
NA	362	77	85,0	35

Az eredmények azt mutatták, hogy a szakirodalomnak megfelelően az alsóbb nyelvállású magánhangzók erősebben nazalizálódnak, így osztályozásuk eredményesebb.

A negyedik tanítás során a nazalizált magánhangzókat (NV) két csoportra osztottuk a nazális-magánhangzó hangkapcsolatban a nazális pozíciója szerint: progresszív nazalizáció (VN) és regresszív nazalizáció (NV). Az akusztikai elemzésekből kiindulva azt vártuk, hogy a regresszív nazalizáció jobban osztályozható. A legjobb eredményt a 16 Gauss-függvénnel leíró rejtett Markov-modell adta.

12. táblázat: Az orális és a progresszív-regresszív nazalizált magánhangzók osztályozási aránya

Magánhangzók	Összes	Helyes találat (%)	Helyes találat törlés nélkül (%)	Törlés
Összes	844	81,39	84,99	227
Orális mgh.	573	71,80	86,55	136
Progresszív naz.	96	56,47	70,00	33
Regresszív naz.	175	69,44	90,20	58

A progresszív és a regresszív nazalizáció közül a regresszív nazalizáció képzett homogénebb csoportot, ez a folyamat hasonlóan hat minden egyes magánhangzóra a nazális minőségétől függetlenül (12. táblázat). A nemzetközi szakirodalomban a felolvasott beszédben a progresszív nazalizációt tud-

ták jobban detektálni (Raja 2006). Feltételezésünk szerint ez a különbség a beszéd típusából fakad: a spontán beszédben nagyobb gondot fordítunk a beszédtervezésre, mint a kivitelezésre.

Az ötödik HMM tanítása során a nazalizált magánhangzókat tanítottuk be a magánhangzó minőségétől függően. A legjobb eredményt a 16 Gauss-függvénnyel leíró rejtett Markov-modell adta.

Az eredmények azt mutatták, hogy az [ε] magánhangzó nazalizálódik a legjobban, amit az akusztikai elemzések is alátámasztanak (13. táblázat). Hipotézisünkkel ellentétben az [i] magánhangzó osztályozása igen gyenge, aminek az lehet az oka, hogy a spontán beszédben gyakran más magánhangzóként realizálódik, így nem alkot homogén csoportot.

13. táblázat: A nazalizált magánhangzók osztályozási arányai a magánhangzó minőségétől függően

Magánhangzók	Összes	Helyes találat (%)	Helyes találat törlés nélkül (%)	Törlés
Összes	237	82,00	84,04	7
AN	56	77,77	80,00	2
EN	120	94,48	95,23	1
ÁN	17	85,00	89,47	1
ON	20	71,42	76,92	2
IN	24	57,14	58,53	1

Következtetések

A jelen kutatásban a magyar spontán beszédben elsőként vizsgáltuk a magánhangzók nazalizációjának megvalósulását, vagyis hogy a nazalizált vokálisok milyen akusztikai paraméterekben különböznek az orálisaktól. A formánsértékekre vonatkozó elemzések – hipotézisünkkel ellentétben – azt mutatták, hogy a spontán beszédben a különféle nazális mássalhangzók nem befolyásolják eltérően a magánhangzók F_1 és F_2 értékeit, vagyis a formánsokra ható nazális hatás független a mássalhangzó minőségétől. Az A_1 (első formáns amplitúdó értéke) és a P_1 (a nazális csúcs) különbségének kontextusfüggő értékére vonatkozó elemzések is megerősítették, hogy a nazális hatás nem különbözik jelentősen attól függően, hogy [m] vagy [n] szerepel a hangkapcsolatban (csak a [ε] magánhangzónál találtunk eltérést a két nazális között). Ezt a jelenséget az automatikus osztályozás is megerősítette, hiszen az osztályozási arány romlott, ha a nazalizált magánhangzókat a nazális minősége szerint csoportokra bontottuk.

Az eredmények szerint a nazalizáció sokkal inkább a magánhangzó-minőség függvénye: az egyes vokálisok eltérő mértékben módosultak nazális környezetben. Az F_1 és F_2 frekvenciaértékének szempontjából a velárisok közül az [a:] egyáltalán nem, az [ɔ] a legjobban nazalizálódik. A nazális-

vokális koartikuláció során az [ɛ] formánsértékei módosultak a legnagyobb mértékben a nem nazális környezethez képest.

Az [ɔ], [ɛ] és [i] magánhangzók F_1 értéke nazális kontextusban szignifikánsan magasabb frekvenciaértéken realizálódott a nazálist nem tartalmazó hangkapcsolatban meghatározott átlagértékhez képest, mindkét nem ejtésében (a nőknél az [o] is). Ez arra utal, hogy a spontán beszédben a magánhangzó-nazalizáció a nyelv függőleges mozgását érinti: amint a lágyszájpad leereszkedik a nazális képzéséhez, a magánhangzó artikulációjakor a nyelv is leereszkedik, tehát spontán beszédben a lágyszájpad és a nyelv egyfajta szinkron mozgással működik a gyors és kevésbé gondozott artikuláció miatt. Ezeknél a magánhangzóknál nemcsak az első formáns frekvenciaértéke változik a nazalizációban, hanem az intenzitása is. A jelen kutatásban igazoltuk az A_1 – P_1 csökkenését a nazalizált magánhangzókban, a nemzetközi szakirodalomnak megfelelően (vö. Huffman 1990, Pruthi 2007). A magánhangzó első formánsának amplitúdója (A_1) csökken, míg az extra nazális csúcs – ami az első formánsamplitúdó után jelenik meg (P_1) – amplitúdója nő (Pruthi 2007). A nazális hatás irányára vonatkozó elemzéseink árnyalt képet mutattak. A progresszív és a regresszív nazalizáció csak az [ɛ] második formánsára volt eltérő hatással, a spontán beszédben nincs lényeges különbség a koartikuláció iránya szempontjából a magánhangzók formánsértékeiben. Az [ɔ] és az [i] A_1 – P_1 értékét a progresszív nazalizáció nagyobb mértékben csökkenti, ugyanakkor a progresszív/regresszív hatás különbsége függ a mássalhangzó minőségétől. Az alsó nyelvéllású magánhangzók A_1 – P_1 paraméterében a bilabiális nazális progresszív hatása eredményezett nagyobb csökkenést, a felső nyelvéllású vokálisban az alveoláris nazális idézett elő jelentősebb változást.

A kutatás eredményei igazolták a nemzetközi szakirodalom megállapításait (Pruthi 2007): a nazális hatás a magánhangzók spektrumának alsóbb frekvenciatartományában jelentkezik a legnagyobb mértékben (az első formáns frekvenciaértékében, illetőleg az F_1 amplitúdójának és a nazális csúcs különbségében); a nazalizáció elsősorban a magánhangzó-minőség függvénye.

Az automatikus osztályozás eredményei azt mutatták, hogy a spontán beszédben a nazalizált magánhangzók akusztikailag jól elkülöníthetők az orális magánhangzóktól. Ennek eredménye 75%. Azt is megállapítottuk, hogy a progresszív és a regresszív nazalizáció közül a regresszív nazalizáció képzett homogénebb csoportot, azaz hasonlóan hat minden egyes magánhangzóra a nazális minőségétől függetlenül. A magyar spontán beszédben a koartikulációs folyamatok döntő többsége regresszív (Gósy 1998, Bóna et al. 2008) ezt a tendenciát a nazálisok hatásával kapcsolatban is igazoltuk az osztályozás során. A regresszív nazalizáció hatása a spontán beszédben nagyobb, a progresszív koartikuláció pedig a felolvasásban (Raja 2006, Horváth 2008). A további kutatások fontos kérdése, hogy a magyar spontán beszédben hogyan realizálódnak a nazalizált magánhangzók a nazálist követő mássalhangzó függvényében (vö. Chen 1996).

Az osztályozás eredményei szerint a nazális hatás megvalósulása függ a magánhangzó minőségétől, ami megegyezik a nemzetközi szakirodalomban olvasható eredményekkel. A nazalizált magánhangzók automatikus osztályozásának eredményei fontosak lehetnek a beszédszintézisben, illetve az érzelem automatikus felismerésében.

Irodalom

- Beddor, Patrice Speeter 1993. The perception of nasal vowels. In Huffman, Marie K. – Krakow, Rena A. (eds): *Nasals, nasalization and the velum*. Academic Press, San Diego, 171–196.
- Beddor, Patrice Speeter 2007. Nasals and nasalization: The relation between segmental and coarticulatory timing. In Trouvain, Jürgen – Barry, William J. (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken 6–10 August 2007*. 249–254.
- Beke András 2009. A veláris magánhangzók stabilitása a spontán beszédben. In: Gecső Tamás – Sárdi Csilla (szerk.) *A kommunikáció nyelvészeti aspektusai*. Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest (megjelenőben).
- Bóna, Judit – Grácz, Tekla Etelka – Markó, Alexandra: Coarticulation rules and speaking style dependency. In Sock, Rudolph – Fuchs, Susanne – Laprie, Yves (eds.): *Proceedings of the International Speech Production Seminar*. Strasbourg, 245–248.
- Bunel, H. Timothy – Lilley, Jason 2008. Schwa variants in American English. *Interspeech 2008*. 1959–1963.
- Chen, Marilyn Y. 1995. Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers. *Journal of the Acoustical Society of America* 98 (5). 2443–2453.
- Chen, Marilyn Y. 1996. *Acoustic correlates of nasality in speech*. PhD thesis. MIT, Cambridge.
- Chen, Marilyn Y. 1997. Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America* 102 (4). 2360–2370.
- Chen, Marilyn Y. 2000. Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics* 28. 43–67.
- Chen, Nancy F. – Slifka, Janet L. – Stevens, Kenneth N. 2007. Vowel nasalization in American English: Acoustic variability due to phonetic context. In Trouvain, Jürgen – Barry, William J. (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken 6–10 August 2007*. 905–908.
- Delvaux, Veronique – Amelot, Angélique – Rossato, Solange 2007. Nasal processes at the interface between phonetics and phonology. In Trouvain, Jürgen – Barry, William J. (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken 6–10 August 2007*. 245–248.
- Fant, Gunnar 1960. *Acoustic theory of speech production*. Mouton, The Hague.
- Földi Éva 1989. Magánhangzó-nazalizáció, hosszú mássalhangzó vagy gemináta? *Egyetemi Fonetikai Füzetek* 2. 98–120.
- Gósy Mária 1998. A zöngésségi folyamat a (spontán) beszédben. *Beszédkutatás* '98. 1–20.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.

- Gósy Mária 2008. Magyar spontán beszéd adatbázis – BEA. *Beszédkutatás* 2008. 116–128.
- Gósy, Mária – Beke, András 2008. Anticipatory coarticulation in a specific context. In Sock, Rudolph – Fuchs, Susanne – Laprie, Yves (eds.): *Proceedings of the International Speech Production Seminar*. Strasbourg, 253–256.
- Greenberg, Steven 1999. Speaking in shorthand: A syllable-centric perspective for understanding pronunciation variation. *Speech Communication* 29 (2). 159–176.
- Hardcastle, William J. – Hewlett, Nigel (eds.) 1999. *Coarticulation*. Cambridge University Press, Cambridge.
- Hawkins, Sarah – Stevens, Kenneth N. 1985. Acoustic and perceptual correlates of the nonnasal-nasal distinction for vowels. *Journal of the Acoustical Society of America* 77 (4). 1560–1575.
- Horváth Viktória 2005. A magánhangzók nazalizációjáról. *Beszédkutatás* 2005. 51–62.
- Horváth Viktória 2008. Nazális hatás a magánhangzóképzés függvényében. *Beszédkutatás* 2008. 22–32.
- Huffman, Marie K. 1990. The role of F_1 amplitude in producing nasal percepts. *Research Laboratory of Electronics* 3. MIT, Cambridge, 6–511.
- Johnson, Keith 2003. *Acoustic and auditory phonetics*. Blackwell, Cambridge–Oxford.
- Krakow, Rena A. 1993. Nonsegmental influences on velum movement patterns: Syllables, sentences, stress and speaking rate. In Huffman, Marie K. – Krakow, Rena A. (eds): *Nasals, nasalization and the velum*. Academic Press, San Diego, 87–116.
- Ladefoged, Peter 2005. *Vowels and consonants*. Blackwell Publishing, Oxford.
- Maeda, Shinji 1982. Acoustic cues for vowel nasalization: A simulation study. *Journal of the Acoustical Society of America* 72 (S1). S102.
- Malécot, Andre 1960. Vowel nasality as a distinctive feature in American English. *Language* 36. 222–229.
- Pickett, James M. 1999. *The acoustics of speech communication*. Allyn&Bacon, London.
- Pruthi, Tarun 2007. *Analysis, vocal-tract modeling and automatic detection of vowel nasalization*. PhD thesis. http://www.isr.umd.edu/Labs/SCL/publications/theses/Pruthi_PhD.pdf. (Az utolsó letöltés ideje: 2009.10.02.)
- Raja, Bilal A. 2006. Recognition of nasalized and non-nasalized vowel. http://www.ece.umd.edu/RITE/archives/merit2006/merit_fair06_papers/Paper_10_Raja.pdf. (Az utolsó letöltés ideje: 2009.07.24.)
- Rose, Philip 2002. *Forensic speaker identification*. Taylor&Francis, London–New York.
- Stevens, Kenneth N. – Fant, Gunnar – Hawkins, Sarah 1987. Some acoustical and perceptual correlates of nasal vowels. In Channon, Robert and Shockey, Linda (eds.): *In honor of Ilse Lehiste*. Foris Publications, Dordrecht, 241–254.
- Szaszák György – Vicsi Klára 2007. Speech recognition supported by prosodic information for fixed stress languages. *Proceeding of TSD conference*. Brno, 262–269.

SZÓEJTÉS ÉS SZÓÉSZLELÉS: VÁLTOZATOSSÁG ÉS ADAPTÁLÓDÁS

Gósy Mária

Bevezetés

A nyelv mentális reprezentációja az egyén nyelvi ismereteinek összessége, amely különféle tényezők következtében folyamatosan módosul. A szavak mentális reprezentációja különféleképpen határozható meg tekintettel a szemantikai sajátosságokra, a grammatikai formára, avagy a fonológiai és a fonetikai szerkezetekre. A szó sokféle meghatározása (fonológiai, lexikai, szemantikai, vö. Zwitserlood 2003, Kenesei 2007) a mentális reprezentáció definícióját is nehezíti. Feltételezik, hogy a jelentés az emberben mentálisan kódolva van (Jackendoff 2002); Libben és Jarema pedig úgy fogalmaz, hogy a mentális reprezentációk metaforák, amelyek lehetőséget adnak arra, hogy megragadjuk a lexikális ismeretek sajátosságait; az elsajátításukra, a szerveződésükre és az alkalmazásukra vonatkozó hipotéziseket tipikus nyelvhasználatban, valamint különböző nyelvi zavarokban (2002: 8). A „képzetsémák” elképzelés (1987-től) hasznosnak tűnhet, mert ennek alapján új hipotézisek alakíthatók ki arról, hogy a fogalmak, illetve fogalmi egységek hogyan szerveződnének a tudatunkban (Grady 2005). A lexéma a fogalom szemantikailag, szintaktikailag és fonológiailag meghatározott formája. A szavak feltételezteően úgy tárolódnak a tudatban mint absztrakt fonológiai reprezentációk vagy mint részletes akusztikai fonetikai reprezentációk. (E tanulmányban az egyszerűség kedvéért a *szó* terminust fogjuk használni.)

A beszélő artikulációs gesztusai, amelyekkel egy szót szándékozik kiejteni, óriási változatosságot mutatnak a különböző beszélők, de egyetlen beszélő esetében is (Rose 1999, Keating et al. 1994, Gósy 2004). Az emberek hol túlzottan jól igyekezzenek artikulálni, hol meglehetősen pongyola az ejtésük (vö. erről részletesebben: Lindblom 1987), de a vizsgálatok szerint az egyes beszélők közötti különbségek mindig nagyobbak, mint az ugyanazon beszélő különféle ejtései közötti eltérések. Az eltérő artikuláció a hangszalagok, az artikulációs csatorna anatómiájában, a beszélés egyéni sajátosságaiban, nyelvjárási és egyéb jellemzőkben érhető tetten. Mindezek a tényezők együttesen jelentkeznek a felszínen mint beszédváltozatok. Rose azt a tényt, hogy ugyanazon beszélő ismételt közlései még azonos szándék és hasonló körülmények esetén sem teljesen azonosak, fonetikai közhelynek tekinti (1999). Különböző fonetikai tényezők elbizonytalaníthatják a beszélő eredeti beszédprodukciónak, miközben az azonban továbbra is megfelel a fonológiai szó mentális repre-

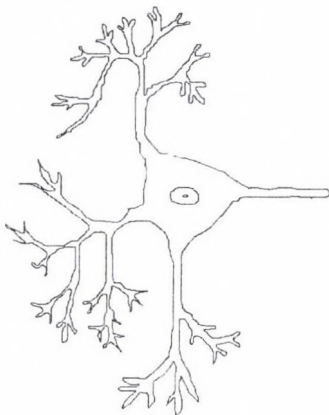
zentációjának. Az angol határozott névelő, a *the* magánhangzóját izolált mondatok felolvasása során igen nagy változatossággal realizálták a beszélők: öt-féle magánhangzót ejtettek (ugyanazon fonéma realizációiként), különböző volt a magánhangzók időtartama, továbbá változott a glottális zár, illetve a laringalizáció jelenléte (Keating et al. 1994). A szakirodalom bőven foglalkozott az artikulációs gesztusok egyéni változatosságával; ebből a szempontból elemezték a dialógusokban és a narratívákban előforduló szavakat, valamint a szóismétléseket (pl. Clark–Wasow 1998, Kohler 2000, Pluymaekers et al. 2005, Horga 2008). A meglehetősen nagy ejtési változatosság ellenére a közlések mégsem okoznak gyakori észlelési problémákat, és ezt azzal magyarázzák, hogy a nyelvi kontextus és a beszéd információs redundanciája segíti a hallgatót a biztos beszédfeldolgozásban. Mindehhez a normalizációs folyamatok is hozzájárulnak, úgy az egyes beszélő esetében, mint a különböző beszélők beszédének hallgatásakor. Ezeknek a folyamatoknak az eredménye ugyanis biztosítja az azonos mentális reprezentációk aktiválását mind a beszélő, mind a hallgató számára (Nusbaum–Magnuson 1997).

A neurolingvisztikai eredmények alapján ismeretesek bizonyos mentális reprezentációk helyei az agykéregben (pl. Scott et al. 2000, Crinion et al. 2003, Dodge–Lakoff 2005, Démonet et al. 2005), de nincs egységes nézet arról, hogy voltaképpen miként fejezhetők ki neurológiai terminusokkal a szavak mentális reprezentációi (mint neurális tevékenységek vagy aktivitások). Egy korábbi, a fonetikai szakirodalomban használatos terminus, a *neurális spektrogram* a beszédhangok akusztikai fonetikai tulajdonságai és a nekik megfelelő mentális reprezentációk kapcsolatára utalt. Bishop ezt a neurális spektrogramot mint neurális reprezentációt értelmezi, amely a beszédhang különböző frekvenciasávjainak energiaösszességéről, valamint ezek változásairól tartalmaz meghatározó információkat, és az agy ezzel a neurális reprezentációval veti össze a beszédhangot (1997: 4). Az alapkérdés azonban az, hogy a szó (amennyiben ez ténylegesen egy „szó”) mentális reprezentációja stabil jelenség-e, amely néhány invariáns jeggyel jellemezhető, avagy éppen ellenkezőleg, rendkívül rugalmas. Ha a mentális reprezentáció kellőképpen rugalmas, akkor képes alkalmazkodni a (bejövő) szó akusztikai fonetikai sajátosságaihoz, és így válik lehetővé a lexikális hozzáférés (a beszédprodukcióban), illetőleg a lexikális előhívás (a beszédmegértésben). A szakirodalomban megkülönböztetik a mentális lexikon aktiválását az aktuális feladattól függően, mivel a végbemenő folyamatok nem azonosak (Frauenfelder–Lahiri 1987, Levelt et al. 1988, Marslen-Wilson 1989 stb.). Ezért beszélünk lexikális hozzáférésről az egyik és lexikális előhívásról a másik esetben. Ez utóbbi feltételezést erősíti az is, hogy valószínűtlen lenne minden szóalakra egy önálló reprezentáció a mentális lexikonban (Greenberg 2006).

Hosszú ideje ismert tény, hogy a beszédakusztikai jel és a szó fonológiai struktúrája között nincs egy-az-egynek megfelelés. Stevens kvantális teóriája (1972) azt feltételezi, hogy a hallgató nem érzékeny a beszédelemek kis

akusztikai változásaira, ezek rejtve maradnak előtte, és így a meghatározó jegek alapján történnek a döntések az észlelési folyamatban. Ezért nem okoznak feldolgozási nehézséget például különböző beszélők ejtésvariációi ugyanazon szó esetén. Más elméletek másként oldják fel ezt a problémát. Többen a beszédjel akusztikai reprezentációja és a lexikális reprezentáció között tételeznek fel közvetlen megfelelést (pl. Andruski et al. 1994, McQueen–Cutler, 2002). Amikor a beszélő igyekszik létrehozni a szó „ideális” (vagy annak vélt) kiejtését, tulajdonképpen valamiféle prototipikus mentális reprezentációt próbál megközelíteni. Hasonlóképpen, a hallgató a szót jellemző akusztikai jelsort igyekszik a mentális reprezentációjával összevetni. Mivel a beszélő artikulációja igen változatos, különösen a spontán beszédben, a szó mentális reprezentációjának tehát nagyon rugalmasnak kell lennie.

Növekvő igény mutatkozik – főként a neurolingvisztikában, illetve a pszicholingvisztikában – a nyelvi tények neurális összefüggéseinek meghatározására, és így számos elmélet fiziológiai alátámasztására. Poeppel és Embich (2005) a fonéma disztinktiv jegyeit a dendriteknek (idegsejtnyúlvány), a szótagot a neuronnak (idegsejt), a morfémat pedig a sejtegyütteseknek felelteti meg (1. ábra).



1. ábra

A neuron sematikus rajza (sejtmag, dendrit, axon)

A feltételezés az ugyanis, hogy két önálló neuron felelős egy szó mentális reprezentációjáért. Az egyikük a fonológiai szerkezetet tartalmazza, míg a másikuk a szójelentést. Az így együttesen aktivált neuronokból jönnek létre a sejtegyüttesek. Egy szó motoros és akusztikus reprezentációi tehát nem függetlenek egymástól, hanem szoros a kapcsolatuk, és ennek alapján formálódnak a különálló funkcionális egységek (Pulvermüller 1999). A szóért felelős két neuron információja eredményezi a szó mentális reprezentációját, amit

egy speciális hologramként¹ képzelhetünk el. A hologramelmélet felhasználása az emlékezet neuronális működéseiben Karl Pribram nevéhez fűződik (vö. 1991). Ugyanannak a szónak a különböző kiejtései tehát egyfajta átfedést mutató sejtegyüttesek lehetnek. Ennek megfelelően egy vagy több sejtegyüttes ugyanazt a fonológiai, de eltérő fonetikai struktúrát tartalmaz. Ez a neurolingvisztikailag támogatott hipotézis felelős a lexikális hozzáférésért a beszédmegértésben, valamint a lexikális előhívásért a beszédprodukcióban. A lexikális hozzáférést és előhívást ugyanakkor két különböző mechanizmusnak tekintik (McQueen 2005).

Ezek az elméletek további kérdéseket vetnek fel. Van-e a beszélőnek tudatos vagy nem tudatos kontrollja a szó artikulációs gesztusai felett? Elméletileg ugyanis a beszélők csökkentik az artikulációs ráfordítási energiát a szavak kiejtésekor abban az esetben, ha azok előjelezhetők a hallgató számára (Lindblom 1990). Biztosítja-e vajon a fonológiai előfeltételezett percpációs korrekció a hangsorok és a szó mentális reprezentációja közötti megfelelést a spontán beszédben? Mennyire használják fel a hallgatók a kontextusból adódó kikövetkeztetés (előfeltevések) lehetőségét a spontán beszédben?

A szavak kiejtési sokfélesége a spontán beszédben mindenekelőtt az artikulációt megelőző beszédtervezési folyamatokkal magyarázható. Amikor a beszélők nem képesek egy közlést azonnal megfelelően megformálni, megszakíthatják a közlés folyamatosságát, szünetet vagy hezitálást iktathatnak közbe (pl.: Levelt 1983, Shriberg 2001, Fox Tree–Schrock 2002, Horga 2008, Horváth 2009). A töltelékszavak előnye, hogy – ellentétben a szünetekkel – nem szakítják meg a beszéd folyamatosságát (Clark–Wasow 1998, Gósy–Horváth 2009). A töltelékszavak elhangzásakor a hallgatónak nem tűnik fel feltétlenül, hogy a beszélőnek pillanatnyi megformálási, bizonytalansági (vagy egyéb) gondja támadt. Azt gondoljuk, hogy a beszélők artikulációs gesztusai kevésbé kontrolláltak, amikor egy töltelékszót ejtenek, mivel a beszédtervezési folyamat eközben egy másik feladatot is véghezvisz, például keresi a megfelelő szót a mentális lexikonban. Másfelől azonban az is elképzelhető, hogy egy adott szó gyakori használata éppen a stabilabb, automatizáltabb, tehát kevesebb eltérést mutató artikulációt eredményezi. Ekkor tehát a kiejtés változatosságának csökkenése következik be. A töltelékszavak gyakoriak (lehetnek) a spontán beszédben, és így lehetőséget kínálnak az artikulációs változatosság vizsgálatára.

A jelen tanulmány célja, hogy bemutassa egyetlen szó, az *akkor* akusztikai fonetikai sajátosságait függetlenül az adott funkciójától, ami e szó esetében egyfelől határozószó, másfelől töltelékszó lehet. Főként a fiatalok beszédében

1. A hologram mint metafora esetünkben azért jelent jó szemléltetést, mivel a hologram nem a kép egyes pontjait rögzíti, hanem a tárgyat letapogató lézerfény nyomait. Egy-egy pontja egyszerre tartozik a kép sok pontjához, hasonlóan a mentális lexikonban feltételezett szóelrendeződéshez. Az egyes emléknymok több idegsejtcsoportban rögzülhetnek, s ez a nyelvi „emlékekre” is alkalmazható.

hallható az *akkor* töltelékszóként. Azt feltételezzük, hogy nagy gyakorisága miatt ez a szó fonetikailag igen változatosan jelenik meg a spontán beszédben, vagyis nem tartjuk valószínűnek, hogy a gyakori használat az ejtés stabilitásához vezet. Négy fő hipotézisünk a szó akusztikai fonetikai paramétereit és percepció aspektusát érinti. 1. Az elemzett szó fonetikai sajátosságai igen rugalmas és alkalmazkodó (adaptív) mentális reprezentációt valószínűsítene. 2. Mind az egyes beszélők esetében, mind a különböző beszélők között nagy artikulációs különbségeket találunk. 3. Néhány invariáns jegy biztosítja az adott hangsor és a szándékolt fonológiai szerkezet közötti kapcsolatot. 4. A hallgatók különböző stratégiákat alkalmaznak a különböző akusztikai fonetikai struktúrájú, de azonos jelentésű szavak észlelésében.

Anyag és módszer, kísérleti személyek

1. Egyetlen szó, az *akkor* ejtéseinek **akusztikai fonetikai** szerkezetét elemeztük (nem választva szét a fentebb említett két funkciót²). A BEA (**besz**élt nyelvi **adat**bázis, vö. Gósy 2008) anyagából 10 fiatal adatközlőt választottunk ki véletlenszerűen, 5 nőt, 5 férfit. Életkoruk 22 és 28 év között szóródott. A hangfelvétel során a beszélők a munkájukról, a hobbijukról, illetve egy kurrens témáról beszélgettek, ez utóbbi illeszkedett az életkorukhoz és a mindennapi tevékenységükhöz. A teljes elemzett anyag időtartama 255 perc (4,25 óra). A női adatközlők anyaga 136 perc (2,26 óra), a férfiaké 119 perc (1,98 óra).

Az akusztikai fonetikai elemzéseket a Praat program 4.2-es verziójával (Boersma–Weenink, 2004) végeztük (44,1 kHz-es mintavételezéssel). Adatoltuk a szavak, az alkotó beszédhangok időtartamát, a [k:] zöngekezdési idejét, a két magánhangzó első két formánsát, valamint a zárhang zárfelpattanásának frekvenciáját. A szó időtartamát a kezdő magánhangzó első hangszalagregzésétől a pergőhang utolsó hangszalagregzéséig mértük; a magánhangzók időtartamát az első és az utolsó hangszalagregzés között; míg a zöngétlen zárhangét a megelőző magánhangzó utolsó hangszalagregzésétől a követő magánhangzó első hangszalagregzésének megindulásáig. Az /o/ realizáció utolsó hangszalagregzésének lecsengésétől adatoltuk a pergőhang időtartamát a hangszalagregzésének lecsengéséig. Ha az ejtés következtében a pergőhang tartalmazott egy svá hangot is, akkor az időtartamot ennek a svának az utolsó hangszalagregzéséig mértük. A zöngekezdési idő (az angol megfelelő, *voice onset time* rövidítése: VOT) mérése a szakirodalom szerinti egyik szokásos gyakorlat szerint történt: a zárhang zárfelpattanásának kezdetétől a követő magánhangzó első hangszalagregzéséig. A hangszínképek és az oscillogrammok vizuális információját hallás alapú döntésekkel egészítettük ki, és együt-

2. A jelen elemzés szempontjából ugyan nincs jelentősége, de fontos információ lehet, hogy a teljes anyagunk 11,88%-ában (34 előfordulás) az elemzett szó mint határozószó fordult elő, míg 252 esetben töltelékszóként funkcionált.

tesen vettük figyelembe. A formánsértékeket a teljes magánhangzó-időtartam közepén mértük (széles sávú spektrogram alapján). A zöngétlen zárhang zár-felpattanásának frekvenciaértékét a 4 kHz alatti legintenzívebb zörejösszetevő közepén mért értékében határoztuk meg. Összesen 11 paramétert adatoltunk minden egyes szónál, így több mint 3000 mérést végeztünk.

2. A percepciók tesztéhez két adatközlő (egy nő, egy férfi) első 24-24 *akkor* szavát használtuk fel (összesen 48-at, függetlenül attól, hogy a szóvégi mássalhangzó realizálódott-e vagy nem). Mindkét adatközlő szövegéből kiválasztottunk egy rövid virtuális mondatot, amely eredetileg is tartalmazta az elemzendő szót, és alkalmas volt arra, hogy további *akkor*-okat be tudjunk illeszteni a helyére úgy, hogy hallás alapján ne legyen érzékelhető ez a változtatás. Mind a női, mind a férfi adatközlő esetében 24 mondatot hoztunk létre, amelyek csak az *akkor* szó eltérő ejtésében különböztek egymástól. Véletlenszerű sorrendben rögzítettük ezeket úgy, hogy az egyes mondatok között 4 másodperc szünetet hagytunk. A női adatközlő mondata: *ha túlszárnyalod saját magad akkor az csak pozitívum*. A férfi tesztmondata: *ötéves koromban kezdtem el hegedülni és akkor megszerettem a hegedűt*.

Egyetemi hallgatók 8 csoportban (20–22 évesek, főként nők) vettek részt a percepciók kísérletben. Minden csoportban 10 vagy 11 hallgató volt. Négy csoport – összesen 42 kísérleti személy – a női adatközlő mondatait hallgatta, míg a másik négy csoport – összesen 40 személy – a férfi adatközlő mondatait. A feladatuk az volt, hogy ítélik meg és minősítsék a mondatokban elhangzó *akkor* kiejtését egy 5 pontos skála segítségével. A skála 1-es pontja a „nem érthető”, az 5-ös a „kiválóan érthető”, a közbülsők közül a 2-es az „alig érthető”, a 3-as a „közepesen érthető”, a 4-es pedig a „jól érthető” minősítést jelentette. A résztvevők a tesztlapra írásban is elolvasták a feladatot, de a kísérletvezető minden csoportban szóban is megfogalmazta azt. A kísérletet megelőzően „bemelegítésként” három mondatot hallottak a kísérleti személyek, amelyeket a tesztelendő anyag nem tartalmazott. A tesztlap példát mutatott a kitöltésre: a megfelelő szám bekarikázásával kellett jelölni a szubjektív ítéletet. A kísérletvezető minden csoportban felhívta a figyelmet az észlelés nehézségére, tekintettel arra, hogy egyetlen 3 perces szünettel (a teszt közepén) 24-szer ugyanazt a mondatot fogják hallani, miközben a változó ejtésű *akkor* szóra kell figyelniük. A teszt mintegy 10 percig tartott minden csoportban. A statisztikai vizsgálatokhoz varianciaanalízist (ANOVA), *t*-próbát, Pearson-féle korrelációelemzést, illetve lineáris regressziós elemzést végeztünk a szokásos, 95%-os szinten (SPSS 14.0).

Eredmények

A kísérleti anyagban 286 *akkor* szó fordult elő, 143 a nőknél és ugyancsak 143 a férfiaknál. A szóvégi pergőhangot az adatközlők anyagunkban 11 esetben (3,92%) nem ejtették (illetve a mássalhangzó nem volt egyértelműen adatolható). Az *akkor* szót megelőző szöveggörnyezet nem mutatott nagy el-

téréseket az egyes beszélők között, és nem adódott koartikulációs jelenség egyetlen esetben sem. Az *akkor* 41,25%-ban az *és* kötőszó után hangzott el, 37,06%-ban pedig néma szünet előzte meg. 13,28%-ban két további kötőszó, a *mert* és a *tehát* után következett az *akkor*. A maradék 8,41%-ban 3 magánhangzóra és 2 mássalhangzóra végződő szó után ejtették ki a beszélők a szót valamelyik funkciójában. Az *akkort* követő szövegkörnyezet már jóval változatosabb volt, a szóvégi pergőhangot különféle magán- és mássalhangzók követték, továbbá néma és kitöltött szünetek (az *akkor* után megjelenő szavak kezdő hangjai között nem volt kiemelkedő előfordulása).

1. Az *akkor* szó akusztikai fonetikai jellemzői

Az adatközlők 1,12 *akkor* szót ejtettek percenként. Az átlagos előfordulás a nőknél percenként 1,05 szó volt (a legkevesebb 19, a legtöbb 35 db), a férfiaknál 1,2 szó (a minimum 20, a maximum 38). A következő példákban az elemzett szó a kétféle – határozószói és töltelékszoói – funkcióban szerepel.

Határozószói funkcióban:

a) amikor leérettségiztem *akkor* $\square^3$ műszaki rajzolóknak mentem

b) a könyvet én a nagynénemtől kaptam és *akkor* még a nyolcadikos írással beírtam neki valami ajánlást

Töltelékszoói funkcióban:

a) és minden évben így hívnak hova is Balatonon nem Ábrahámháza igen oda és onnan kiindulunk és *akkor* így körbe így majdnem körbe a Balatonon

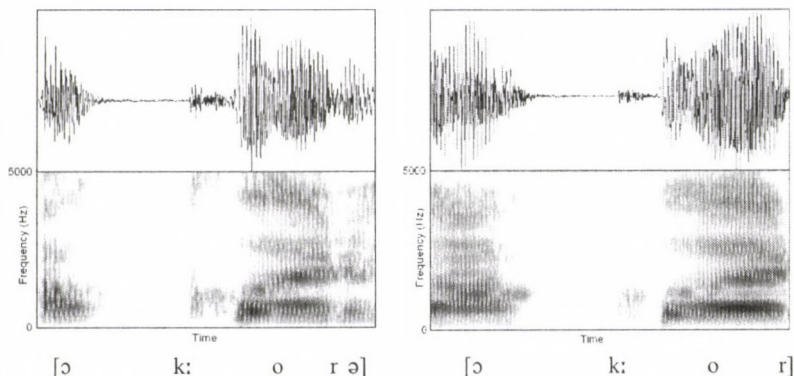
b) ott állt mellettem a srác és $\square$ így *akkor* ilyen ragasztóval így firkálta össze az üveget

A 2. ábra ugyanazon női adatközlő *akkor* ejtése alapján készült regisztrátumokat mutatja. Minden paraméter az akusztikai sajátosságok relatív nagy eltéréseit szemlélteti mindkét beszélő esetében.

Mit láthatunk az ábrákon? Vizuálisan is érzékelhetően különböznek az időviszonyok, mind a teljes szó, mind az egyes beszédhangok tekintetében; változik a zármássalhangzó zárfelpattanásának intenzitása; alapvetően eltérnek egymástól a magánhangzók formánsai, különösen a hangsúlyos [o] hangé. A pergőhang az adatközlők ejtésében részben vokalizálódik, részben látható felpattanásokat tartalmaz (ezek a nyelv perdületeit szemléltetik), esetenként megjelenik a svá is a mássalhangzó részeként.

Az *akkor* szó kiejtésének objektív időtartamai rendkívül nagy szóródást mutatnak az egyes beszélők között, a különbség statisztikailag szignifikáns [egytényezős ANOVA a szóidőtartamra: $F(9, 285) = 3,615, p = 0,001$]. Noha a női adatközlőknél kissé nagyobb időtartambeli változatosságot tapasztaltunk, mint a férfiaknál, a nemek közötti különbség matematikailag nem igazolható. A legrövidebb *akkor* a nőknél 136 ms-os, a férfiaknál 144 ms-os, a leghosszabb pedig a nőknél 580 ms-os, a férfiaknál pedig 430 ms-os volt.

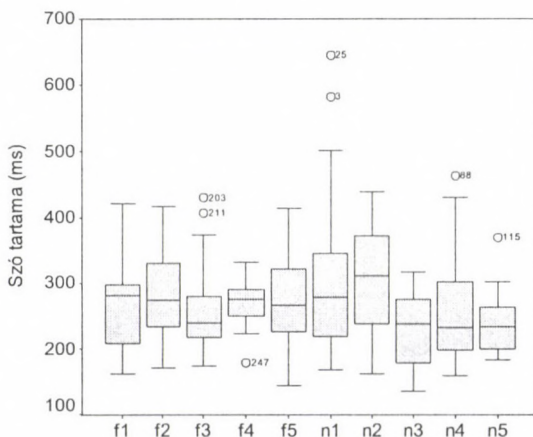
3. A négyzet jele a néma szünetet szemlélteti.



2. ábra

Két *akkor* szóról készült rezgésekép (felül) és hangszínképek (alul) ugyanazon női adatközlő ejtésében különböző szövegekörnyezetben

Kisebb változatosságot feltételeztünk a szó temporális viszonyaiban az egyes beszélők saját ejtéseit tekintve (vö. Lee et al. 1996, Dankovičová–Nolan 1999). Adataink azonban nem erősítették meg ezt a feltételezést: mind a női, mind a férfibeszélők ejtészváltozatai hasonlóan nagy különbségeket mutatnak, mint a beszélők közöttiek (3. ábra).

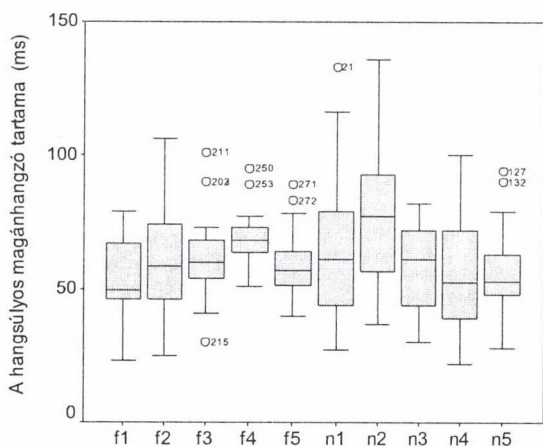


3. ábra

Az *akkor* szóidőtartamainak alakulása az egyes beszélőknél (medián és szóródás; f = férfi, n = nő)

A legnagyobb különbség két női beszélő ejtése között a teljes szó időtartamát tekintve 444 ms volt. A női beszélők saját ejtéseit elemezve, a legnagyobb eltéréseket mutatók 410 ms-os és 304 ms-os különbséggel produkálták az *akkor* szót. A férfiaknál 286 ms-nak adódott a legnagyobb különbség két beszélő szóidőtartama között, míg az egyes beszélők saját ejtése közötti legnagyobb eltérés 270 ms-os és 258 ms-os volt. Az adatok tehát azt igazolják, hogy az egyes beszélők ejtésváltozatai is nagymértékben különböznek egymástól a spontán beszédükben.

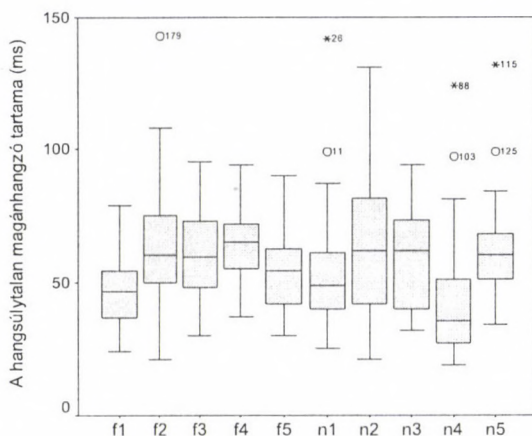
Az elemzett szó két (fonológiaiag rövid) magánhangzójának időtartamai – ha egymással összevetjük őket – szignifikánsan különböznek az egyes beszélőknél [egytényezős ANOVA: $F(1, 571) = 4,368$, $p = 0,037$]. A hangsúlyos magánhangzó tehát hosszabban realizálódik, mint a hangsúlytalan, de a mért időtartamok különbsége nem túlzottan nagy (4. és 5. ábra).



4. ábra

Az /o/ fonéma realizációinak időtartamai a női (n) és a férfi (f) adatközlőknél (medián és szóródás)

Elemeztük, hogy az /o/ és az /o/ realizációinak időtartamai mennyire térnek el a különböző beszélők ejtésében. Az adatok szerint mindkét magánhangzó időtartama szignifikáns különbséget mutat a beszélők között (az [o]-ra: $F(9, 285)=4,029$, $p=0,001$; az [o]-ra: $F(9, 285)=2,751$, $p=0,004$). Hasonlóképpen matematikailag igazolható a különbség a zöngétlen veláris zármássalhangzó időtartamában is az egyes beszélők között ($F(9, 285)=8,478$, $p=0,001$); az artikuláció tartama nagyobb változatosságot mutat a női, mint a férfi adatközlőknél (1. táblázat).



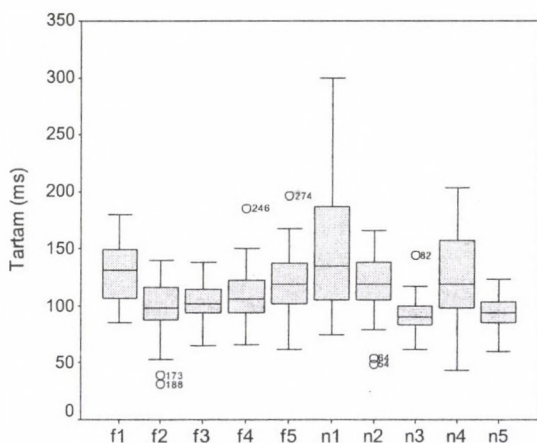
5. ábra

Az /o/ fonéma realizációinak időtartamai a női (n) és a férfi (f) adatközlőknél (medián és szóródás)

1. táblázat: Az *akkor* szó temporális jellemzőinek adatai (SD = átlagos eltérés)

akkor	Nők		Időtartamok (ms) Férfiak		Összes adatközlő	
	Átlag	SD	Átlag	SD	Átlag	SD
szó	277,52	109,38	271,37	60,44	274,45	88,26
o	64,34	25,29	61,16	15,41	62,75	20,96
o	58,90	27,53	58,77	19,26	58,84	23,72
k	119,75	46,57	110,23	26,71	114,99	38,19
VOT	47,87	16,42	34,20	11,42	41,03	13,92
r	34,53	29,11	37,87	26,52	36,20	27,81

A fonológiailag hosszú zöngétlen zárhang időtartama 39 ms és 300 ms közötti értékekkel realizálódott anyagunkban; a mássalhangzók egy részének objektív időtartama a fonológiailag rövid magánhangzók realizációjának felelt meg. A nők és a férfiak között szignifikánsan eltértek az időtartamok [egytényezős ANOVA: $F(1, 285) = 4,499, p = 0,035$]. Az előbbieket ejtésében az időviszonyok szélesebb tartományban realizálódtak. A zárhang temporális sajátosságai az egyes beszélők saját ejtéseiben is nagy eltérést mutattak (6. ábra). A legnagyobb különbség a női beszélőknél 225 ms volt, a férfi beszélőknél 115 ms volt.



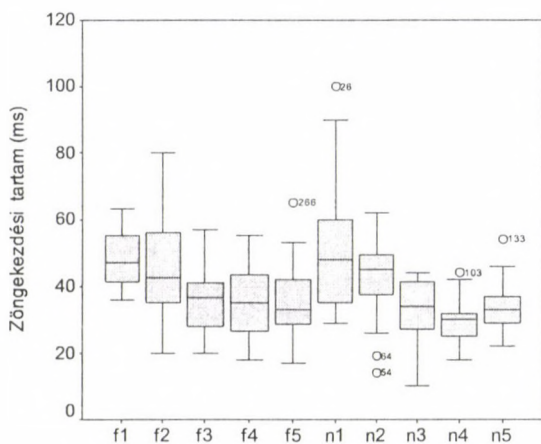
6. ábra

A zöngétlen zármássalhangzó időtartamai az összes beszélőnél
(medián és szóródás; n = nő, f = férfi)

A zöngekezdési időt a zöngétlen zárhangok invariáns jegyeként tartják számon a zöngétlenség tekintetében. A magyarban női adatközlők spontán beszédében mért [k] mássalhangzó VOT-értékének átlaga 35,31 ms-nak adódott (Gósy 2000); nincs azonban adat a fonológiailag hosszú, illetve férfiejtésű, zöngétlen veláris zármássalhangzó zöngekezdési idejére. A jelen anyagban az összes [k:] mássalhangzó VOT-jének átlaga 41,03 ms volt (7. ábra). A férfiak és a nők adatai között nincs szignifikáns különbség; az egyes beszélők között azonban igen [egytényezős ANOVA: $F(9, 285) = 10,880$, $p = 0,001$]. A pergőhang időtartama nem mutatott szignifikáns eltérést a beszélők között (1. táblázat). A svá a pergőhang részeként (vö. Vago–Gósy 2007) mindössze 9,44%-ban jelent meg, összesen 27 előfordulást adatoltunk.

A hipotézisünket igazoló eredményeket kaptunk a magánhangzók formánsszerkezetének vizsgálatakor. Az /o/ realizációi a női adatközlők ejtésében jellegzetes eltéréseket mutattak, az első két formáns frekvenciaértékében szignifikáns különbséget találtunk az egyes beszélők között [egytényezős ANOVA az F_1 esetében: $F(4, 142) = 5,390$, $p = 0,001$; az F_2 esetében: $F(4, 142) = 7,352$, $p = 0,001$]. Hasonló eredményeket kaptunk az /o/-realizációk esetében is, a különbségek itt is statisztikailag alátámasztottak [az F_1 esetében: $F(4, 142) = 8,250$, $p = 0,001$; az F_2 esetében: $F(4, 142) = 6,068$, $p = 0,001$]. Ez arra utal, hogy mindkét magánhangzó artikulációja beszélőnként változó, a magánhangzók képzése során a nyelv függőleges és vízszintes irányban is változik az adott beszélők ejtésének függvényében. A férfiak ejtésében elemezve ugyanezeket a paramétereket, némiképpen más eredménye-

ket látunk. Az /ɔ/ realizációinak első két formánsa szignifikánsan különbözik az egyes beszélők között [az F_1 esetében: $F(4, 142) = 2,857, p = 0,026$; az F_2 esetében: $F(4, 142) = 13,691, p = 0,001$]. A hangsúlytalan /o/ realizációinak első formánsa nem mutatott jelentős frekvenciaeltéréseket, míg a második formáns igen [$F(4, 142) = 3,034, p = 0,020$]. A férfiak ejtésében a hangsúlytalan magánhangzó csak részben módosul a különböző beszélők ejtésében; a nyelv függőleges irányú változása jelentős, vagyis a középső nyelválláshoz képest elmozdul az artikulációs gesztus. A 2. táblázat összegzi az *akkor* szó két magánhangzójára kapott formánsfrekvencia-értékeket.



7. ábra

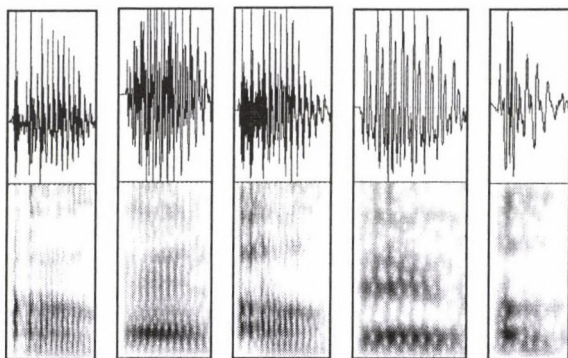
A [k:] mássalhangzó zöngekezdési idejének temporális adatai az összes adatközlőnél (medián és szóródás; n = nő, f = férfi)

2. táblázat: Az *akkor* szó magánhangzóinak formánsfrekvenciái, átlag és átlagos eltérés (SD)

Formánsok		Formánsfrekvencia-értékek (Hz)			
		Nők		Férfiak	
		Átlag	SD	Átlag	SD
ɔ	F ₁	623,13	94,44	567,37	79,76
	F ₂	1444,68	219,40	1261,00	149,90
o	F ₁	546,03	57,99	476,59	44,54
	F ₂	1298,90	199,48	1123,21	139,70

Mind a hangsúlyos, mind a hangsúlytalan magánhangzó ejtése a vizsgált szavakban a szakirodalmi adatokhoz képest nyiltabb ejtést igazol, és a szájüreg első felében képzett hangok akusztikus vetületét mutatja. A formáns-

frekvencia-értékek a női adatközlőknél nagyobb változatossággal realizálódnak (8. ábra: példák a hangsúlyos magánhangzóra), mint a férfi adatközlőknél.



8. ábra

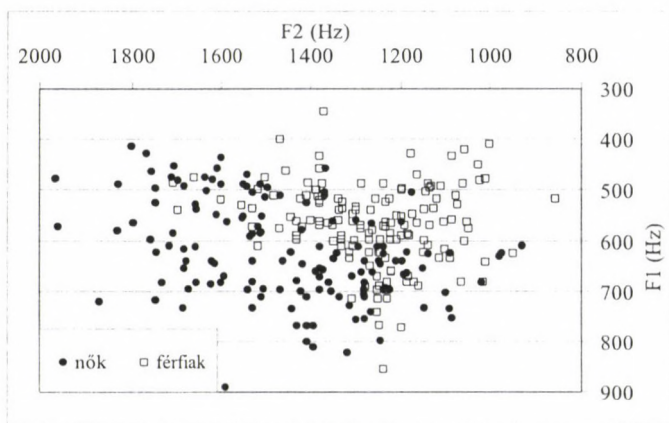
Öt /o/-realizáció ugyanazon női adatközlő *akkor* ejtéseiben
(felül: rezgéskép, alul: hangszinkép, 5 kHz-es tartományban)

A férfiak a hangsúlytalan magánhangzókat nagyobb mértékben ejtették hasonlóan, mint a hangsúlyosakat. Ejtésükben az /o/ magánhangzó artikulációja a semleges magánhangzóét, a sváét közelíti, nemegyszer megegyzik azzal. Ez azt jelenti, hogy a férfi adatközlők hajlamosak voltak a hangsúlytalan helyzetű magánhangzót sváként realizálni. A 9. és a 10. ábra a két magánhangzó F_1/F_2 realizációit mutatja valamennyi adatközlő artikulációjában. Az *akkor* szóban megvalósuló két magánhangzó formánsfrekvenciái más magánhangzók jellemző formánsértékeivel mutatnak átfedést, különösen az [a:, ɛ, ø, ə] hangokéival.

A hangsúlyos magánhangzó formánsai valamennyi beszélő ejtésében nagyobb szóródást mutatnak, mint a hangsúlytalanéi. A hangsúlytalan helyzetű magánhangzók – spontán beszédben általában is (vö. Gósy 2004, Beke 2009) – gyakrabban realizálódnak semleges magánhangzókként, és az adatok az mutatják, hogy ezek ejtése kisebb variabilitást mutat, mint a hangsúlyos helyzetű magánhangzóké. A 11. a)–b) és a 12. a)–b) ábra azt szemlélteti, hogy milyen különbségek figyelhetők meg az adott magánhangzók artikulációjában az egyes beszélőknél. Jól láthatók a két magánhangzó ejtésének eltérései is.

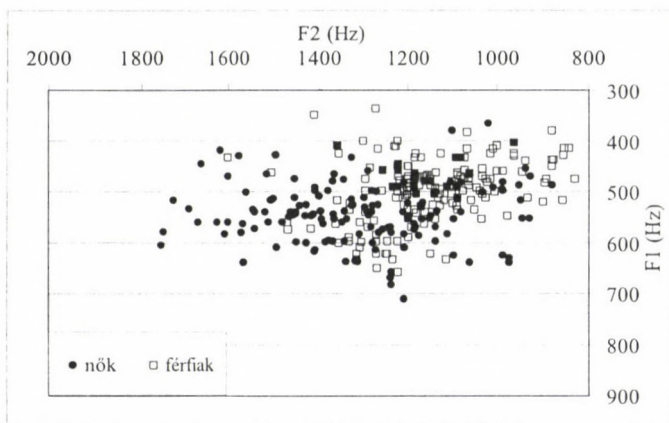
A formánsértékek alapján az alábbi megállapítások tehetők. a) Az tehát, hogy a hangsúlyos magánhangzók szélesebb akusztikai tartományban realizálódnak, mint a hangsúlytalanok, azt jelenti, hogy a hangsúlyos helyzet nem kedvez a magánhangzóminőség stabilitásának. b) A hangsúlytalan magánhangzók szűkebb F_1/F_2 tartománya a fonéma újabb realizációját, jelen esetben a semleges magánhangzó nagyobb mértékű megjelenését igazolja. A hangsúlytalan helyzet tehát itt nem a fonéma alapváltozatának stabilizálódását,

hanem egy másik hangrealizáció kialakulását idézi elő. c) A hangsúlytalan magánhangzók második formánsa szélesebb frekvenciatartományt mutat, mint az első formáns. Ez azt jelenti, hogy a nyelv függőleges mozgása variábilisabb, mint a vízszintes mozgása.



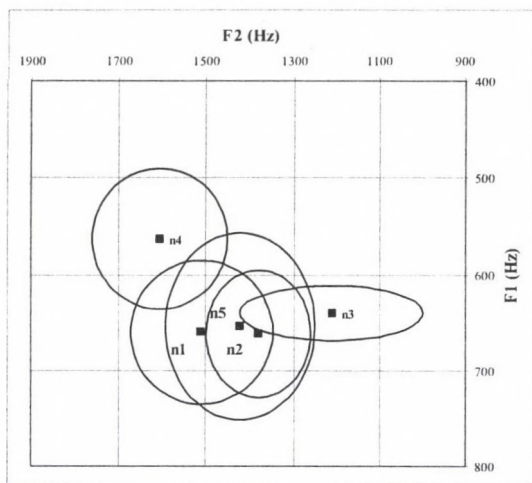
9. ábra

Az /o/ fonéma realizációinak F_1 és F_2 értékei az *akkor* szó előfordulásaiban



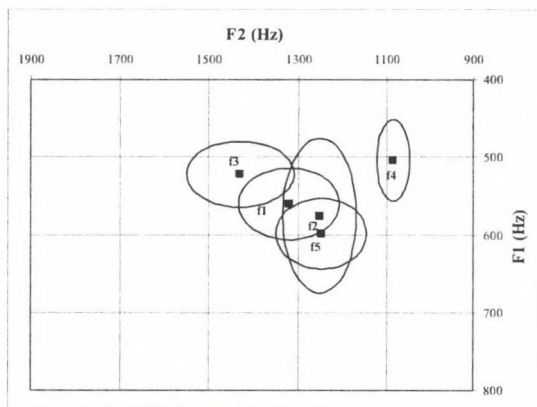
10. ábra

Az /o/ fonéma realizációinak F_1 és F_2 értékei az *akkor* szó előfordulásaiban



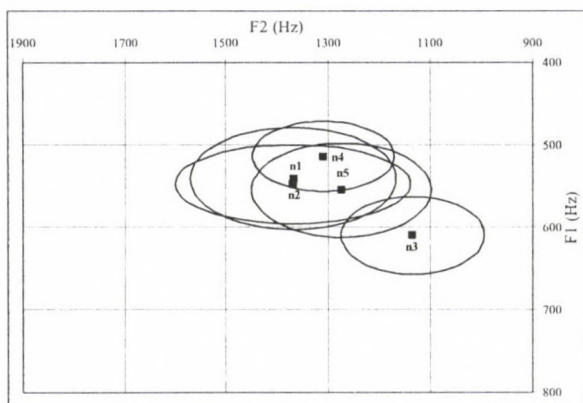
11. a) ábra

Az /ɔ/ fonéma realizációinak F_1/F_2 értékei a női adatközlőknél (a fekete négyzetek a mediánt szemléltetik)



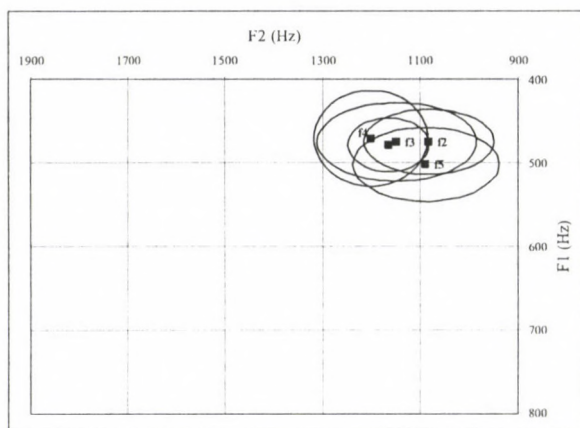
11. b) ábra

Az /ɔ/ fonéma realizációinak F_1/F_2 értékei a férfi adatközlőknél (a fekete négyzetek a mediánt szemléltetik)



12. a) ábra

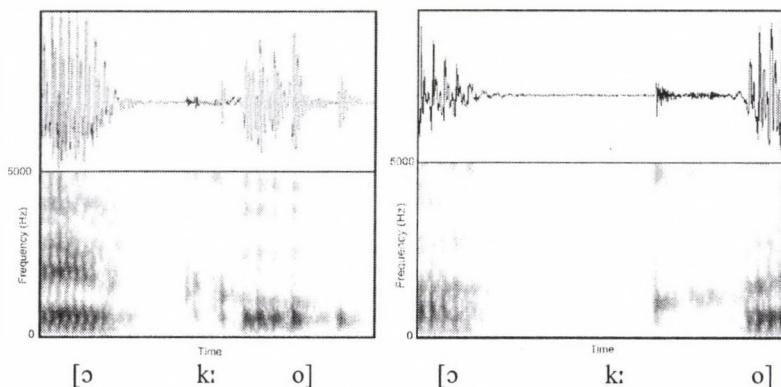
Az /o/ fonéma realizációinak F_1/F_2 értékei a női adatközlőknél (a fekete négyzetek a mediánt szemléltetik)



12. b) ábra

Az /o/ fonéma realizációinak F_1/F_2 értékei a férfi adatközlőknél (a fekete négyzetek a mediánt szemléltetik)

Elemeztük a zöngétlen zárhang zár-felpattanását. Az összes zöngétlen zárhang 18,9%-ában azt tapasztaltuk, hogy két zörejfelpattanás jellemezte a mássalhangzót (13. ábra; a regisztrátumok a megelőző és a követő magánhangzóknak csak egy részét mutatják).



13. ábra

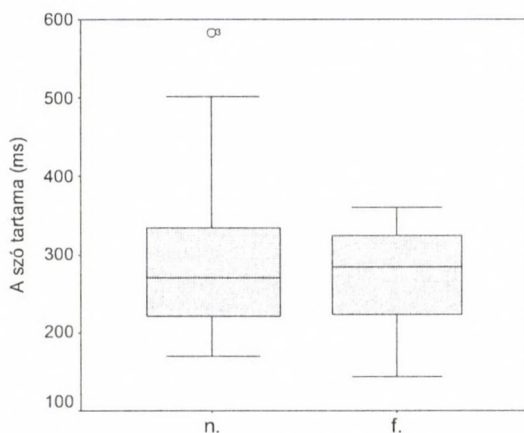
A zöngétlen veláris zárhang két (bal oldalon), illetve egy (jobb oldalon) zárfelpattanással ugyanazon női beszélőnél

Nem találtunk összefüggést azonban a zárfelpattanások száma és a zárhang objektív időtartama között. Úgy tűnik, ez beszélőspecifikus jelenség; noha minden adatközlőnél előfordult, volt, akinél csupán egyetlen ilyen ejtést látunk, és akadt, akinek 45,94%-ban két zárfelpattanás jellemezte a [k:] más-salhangzóját.

A zöngétlen zárhang zárfelpattanásának átlagos frekvenciaértéke a női adatközlők ejtésében 1023,41 Hz (átlagos eltérés: 136,37 Hz) volt, míg a férfi adatközlőknél alacsonyabb, az átlag 850,93 Hz (átlagos eltérés: 117,60 Hz). A nemek között – várhatóan – szignifikáns eltérés mutatkozott [egytényezős ANOVA: $F(1, 285) = 79,812$, $p = 0,001$]. Az egyes beszélők adatai ismét nagy különbségeket mutattak mind az egyes beszélők ejtészváltozataiban, mind a női és a férfi adatközlőket egymáshoz hasonlítva. A CV típusú hangkapcsolatokban a zárfelpattanás frekvenciája a veláris zárhangoknál az amerikai angolban tükrözi a követő magánhangzó hatását (Keating et al. 1994), és ezzel magyarázható a frekvenciaérték szóródása. Esetünkben a fonetikai kontextus e tekintetben állandó volt, így a variabilitást egyértelműen az ejtési sajátosságokkal magyarázhatjuk.

2. Az akkor percepció minősítése

A percepció teszt eredményeinek bemutatása előtt elemezzük a tesztszavak akusztikai fonetikai sajátosságait. A kísérlethez kiválasztott mindkét adatközlő akkor ejtései igen változatosak. Noha nincs statisztikailag igazolható különbség a két adatközlő szóidőtartamait tekintve, a női beszélő nagyobb időkülönbségekkel realizálja az adott szót (14. ábra).



14. ábra

A percepció teszt anyagának szavai: teljes időtartam a női (n.) és a férfi (f.) adatközlőnél (medián és szóródás)

A női beszélő legrövidebb *akkor* szava 170 ms-os volt, a leghosszabb 580 ms időtartamú; a férfi adatközlő ejtésében a legrövidebb szó 148 ms-os, a leghosszabb pedig 360 ms-os volt (3. táblázat). Nem találtunk szignifikáns eltérést az *akkor* két magánhangzójának időtartamában a női, illetve a férfi adatközlő ejtését összehasonlítva. A férfibeszélő a két magánhangzót hasonló időtartamban realizálta. Hasonlóképpen nem volt eltérés a [k:] mássalhangzó és a pergőhang fizikai időtartamában sem a két beszélő között. Szignifikáns különbséget láttunk ugyanakkor a zöngétlen zárhang zöngékezdési idejében [egytényezős ANOVA: $F(1, 47) = 11,94$; $p = 0,002$]. A női adatközlő VOT-je általában hosszabb volt, mint a férfié, úgy tűnik tehát, hogy a zöngétlenség jegye a női beszélőnél erőteljesebben érvényesült. Minél hosszabb ugyanis a zöngékezdési idő, annál jobban mutatkozik a hangszalagrezgés hiánya a zár-mássalhangzó képzése alatt (3. táblázat). A zárhang zörejfrequenciája várhatóan a női adatközlőnél volt magasabb (átlag: 1022,12 Hz, átlagos eltérés: 182,15 Hz, a férfi adatközlő átlaga: 813,95 Hz, átlagos eltérés: 135,24 Hz). A két beszélő között e tekintetben a különbség szignifikáns [egytényezős ANOVA: $F(1, 47) = 24,435$, $p = 0,001$].

A formánsfrekvenciák igen változatosak, különösen a női adatközlő ejtésében. A 4. táblázat az első két formáns értékeinek átlagát és az átlagos eltérést összegzi a két beszélőnél.

Elemeztük a hangsúlyos és a hangsúlytalan magánhangzó formánsfrekvenciáit a két beszélőnél. Az eredmények szignifikáns különbséget igazoltak a két különböző magánhangzó első formánsai között mindkét adatközlőnél [párosi-

tott t -próba, női adatközlő: $t(23) = 6,064$, $p = 0,001$; férfi adatközlő: $t(23) = 3,408$, $p = 0,002$. A második formánsok a női adatközlő ejtésében mutattak szignifikáns különbséget [$t(23) = 2,758$, $p = 0,011$], a férfi adatközlő esetében azonban nem. Ez azt jelenti, hogy a férfibeszlő elsősorban a nyelve függőleges mozgásával különítette el az [ɔ] és az [o] magánhangzókat. A női beszlő esetében a nyelv mindkét irányú mozgása meghatározó jegy volt a két fonéma realizációiban. Ennek következtében a női beszlő artikulációs gesztusai nagyobb különbséget eredményeztek a két magánhangzó minőségében a férfi ejtéséhez viszonyítva. A hangsúlytalan magánhangzó formánsfrekvenciái valamivel nagyobb mértékű átfedést mutatnak mindkét beszlőnél, mint a hangsúlyos magánhangzó esetében. Ez – mint korábban láttuk a többi adatközlőnél is – a semleges magánhangzó képzésére utal. Az 5. és 6. táblázat a két beszlő tesztszavainak valamennyi elemzett paraméterét összegzi.

3. táblázat: A tesztszavak temporális adatai (SD=átlagos eltérés)

akkor	Időtartam adatok (ms)			
	Női beszlő		Férfi beszlő	
	Átlag	SD	Átlag	SD
szó	295,33	101,11	272,45	68,17
ɔ	63,95	24,96	57,66	14,39
o	54,29	28,73	55,41	16,63
k:	139,66	49,32	120,87	30,11
VOT	47,87	16,42	34,20	11,42
r	36,66	27,60	38,50	25,43

4. táblázat: A tesztszavak magánhangzóinak formánsértékeinek átlaga és átlagos eltérése (SD)

Formánsok		Formánsfrekvencia-értékek (Hz)			
		Női adatközlő		Férfi adatközlő	
		Átlag	SD	Átlag	SD
ɔ	F ₁	655,83	89,36	580,91	55,34
	F ₂	1537,41	199,47	1245,04	119,94
o	F ₁	531,95	52,95	482,41	49,44
	F ₂	1365,25	223,77	1125,79	163,73

5. táblázat: A női adatközlő tesztszavainak paraméterei
(z. i. = zönggekezdesi idő)

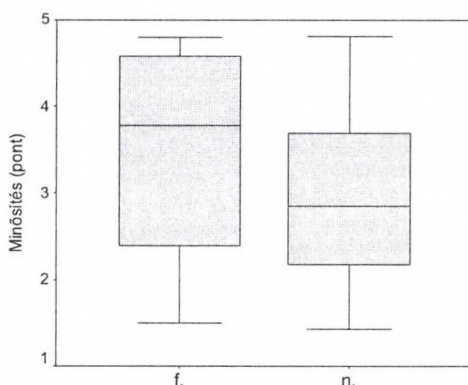
Szóidő (ms)	/ɔ/ idő (ms)	/ɔ/ F ₁ (Hz)	/ɔ/ F ₂ (Hz)	/o/ idő (ms)	/o/ F ₁ (Hz)	/o/ F ₂ (Hz)	/k/ idő (ms)	z.i. (ms)	/r/ idő (ms)	/k/ zö- rej (Hz)
219	42	732	1528	35	561	1625	110	32	32	1256
323	43	512	1370	57	597	1344	195	74	28	913
583	90	700	1356	166	487	952	273	82	54	776
384	64	660	1390	66	560	1510	197	45	57	1164
219	49	671	1380	40	573	1540	105	32	25	982
276	65	597	1760	39	430	1576	131	41	41	958
255	86	683	1600	40	524	1200	111	49	20	1256
238	29	573	1515	80	380	1100	129	54	0	1324
356	35	475	1638	49	560	1600	208	90	64	776
177	41	720	1870	26	548	1417	98	40	12	845
502	90	756	1300	99	487	879	206	48	107	982
222	63	769	1430	46	548	1197	98	33	15	1393
196	66	890	1590	25	570	1283	93	43	12	822
184	42	573	1960	60	499	989	73	35	9	1187
345	69	585	1711	54	548	1148	142	32	80	954
279	75	732	1687	32	540	1540	138	63	34	973
170	27	646	1417	42	561	1670	75	29	6	734
258	79	640	1460	50	502	1560	97	34	32	1280
266	100	622	1442	48	599	1450	118	48	0	996
322	56	622	1748	50	597	1369	159	53	57	1005
454	133	683	1736	58	548	1430	181	31	82	988
310	61	579	1420	49	512	1454	161	60	39	945
292	79	695	1503	46	475	1454	138	48	29	988
258	51	625	1087	46	561	1479	116	53	45	1034

6. táblázat: A férfi adatközlő tesztszavainak paraméterei
(z. i. = zönggekezdesi idő)

Szóidő (ms)	/ɔ/ idő (ms)	/ɔ/ F ₁ (Hz)	/ɔ/ F ₂ (Hz)	/o/ idő (ms)	/o/ F ₁ (Hz)	/o/ F ₂ (Hz)	/k/ idő (ms)	z. i. (ms)	/r/ idő (ms)	/k/ zörej (Hz)
323	62	524	1330	82	475	1063	105	28	74	708
227	59	540	1327	54	423	1045	77	32	37	592
179	52	536	1330	35	475	1026	84	42	8	868
266	59	585	1270	62	480	1100	110	23	35	525
159	39	570	1290	32	350	1410	88	50	0	799
267	46	590	1332	42	536	1185	156	42	23	776
314	53	610	1307	78	548	977	116	27	67	982
247	54	597	1222	42	499	1038	142	30	9	913
166	46	573	1417	39	561	1469	81	23	0	799
358	57	561	1307	66	480	1063	128	38	107	913
347	50	570	1190	70	512	1087	135	35	92	616
320	59	622	1246	60	499	1075	147	41	54	639

Szóidő (ms)	/ɔ/ idő (ms)	/ɔ/ F ₁ (Hz)	/ɔ/ F ₂ (Hz)	/o/ idő (ms)	/o/ F ₁ (Hz)	/o/ F ₂ (Hz)	/kz/ idő (ms)	z. i. (ms)	/r/ idő (ms)	/kz/ zörej (Hz)
257	70	590	1197	42	400	1222	125	65	20	868
360	66	622	1270	82	450	879	136	42	76	1005
219	46	597	1332	45	499	1258	107	37	21	980
230	69	597	1271	40	512	1026	117	25	4	1005
214	63	634	1112	45	426	1014	93	17	13	913
355	89	597	1246	54	536	1087	140	31	72	730
318	83	622	1148	59	487	1063	132	33	44	945
144	43	410	1000	30	434	1600	62	20	9	824
339	64	517	857	58	493	1045	196	20	21	688
304	40	634	1315	60	470	1033	132	29	72	798
301	51	622	1315	63	528	1127	157	53	30	837
325	64	622	1250	90	505	1127	135	38	36	812

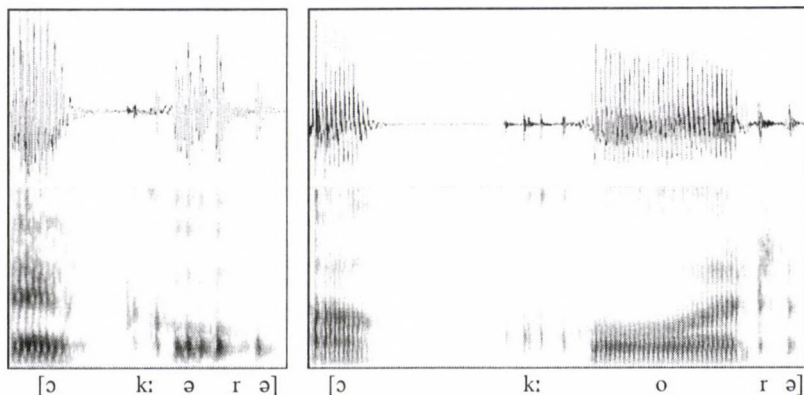
A percepciós teszt eredményei azt mutatják, hogy a férfi adatközlő szavait jobbnak minősítették, mint a női adatközlő *akkor* szavait (15. ábra).



15. ábra

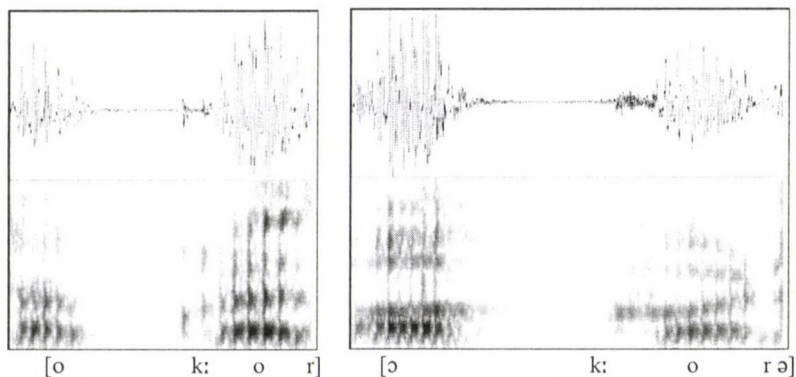
A tesztszavak minősítésének átlagai (medián) és a szóródás a két adatközlőnél (n = nő, f = férfi) az ötpontos skála használata alapján

A pontozás átlaga a férfi esetében 3,49 pont (átlagos eltérés: 1,16 pont), a női beszélő esetében 2,95 pont (átlagos eltérés: 0,91 pont); a minősítések a két adatközlő között azonban nem mutatnak szignifikáns különbséget. A kérdés az, hogy vajon mely paraméterek alapján hozták a kísérleti személyek a döntéseiket. A következő két ábra (16. és 17.) annak a négy *akkor* szónak a regisztrátumait mutatja, amelyeket a kísérleti személyek a legrosszabb és a legjobb minőségűnek ítélték a női és a férfi adatközlő ejtésében.



16. ábra

A női adatközlő legrosszabbul (bal) és legjobban érthetőnek (jobb) minősített *akkor* ejtéséről készült rezgéskép (fölül) és hangszínkép (alul)



17. ábra

A férfi adatközlő legrosszabbul (bal) és legjobban érthetőnek (jobb) minősített *akkor* ejtéséről készült rezgéskép (fölül) és hangszínkép (alul)

Az akusztikai fonetikai adatok utalnak arra, hogy a négy fenti tesztszó esetében milyen paraméterek együttese eredményezett jó, illetve rossz percepció minősítést. Ha egyenként megnézzük ezeket a paramétereket, és hozzárendeljük a percepció minősítést, az egyfajta kiindulást jelent az adott paraméternek a percepció döntésekben betöltött szerepének megismeréséhez. A szóidőtartamokat összevetve a percepció minősítésekkel, azt látjuk, hogy a kísérleti személyek annál jobbnak ítélték a szót, minél hosszabb volt az időtartama. A magánhangzók pontos (köznyelvi normának tekinthető) kiejtése is (vö. a következményes formánsszerkezetet) a pozitív minősítések növekedését ered-

ményezte. Megjegyezzük, hogy a legrosszabbnak ítélt női ejtésben az [o]-nak irreguláris zöngéje volt. A zárhangok zárfelpattanásában nem volt lényeges különbség a négy ejtésben. A pergőhangot követő svá mindkétféle minősítésű szóban előfordult, így ennek a hatása sem tekinthető meghatározónak.

Az elemzések során első megközelítésben az összes tesztszót két csoportba soroltuk a percepcióos tesztben résztvevők minősítése alapján. Az egyik csoportba azok a szavak kerültek, amelyek 1-es vagy 2-es, tehát igen gyenge minősítést kaptak, a másikba azok, amelyeket jónak ítélték, 4-es vagy 5-ös ponttal. (Értelemszerűen tehát a 3-as minősítésűeket nem vettük figyelembe; ezek aránya a teljes anyagban 23,42%, a női beszélőnél ez az arány 27,87%, a férfinnél 18,75%.) Elemeztük a két csoportba került szavak akusztikai fonetikai tulajdonságait.

A férfi adatközlő *akkor* szavait alig érthetőnek minősítették, amikor az átlagos szóidőtartamuk 172,5 ms volt, a hangsúlytalan magánhangzóé 38,66 ms, a zöngétlen zárhang ideje 79,5 ms és a pergőhangé 12,33 ms. A zárhang zörejfrequenciájának átlaga 791,66 Hz. A hangsúlyos magánhangzó első formánsainak átlaga ekkor 536,5 Hz-nek, a második formánsoké 1260,16 Hz-nek adódott; míg a hangsúlytalan magánhangzó F_1 -ének átlaga 471,2 Hz, a második formánsé pedig 1451,5 Hz volt. Kiválónak minősítették az ejtést, ha az időtartamok hosszabbak voltak; ekkor a szóidőtartamok átlaga 315,5 ms, a hangsúlytalan magánhangzóé 63,35 ms, a zöngétlen zárhang időtartamának átlaga 142,3 ms, a pergőhangé pedig 57,9 ms volt. A zárhang zörejfrequenciájában nem volt jelentős eltérés az előző csoport adatához képest, az átlag itt 795,8 Hz. A hangsúlyos magánhangzók első formánsának átlaga kissé magasabb a kiváló minősítéskor, 593,9 Hz, míg a második formánsoké alacsonyabb, 1096 Hz. A hangsúlytalan magánhangzó F_1 értéke 507,3 Hz, nem mutat jelentős különbséget az alig érthetőnek minősített szavak /o/-realizációitól, az F_2 értéke azonban igen, kiváló minősítés esetén jóval alacsonyabb, 1108,6 Hz (a különbség közel 350 Hz).

A női beszélő esetében az *akkor* szavakat alig érthetőnek minősítették, ha az átlagos szóidőtartamuk 252,5 ms volt, a zárhang időtartama 125,2 ms, a pergőhangé 34 ms. A zárhang zörejfrequenciájának átlaga 1032,12 Hz. A hangsúlyos magánhangzó első formánsainak átlaga 611,87 Hz, a második formánsoké 1700,11 Hz. A hangsúlytalan magánhangzók átlagos F_1 -e 500,87 Hz, F_2 -je pedig 1242,62 Hz. Kiválónak minősítették a kísérleti személyek a szóejtést az időtartamok növekedése esetén, ekkor a szóidőtartamok átlaga 489,6 ms volt, a zárhangé 225,3 ms, a pergőhangé pedig 72,6 ms. A zárhang átlagos zörejfrequenciája nem különbözött a gyenge minősítésű szavakban mérttől (az átlag itt: 974 Hz). A hangsúlyos magánhangzó F_2 -értékei jelentősen különböztek a két minősítésben. Kiváló minősítéskor az F_2 -átlag alacsonyabbnak, 1348,6 Hz-nek adódott (a különbség minegy 350 Hz); az F_1 átlaga pedig kismértékben magasabb volt (705,3 Hz), mint a gyenge ejtésű szavak esetében. A hangsúlytalan magánhangzó első formánsai csaknem megegyeztek a

másik kategóriába került magánhangzókéival, az átlag itt 511,3 Hz, míg a második formánsok jelentősen alacsonyabbak, az átlag ebben a kategóriában 1113,6 Hz (a különbség mintegy 130 Hz).

A percepciós döntések természetesen nem az egyes paraméterek hatását, hanem a szavak összes sajátossága alapján kialakult benyomást tükrözik. Korrelációs elemzésekkel igyekeztünk meghatározni az objektív akusztikai paraméterek és a percepciós minősítések közötti összefüggéseket. Arra kerestünk választ, hogy a percepciós döntések szerint melyek lehetnek a meghatározó és a kevésbé fontos fonetikai paraméterek.

A **férfi beszélőnél** erős korrelációt találtunk a szavak teljes időtartama és az észlelési minősítés között (Pearson-féle teszt: $r = 0,742$; 99%-os szinten). [A szó teljes időtartama továbbá erős korrelációt mutat a hangsúlytalan magánhangzó ($r = 0,815$), a zárhang ($r = 0,741$) és a pergőhang ($r = 0,747$) időtartamával.] Ez azt jelenti, hogy a szóidőtartamot a vizsgált szavakban elsősorban a hangsúlytalan magánhangzó és a két mássalhangzó időtartama jellemzi. Nem meglepő tehát ennek alapján, hogy közepes, illetve erős korrelációt találtunk a percepciós minősítések és a hangsúlytalan magánhangzó ($r = 0,564$), valamint a zárhang időtartama ($r = 0,71$) között. Meghatározó a hangsúlytalan magánhangzó első formánsa is ($r = 0,588$). Noha a teljes szó időtartamának alakulásában fontos a pergőhang, időtartama mégsem meghatározó a percepciós minősítésekben. Mindezek alapján kijelenthetjük, hogy az elemzett, férfi ejtésű *akkor* szavak esetében négy paraméter, illetve az ezeket eredményező artikulációs gesztusok felelősek a percepciós minősítésért. Ezek: a teljes szó, a hangsúlytalan magánhangzó és a zárhang időtartama, valamint a hangsúlytalan magánhangzó első formánsa. Artikulációs szempontból ez elsősorban az időzítés, valamint a hangsúlytalan magánhangzók pontos ejtésének fontosságára utal a férfi beszélő ejtésében.

A **női adatközlő** ejtése alapján kapott korrelációs adatok szerint a szó teljes időtartama szintén meghatározó, bár az összefüggés kevésbé erős, mint a férfi adatközlő esetében ($r = 0,571$). A szó teljes időtartama itt valamennyi beszédhang időtartamával jól korrelál (/ɔ/-realizációkra: $r = 0,567$, /o/-realizációkra: $r = 0,615$, a /k/-realizációkra: $r = 0,928$, a /r/-realizációkra: $r = 0,669$). Erős, illetve közepes korrelációt találtunk továbbá a percepciós döntések és a zárhang időtartama ($r = 0,562$), valamint a hangsúlyos magánhangzók második formánsának frekvenciája ($r = 0,755$) között.

Úgy tűnik, hogy minél nagyobbak a szó teljes időtartamának eltérései, mint azt a női adatközlő esetében látjuk, a kísérleti személyek annál kisebb mértékben használnak fel további paramétereket a döntéseikben. Ez lehet a magyarázata annak, hogy a női beszélő esetén három, a férfi beszélőnél négy paraméter volt meghatározó a percepciós minősítésekben. Mindkét beszélő esetén a szó teljes időtartama és a zárhang időtartama, valamint a hangsúlytalan magánhangzó akusztikai fonetikai sajátosságai gyakoroltak hatást a minősítésekre. Az utóbbi esetében azonban eltérő paraméterek voltak fontosak:

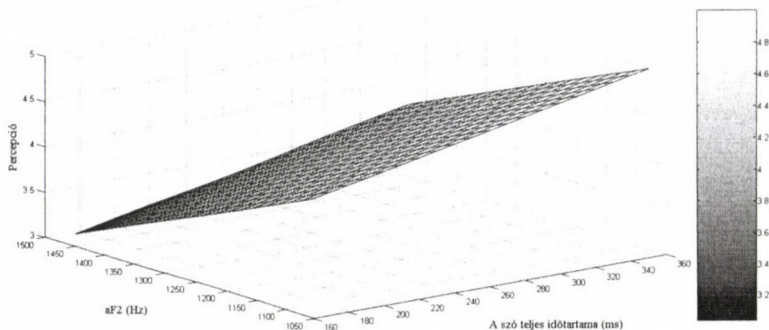
férfiejtésben az /o/-relizációk időtartama és első formánsa, a női ejtésben csupán az /o/-realizációk második formánsa.

Említettük, hogy a percepciók döntései kétségkívül a hallott szavak többé-kevésbé összes akusztikai fonetikai sajátossága alapján történnek (valószínűsíthetően egyéb, itt nem vizsgált további paraméterek is szerepet játszhatnak a minősítésekben, például a dallammenet, a hangszínezet). Adataink azt mutatják, hogy a percepciók döntéseiben kitüntetett paraméterek nem tekinthetők állandóknak, beszéd- és beszélőfüggetleneknek. Úgy tűnik, hogy ezek a paraméterek a mindenkori beszédjel sajátosságai alapján változnak (ahogy ezt a női és a férfi adatközlő esetében láttuk); bár egyetlen szó minősítése alapján csupán óvatosan fogalmazhatunk.

Az utolsó elemzéssorozatban arra a kérdésre igyekeztünk választ kapni, hogy melyek azok a paraméterek, amelyek másoknál jobban vesznek részt a hallgatók percepciók minősítéseiben. A válaszhoz lineáris regressziós modellt (Tacq 1997) alkalmaztunk. Ebben a modellben a függő változó viselkedése néhány független változóval magyarázható. A módszer (stepwise) segítségével lehetőség nyílt arra, hogy valamennyi független változót elemezhessük.

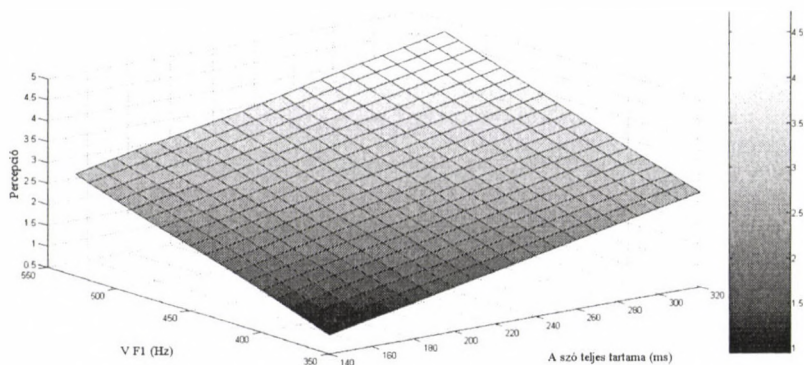
Az eredmények azt mutatják, hogy a női beszélő szavainak percepciók minősítése csupán a teljes szóidőtartam alapján már 51%-osan magyarázható. Ha ehhez az [ɔ] magánhangzó második formánsát is hozzávesszük, akkor a minősítések 68%-ban magyarázhatók. Ez a két paraméter tehát döntő hatással a percepcióra. Egy további paraméter – ami esetünkben az /ɔ/-realizációk első formánsa –, az arányt csupán 6%-kal növeli, vagyis ennek a paraméternek a hatása már csaknem elhanyagolható. A percepciók minősítéseket a férfibeszélő esetében 54%-ban magyarázza a szóidőtartam. Az /o/-realizációk első formánsával együtt azonban e két paraméter a percepciók minősítések 66%-áért felelős. Ez azt jelenti, hogy az *akkor* szavak kiejtésének minősítése a férfi adatközlőnél döntően e két változóval magyarázható. A lineáris regressziós egyenlet segítségével interpolációt végeztünk az eredmények szemléltetésére. A 18. ábra a női beszélő, a 19. ábra a férfibeszélő ejtése alapján adatolt paraméterek és a kapott minősítések összefüggéseit mutatja. Minél világosabb az árnyékolás, annál jobb a minőség.

A hallgató tehát képes arra, hogy a beszédmegértés során függetlenítse a feldolgozást a beszélő személy egyéni kiejtési sajátosságaitól. Ez úgy lehetséges, hogy egy vagy több sejtgyüttes valóban ugyanazt a fonológiai, de eltérő fonetikai struktúrát tartalmaz. A hallgató pedig a hasonlókat aktiválja a beszédészlelés során.



18. ábra

Az *akkor* kiejtésének beszédpercepció minősítését meghatározó paraméterek háromdimenziós képe női ejtés alapján



19. ábra

Az *akkor* kiejtésének beszédpercepció minősítését meghatározó paraméterek háromdimenziós képe férfi ejtés alapján

Következtetések

A jelen kutatás egyetlen szó, az *akkor* ejtési variációit elemezte spontán beszédben a BEA-adatbázis véletlenszerűen kiválasztott tíz fiatal beszélőjének 286 szava alapján. Az egyik fő kérdés az volt, hogy mely paraméterek változnak meg az elemzett szavak kiejtésének variabilitása következtében. Noha feltételeztük, hogy lesznek ejtési különbségek, elsősorban a különböző beszélők között tételeztük fel a nagyobb variabilitást. Az eredmények azonban sokkal nagyobb artikulációs különbségeket igazoltak a vártnál, mind a beszélők között, mind az ugyanazon beszélő ejtéseiben.

Az adatok megerősítették a hipotézisünket, miszerint a lexikális reprezentációnak rugalmasnak és alkalmazkodónak kell lennie a mindenkori beszédjel függvényében. Az akusztikai fonetikai elemzések és a percepciók minősítési eredmények megerősíteni látszanak a „szóhologram” elméletet a lexikális egységek tárolásával kapcsolatosan. Az adatok ugyanis egy rendkívül rugalmas neurális reprezentációt igazolnak, amely – jelen ismereteink szerint – legszemléletesebben a szóhologrammal képzelhető el.

Eredményeink szerint egy szó gyakori előfordulása a beszédben a kiejtés nagymértékű változatosságához vezet. Jól ismert tény, hogy a beszéd igen redundáns, jóval több akusztikai fonetikai információt tartalmaz, mint amennyire a beszédfeldolgozásnak szüksége lenne (pl. Scott 2005). Az *akkor* szó invariánsnak tekinthető akusztikai fonetikai paraméterei az alábbiak voltak. Relative stabilnak mutatkozott a zöngétlen zárhang zöngékezdési ideje, ami a zöngétlenség jegyének biztosítására szolgál. A zárfelpattanás specifikus frekvenciája 1500 Hz alatt realizálódott, ami megfelel a hátul képzettségre utal. A hangsúlyos magánhangzó formánsainak többsége egyértelműen egy hátul képzett és alsó nyelvvállású magánhangzóra volt jellemző. A szóvégi pergőhang jelenléte a versengő szóformák lehetőségét szűkítette, függetlenül annak aktuális fonetikai realizációjától. A hangsúlytalan /o/ magánhangzó anyakban nagymértékben sváként realizálódott, ez a tény nem segítette a lexikális hozzáférési folyamatot. Az /ɔk:or/ fonológiai struktúrája – az ejtészvariációk következtében – sokféle fonetikai hangsorként realizálódott. A kontextus természetesen szűkíti a lehetséges szavak aktiválását, továbbá a hallgató előfeltevései és beszédtapasztalata is a versengő szavak szűkülésének irányába mutat.

A hangsúlyos magánhangzó minőségének változását magyarázhatná a kontextushatás; esetünkben azonban ez nem lehetséges, mivel az akkor közel 80%-ban vagy a [j] mássalhangzót, vagy néma szünetet követően fordult elő. Ezért azt feltételezzük, hogy az elemzett szó akusztikai fonetikai változatossága a beszédtervezési folyamat sajátosságaiban keresendő. Tekintve, hogy az *akkor* többségében töltelékszóként funkcionál a vizsgált spontán szövegekben, ezért az artikulációs gesztusaira kevesebb figyelem fordítódik. Ejtése közben a beszédtervezési folyamat már más „feladatra” koncentrál. Az *akkort* a beszélő nem a hallgató számára fontos információként aktiválja, hanem a saját beszédtervezési folyamatai könnyítésére, egyfajta időnyerésre. A percepciók kísérlet résztvevői különböző paraméterek alapján minősítették az *akkor* szavakat az adott beszélő kiejtésének függvényében. Ez arra utal, hogy a percepciók döntésekért felelős akusztikai fonetikai tényezők nem feltétlenül előre jósolhatók, mivel azok az adott beszédjel sajátosságainak függvényében változhatnak.

A hallgató percepciók stratégiái rugalmasak, alkalmazkodnak a változó beszédjelhez. Ha a beszélők mind egyformán artikulálnának, akkor a beszédpercepciók mechanizmusnak nem kellene rugalmasnak lennie (McQueen

2005). A mechanizmus összevető műveletei során megtörténik a bejövő beszédjel akusztikai tulajdonságainak összehasonlítása a tárolt (mentális) reprezentációval. Ez az összevető művelet tekintetbe veszi az egyéni kiejtési jellemzőket, ahogyan azt a kísérleti eredményeink megerősítették. A neuronális reprezentációk elméletét alkalmazva, azt mondhatjuk, hogy valamennyi, az *akkor* szónak megfelelő neuronegyüttes aktiválódik. Az elhangzó szavakat a hallgató összeveti a mentális lexikonában meglévő „szóhologramokkal”, és ezért képes választásra és minősítésre.

A jelen kutatás eredményei megerősítették azt, hogy a beszélők bizonyos mértékig kontrollálják ugyan az artikulációs gesztusaikat, mégis igen változatosak a szórealizációk akusztikai fonetikai megvalósulásai. Bizonyos invariáns jegyek megőrződnek, lehetővé téve a (majdnem) biztos lexikális hozzáférést. Adataink, illetőleg a felvetett kérdések további kutatást igényelnek mind fonetikai, mind pszicholingvisztikai szempontból. A mesterséges beszédfelismerő rendszerek számára felhasználható eredmény, hogy a gyakori szavak ejtészvariációival érdemes célzottan foglalkozni a jobb felismerési eredmény elérése érdekében.

Irodalom

- Andruski, J. E. – Blumstein, Sheila E. – Burton, M. 1994. The effect of subphonetic differences on lexical access. *Cognition* 52/3. 163–187.
- Beke András 2009. A veláris magánhangzók stabilitása a spontán beszédben. In: Gecső Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai*. Kodolányi János Főiskola–Tinta Könyvkiadó, Székesfehérvár–Budapest (megjelenőben).
- Bishop, Dorothy V. M. 1997. *Uncommon understanding. Development and disorders of language comprehension in children*. Psychology Press, London.
- Boersma, Paul – Weenink, David. 2005. *Praat: doing phonetics by computer*, <http://www.praat.org/> [visited March 12, 2005].
- Clark, Herbert H. – Wasow, T. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.
- Crinion, Jennifer T. – Lambon-Ralph, A. – Warburton, Elisabeth A. – Howard, David – Wise, Richard S. 2003. Temporal lobe regions engaged during normal speech comprehension. *Brain* 126. 1193–1201.
- Dankovičová, J. – Nolan, Francis 1999. Some acoustic effects of speaking style on utterances for automatic speaker verification. *Journal of the International Phonetic Association* 29. 115–229.
- Démonet, Jean-Françoise – Thierry, Guillaume – Cardebat, Dominique 2005. Renewal of the neurophysiology of language: Functional neuroimaging. *Physiological Review* 85. 49–95.
- Dodge, Ellen – Lakoff, George 2005. Image schemas: from linguistic analysis to neural grounding. In: Hampe, Beate (ed.): *From perception to meaning. Image schemas in cognitive linguistics*. Mouton de Gruyter, Berlin–New York, 57–92.
- Fox Tree, Jean E. – Schrock, J. C. 2002. Basic meanings of *you know* and *I mean*. *Journal of Pragmatics* 34. 727–747.

- Frauenfelder, Uli H. – Komisarjevsky Tyler, L. (eds.) 1987. *Spoken word recognition*. Cambridge, Massachusetts–London.
- Gósy Mária 2000. A [p, t, k] mássalhangzók zöngékezdési ideje. *Magyar Nyelvőr* 124/2. 195–204.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszéd kutatás 2008*. 194–207.
- Gósy Mária – Horváth Viktória 2009. Hogyan tükrözi a kiejtés a nyelvi funkció változását? In: Keszler Borbála és Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Kiadó, Budapest, 37–45.
- Grady, Joseph E. 2005. Image schemas and perception: Refining a definition. In: Hampe, Beate (ed.): *From perception to meaning. Image schemas in cognitive linguistics*. Mouton de Gruyter, Berlin–New York, 35–56.
- Greenberg, Steven 2006. A multi-tier framework for understanding spoken language. In: Greenberg, Steven – Ainsworth, William A. (eds.): *Listening to speech. An auditory perspective*. Erlbaum, Mahwah, New Jersey–London, 411–434.
- Heilman, Kenneth M. – Valenstein, Edward (eds.) 2003. *Clinical neuropsychology*. Oxford University Press, Oxford.
- Horga, Damir 2008. Repetitions in interrupted speech production. *Beszéd kutatás 2008*. 157–171.
- Horváth Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben*. PhD-disszertáció. ELTE, Budapest.
- Jackendoff, R. 2002. *Foundations of language*. Oxford University Press, Oxford.
- Keating, Patricia A. – Byrd, Dani – Flemming, Edward – Todaka, Yuichi 1994. Phonetic analyses of word and segment variation using the TIMIT corpus of American English. *Speech Communication* 14. 13–142.
- Kenesei, István 2007. Semiwords and affixoids: The territory between word and affix. *Acta Linguistica Hungarica* 54. 263–293.
- Kohler, Klaus 2000. Investigating unscripted speech: Implications for phonetics and phonology. *Phonetica* 57. 85–94.
- Lee, Chin-Hui – Soong, Frank K. – Paliwal, Kuldip K. (eds.) 1996. *Automatic speech and speaker recognition*. Kluwer Academic Publisher, Boston–Dordrecht–London.
- Levelt, Willem J. M. 1983. Monitoring and self-repair in speech. *Cognition* 14. 41–104.
- Levelt, Willem J. M. – Roelofs, A. – Meyer, Antje S. 1988. *A theory of lexical access in speech production*. Cambridge University Press, Cambridge, Massachusetts.
- Libben, Gary – Jarema, Gonia 2002. Mental lexicon research in the new millenium. *Brain and Language* 81. 2–11.
- Lindblom, Björn 1986. On the origin and purpose of discreteness and invariance in sound patterns. In: Perrell, Joseph S. – Klatt, Dennis H. (eds.): *Invariance and variability of speech*. Erlbaum, Hillsdale, New Jersey, 493–510.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, William J. – Marchal, A. (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–440.
- Marslen-Wilson, William D. 1989. *Lexical representation and process*. Cambridge University Press, Cambridge, Massachusetts–London.

- Nusbaum, Howard – Magnuson, James 1997. Talker normalization: Phonetic constancy as a cognitive process. In Johnson, Keith – Mullenix, John W. (eds.): *Talker variability in speech processing*. Academic Press, San Diego, 109–132.
- Pribram, Karl H. 1991. *Brain and perception: Holonomy and structure in figural processing*. Erlbaum, Hillsdale, New Jersey.
- Pluymaekers, M. – Ernestus, M. – Baayen, H. R. 2005. Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62. 146–159.
- Pulvermüller, Friedemann 1999. Words in the brain's language. *Behavioral and Brain Sciences* 22. 253–336.
- Pulvermüller, Friedemann 2003. *The neuroscience of language*. Cambridge University Press, Cambridge.
- Rose, Phil 1999. Long- and short-term within-speaker differences in the formants of Australian hello. *Journal of the International Phonetic Association* 29. 1–33.
- Scott, Sophie K. – Blank, Catrin, C. – Rosen, Stuart – Wise, Richard J. S. 2000. Identification of a pathway for intelligible speech in the frontal lobe. *Brain* 123. 2400–2406.
- Shriberg, Elisabeth 2001. To 'errrr' is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31. 153–169.
- Stevens, Kenneth N. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In David, E. E. – Denes, Peter B. (eds.) *Human communication: a unified view*. McGraw Hill, New York, 51–56.
- Tacq, Jacques 1997. *Multivariate analysis techniques in social science research. From problem to analysis*. Sage Publications, London.
- Vago, Robert – Gósy, Mária 2007. Schwa vocalization in the realization of /r/. In Trouvain, Jürgen – Barry, William J. (eds): *16th International Congress of Phonetic Sciences. Proceedings*. Saarbrücken University, Saarbrücken, 505–509.
- Zwitserslood, Pienie 2003. The internal structure of words: Consequences for listening and speaking. In Schiller, Niels O. – Meyer, Antje S. (eds.): *Phonetics and phonology in language comprehension and production*. Mouton de Gruyter, Berlin–New York, 79–114.

A kutatás az OTKA 78315 sz. pályázat keretében folyt.

AZ IDŐS ÉLETKOR TÜKRÖZŐDÉSE A MAGÁNHANGZÓK EJTÉSÉBEN

Bóna Judit

Bevezetés

Életünk során beszédünk folyamatosan változik. Az öregedés nemcsak a beszéd, hanem a nyelvhasználat minden területén megnyilvánul. Az idők szóhasználatát az életkor előrehaladtával eltér a fiatalok szóhasználatától; illetve a fiatalabbak számára régiesnek tűnhet. Az idősekre jellemző az időben korábbi nyelvhasználat szavainak pontosabb ismerete, és gyakran nem értik meg az új keletkezésű szavakat (Libárdi 2001; Kiss 2002). Nemcsak a szóismeretben, hanem a lexikális előhívásban is másképp teljesítenek az eltérő életkorú személyek: az idősek gyorsabbak az igék előhívásában, mint a fiatal felnőttek, míg a köznevek esetében többször fordul elő szóelőhívási nehézségük (Gósy 2001). Az idősek, hasonlóan a fiatalokhoz, a lexikális hozzáférés során nagyobb mértékben használják fel az aktuális hangzást, mint a fiatal és a középkorú felnőttek; azaz ha fonetikailag torzított hangsort hallanak, pontosabban képesek az eredeti szavakat kontextus nélkül is felismerni. Ennek valószínűleg az lehet a magyarázata, hogy az idősek a bizonytalanabb beszédészlelésük miatt gyakrabban kényszerülnek pontatlanul észlelt szavak azonosítására (Gósy 2000).

Az életkor befolyásolja a megakadásjelenségek gyakoriságát is (Menyhárt 2003). Egy a „nyelvem hegyén van” jelenséget vizsgáló kísérlet eredménye azt mutatja, hogy az idősek hosszabb reakcióidővel képesek a kívánt szót előhívni, illetve gyakrabban hívnak elő téves szót (Horváth 2006).

A hormonális változások, illetve a szervezet öregedése, elhasználódása mind az alaphangmagasságra és a zöngéképzésre, mind az artikuláció pontosságára hatással vannak. A gyermekhang a pubertáskori hormonális változások hatására válik felnőtthanggá, majd a 60. életév után (szintén hormonális okok miatt) a hangminőség ismét nagy változáson megy keresztül, és kialakul az öreghang (Balázs 1993).

A szervezet előregedése minden beszédképző szervet érint. A légző rendszer öregedésének következtében csökken a tüdőkapacitás, ami a hangerő csökkenését vonja maga után. Megrövidül a hangtartás is, a hang reszketővé válik. A hangszalagok rugalmatlanabbá válnak, a gégeizomzat leépül, az alaphangmagasság a férfiaknál emelkedik, a nőknél alacsonyabb lesz. Beszűkül a hangterjedelem is: a felnőtthangra jellemző, átlagosan 2 oktáv hangterjedelemmel szemben az időseket bő 1 oktáv hangterjedelem jellemzi. Nehezi-

tetté válik a nyelv mozgása, ami az artikuláció pontatlanságát vonja maga után. A lány szájpad mozgásának renyhülése pedig orrhangzós színezetet okozhat az idős emberek beszédében (Balázs 1993).

Több kutatás igazolta a beszédtempó lassulását is az életkor előrehaladtával. A felnőttekre jellemző magyar köznyelvi átlagos artikulációs tempó 12,5–14 hang/s, idős korban ez átlagosan 10,0 hang/s (Gósy 1997). Két színész, Páger Antal és Dajka Margit fiatalkori és időskori hangjának összevetése a következő eredményt hozta: Páger Antal idős korában 3,7 hanggal, Dajka Margit pedig 2 hanggal ejtett kevesebbet másodpercenként (ez mintegy 40, illetve 17 szó különbséget jelent percenként). Páger idős korában nyolcszor, Dajka ötször annyi szünetet tartott, mint fiatalon (Balázs 1993). Hasonló eredményekre jutott Menyhárt Krisztina (2000) és Gocsál Ákos is (2000), utóbbi adatai azonban nem mutatnak szignifikáns összefüggést az életkor és a szünettartás között.

Egy magyar nyelvű longitudinális vizsgálatban arra keresték a választ, hogy a szegmentális szinten milyen változások következnek be 25 év elteltével ugyanazon beszélők beszédében. A kutatás tárgya egyetlen magánhangzó, az [e] volt, amelynek az első 7 formánsát, formáns-sávszélességét, illetve időtartamát vizsgálták, és elemezték mindegyik beszélő alaphangmagasságát is. A beszélők az első felvétel készítésekor 23, 30 és 32 évesek voltak; a második felvételkor 48, 55 és 57 évesek. Az eredmények azt mutatták, hogy az alaphangmagasság és az életkor összefüggése egyénenként változó mértékű: a beszélőknél részben erőteljes csökkenést, részben változatlanságot tapasztaltak. Az F_1 és az F_2 (a hangminőséget biztosító formánsok), illetve az F_4 értéke állandóságot mutatott, míg az F_3 , az F_6 és az F_7 helye jellegzetesen megváltozott. A formáns-sávszélességek átlagosan 20–40%-kal nőttek 25 év elteltével, ugyanakkor az F_4 esetében a sávszélesség minden helyzetben és minden beszélőnél szűkült. A hangidőtartamok az idősebb korban egyértelműen megnyúltak (Gósy–Nikléczy 2000).

A magánhangzók akusztikai szerkezetét vizsgálták egy amerikai angol anyanyelvűekkel végzett kutatásban fiatalok és idősek beszédében. Az eredmények szerint az F_1 szignifikáns eltérést mutatott az idősek és a fiatalok ejtése között minden magánhangzónál; míg az F_2 -ben a hátul képzett magánhangzók esetén találtak szignifikáns eltérést (Watson–Munson 2007).

Egy másik, szintén amerikai angol anyanyelvűekkel végzett kutatás azt hangsúlyozza, hogy az életkori változásokat a nemek, illetve a hallásállapot függvényében kell vizsgálni. 27 fiatal és 59 idős személy beszédében elemezték az alapfrekvenciát, az első három formáns, illetve a zöngeskedési időt (VOT). A nemek és életkorok szerint differenciált csoportok között az F_0 , az F_1 és a VOT értékeiben találtak szignifikáns eltérést, de az egyes paraméterek erősen függtek a vizsgált beszédhangok minőségétől is (Torre–Barlow 2009).

A jelen tanulmányunk célja 1. annak bemutatása, hogy hogyan változik a magánhangzók frekvenciaszerkezete az időskori beszédben; 2. kimutathatók-

e az angolhoz hasonló változások a magyarban is, és 3. ha igen, mely akusztikai paraméterekben jelennek meg. Mivel a magyarban a magánhangzók (szinte) mindig zöngések, ezért további célunk 4. a magánhangzók formánsai mellett a zöngé minőségének elemzése is. Hipotéziseink szerint a fiziológiai változásoknak megfelelően a vizsgált paraméterek mindegyike mutat változást az életkor előrehaladtával.

Anyag, módszer, kísérleti személyek

Vizsgálatunkhoz a BEA (vö. Gósy 2008) hanganyagaiból választottuk ki nyolc idős (életkoruk 70–80 év) és nyolc fiatal (életkoruk 20–32 év) női adatközlő spontánbeszéd-felvételét. Mindegyik beszélő az életkorának megfelelő ép hallással rendelkezett, beszédhibája egyiknek sem volt. A felvételek csendesített szobában, digitálisan lettek rögzítve. A kutatáshoz adatközlőnként mintegy egy perces hanganyagot elemeztünk; a beszédprodukciókban összesen 3971 magánhangzót adatoltunk. Mivel a magánhangzók gyakorisága igen eltérő a spontán beszédben, ezért kutatásunkban csak azokat a magánhangzókat vettük figyelembe, amelyekből összesen legalább 100 db szerepelt a vizsgált beszédprodukciókban. Így az [ɔ, a:, ɛ, e:, i, o, ø, u] magánhangzók formánsait elemeztük. A vizsgált magánhangzók darabszáma az 1. táblázatban olvasható.

1. táblázat: A vizsgált magánhangzók darabszáma

V	[ɔ]	[a:]	[ɛ]	[e:]	[o]	[i]	[ø]	[u]
db	737	394	1089	500	498	313	114	103

A hanganyagokban a Praat 5.0 szoftverrel (Boersma–Weenink 1998) manuálisan felcímkéztük a magánhangzókat, majd egy erre a célra írt script segítségével automatikusan megmértük a beszédhangok F_1 , F_2 és F_3 értékét, a beszédhangokban mért alaphangmagasságot (F_0), illetve a zöngé minőségére vonatkozó adatokat: a jittert, a shimmert és a jel/zaj viszonyt. (A jitter a hangszalagrezgések frekvenciaingadozását, a shimmer a hangszalagrezgés amplitúdóingadozását, a jel/zaj viszony pedig a zöngé és a levegőáramlási zörej arányát jelzi.) A beszélőkre jellemző, átlagolt F_0 -értékek meghatározásánál nem vettük figyelembe a glottalizált ejtésű magánhangzók alapfrekvenciáját.

A script 20 ms-ként adatolta a hangminta tulajdonságait, majd az annotált időtartamokba eső eredményeket átlagolta. Az automatikus mérések eredményeit ezután manuálisan ellenőriztük. Az automatikus mérésnek köszönhetően egy hangot nem egyetlen időpillanatban mértünk, így pontosabb eredményeket kaptunk. Az adatokon statisztikai elemzést végeztünk az SPSS 13.0 verziójával (egytényezős varianciaanalízis és Tukey post hoc teszt). (A kis elemszám miatt az egyéni különbségek statisztikai elemzését az [ø] és az [u] hangok formánsainak esetében nem tudtuk elvégezni.)

Eredmények

Először a zöngé minőségére vonatkozó adatokat elemeztük. A 2. táblázatban az egyes adatközlők beszédében mért jitter, shimmer és a jel/zaj viszony értékei olvashatók.

2. táblázat: A zöngéminőség jellemzői (i = idős; f = fiatal)

Beszélő	Alaphangmagasság (Hz)	Jitter (%)	Shimmer (dB)	Jel/zaj (dB)
i1	168	4,39	0,15	4,63
i2	166	2,44	0,09	7,68
i3	163	2,31	0,12	7,48
i4	183	1,94	0,10	8,37
i5	180	5,53	0,17	3,11
i6	145	4,54	0,13	2,96
i7	209	3,10	0,14	7,45
i8	155	5,35	0,16	1,32
f9	202	1,10	0,11	12,39
f10	151	1,65	0,09	8,64
f11	150	1,78	0,14	9,37
f12	187	1,83	0,12	10,28
f13	238	3,16	0,14	7,92
f14	184	2,05	0,11	10,56
f15	198	2,70	0,16	7,48
f16	206	1,20	0,10	11,25

Már a zöngére jellemző átlagos adatok is jelzik az életkori különbségeket. Az alaphérfekvencia átlagos értékei a fiataloknál tendenciaszerűen magasabbak, mint az idősek beszédében, de két fiatal beszélőnél igen alacsony F_0 -értékeket adatloltunk. A legnagyobb eltérést a jel/zaj viszony értékeiben találjuk a két csoport között: az idősek beszédében alacsonyabb értéket mutat, azaz az ő beszédük zörejesebb, mint a fiataloké. (Minél alacsonyabb a jel/zaj viszony értéke, annál jobban elfedi a levegőáramlási zörej a zöngét.) A jitter átlagos értékei az idősebbeknél a magasabbak; ez azt jelenti, hogy a zöngéképzésük „szabálytalanabb”, mint a fiataloké. Az egytényezős varianciaanalízis eredményei mind a négy értékben szignifikáns különbséget mutatnak az idősek és a fiatalok között: az F_0 tekintetében $F(1, 3496) = 345,904$; $p \leq 0,001$ a jel/zaj viszony esetén $F(1, 3613) = 913,893$; $p \leq 0,001$; a jitter esetén $F(1, 3911) = 264,604$; $p \leq 0,001$; a shimmer esetén $F(1, 3887) = 8,252$; $p = 0,004$.

Bár a két életkori csoport adatai szignifikánsan különböznek a zöngé akusztikai sajátosságainak tekintetében, nagyok az egyéni eltérések. Az idősek között vannak a zöngéképzés szempontjából kimondottan „fiatalos hangúak” (például i4), és a fiatalok között is vannak olyanok, akiknek a zöngéképzése „szabálytalanabb” a kortársaikénál (f13 és f15).

A két életkori csoportban az egyes hangokra jellemző formánsok átlagértékeit a 3. táblázatban, a formánsok szóródását a 4. táblázatban foglaltuk össze. Az F_1 átlagos értékei (a nyelv függőleges mozgására vonatkozó adatok) – a legalsó nyelvállású magánhangzó kivételével – a fiataloknál magasabbak. Az F_2 -értékek (amelyek a nyelv vízszintes mozgásától függnek) a fiatalok ejtésében a beszédhangtól függően vagy magasabbak, vagy közel azonosak az időseknél adatolt értékekkel; a felső nyelvállású hátul képzett magánhangzó ([u]) esetében az átlagos F_2 alacsonyabb a fiataloknál, mint az időseknél. Az F_3 átlagértékeit tekintve csak az [i, ø, u] hangok ejtésében találunk jelentős különbséget a két csoport között.

3. táblázat: A formánsok átlagértékei a két életkori csoportban

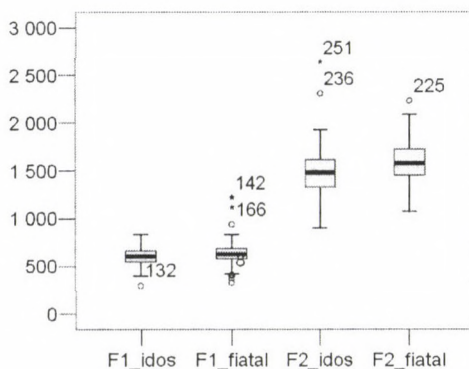
Magánhangzó	Idősek			Fiatalok		
	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
[ɔ]	606	1474	2790	639	1593	2806
[a:]	738	1664	2820	725	1876	2812
[ɛ]	608	1917	2853	628	1941	2882
[e:]	437	2288	2936	474	2182	2916
[i]	399	2231	2931	437	2063	2829
[o]	480	1430	2727	534	1475	2746
[ø]	466	1790	2688	480	1818	2717
[u]	413	1663	2672	451	1535	2737

4. táblázat: A formánsok szórása a két életkori csoportban

Magánhangzó	Idősek			Fiatalok		
	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
[ɔ]	300–838	906–2637	1862–3655	333–1225	1075–2230	2070–3459
[a:]	375–940	1075–2110	2211–3610	384–1029	1014–2303	1901–3394
[ɛ]	373–828	1531–2383	2312–3668	363–1011	913–2374	2159–3553
[e:]	257–567	1538–2958	2280–3781	332–986	831–2809	2274–3381
[i]	335–622	1492–2625	2324–3457	259–1121	895–2791	2258–3348
[o]	324–659	791–2691	2189–3749	351–847	877–2170	2053–3644
[ø]	410–559	1598–1955	2297–3103	382–603	1431–2104	2216–3237
[u]	342–510	1049–2356	2225–3259	355–1620	851–2189	2177–3262

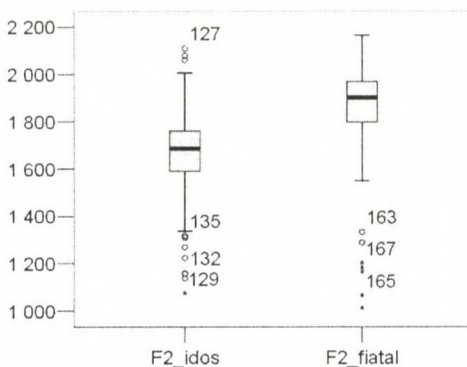
A magánhangzók minősége jelentősen befolyásolta azt, hogy statisztikailag volt-e eltérés a két csoportban mért formánsértékek között. A hátul képzett magánhangzók közül az [ɔ] hang esetében az F_1 és F_2 (1. ábra), az [a:] magánhangzó formánsai közül az F_2 (2. ábra), az [o] hang ejtésében pedig az F_1 értékei között találtunk szignifikáns eltérést (3. ábra). Az [u] formánsérté-

keiben nincs szignifikáns különbség az idősek és a fiatalok között. A statisztikai elemzés szerint az [ɔ] magánhangzó F_1 értéke esetén az idősek és a fiatalok között $F(1, 735) = 26,412, p \leq 0,001$; az F_2 esetén $F(1, 735) = 58,014, p \leq 0,001$. Az [a:] magánhangzó F_2 -értékeire vonatkozóan $F(1, 392) = 151,532, p \leq 0,001$. Az [o] hang ejtésében pedig az F_1 értékei között $F(1, 496) = 97,762, p \leq 0,001$.



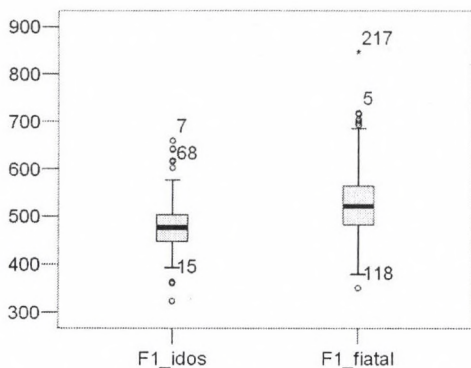
1. ábra

Az [ɔ] magánhangzó F_1 és F_2 értékeinek szórása (Hz)



2. ábra

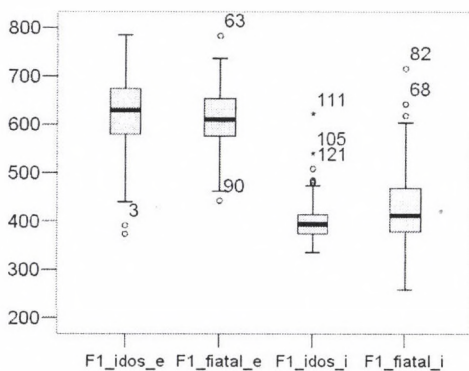
Az [a:] magánhangzó F_2 értékeinek szórása (Hz)



3. ábra

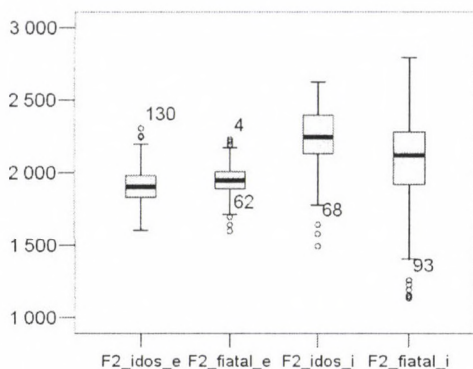
Az [o] magánhangzó F_1 értékeinek szórása (Hz)

Az elől képzett magánhangzók közül az [ε] és az [i] magánhangzók mindhárom formánása szignifikánsan különbözik az idősök és a fiatalok ejtésében (4., 5. és 6. ábra). Az [ε] esetében F_1 -re $F(1, 1087) = 15,897, p \leq 0,001$; F_2 -re $F(1, 1087) = 6,884, p = 0,009$; F_3 -ra $F(1, 1087) = 4,476, p = 0,035$. Az [i] esetében F_1 -re $F(1, 311) = 20,695, p \leq 0,001$; F_2 -re $F(1, 311) = 22,998, p \leq 0,001$; F_3 -ra $F(1, 311) = 16,356, p \leq 0,001$. Az [e:] magánhangzó esetében az F_1 és az F_2 értékek (7. ábra) mutattak szignifikáns eltérést az idősök és a fiatalok között: F_1 -re $F(1, 498) = 56,173, p \leq 0,001$; F_2 -re $F(1, 498) = 18,252, p \leq 0,001$. Az [ø] formánsértékeiben nem találtunk szignifikáns eltérést a két életkori csoportban.



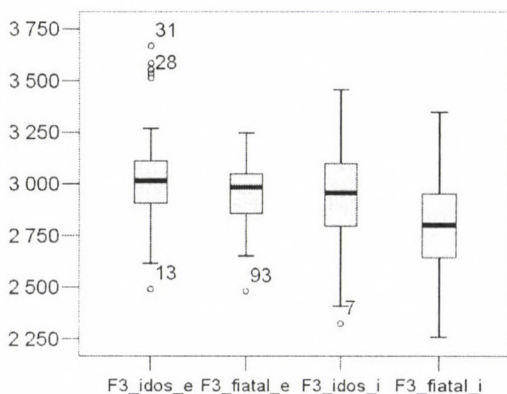
4. ábra

Az [ε] és az [i] magánhangzók F_1 értékeinek szórása (Hz)



5. ábra

Az [ɛ] és az [i] magánhangzók F₂ értékeinek szórása (Hz)

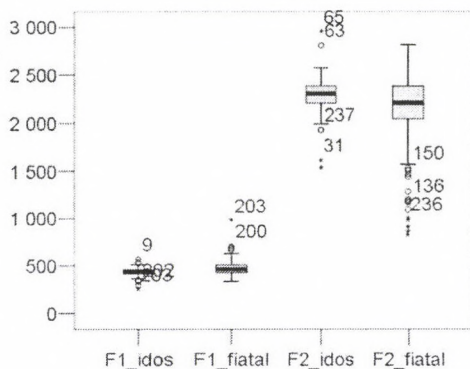


6. ábra

Az [ɛ] és az [i] magánhangzók F₃ értékeinek szórása (Hz)

Az egyes beszélők között nemcsak a zöngeminőség, hanem a formánsok tekintetében is nagy különbségek voltak mindkét csoporton belül. Az egytényezős varianciaanalízis szerint mindegyik vizsgált magánhangzó (kivéve az [ø] és az [u]) mindhárom formánsában szignifikáns eltérést találtunk a beszélők között (minden esetben $p \leq 0,001$). A Tukey post hoc tesztek eredményei szintén azt mutatták, hogy voltak idős beszélők, akiknek a formánsai majdnem minden fiatalétól szignifikánsan különböztek, és voltak olyanok is, akiknél csak kis számban találtunk szignifikáns eltérést a másik életkori csoport beszélőitől. Összevettük a legidősebb (80 éves) és a legfiatalabb (20

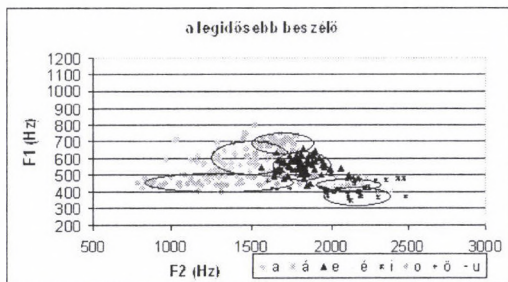
éves) beszélő magánhangzóinak első és második formánsait. A 8. és a 9. ábrán ezeket az értékeket ábrázoltuk, jelöltük a sűrűsödési ellipsziseket is.



7. ábra

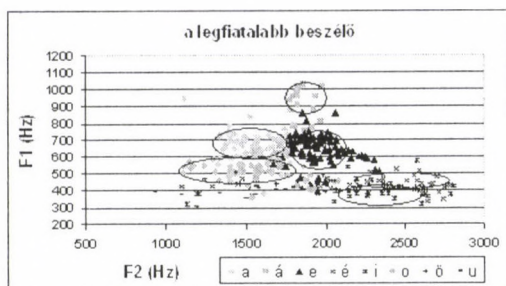
Az [e:] magánhangzó F_1 és F_2 értékeinek szórása (Hz)

Az adatok között vizuálisan is igen nagy az eltérés. Bár mindkét beszélő formánsértékei igen széles tartományban szóródnak, a fiatal beszélő magánhangzóit jobban elkülönülnek egymástól, mint az idős beszélőéi. Az idős beszélő sokkal centralizáltabban képi a magánhangzóit, amelyek a nyelv függőleges mozgása szerint (F_1 értékek) lényegesen kisebb különbségeket mutatnak a fiatal beszélőnél. A nyelv vízszintes mozgására vonatkozó adatok (F_2 értékei) azt mutatják, hogy az idős beszélő magánhangzóit valamivel hátrébb képzetebbek, mint a fiatalabb beszélő hangjai.



8. ábra

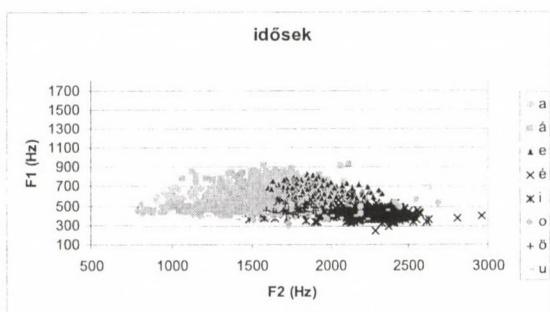
A legidősebb beszélő magánhangzóinak formánsai



9. ábra

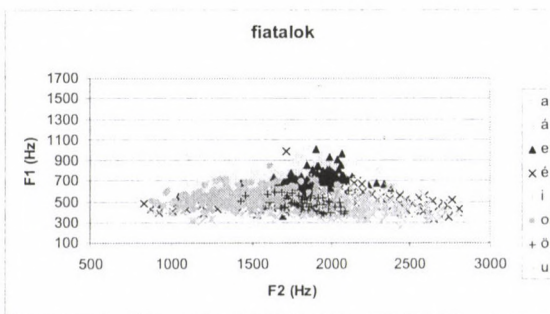
A legfiatalabb beszélő magánhangzóinak formánsai

Ezt az összevetést elvégeztük minden idős és minden fiatal beszélő összes magánhangzójára is (10. és 11. ábra), amelyben hasonló eredményre jutottunk.



10. ábra

Az összes idős beszélő formánsai



11. ábra

Az összes fiatal beszélő formánsai

Összegzés

Vizsgálatunkban nyolc 70 évesnél idősebb női beszélő magánhangzóinak akusztikai szerkezetét vetettük össze nyolc 20–32 éves női beszélő magánhangzóival. Kiinduló hipotézisünk, amely szerint mind a zöngéképzésben, mind a formánsértékekben különbséget találunk a két életkori csoport között, igazolódott. Az időskori hormonális változások hatására a zöngemínőség jelentősen megváltozik, a hang erőtlenebb, zöreijesebb (ezt bizonyítják a jel/zaj viszony alacsonyabb értékei), „szabálytalanabb” lesz (amint azt a jitter és a shimmer eltérései is mutatják) – ezt statisztikai adatokkal is alátámasztottuk.

A nyelv mozgásának lassulása, nehézkesebbé válása pontatlanabb artikulációhoz vezet, ezt az objektív akusztikai adataink, azaz az F_1 , F_2 és F_3 értékei is bizonyították. A vizsgált magánhangzók közül a legritkább előfordulású hangokban [ø, u] nem találtunk különbséget a két csoport között. Ennek valószínűleg elsősorban az az oka, hogy ez a két magánhangzó igen kis számban fordult elő az elemzett anyagban. A többi beszédhang esetében – a magánhangzó képzési jegyeitől függően – találtunk szignifikáns eltérést az egyes formánsértékek között. A nagy egyéni különbségek ugyanakkor azt jelzik, hogy az életkori változások eltérő módon jelentkezhetnek az egyes beszélők ejtésében.

Irodalom

- Balázs Boglárka 1993. Az időskori hangképzés jellemzői. *Beszéd kutatás '93*. 156–165.
- Boersma, Paul – Weenink, David 1998. *Praat: doing phonetics by computer (Version 5.0.1)*. http://www.fon.hum.uva.nl/praat/download_win.html.
- Gocsál Ákos 2000. A beszéd időviszonyai különböző életkorú személyeknél. *Beszéd kutatás 2000*. 39–50.
- Gósy Mária 1997. A magyar beszéd tempója és a beszédmegértés. *Magyar Nyelvőr* 121. 129–139.
- Gósy Mária 2000. Az életkor hatása a mentális lexikon működésére. *Magyar Nyelvőr* 124. 410–423.
- Gósy Mária 2001. Szóasszociációs műveletek az életkor függvényében. *Alkalmazott Nyelvtudomány* 1/1. 17–30.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszéd kutatás 2008*. 194–207.
- Gósy Mária – Nikléczy Péter 2000. Az idő változásának és a beszéd állandóságának paradoxona. *Beszéd kutatás 2000*. 132–143.
- Horváth Viktória 2006. The tip of the tongue phenomenon with elderly. In Lengyel, Zsolt – Navracsics, Judit (eds.): *Selected papers of 8th Summer School of Psycholinguistics*. Veszprém. (CD.)
- Kiss Jenő 2002. *Társadalom és nyelvhasználat*. Nemzeti Tankönyvkiadó, Budapest.
- Libárdi Péter 2001. A nyelvi változás tükröződése az egyén mentális lexikonában. *Alkalmazott Nyelvtudomány* 1/1. 39–49.
- Menyhárt Krisztina 2000. A beszéd temporális sajátosságai kétnyelvűeknél (kisiskoláskortól időskorig). *Beszéd kutatás 2000*. 51–62.

- 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.
- Torre, Peter – Barrow, Jessica A. 2009. Age-related changes in acoustic characteristics of adult speech. *Journal of Communication Disorders* 42. 324–333.
- Watson, Peter J. – Munson, Benjamin 2007. A comparison of vowel acoustics between older and younger adults. In Trouvain, Jürgen – Barry, William J. (eds): *Proceedings of the 16th International Congress of Phonetic Sciences*. 561–564.

Köszönöm Beke Andrásnak a mérések automatizálásában nyújtott segítségét.
A kutatás a Bolyai János Kutatási Ösztöndíj támogatásával készült.

STIGMATIZÁLT HANGLEJTÉSFORMA A SPONTÁN BESZÉDBEN

Markó Alexandra

Bevezetés

A nyelvi viselkedés rendszeresen válik bíráló megjegyzések célpontjává. A magyar beszéd hangtani, s ezen belül szupraszegmentális sajátosságai közül ugyanakkor viszonylag kevés olyan van, amellyel kapcsolatban gyakoriak a kritikák, bizonyos formák helytelennek/rossznak minősítése. Ebben a tekintetben a leggyakrabban azt a jelenséget minősítik és kommentálják a véleményformálók, amelyet úgy írnak körül, hogy a beszélő „a mondat végén felviszi a hangsúlyt”. Ez a közbeszédbeli állítás általában a spontán beszédre vonatkozik, és a záró dallamszakasz emelkedő/szökő voltára utal (valójában tehát nem a hangtani értelemben használatos *hangsúly*-ként értelmezhető – függetlenül attól, hogy a hangsúlyérzetet befolyásolja az alaphangmagasság realizációja).

A fenti jelenség erősen stigmatizált a közbeszédben. Ennek nyomait a virtuális forrásokból nagy számban adatolhatjuk – néhány példa az internetről (blogokból, fórumhozzászólásokból): „sajnos [...] valóban magyartalanul mondja a szöveget [...], a mondat végén felviszi a hangot, ahelyett, hogy levinné”; „szerintem is borzasztó [...], állandóan felviszi a hangsúlyt a mondat végén”; „ellentmond a magyar nyelv helyes hangsúlyozásának, ha valaki a mondat végén felviszi a hangsúlyt, ezért nagyon zavaró azok számára, akiknek még fontos a helyes magyar beszéd”.

Ugyancsak gyakori, hogy ezt a beszéd-sajátosságot beszélői csoportokhoz kötik. Jellemzően a nők beszédével kapcsolatosan hangzik el ez a bíráló megjegyzés (a fenti, internetről származó példák mind nők beszédét minősítik). Egy rögtönzött (interneten közzétett kérdőívvel végzett) közvélemény-kutatásban 20 adatközlő (10 férfi, 10 nő, értelmiségiek, különböző végzettségűek és foglalkozásúak; az életkoruk 24 és 64 év között szóródott, átlaga 29,5 év) közül 12 szerint a nőkre jellemző(bb), hogy „a mondat végén felviszik a hangsúlyt”, a többiek szerint mindkét nemre vagy egyikre se jellemző, illetve volt, aki nem válaszolt; de egyetlen olyan válasz sem érkezett, amely szerint a férfiakra lenne jellemzőbb ez a dallamhasználati mód.

Természetesen a nyelvészeti, hangtani szakirodalomban is megjelennek a közlés végi dallamemelésre vonatkozó (néha minősítő, máshol csak konstataáló) megjegyzések. A magyar köznyelvi beszéddallam vizsgálatában hosszú ideig uralkodó volt a felolvasott (esetleg eljátszott) mondatok és szövegek

elemzése, ezek mellett azonban a 20. század eleje óta több közlemény látott napvilágot a spontán beszéd hanglejtéséről. Mintegy az ezredfordulóig a hallási alapú leírások voltak a jellemzők, ez után egyre nagyobb számban jelennek meg objektív műszeres vizsgálatok eredményei is. A következőkben a magyar (főként) spontán beszéd dallamát vizsgáló fontosabb munkákat tekintjük át, kiemelve azokat az állításokat, amelyek akár a hivatkozott jelenséggel, akár a női-férfi dallamhasználat különbségével foglalkoznak. A bemutatás sorrendje alapvetően kronologikus, de a tematikusan érintkező forrásokat egy helyen tárgyaljuk az esetleges időbeli távolság ellenére.

A hatvanas évek a magyar szupraszegmentális fonetika és ezen belül a dallamvizsgálatok fellendülésének időszaka. Az évtized elején jelent meg Elekfi László munkája (*Vizsgálatok a hanglejtés megfigyelésének módjaihoz*, 1962), amely bizonyos értelemben módszertani kézikönyv a hanglejtésvizsgálatokhoz. Fontosnak tartjuk megemlíteni, hogy az ugyancsak ekkor megjelent kétkötetes akadémiai leíró nyelvtan (Tompá szerk. 1961–1962) a grammatika szerves részeként tárgyalja a szupraszegmentumok (hangsúly, hanglejtés, szünet) nyelvileg funkciós formáit (az e tárgyú fejezetek Deme László munkái). A hanglejtés elemzésében, minősítésében elterjedté válik az a módszer, hogy a frekvenciaváltozás iránya és meredeksége alapján ötféle hanglejtésformát különböztetnek meg: eső, ereszkedő, lebegő, emelkedő, szökő (vö. pl. Deme 1962) – elsősorban percepciós alapon.

Fónagy Iván és Magdics Klára munkája *A magyar beszéd dallama* (1967) című (hallási alapú elemzésekre épülő) monográfia, amely többek között spontánbeszéd-adatokat is tartalmaz (bár ezek nem mindig különülnek el egyértelműen az „eljátszott” mintáktól). A szerzők elsőként vetik fel, hogy vannak „nőies” és „férfias” dallamformák: percepciós tesztek alapján „női-es”-nek minősítették a szöközárral (terc vagy kvart fellépéssel) végződő kijelentő mondatokat és a magas szinten hullámzó (ereszkedő-emelkedő) felkiáltó mondatokat. A Fónagy(ék)-féle elkülönítésben nincs értéktétele, ezeket a nemek szerint eltérő formákat semlegesnek tartják.

A magyar szakirodalomban a spontán beszéd akusztikus jellemzőiről Wacha Imre közölt először átfogó leírást (1974). Elekfi és Wacha közösen írt monográfiája (é. n. [2003]) ugyancsak tartalmaz a spontán beszédéről szóló fejezetet. A spontán beszéd dallamstruktúrájára vonatkozó hallásalapú elemzések jellemző jegyeket találják a „befejezetlenséget jelölő gyakori szöközáras formák”-at, de ezeket nem kötik a beszélő neméhez (vö. Wacha 1999, Elekfi–Wacha é. n. [2003]).

A *Nyelvművelő kézikönyv* jelentős terjedelmet szentel a hanglejtésbeli hibáknak (Elekfi 1983), amelyek között megemlíti a glottalizáció jelenségét (amikor a beszélő „hangja kellemetlen nyekergésfélébe vész el” – i. m. 774) és a szakasz végi dallamemelést, amelyet „énekítő”, „pesties” stb. beszédnek is nevez (i. m. 775). Boros Rezső (1971) és mások ugyancsak hibáztatják ezt a „pesties”-nek nevezett dallamformát (a további forrásokra vö. Gósy 1993).

Az eddig áttekintett közlemények többnyire általános megállapításokat fogalmaznak meg a spontán beszéd dallamáról. A beszéddallam kísérleti vizsgálatán (is) alapuló források nagyobbik része nem a spontán beszédet, hanem felolvasott, elicitált mondatokat, szövegeket állítanak a kutatás fókuszába. A mondatfajták és beszédaktusok szupraszegmentális leírása több tanulmány tárgya volt (pl. Bolla 1992, Olaszky 1995, Gósy 2000), különösen a kérdő mondatformákkal kapcsolatos tanulmányok készültek nagy számban (Gósy 1993, Olaszky 2002, Varga 2002a). Fonológiai keretben született meg Hunyadi (2002) és Varga (2002b) összefoglaló monográfiája.

Ezzel szemben a spontán beszéd dallamára (is) irányuló (kevésnek mondható) eddigi műszeres, korpuszalapú vizsgálatok konkrét kérdésekből-hipotézisekből indul ki, így a tanulmányok a spontán beszéd hanglejtésének egy-egy sajátos vonatkozását tárgyalják. Gósy Mária (2003) az ún. virtuális mondatok szerepét vizsgálja a spontán beszéd percepciójában, s a szerző számot ad a hanglejtés megvalósulásának sajátosságairól is. A virtuális mondatok műfajfüggő akusztikai és percepciós jellemzőit elemezte Váradi (2008), és azt találta, hogy a narratív és leíró szövegekben eltérő a prozódiai határjelzés. Mindkét kutatásban nagy számban adatoltak a virtuális mondatok végén emelkedő dallammenetet (csak női adatközlőtől származó szövegek alapján, de nem merült fel a nemi különbség kérdése).

Közel kétszáz spontán kérdő megnyilatkozás elemzésének eredményei alapján megállapítottuk, hogy ezek sokszínűbbek, mint amilyenek a felolvasott, elicitált vagy eljátszott mintapéldáik alapján a szakirodalom bemutatja őket. A változatosság elsősorban az F_0 modulációjának lehetőségeiben, másrészt pedig a hangköz értékeiben mutatkozik meg (Markó 2007).

Felolvasott és spontán szövegek dallami realizációnak összevetését végezte el Beke (2008), a teljes közlés hanglejtését vizsgálta, általánosító célzattal, így a szakasz végi dallami sajátosságokról nem közöl adatokat.

Négy beszélő (két férfi, két nő) különböző műfajú spontánbeszéd-produkcióiban (monologikus, dialogikus helyzet, képsorozat leírása) a közlés végi dallamrealizációk között eltérő arányban adatoltunk emelkedő/szökő változatokat: a legritkábban a társalgásban, a leggyakrabban (a legnagyobb kognitív erőfeszítést követelő) képleírásban. Az adatközlők kis száma miatt az egyéni különbségek felülmúlták a nemek szerinti eltéréseket, de az auditív és vizuális elemzés együttese alapján a nőknél gyakrabban tapasztaltunk szökőzárt, mint a férfiaknál (Markó 2007–2008).

Összegezve a szakirodalombeli megállapításokat: a spontán beszéd dallami sajátosságai között a legtöbb forrás megemlíti a szakasz/szólam/mondat végi frekvencianövekedést. A szubjektív vizsgálatok körében ezt a jelenséget gyakran szűkebb tartalmú terminussal, *szökőzár*-ként illetik, azaz nem pusztán növekedő, hanem meredeken (rövid idő alatt nagymértékben) emelkedő frekvenciáról beszélnek. A nyelvművelő szándékú munkák hibáztatják ezt a

formát. Bizonyos források (akárcsak a közvélekedés) különösen gyakran tartják a nők beszédében.

Ezek a tapasztalatok motiválták a jelen, korpuszalapú kutatást, amelyben a megnyilatkozás végi frekvencianövekedéssel kapcsolatban három kérdésre keressük a választ: 1. Valójában (objektíve) mennyire gyakori az emelkedő/szökő dallamzár a spontán beszédben? 2. Igazolható-e a szökőzárként való gyakori megjelenés, azaz valóban meredeken emelkedő dallamformák realizálódnak-e a megnyilatkozások végén? 3. Jellemzőbb-e a nők beszédére akár az emelkedő, akár a szökő zárlat?

A kutatás akusztikai elemzések és percepciók ellenőrző tesztek sorozatából állt. Arra törekedtünk, hogy ne egyszemélyi szubjektív ítéletek, hanem validált adatok alapján adjunk választ a megfogalmazott kérdésekre.

Kísérleti személyek, anyag és módszer

A vizsgálat anyagát az ún. BEA magyar beszélt nyelvi adatbázisból válogattuk. (Az adatbázis az MTA Nyelvtudományi Intézetének Fonetikai Osztályán készül, jellemzőiről lásd Gósy 2008.) A beszélőket úgy választottuk ki, hogy három, jól elkülöníthető korosztályt reprezentáljanak, egyenlő nemi arányban. A különböző végzettségű beszélőket koronként és nemenként illesztettük, arra ügyeltünk, hogy ún. hivatásos vagy vélhetően tudatos beszélő (pl. tanár, nyelvész, színművész) ne legyen a kiválasztott adatközlők között. Mindezen paraméterek alapján az adatbázisból 12 beszélő (három korosztályból 2-2 nő és férfi) hanganyaga bizonyult alkalmasnak a vizsgálatra (a válogatás időpontjában). A fiatal beszélők életkora 20–22 év, a középkorúaké 39–45 év, az idősebbeké pedig 57–64 év között szóródik.

Minden beszélőtől kiválasztottunk egy körülbelül 4 perces interjúrészletet, amelyben főleg magukról (munkájukról, hobbijukról) beszélnek, valamint egy mintegy 2 perces felolvasást, amely egy (a cím nélkül) 12 mondatot tartalmazó ismeretterjesztő szöveg meghangosítása. A spontán szövegrészleteket lejegyeztük, majd mindegyiket lejátszottuk öt-öt, a kutatás célját nem ismerő személynek. A minősítésben 28 nyelvész és (elsősorban magyar szakos) bölcsészhallgató vett részt, életkoruk 18–35 év között szóródott. Az volt a feladatuk, hogy jelöljék a központosítás nélkül lejegyzett szövegben a megnyilatkozások végét (a módszerre lásd Gósy virtuálismondat-kísérletét, 2003). Egy-egy minősítőnek 1–5 szöveget játszottunk le, számítógépről, aktív hangfal segítségével – egyenként és csoportosan is végeztünk tesztelést. Azokat a szöveghelyeket vettük tekintetbe a további vizsgálatban, amelyeknek az esetében ötből legalább négyen megnyilatkozásvéget jelöltek; illetőleg automatikusan megnyilatkozásvégeként kezeltük a beszédlépéseket¹ záró közléseket (amelyek után megszólalt az interjúkészítő).

1. Társalgási egység, forduló („turn”): egy beszélő által egyben elmondott szövegegység.

Mind a spontán, mind a felolvasott szövegben megvizsgáltuk a megnyilatkozásokat záró intonációs egységeket a Praat 5.0 (Boersma–Weenink 1998) segítségével a zöngeminőség és a frekvenciaváltozás tekintetében. A hangszalagműködést vizuális és auditív jellemzők együttese alapján minősítettük normálnak (modal), illetve glottalizáltnak (creaky). Glottalizáltnak tekintettük azt a szakaszt, amelyben a dallamgörbe eltorzul (megszakad, vagy jelentősen a beszélő jellemző hangterjedelme alá csökken), és a periódusok láthatóan megritkúlnak, miközben hallhatóan megváltozik a hangszínezet (a glottalizáció vizuális megjelenésére lásd Böhm–Ujváry 2008: 108–109, 1. és 2. ábra).

Csak a normál (kvázireguláris) hangszalagrezgéssel megvalósuló megnyilatkozásvégek alaphangmagasságát adatoltuk. A hanglejtésben jelentkező utolsó töréspontonál és a közlés végén megmértük a frekvenciát, valamint megmértük ennek a szakasznak az időtartamát. Arra vonatkozóan, hogy az alaphangfrekvencia-változás objektíven hogyan adatolható, több módszert is találunk a szakirodalomban.

Fónagy és Magdics (1967) zenei hangközöknek feleltetik meg a beszédbeli frekvenciaváltozások mértékét. A zenei hangközök a frekvenciaértékek arányát adják meg – például a kvint hangköz esetében a frekvenciaértékek 2 : 3 arányban viszonyulnak egymáshoz, így a hangköz megadható a két érték hányadosaként: 1,5. Ugyanígy a kis terc 1,2, az oktáv hangköz 2,0 stb. Ezt a módszert némi módosítással úgy alkalmaztuk, hogy mindig a Hz-ben megadott záró és kezdő frekvenciaérték hányadosaként határoztuk meg a megnyilatkozás végi intonációs egység hangközét: záró érték/kezdő érték. Ezáltal a hányados az alaphangmagasság-változás irányát is jelzi: ha az értéke 1,00-nél nagyobb, a hanglejtés emelkedő jellegű, ha az érték 1,00-nél kisebb, ereszkedő típusú a hanglejtés. Például: ha a frekvencia az utolsó szakaszban 199-ről 228 Hz-re emelkedik, a két érték hányadosa ($228/199$), azaz a hangköz 1,15; ha 264 Hz-ről 186-ra csökken az alaphangfrekvencia, a hangköz $186/264 = 0,70$. Ezzel a mérési módszerrel tehát csak 1,00 hangközérték esetében minősítettük lebegőnek a záró intonációs egységet, szemben más irodalmi forrásokkal (Kassai 1998: 213 alapján – az alaphangmagasságtól függően – nagyjából 10%-nál kisebb frekvenciaeltérés esetén a dallam változatatlannak minősül). Ez, a frekvenciaértékek arányát megállapító számítási mód a frekvenciaváltozás időtartamára nincs tekintettel.

Egy másik számítási módszert sugall Gósy (1989) vizsgálata, amelyben a dallam lefutását az időegység alatt lezajlott frekvenciaváltozással összhangban minősítették az adatközlők. Ez alapján kiszámítottuk tehát, hogy az egyes megnyilatkozásvégek esetében milyen relatív frekvenciakülönbség mutatható ki, azaz a maximális és a minimális frekvenciaérték (Hz) különbségét elosztottuk a közlésrészlet másodpercben mért időtartamával. Az így kapott értékek két példán illusztrálva ekképpen értendők: Ha a frekvencia és az időtartam hányadosa 100, az olyan realizációra vonatkozhat, amelyben 500 ms alatt 200 Hz-ről 250 Hz-re 50 Hz-nyi növekedés történt ($50 \text{ Hz}/0,5 \text{ s} = 100$). Ha a

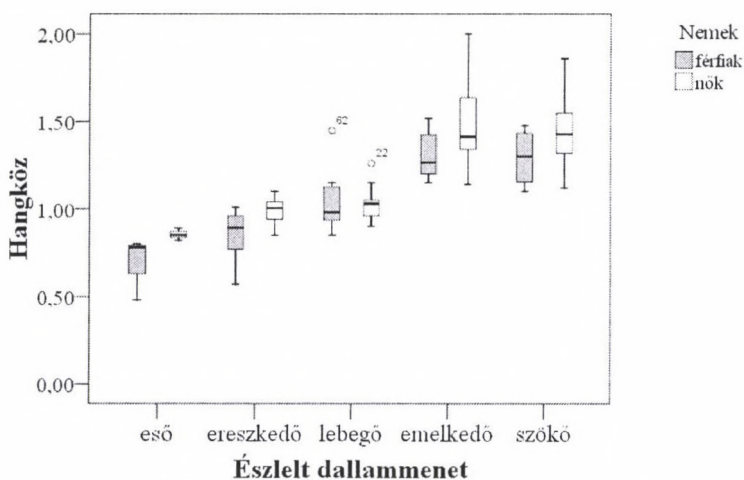
hányados -150 , az azt jelentheti, hogy 200 ms alatt 120 Hz-ről 90 Hz-re 30 Hz-nyit csökkent a frekvencia ($-30 \text{ Hz}/0,2 \text{ s} = -150$). Az előjel tehát azonosíthatóvá teszi a frekvenciaváltozás irányát, illetőleg ha a frekvenciakülönbség értéke 0 („tökéletesen lebegő” dallam esetén), a mérőszámként használt hányados is 0 lesz.

Hagyományosan a frekvenciaváltozás iránya és meredeksége alapján ötféle hanglejtésformát szokás megkülönböztetni: eső, ereszkedő, lebegő, emelkedő, szökő. Ugyanakkor nincs arra vonatkozó forrás, hogy milyen objektív értékek vagy sávok határolják el egymástól a különböző hanglejtésformákat, például milyen hangközértékig tekinthető emelkedőnek a hanglejtés, és honnan számítjuk szökőnek. Ily módon ez a minősítés valamilyen mértékig mindenkor szubjektív, különösen mivel általában egy személy percepcióis ítéletét tükrözi. Gósy Mária az említett kísérletében 100 ms alatt történt 100 Hz-es változásig tartotta emelkedőnek/ereszkedőnek a dallamot, ennél nagyobb változást minősített szökőnek/esőnek – ahogy fogalmazott – „önkéntesen” (1989: 168). Ez az elkülönítés számunkra azonban nem követhető, mert ez alapján a korpuszunkban sem szökő, sem eső típusú dallamforma nem jelentkezik a megnyilatkozások végén. Ahhoz, hogy a szakaszvégi dallamemelésre vonatkozó megállapításokat (például a szökőzár gyakoriságáról és nemfüggő voltáról) igazolni vagy cáfolni tudjuk, független, valamilyen mértékig objektivizálható hanglejtés-minősítésre van szükség.

Erre a célra alkalmaztuk a következő minősítési eljárást. 116 darab normál zöngével megvalósult záró szakaszt a szövegkörnyezet nélkül kivágtunk a spontánbeszéd-felvételekből. A válogatásban minden beszélő közel azonos arányban (9 - 10 mintával) szerepelt, a női és a férfi adatközlők mintái összekeverten jelentek meg. A közlésrészleteket véletlenszerű sorrendben lejátszottuk 10 , beszéddel foglalkozó személynek (nyelvész, mérnök), akiket arra kértünk, hogy minősítsék azokat eső, ereszkedő, lebegő, emelkedő vagy szökő dallamúnak. A hangfelvétel elején 8 , különböző beszélőktől származó közlésminta hangzott el bemelegítésképpen, így a hallgatóknak volt lehetőségük hozzászokni a feladathoz, a beszélők hangjához, és optimalizálhatták a lejátszási hangerőt. A bemelegítésben szereplő minták az éles tesztben különböző helyeken megismétlődtek; a teszt eleji előfordulásokat a kiértékelésnél természetesen nem vettük figyelembe. A tesztanyagot a hangerő szempontjából normalizáltuk, így a különböző beszélőktől származó minták intenzitása nem tért el egymástól (az intenzitáskülönbség tehát nem befolyásolhatta a dallam megítélését). A résztvevők tesztlapot kaptak, amelyen a helyesírási formában lejegyzett közlésrészletek szerepeltek egymás alatt, valamint mindegyik mellett az ötféle dallamminősítésre (eső, ereszkedő, lebegő, emelkedő, szökő) vonatkozó cella. Azt kértük az adatközlőktől, hogy az észlelt dallamtípusnak megfelelő cellába tegyenek ikszet. Az egyes közlésrészletek közé 5 másodperces szüneteket iktattunk be, ez alatt kellett a tesztlapon beikszelni az adott mintára vonatkozó ítéletet. A hanganyag teljes időtartama

mintegy 13 perc volt, a meghallgatását megismételheték az adatközlők, az ítéleteket javíthatták. A további feldolgozásban azokat a megnyilatkozásvégeket vettük tekintetbe, amelyeknek a hanglejtését legalább az adatközlők 60%-a azonosan minősítette, és a többi minősítés valamely szomszédos tartományba esett. Így végül 92 (43 női, 49 férfi beszédből való) részlet adatai szolgálták az objektív hanglejtés-minősítési eljárás kidolgozásához.

Az adatközlőknek az egy adott közlésrészletre adott ítéleteit (eső, ereszkedő, lebegő, emelkedő, szökő) először a hangközértékre vetítve vizsgáltuk meg. Az 1. ábrán láthatjuk, hogy az egyes minősítések között nem mindenhol húzhatók meg határok.



1. ábra

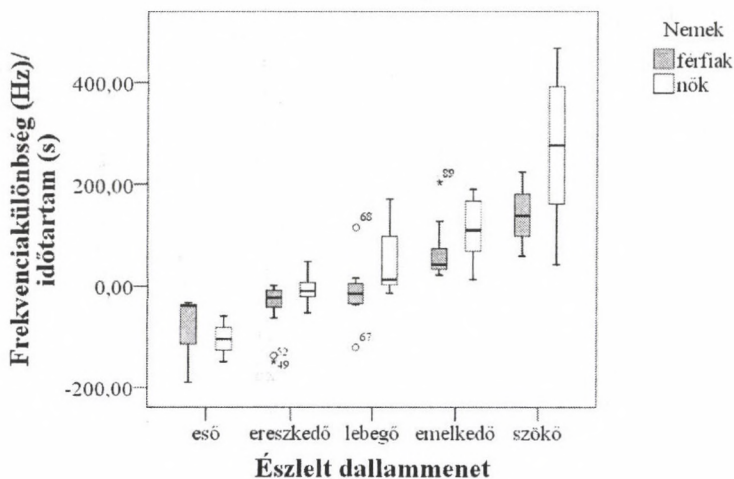
A dallammenet percepciósi minősítése a hangközérték függvényében (a 92 sikeresen azonosított szakaszvég alapján)

Bár az egytényezős varianciaanalízis mind a nők, mind a férfiak esetében szignifikáns eltérést mutat az egy-egy típusba sorolt dallammenetek között a hozzájuk tartozó hangközérték alapján [a nőknél $p < 0,001$, $F(4, 36) = 18,111$; a férfiaknál $p < 0,001$, $F(4, 36) = 22,081$], a Tukey-féle post hoc teszt csak az alábbi típusok között adott szignifikáns eredményt. A nőknél: eső vs. emelkedő, eső vs. szökő, ereszkedő vs. emelkedő, ereszkedő vs. szökő, lebegő vs. emelkedő, lebegő vs. szökő. A férfiaknál: eső vs. lebegő, eső vs. emelkedő, eső vs. szökő, ereszkedő vs. emelkedő, ereszkedő vs. szökő, lebegő vs. emelkedő (az 1. táblázatban sötéttel jelölt cellák). Természetesen ennek háttérében állhat az, hogy a meredekség érzete nem pusztán a hangközből, a lezajlott frekvenciaváltozásból adódik, hanem nagymértékben befolyásolja az is, hogy ez a változás milyen időtartam alatt történt meg.

1. táblázat: Az észlelt dallamtípusokhoz tartozó hangközértékek eltérése a Tukey-féle post hoc teszt alapján
(szignifikanciaértékek 95%-os szinten; a vonal fölött a férfiak, a vonal alatt a nők beszédében mért adatok)

p =	eső	eresz kedő	lebe- gő	emel kedő	szőkő	
eső		0,432	0,009	0,000	0,000	férfiak
eresz kedő	0,808		0,077	0,000	0,000	
lebegő	0,597	0,982		0,002	0,075	
emel kedő	0,000	0,000	0,000		1,000	
szőkő	0,000	0,000	0,000	1,000		
	nők					

Kiszámítottuk tehát, hogy az egyes megnyilatkozásvégek esetében milyen relatív frekvenciakülönbség mutatható ki, azaz a maximális és a minimális frekvenciaérték (Hz) különbségét elosztottuk a közlésrészlet másodpercben mért időtartamával. A 2. ábrán a percepciók minősítések és a frekvenciaváltozás/időtartam értékének összefüggését látjuk. Mindkét nem esetében szignifikáns a hanglejtéstípusok közötti variancia az ANOVA szerint: $p < 0,001$ [nők: $F(4, 39) = 22,577$; férfiak: $F(4, 43) = 14,381$]. A Tukey-féle post hoc teszt az alábbi vonatkozásokban adott szignifikáns eredményt a nőknél: eső vs. emelkedő, eső vs. szőkő, ereszkedő vs. emelkedő, ereszkedő vs. szőkő, lebegő vs. szőkő, emelkedő vs. szőkő. A férfiaknál: eső vs. emelkedő, eső vs. szőkő, ereszkedő vs. emelkedő, ereszkedő vs. szőkő, lebegő vs. emelkedő, lebegő vs. szőkő (a 2. táblázatban sötéttel jelölt cellák). A statisztika alapján tehát nem különül el markánsan az eső, az ereszkedő és lebegő dallam, illetve a nőknél a lebegő és az emelkedő, a férfiaknál pedig az emelkedő és a szőkő.



2. ábra

A dallammenet percepciós minősítése a frekvenciaváltozás/időtartam tekintetében (a 92 sikeresen azonosított szakaszvég alapján)

2. táblázat: Az észlelt dallamtípusokhoz tartozó frekvenciaváltozás/időtartam értékek eltérése a Tukey-féle post hoc teszt alapján (szignifikanciaértékek 95%-os szinten; a vonal fölött a férfiak, a vonal alatt a nők beszédében mért adatok)

p =	eső	eresz kedő	lebe- gő	emel kedő	szökő
eső		0,624	0,256	0,001	0,000
eresz kedő	0,383		0,782	0,000	0,000
lebegő	0,065	0,554		0,021	0,001
emel kedő	0,002	0,012	0,595		0,146
szökő	0,000	0,000	0,000	0,000	

férfiak

nők

Összegezve a percepciók teszt eredményeinek kétféle módszerrel történt kiértékelését: mind az egyszerű hangközszámitás (a legnagyobb és a legkisebb frekvenciaérték hányadosa), mind a tartamra vetített frekvenciakülönbség esetében összesen hat dallamtípuspár különböztethető meg a fenti statisztikai elemzéssel (a lehetséges tízből). A két módszer eredménye bizonyos mértékben különbözik abban a tekintetben, hogy mely párok adnak ki markáns eltérést. A fenti analízis azonban nem ad választ arra kérdésre, hogy a szomszédos dallamminőségek hogyan különíthetők el egymástól, azaz hol húzódik a határ az emelkedő és szökő, a lebegő és emelkedő stb. dallamminőségek között.

A szomszédos dallamminőségek elkülönítéséhez, a köztük lévő matematikai határ megállapításához ROC (Receiver Operating Characteristic) analízist alkalmaztunk. Az ROC görbe segítségével meghatározhatjuk, hogy egy osztályozó rendszerben egy ismertetőjegy milyen átlagos valószínűséggel tud két csoportot elkülöníteni a választott küszöbérték megadásától függetlenül (Beke megj.). Az analízis azt az eredményt adta, hogy az időtartamra vetített frekvenciakülönbség alapján elkülöníthetők a (jelen kutatás kérdésfeltevése szempontjából jelentős) lebegő-emelkedő és emelkedő-szökő típusú hanglejtések, míg a hangközértékek alapján nem. Az ROC analízis eredményeit a 3. táblázat foglalja össze. A határértékekben jelentős különbség látható a nemek között, azaz a nők tágabb hangköze a hanglejtésük minősítésében is szerepet játszik: a frekvenciakülönbség/tartam nagyobb hányadosától minősítik a hallgatók emelkedőnek, illetve szökőnek a nők hanglejtését, mint a férfiakét (ahogy a 2. ábra is mutatja).

3. táblázat: A szomszédos dallamminőségek elkülönülése az ROC analízis alapján

	Lebegő-emelkedő		Emelkedő-szökő	
	Nők	Férfiak	Nők	Férfiak
Az ROC analízis felismerési aránya	77%	89%	81%	85%
Kategóriahatár [frekvenciakülönbség (Hz)/tartam (s)]	33,84	16,70	189,86	127,49

A frekvenciakülönbség/időtartam szerint tehát a vizsgált dallamminőségekhez objektív értékskálát tudtunk rendelni, ami lehetővé tette, hogy a hagyományos terminusok szerint minősítsük a korpuszban megjelent megnyilatkozásvégek hanglejtését, és ellenőrizhessük a szakirodalmi állítások helyességét.

Az adatok feldolgozása és különböző szempontú összevetése t -próbával és egytényezős varianciaanalízissel (95%-os szignifikanciaszinten) történt az SPSS 16.0 verziójával.

Eredmények

A felolvasások esetében 144 (12×12) mondatból 141 volt elemzésre alkalmas, a kihagyott három esetben tévesztett a felolvasó, ezért spontán szöveget szúrt be, vagy elnevette magát. A spontán szövegekben a tesztelők egybehangzó ítélete és a beszédlépésváltások figyelembevétele alapján 245 megnyilatkozás volt azonosítható (szövegenként 12–34, beszélőnként átlagosan mintegy 20), ezek közül 231 esetben volt vizsgálható az alaphérfkvencia (egyszerre beszélés, nevetés, torokköszörülés miatt zártuk ki a többi megnyilatkozásvéget).

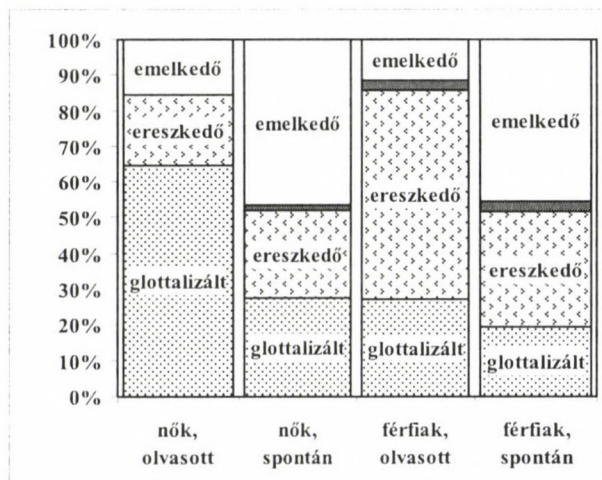
A zöngeminőség vizsgálata azt az eredményt hozta, hogy a glottalizáció beszélőfüggő (vö. pl. az amerikai angolra Dilley–Shattuck–Hufnagel 1995, a magyarra Böhm–Ujváry 2008), de ugyancsak nagymértékben függ a beszéd módtól. A felolvasásban mintegy kétszer akkora arányban adatoltunk glottalizált szakaszvégeket (46,10%), mint a spontán beszédben (23,81%). Volt olyan adatközlő, aki a felolvasásában egyszer sem, a spontán beszédében pedig egyszer (a megnyilatkozásainak 8,33%-a) glottalizált. Volt azonban olyan beszélő is, aki egy kivétellel minden felolvasott mondatot irreguláris zöngével zárt (91,67%), a spontán beszédére mégsem volt ennyire jellemző ez a sajátosság (itt csak 23,53%-ban jelent meg). Az az eredmény, hogy a felolvasásban a mondatok közel fele irreguláris zöngével végződött, arra utal, hogy a glottalizáció a határjelzés markáns eszköze a magyar beszélők körében (hasonló eredményeket kaptak a svédre, a brit és az amerikai angolra stb. – vö. Redi–Shattuck–Hufnagel 2001, Henton–Bladon 1988 stb.). Ezt a sejtést erősíti az is, hogy a beszédlépések végén az esetek mintegy felében tapasztaltunk glottalizációt (természetesen itt is beszélőfüggő: volt, aki soha, de olyan is, aki mindig így adta át a szót). Hasonló eredményeiket Böhm és Ujváry fiziológiai okokkal magyarázzák (2008: 116), további vizsgálatoknak kell azt kideríteniük, hogy a glottalizáció mely eseteiben számolhatunk okként pusztán a zöngékepzés biológiai hátterével, és mikor tekinthető pragmatikai funkciójúnak.

A nemek szerinti összehasonlításból az derül ki, hogy a vizsgálatban részt vevő nők felolvasására a legjellemzőbb a glottalizáció, az általuk irreguláris zöngével befejezett megnyilatkozások aránya (64,79%) több mint kétszerese akár a férfiak felolvasásában adatolt glottalizált megnyilatkozásvégeknek (27,14%), akár a nők spontán beszédében megjelenő közlés végi glottalizáció arányának (27,64%).

Mivel a glottalizáció egyik akusztikai következménye az alaphangmagasságban mérhető nagyfokú frekvenciacsökkenés, amely gyakran a dallam lépéseként érzékelhető, a megnyilatkozásokat lezáró alaphérfkvencia-változások vizsgálatában továbbra is figyelembe vettük az ilyen zöngeminőséggel jellemezhető szakaszokat.

Az olvasott és a spontán beszéd között jellegzetes eltéréseket találtunk az alaphérfkvencia realizációinak tekintetében (3. ábra). Míg a felolvasásban az

emelkedő dallam (azaz a Fónagyékat követő számítással az 1,00 értéknél nagyobb hangköz) előfordulási aránya nem haladja meg a 16%-ot, a spontán beszédben ez az arány mindkét nem esetében megközelíti az 50%-ot (nők: 46,34%, férfiak: 45,37%). A lebegő hanglejtés megjelenése a mérési módszerből adódóan (azaz amiatt, hogy csak az 1,00 hangközű szakaszokat soroltuk ide) minden beszédmódban és mindkét nemnél elhanyagolható (0,00–2,86%). A glottalizált és az ereszkedő hanglejtésű megnyilatkozásvégek együttese nemtől függetlenül mintegy 85%-ot ad ki az olvasásban, a spontán beszédben pedig 50% körüli arányt mutat. Ez alapján azt valószínűsíthetjük, hogy a szakaszvégi glottalizáció funkciója részben valóban azonos az ereszkedő hanglejtésével: határjelölő szerepet (is) tölt be. Annak a jelenségnek, hogy a nők esetében éppen a glottalizáció a domináns, szociolingvisztikai és fiziológiai okai is lehetnek – a kérdés mindenképpen további vizsgálatot igényel.



3. ábra

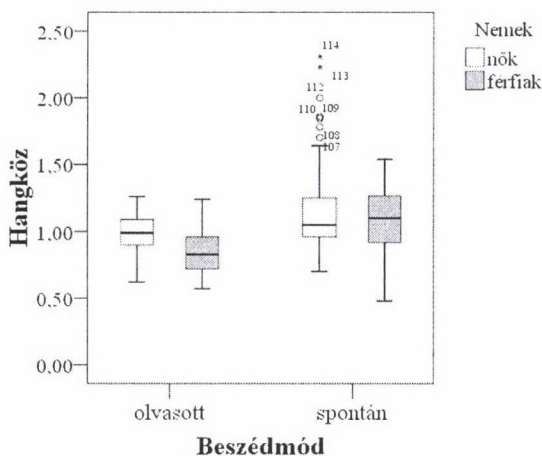
A megnyilatkozásvégi alapfrekvencia alakulása nemek szerint felolvasásban és spontán beszédben (a hangközértékek alapján)

A megnyilatkozás végi alapfrekvencia-változás összevetésének legfontosabb tanulságai a kutatási kérdések vonatkozásában úgy fogalmazhatók meg, hogy valóban jelentősen (a jelen vizsgálat adatai alapján mintegy háromszor) gyakoribb az emelkedő hanglejtés a spontán beszédben, mint a felolvasásban, ugyanakkor ez a jelenség nem jellemzőbb a nők beszédére, mint a férfiakéra – a jelen vizsgálati anyagban. Itt érdemes megemlíteni a központozás „tudatalatti” hatását a felolvasás szupraszegmentális szerkezetére, valamint azt a

tényt, hogy az olvasástanításban mindmáig alapvető instrukció a „pontnál vidd le a hangsúlyt” (sic!).

A hangközértékek átlaga az olvasott szövegekben (a nemektől függetlenül) 0,89, szórása 0,16; a spontán szövegekben a hangközátlag 1,12, a szórás 0,29. Az eltérés szignifikáns: független kétmintás t -próbával $p < 0,001$, $t(250) = -6,531$. A spontán beszédre tehát statisztikai értelemben is jellemzőbbek az emelkedő (1,00-nél nagyobb hangközű) szakaszvegek, a felolvasásra pedig az ereszkedő (1,00-nél kisebb hangközű) zárlatok.

A 4. ábra a hangközértékek terjedelmét mutatja a nemek és a beszédmód tekintetében. Mindkét nem esetében jól látszik, hogy a spontán beszéd közlészárlataiban mérhető hangközértékek tágabb tartományban szóródnak, mint a felolvasásban. Ez az eredmény megegyezik a korábbi szakirodalom állításával, mely szerint az élőszó használja ki a legnagyobb mértékben a hangterjedelem adta lehetőségeket (vö. pl. Wacha 1974), ugyanakkor szemben áll Beke (2008) adataival, ahol az automatikus mérések szűkebb hangterjedelmi adatokat detektáltak a spontán beszédben, mint a felolvasásban.



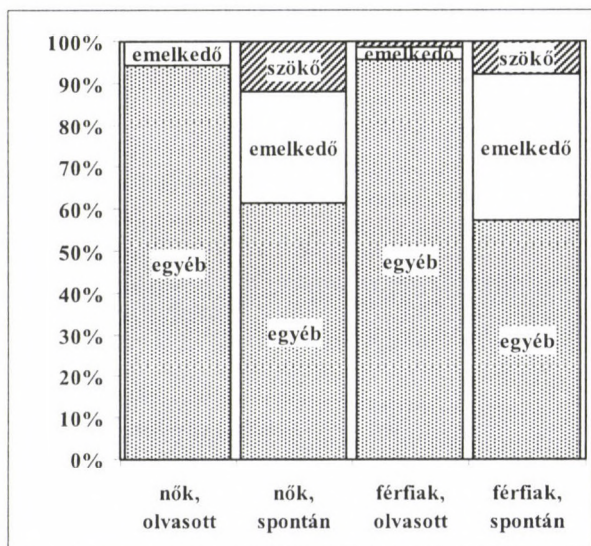
4. ábra

A hangközértékek szóródása (és mediánja) nemenként a beszédmód függvényében

A nemek szerinti eltérések másképp jelentkeznek a felolvasott szöveg mondatvégein, mint a spontán közlészárlatokban. Az olvasott szövegben szignifikáns különbséget találunk a nemek között [$p < 0,001$; $t(74) = 3,861$], ezt azonban az elemszámok jelentős eltérése önmagában is indokolhatja, hiszen a glottalizáció nagyarányú megjelenése miatt a nőknél csak mintegy az esetek felében volt adatolható a zárlat hangköze. A spontán beszédben

ugyanakkor (ahol az elemszámok kiegyenlítettek) nem találtunk szignifikáns eltérést a nemek szerint a hangközértékek megvalósulásában. Ez az eredmény ugyancsak azt sugallja, hogy a dallamzárlat alakulása a spontán beszédben nem mutat jellegzetes eltéréseket a nemek tekintetében.

A módszertani részben bemutatott percepciók teszt eredményei alapján választ adhatunk arra a kérdésre is, hogy mekkora az emelkedő és szökő hanglejtéstípusok aránya a spontán beszédben (a többi hanglejtéstípust a továbbiakban összevonva kezeljük). A kapott határértékeket a teljes korpusz frekvenciakülönbség/időtartam adataira vetítve azt kapjuk, hogy az emelkedő és szökő hanglejtések együttes aránya 40% körüli a spontán és 5% körüli az olvasott szövegekben, nemektől függetlenül (5. ábra). A nők spontán beszédében – ebben a korpuszban – ugyan másfélszer gyakoribb volt a szökőzár (11,97%), mint a férfiakéban (7,92%), de még így is alig haladta meg a 10 százalékot. Az emelkedő megnyilatkozásvég ezzel együtt negyed annyival volt ritkább a nőknél (26,50%), mint a férfiaknál (34,65%).

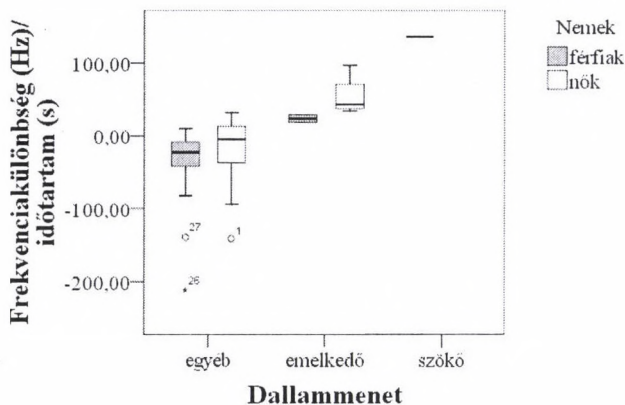


5. ábra

Az emelkedő és szökő hanglejtésű megnyilatkozásvégek aránya az olvasott és a spontán szövegekben nemek szerint – a frekvenciakülönbség/időtartam hányados alapján

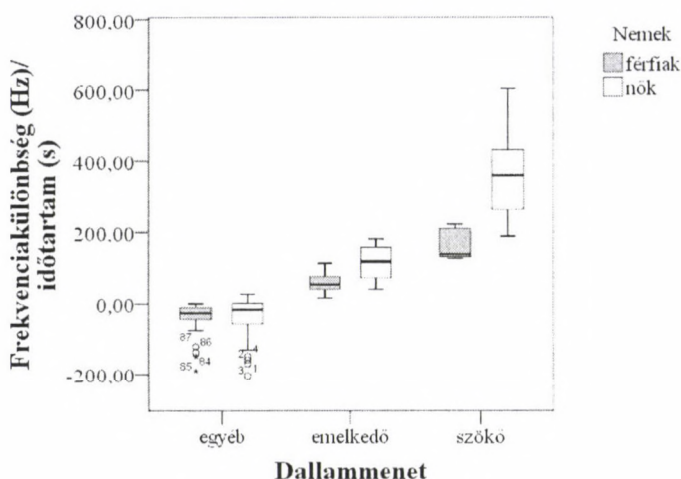
A valódi különbség vélhetően tehát nem az, hogy a nők beszédére jellemzőbb a szökőzár, hanem az, hogy a szökőzár a nők beszédében nagyobb tartományt fog át, tágabb regiszterben valósul meg, mint a férfiakéban.

Ennek ellenőrzésére a teljes korpusz összes megnyilatkozásának záró dallamára elkészítettük a frekvenciakülönbség/időtartam hányadosát ábrázoló dobozdiagramot (6. és 7. ábra). A dobozdiagramon jól látszik, hogy minden dallamkategóriában jelentősen nagyobb terjedelemben realizálódnak a nők adatai, mint a férfiakéi (lásd a dobozok és vonalak függőleges kiterjedését). Az átlagos eltérés számadatai sokatmondók. Az emelkedő hanglejtés SD értéke a nőknél 41,31, a férfiaknál 21,05 a spontán beszédben, azaz mintegy kétszer olyan tág adattartományt fog át ugyanaz a dallamminőség a nők esetében. A szökő hanglejtés tekintetében még nagyobb az eltérés: a nők 106,75-os átlagos eltérés értékéhez képest a férfiaké csak 37,78. Ugyan a szökő dallam kategóriájában a nők adatainak száma is nagyobb (14, szemben a férfiaktól származó 8-cal), a többi dallamminőség esetében azonban mindenhol kiegyenlített az adatok nemek szerinti aránya, sőt az olvasott beszéd eső/ereszkedő/lebegő (az ábrán *egyéb*-bel jelölve) kategóriájában a férfiak adatainak száma a kétszerese a nőkének (a nőknél jóval nagyobb arányban regisztrált glottalizáció miatt). A nemek közötti jelentős eltérést tehát a produkciós sajátosságok különbségéből (a nők tágabb hangköze) eredő percepció benyomás magyarázza, nem pedig a vizsgált paraméterek valós eltérése.



6. ábra

Az emelkedő és szökő hanglejtés eloszlása az olvasott szövegekben nemek szerint – a frekvenciakülönbség/időtartam hányados alapján



7. ábra

Az emelkedő és szökő hanglejtés eloszlása a spontán szövegekben nemek szerint – a frekvenciakülönbség/időtartam hányados alapján

Következtetések

A kutatásban olyan – részben szakirodalmi, részben a közbeszédben megjelenő – állítások igazságtartalmát vizsgáltuk rögzített beszédkorpusz alapján, amelyek bizonyos esetekben egyes beszélői csoportok nyelvi viselkedésére vonatkozó (le)minősítéseként fogalmazódnak meg. A bevezetőben feltett kérdésekre a kísérletsorozat alapján egyértelmű válaszok adhatók. Az emelkedő (most ennek részeként kezelve a szökő minőséget is) dallam közlészáró helyzetben valóban gyakoribb a spontán beszédben, mint az olvasásban. Míg az olvasásban ez a hanglejtésforma csak néhány százalékban jelenik meg a mondatok végén, a spontán szövegekben a hallgatók által megnyilatkozásvégnak érzékelt helyek közel felére jellemző, hogy az alapfrekvencia növekszik az utolsó részben. Ennek megfelelően természetesen az ún. szökőzár is jelentősen gyakoribb a spontán beszédben (hiszen az olvasásban szinte nem is fordul elő), a többi dallamtípushoz képest azonban nem tekinthető kirívónak az aránya, mintegy 10%-ban adatoltuk. A nemek között nem találtunk jelentős különbségeket abban a tekintetben, hogy akár az emelkedő, akár a szökő zárlat jelentősen gyakoribb lenne a nők esetében. Ugyanakkor az eredmények arra a következtetésre vezetnek, hogy a nők közismerten tágabb beszéd-hangterjedelme (vö. pl. Traunmüller–Eriksson 1995) miatt az egyes dallamminőségekre jellemző realizációik is tágabb tartományban szóródnak. Emiatt az emelkedő és a szökő dallamhoz tartozó objektív akusztikai értékek is magasabbak, mint a férfiak esetében. Ez egybevág a szakirodalmi ál-

lítással, hogy „minél magasabb az alaphang, annál nagyobb a dallamemelkedés észlelési küszöbe” (Kassai 1979: 150 – Rossi 1971 alapján).

Az objektív adatok tehát nem támasztják alá sem azt a szakirodalmi állítást, hogy a spontán beszédben gyakoriak lennének a szöközárak (bár az emelkedő zárlatok valóban gyakoribbak), sem azt a (szakmai és laikus körökben egyaránt hangoztatott) vélekedést, hogy az emelkedő frekvenciájú közlésvégek jellemzőbbek volnának a nők beszédére. Ezeknek a megállapításoknak a háttérében nyilvánvalóan percepciós okok állnak: a szökönek nevezett dallam jobban kitűnik. Mivel a nők esetében a tágabb hangterjedelem a szökö dallamnak szélsőséesebb megvalósulásait is lehetővé teszi, ez az ő beszédükben még kirívóbbnak tetszhet. Ez alapján felvetődik, hogy nem pusztán szociolingvisztikai-pragmatikai tényezők állnak a dallamhasználatban tapasztalható nemi különbségek háttérében, hanem fiziológiai természetű okok is.

A kísérletsorozat egyes részeredményei más tanulságokkal is szolgálnak. A beszédlelés végi glottalizációval kapcsolatban a magyarra vonatkozóan megállapítható, hogy ezt a jelenséget a beszédpartner valószínűleg felhasználja a szóátadás detektálásában, elképzelhető ugyanakkor, hogy pusztán a frekvenciacsökkenés ténye idézi ezt elő. Az adatok alapján további kísérletek tervezhetők ebben az irányban. A nemek közötti eltérések ugyancsak jelentősnek mutatkoznak. Az, hogy ezek háttérében anatómiai, szociolingvisztikai vagy más tényezők (esetleg ezek kombinációja) játszóak-e a főszerepet, ugyancsak újabb kutatások kérdése lehet (vö. Redi-Shattuck-Hufnagel 2001).

Elsőként kíséreltük meg a magyar szakirodalomban elterjedt ötfokozatú dallamminősítést kategóriáit objektív, akusztikai mérésekből származó adatokhoz kötni. A dallamminőségek közötti átfedések azt mutatják, hogy a hanglejtés percepcióját nem pusztán a frekvencia- és az időparaméterek befolyásolják (mivel a tesztanyag intenzitása kiegyenlített volt, ez nem befolyásolhatta az észlelést); nyilvánvalóan szerepet játszanak más tényezők, mint a beszélő egyéni hangszínezete, artikulációs tempója, a hallgató percepciós preferenciái stb. Az is kiderül az adatokból, hogy a dallam minősítése függ a beszéd nemétől, azaz az alaphangmagasságától.

A dallam vizsgálata hagyományosan, hosszú ideig felolvasott, és ezen belül is erősen normatív szövegeken történt, az intonációs szabályok megállapításának igénye motiválta. A spontán beszéd természetesen eltér ezektől a szabályoktól, ez azonban nem egyfajta beszélői igénytelenségből, hanem a tervezési folyamatok különbözőségéből fakad. A spontán beszédben a fonetikai forma (és ezen belül a szupraszegmentális szerkezet) tervezése (és kivitelezése) alárendelődik a közlés tartalmi, szintaktikai-morfológiai tervezésének, hiszen a produkciós részfolyamatok egy időben zajlanak. A hangzó kimenetet nagymértékben automatizmusok állítják elő. A felolvasásban (különösen, ha ismerjük a szöveget) az artikulációs kivitelezés jóval nagyobb mértékben tudatos, pontosan megtervezett, jól formált. Mivel a legtöbb, a dallamot érintő

megállapítás felolvasott szövegeken alapul, a spontán beszéd dallamformái ebben a szemléletben könnyen kirívónak, helytelennek tűnhetnek. Holott csak a megközelítésmód nem adekvát a jelenséggel.

Mindemellett pragmatikai tényezőkkel is számolnunk kell. A társalgásban való részvételt szabályok irányítják, amelyek a beszédpartnerek, a helyzet, a téma stb. ismeretében különböző módokon léphetnek életbe, írhatják elő a nyelvi viselkedésformák megválasztását. Egy korábbi kutatásban kimutattuk, hogy például a szöközár előfordulási aránya függ a beszédhelyzettől. További kutatások igazolhatják, hogy más pragmatikai tényezők ugyancsak hatással lehetnek a jelenség előfordulási gyakoriságára.

Irodalom

- Beke András 2008. A felolvasás és a spontán beszéd alaphangszerkezeteinek vizsgálata. *Beszédkutatás* 2008. 93–107.
- Beke András megj. A veláris magánhangzók stabilitása a spontán beszédben.
- Boersma, Paul – Weenink, David 1998. *Praat: doing phonetics by computer (Version 5.0.1)*. http://www.fon.hum.uva.nl/praat/download_win.html.
- Bolla Kálmán 1992. *Szuprasegmentális elemzések*. Egyetemi Fonetikai Füzetek 7. ELTE Fonetikai Tanszék, Budapest.
- Boros Rezső 1971. Beszéddallamok, különösen éneklő vagy idegenszerű jellegűek. *Magyar Nyelv* 67. 460–467.
- Böhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutatás* 2008. 108–120.
- Deme László 1962. Hangsúly, szörend, hanglejtés, szünet. In Tompa József (szerk.): *A mai magyar nyelv rendszere. Leíró nyelvtan II*. Akadémiai Kiadó, Budapest, 457–522.
- Dilley, Laura – Shattuck-Hufnagel, Stefanie 1995. Variability in glottalization of word onset vowels in American English. In: *Proceedings of the XIIIth international congress of phonetic sciences*. Stockholm. 586–589.
- Elekfi László 1962. *Vizsgálatok a hanglejtés megfigyelésének módjaihoz*. Nyelvtudományi Értekezések 34. Akadémiai Kiadó, Budapest.
- E[lekfi] L[ászló] 1983. Hanglejtés. In Grétsy László – Kovalovszky Miklós (főszerk.): *Nyelvművelő kézikönyv. Első kötet. A–K*. Második, változatlan kiadás. Akadémiai Kiadó, Budapest, 755–784.
- Elekfi László – Wacha Imre é. n. [2003]. *Az értelmes beszéd hangzása. Mondatfonetika – kitekintéssel a szövegfonetikára*. Szemimpex Kiadó, Budapest.
- Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama*. Akadémiai Kiadó, Budapest.
- Gósy Mária 1989. *Beszédészlelés*. Linguistica. Series A. Studia et dissertationes 2. MTA Nyelvtudományi Intézete, Budapest.
- 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 aka-*

- démikus 70. születésnapjára. Debreceni Egyetem Kossuth Egyetemi Kiadó, Debrecen, 123–129.
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszéd kutatás* 2003. 19–43.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszéd kutatás* 2008. 194–207.
- Henton, Caroline – Bladon, Anthony 1988. Creak as a sociophonetic marker. In Hyman, Larry M. – Li, Charles N. (eds.): *Language, speech and mind. Studies in honour of Victoria A. Fromkin*. Routledge, London–New York. 3–29.
- Hunyadi László 2002. *Hungarian sentence prosody and Universal Grammar. On the phonology–syntax interface*. (MetaLinguistica 13.) Peter Lang, Frankfurt am Main.
- Kassai Ilona 1979. Gyermeknyelvi dallamminták. *Magyar Fonetikai Füzetek* 4. 147–171.
- Kassai Ilona 1998. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- Markó Alexandra 2007. Kérdő funkciójú hanglejtésformák a spontán beszédben. *Beszéd kutatás* 2007. 59–74.
- Markó Alexandra 2007–2008. Szakaszvégi dallamrealizációk a spontán beszédben. *Nyelvtudomány III–IV*. 127–138.
- Olaszy Gábor 1995. A kérdés, a figyelmeztetés, a felszólítás és a kérdés prozódiaja a kijelentő mondat tükrében. *Beszéd kutatás* 1995. 46–61.
- Olaszy Gábor 2002. A magyar kérdés dallamformáinak és intenzitás szerkezetének fonetikai vizsgálata. *Beszéd kutatás* 2002. 83–99.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.
- Tompa József (szerk.) 1961–1962. *A mai magyar nyelv rendszere. Leíró nyelvtan I–II*. Akadémiai Kiadó, Budapest.
- Traumüller, Hartmut – Eriksson, Anders 1995. The frequency range of the voice fundamental in the speech of male and female adults. Manuscript. http://www.ling.su.se/staff/hartmut/fo_m&f.pdf
- Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszéd kutatás* 2008. 134–147.
- Varga László 2002a. Az egy szótagú magyar eldöntendő kérdések intonációja. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 127–140.
- Varga, László 2002b. *Intonation and stress. Evidence from Hungarian*. Palgrave Macmillan, Houndmills, Basingstoke.
- Wacha Imre 1974. Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok* X. 203–216.
- Wacha Imre 1999. *Szóveg és hangzása: Cikkek, tanulmányok a beszédről*. Magyar Rádió Rt., Budapest.

A kutatás első eredményeit bemutató előadás a Magyar Nyelvtudományi Társaságban hangzott el 2009. április 7-én. Köszönöm Elekfi László, Kassai Ilona és Siptár Péter hozzászólásait, amelyeket hasznosíthattam a további munkában.

Köszönöm Beke Andrásnak a matematikai analízis elvégzésében nyújtott segítségét. A tanulmány a Bolyai János Kutatási Ösztöndíj támogatásával készült.

BESZÉLŐVÁLTÁSOK A BESZÉDPARTNEREK SZEMÉLYES KAPCSOLATÁNAK FÜGGVÉNYÉBEN

Bata Sarolta

Bevezetés

A beszédalkalmazkodás

Beszédünkre számos tényező hathat, a beszélő saját jellemzőin (testi, lelki állapot, nem, kor stb.) kívül a beszédpartner jellemzői és a beszédszituáció egyéb tényezői is jelentős szerepet játszhatnak. A beszédpartnerhez való alkalmazkodásnak a nemzetközi szakirodalom eredményei alapján számos aspektusa van. Bell (1984) hallgatóközpontú elgondolása alapján a beszéd egyes tényezőinek alakulásában a hallgatóság nagysága, valamint a hallgatók által betöltött szerep (címzett, szándéktalan fültanú stb.) játszik a legnagyobb szerepet. Giles és Smith (1972) szociálpszichológiai szempontból értelmezi a beszédalkalmazkodást. Ebben a keretben a beszéd szociális interakcióként értelmeződik, az elméletben jelentős szerepe van a csoporthoz való tartozás kifejezési szándékának. Az alkalmazkodás lehetőségét így befolyásolja az ismeretség, a partnerről tudott információk (pl. a másik személynek tulajdonított presztízs, befolyás), az újabb találkozások valószínűsége, illetve a másik személy cselekedetinek tulajdonított (valós vagy vélt) ok. Ezt a modellt fejlesztette tovább Giles számos empirikus kutatásra támaszkodva (pl. Giles 1977, 1979), létrehozva a beszédalkalmazkodás modelljét. Az alkalmazkodási elmélet kimondja, hogy a beszélő bizonyos körülmények között (pl. siker, társadalmi elismerés elérése érdekében) átalakíthatja a stílusát, beszédváltozatát, beszédének szupraszegmentális jellegzetességeit. A beszélők tudatosan alkalmazkodhatnak az elvárásokhoz, illetve a beszédpartnerük beszédtulajdonságaihoz (konvergencia), de tudatosan éppen el is távolodhatnak azoktól (divergencia). Számos tanulmány született ebben az elméleti keretben, magyarázták vele többek között a szleng működését (Drake 1979), a kódváltás mechanizmusát (Bourhis 1978, idézi Drake 1979), vizsgálták ebben a keretben tizenéveseknél a csoporthoz való tartozás nyelvi kifejeződésének nemi meghatározottságát (Eisikovits 1987), valamint egy óceániai kreol nyelv használatának a címzethez való igazítását (Meyerhoff 1998).

Kifejezetten a beszédpartner vélt vagy valós ismeretének a beszéd jellemzőire gyakorolt hatását vizsgálta Labov (1972): különböző társadalmi osztályokat kiszolgáló üzletek eladóitól *r* hangzókat elicitált, és azok képzési módját elemezte. Az eladók többnyire a várható vásárlói körtől megszokottnak tekinthető ejtési formát alkalmazták. Douglas–Cowie (1978:43,44) egy inter-

jú első és második órájának elejéről vett mintákat hasonlított össze. Azt találta, hogy az idő elteltével az ismerőssé válás hatására az interjúkészítő és az adatközlő beszéde a kevésbé sztenderd felé tolódott.

A magyar szakirodalomban beszédalkalmazkodási témakörben eddig főként az anyák gyermekükhöz intézett beszédét elemezték (pl. Réger 1990, Kátainé Koós 2001), az első célzott (fonetikai aspektusú) beszédalkalmazkodás-vizsgálatok a beszédpartner életkorának függvényében vizsgáltak a beszéd temporális szerkezetét (Bata–Grácz 2008), illetve a társalgás egyes fonetikai jellemzőinek alakulását (Bata 2009). A vizsgálatok interjúhelyzeteket elemeztek, az interjúkészítő azonos személy volt, a felvételek célszemélyei pedig fiatal (20–30 év közötti) és idős (65–80 év közötti) beszélők. Az eredmények szerint a beszédpartner életkora hatással van a beszélő beszédének időszerkezeti mintázataira (beszédtempó, szünetelési jellegzetességek), a különböző beszédműfajokban más-más módon. A legjelentősebb életkorhoz való alkalmazkodás a felolvasás (mondatok felolvasása visszamondatási céllal) és az instrukciók (a feladatok ismertetése az adatközlőnek) esetében jelent meg. A társalgás számos paraméterének alakulásánál szintén meghatározó tényező volt a beszédpartnerek életkora. A társalgás időtartama jelentősen hosszabb volt, valamint témaváltásra és beszédlépésváltásra is gyakrabban került sor fiatal beszédpartnerek esetében. Az idős beszédpartnerek részvételekor csökkent a hallgatások időtartama, rövidebb az az időtartam, amíg a beszédpartnerek elviselik, nem érzik kellemetlennek a jelkimaradást. Az ismeretség hatására a jelen tanulmány mellett Grácz Tekla Etelka (2009) vizsgálta mondatismétlési feladat során. Az ismeretlen beszédpartnernek esetében a felvételvezető beszédének tempóértékei szignifikánsan alacsonyabbak voltak, mint ismerősök esetében. A szünettartási jellegzetességek is hasonló tendenciát mutattak: az ismeretlen beszédpartnerekhez intézett felolvasás során az felvételvezető jelentősen kevesebbszer tartott szünetet, valamint a szünetek időtartama is rövidebb volt, mint a másik két társalgásban. Az egyes eredményeken túl a tanulmányok a beszédalkalmazkodásra vonatkozó kutatások lehetséges kiindulási irányait térképezték fel.

A beszédlépések és a beszélőváltások

A társalgás alapegysége a beszédlépés, amely során a társalgás egyik résztvevője beszél, amíg át nem adja, vagy át nem veszik tőle a megszólalás jogát. A beszélőváltás mechanizmusát a társalgás létrejöttének szabályszerűségei felől közelíti a diskurzuselemzés, illetve a konverzációelemzés (pl. Brown–Yule 1989, Schenken 1978, Iványi 2001). Az elméleti megközelítés szerint rendszerint egyszerre csak egy beszélő beszél (Sacks et al. 1974:706), így a beszédlépések közötti átfedések (egyszerre beszélés), illetve a (hosszabb) szünetek általában hibának minősülnek, amelyek javítására maguk a beszélők is törekszenek (Boronkai 2008). Az átadásra alkalmas helyet a beszélő többnyire jelzi verbális vagy nonverbális eszközökkel. A beszédlépéseknél tehát cél, hogy azok minél kevesebb hiányt és átfedést hozzanak létre.

Ehhez szükséges, hogy a beszédpartnerek fölismerjék, mikor zárul le egy beszédlépés, mikor kezdeményezhetnek újabb beszédlépést; azaz képesek legyenek szintaktikailag és értelmileg tagolni a beszéd folyamatot, és e szempontok alapján beszédlépésként értelmezni bizonyos egységeket (Auer 1996: 57). Az időben előrehaladva az aktuális beszédlépés végpontja egyre inkább megjósolható; vagyis amikor a beszélő beszédlépésének végéhez érkezik, a hallgató már szinte teljes bizonyossággal képes felismerni, hogy a beszéd folyamatot potenciális váltási ponthoz érkezett (Auer 1996: 60). A végpont felismeréséhez hozzájárulnak a hallgató által is érzékelt virtuálismondathatárok (vö. Gósy 2003).

A beszédlépésváltások empirikus vizsgálatának igen gazdag a nemzetközi szakirodalma. A beszédlépések szabályszerűségeit bemutató, empirikus kutatásokat is feldolgozó tanulmány Sacks és szerzőtársai (1974) munkája. Számos kutatás vizsgálja a beszélőváltásokat megelőző beszédlépészáró egységek dallammenetét (Cutler–Pearson 1986, Wells–Peppé 1996), a beszédlépések végének előjelezhetőségét (Stephens–Beattie 1986a, 1986b; De Rulter et. al. 2006), a dallammenetek alakulásának hatását a beszélők félbeszakításának gyakoriságára (Beattie–Cutler–Pearson 1982), a zöngeminőség megváltozásának a beszélőváltásokra gyakorolt hatását (Ogden 2001, Ogden 2004). A beszédlépésváltások időviszonyainak problematikáját is több aspektusból elemezték: kognitív megközelítésű időzítési modellt hoztak létre (Wilson–Wilson 2005), vizsgálták a hallgatások és az egyszerre beszélések időviszonyainak nyelvspecifikus és univerzális vonásait, például japánra (Tanaka 2006), illetve tíz, eltérő paraméterekkel jellemezhető nyelvre vonatkoztatva (Stivers et al. 2009), valamint eredmények születtek a szintaktikai szempontból eltérő szerepű szüneteknek a beszélőváltásokra gyakorolt hatásáról (ten Bosch et al. 2005).

A magyar szakirodalomban is készültek fonetikai aspektusú kutatások, amelyekben vizsgálták a hallgatások és az egyszerre beszélések megjelenési formáit, valamint a váltási pontok szintaktikai jellegzetességeit (Markó 2006), valamint a néma szünetek hatását a beszélőváltások menetére (Evellei 2009).

A jelen vizsgálat célja

A tanulmány a beszédalkalmazkodás jelenségét vizsgálja a felvételvezető és az adatközlő ismeretségének aspektusából, a beszélőváltások alakulását tekintve. Vizsgálatunkban arra keressük a választ, hogy a beszédpartnerek személyes kapcsolata (ismeretlenség vs. barátság) hogyan hat a beszédlépések és beszélőváltások jellemzőire, a beszédlépésváltásokat forma és funkció tekintetében egyaránt vizsgáltuk.

Hipotézisünk szerint a beszédpartnerek ismeretségi foka befolyásolja a beszélőváltások alakulását a társalgásban, eltérő formai és funkcióbeli megoszlás lesz jellemző a személyes kapcsolat mentén. A formát tekintve feltételezésünk szerint az ismerősök közötti társalgás során gyakoribb lesz az egyszerre beszélés, az ismeretlenek közötti társalgásban pedig a hallgatás. A funkció

szerinti megoszlásban az ismeretlen beszédpartnerek esetében jelentősebb mértékben előforduló kérdés-válasz jellegű beszédlépésváltást vártunk.

A szakirodalmi eredmények alapján feltételeztük, hogy a társalgás fonetikai paramétereinek alakulásában különbség lesz a társalgás eleje és vége között az ismerőssé válás jelensége miatt.

Anyag, módszer, kísérleti személyek

A vizsgálatához a BEA-adatbázis (Gósy 2008) négy felvételének társalgási részét választottuk ki. Két felvétel esetében az adatközlő az interjúkészítővel baráti viszonyban állt (a későbbiekben *barátként/ismerősként* hivatkozunk rájuk), míg a másik két adatközlő ismeretlenként érkezett a felvételre, azelőtt nem találkoztak, és várhatóan a későbbiekben sem fognak találkozni a felvételezővel (a továbbiakban *ismeretlenként* utalunk rájuk). (Természetesen, a szó szoros értelmében nem tekinthetjük ismeretlennek azt az embert, akivel a beszélő már interjúhelyzetben vett részt, a megnevezéseket csak a megkülönböztethetőség kedvéért használjuk. Valójában, régi, illetve új ismeretséget értünk alatta.) A beszélgetőpartnereket megkíséreltük a legtöbb változó mentén illeszteni, azonban mivel a már meglévő BEA-korpuszból kívántuk választani a felvételeket, kénytelenek voltunk kompromisszumot kötni. Mind a négy felvételt ugyanazon személy vezette, egy 27 éves nő. Az ismerős és ismeretlen adatközlőket nemben és korban illesztettük, mindannyian férfiak voltak, az ismerősök 38 és 45 évesek, az idegenek pedig 39 és 44 évesek. A társalgás harmadik résztvevője esetében kellett kompromisszumot kötnünk: ő három esetben a felvételező egy 28 éves kolléganője, egy esetben pedig 26 éves kollégája volt.

A BEA-felvételsorozat több részből áll: a mondatismétlés 25 mondatot tartalmaz. Ezeket a felvételező olvassa fel, és az adatközlőnek meg kell ismételni a hallott mondatot minél pontosabban. Ezután egy interjú következik, amelyben a munkájáról, hobbjáról kell beszámolni az adatközlőnek, majd egy véleménynyilvánítás, melyben egy aktuális témáról kell véleményét kifejtenie. Ezen egységek célja lehetőség szerint kvázimonologikus beszéd rögzítése, azonban a felvételező segítség, a beszédkedv továbblendítése céljából további kérdéseket, saját élményeket fűzhet a beszédhez. Ezután az adatközlőnek két szöveget kell meghallgatnia, majd azok tartalmát minél pontosabban összefoglalnia. A tartalmi összefoglalókat követi a hármas beszélgetés. Ekkor egy aktuális témáról kérdezi a felvételező az adatközlőt, illetve kollégáját. A hármas beszélgetéseknek az első, illetve az utolsó négy percét elemeztük, hogy összevethessük, hogyan befolyásolja a beszédpartnerek összeszokása a beszédjellemzőket. Az egyes társalgásfelvételek eleje és vége között eltelt idő a két ismerős beszélő esetében több mint 10 perc, az egyik ismeretlen beszélővel folytatott társalgásban körülbelül 6 perc, a másíknál pedig több mint 14 perc.

A hanganyagokat annotáltuk, címkéztük a beszélőváltásokat azok típusa szerint. Méréseket és számításokat végeztünk a beszélőváltások előfordulási arányaira, valamint az egyszerre beszélések és hallgatások időtartamadataira vonatkozóan. Az annotálást a Praat 5.0 verziójával, a leíró statisztikai elemzéseket pedig az SPSS 13.0 verziójával végeztük.

Eredmények

A beszélőváltások

A vizsgálati korpuszunknál a beszélőváltásokat több szempontból osztályoztuk: a formai megközelítés (az időzítés vizsgálata) mellett figyelembe vettük, hogy milyen funkciója van a beszélőváltásoknak (mi az oka annak, hogy a következő beszélő átveszi a szót az aktuális beszélőtől).

A **funkciók** szerinti csoportosítás:

1. Visszacsatolások:

a) a figyelem/egyvetértés rövid jelzése (*igen, aha*), ide kerültek az ilyen funkciójú hűmmögések is;

b) a beszédpartner beszédlépésének befejezése;

c) a beszédpartner beszédlépés(részlet)ének megismétlése: a hozzá kapcsolódó beszédszándék a helyeslés.

2. A beszédpartnerhez intézett kérdések.

3. A beszédpartner kérdésére adott válaszok.

4. A korábbi beszélő mondanivalójához kevésbé kapcsolódó megszólalás, saját gondolatok, élmények elmondása.

5. A saját korábban elkezdett, de a beszédpartnerek által megszakított beszédlépés folytatása.

A beszélőváltások **formája** alapján öt csoport jött létre (1. ábra):

a) A beszélő korábban elkezdte a saját fordulóját, mint ahogy az előző beszélő befejezte volna (a továbbiakban: vége előtti megszólalás).

b) Két (vagy több) beszélő egyszerre kezd el a beszédlépését (a továbbiakban: egyszerre kezdés).

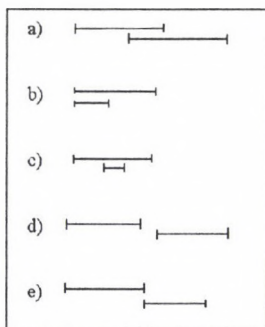
c) A beszélő átmenetileg párhuzamosan beszél beszédpartnerével, a később megszólaló fejezi be előbb a beszédlépését (a továbbiakban: párhuzamos forduló).

d) Hallgatás: a két beszédlépés között eltelik valamennyi idő.

e) Azonnali váltás: az új megszólaló az előző beszélő beszédlépésének befejeztekor kezdi a saját fordulóját.

Összevetettük a társalgások elején és végén előforduló beszélőváltások **előfordulásainak számát**, és azt találtuk, hogy mind az ismerős, mind az ismeretlen beszédpartnerek esetében a társalgás utolsó négy percében több beszédlépésváltásváltás fordult elő (2. ábra), mint az első percekben, a legjelentősebb különbség az *i_2* kóddal jelölt idegenek közötti társalgásban fordul elő, itt a társalgás végén több mint hatszor annyi beszélőváltás történt, mint az elején. Ennek az lehet az oka, hogy a beszélgetés eleje gyakorlatilag az

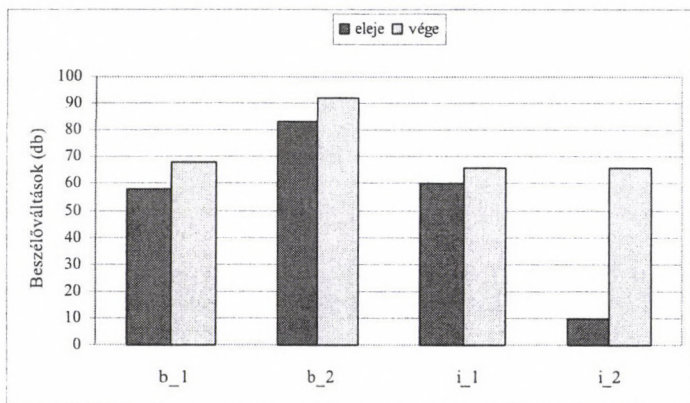
adatközlő monológja, viszont a beszélgetés végére ez a forma fellazul, és több teret kapnak a beszélőpartnerek megszólalásai. Az a tendencia, hogy a társalgás végére nő a beszélőváltások száma, azt jelzi, hogy az idő előrehaladtával a beszélgetésekben a hosszabb monologikus részek csökkennek, pergőbbé válik a társalgás, egymást váltják a rövidebb megszólalások. Az i_1 kóddal jelölt beszélgetés esetében valamivel kevesebb beszélőváltás fordult elő a beszélgetés elején, mint a végén, ennek oka a beszélgetés témájában keresendő: a társalgás elején előforduló témák mélyebben érintették a beszélőket, többször volt igényük a szó átvételére.



1. ábra

A megszólalások lehetséges időbeli elrendeződései

(A felső vonal az elsőként, az alsó vonal a másodikként megszólaló beszélő beszédlépését jelképezi)



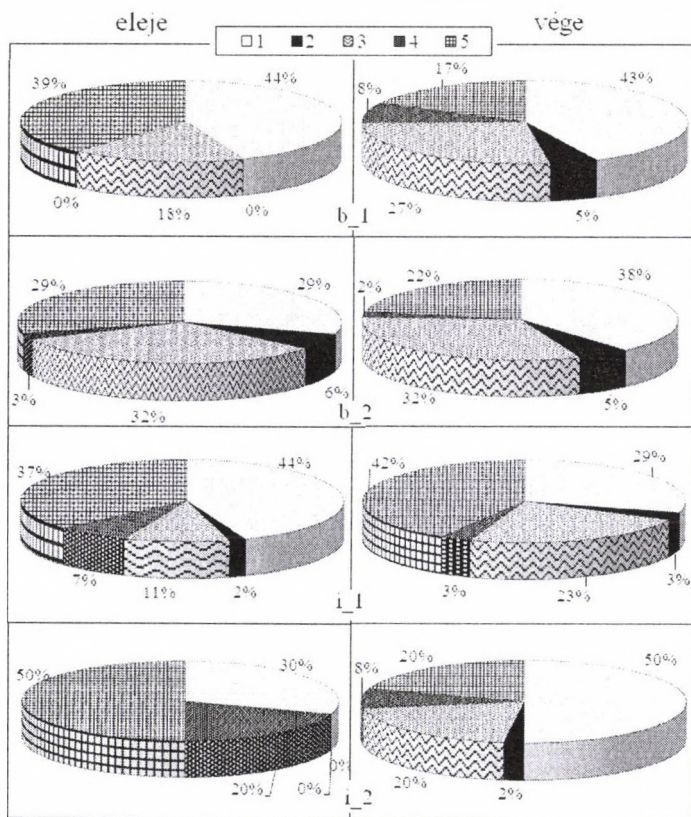
2. ábra

A beszélőváltások számának változása az egyes társalgásokban az első 4 percben és az utolsó 4 percben

Az beszélőváltások **forma** szerinti megoszlását ábráztuk a 3. ábrán. A leggyakrabban előforduló típus a hallgatás, valamint a párhuzamos forduló, azaz a beszélők leggyakrabban úgy kezdeményeznek beszédlépést, hogy előtte megvárják az előző forduló végét, de nem kevésbé jellemző az a jelenség, hogy a beszédpartnerek átmenetileg párhuzamosan beszélnek, majd az utóbb megszólaló beszélő elhallgat. A hallgatás előfordulási aránya a legtöbb esetben nagyobb az ismeretlen beszédpartnerrel folytatott társalgások során. Jelentős különbség adatolható a társalgások eleje és vége között: a társalgás végére átlagosan 34%-kal csökken a hallgatások aránya, ezzel párhuzamosan átlagosan 43%-kal növekszik az egyszerre beszélést eredményező beszélőváltások előfordulási aránya. Az utóbbiak közül kiemelendő a vége előtti megszólalás (a beszélő nem várja meg az előző forduló végét, hanem már korábban belekezd saját beszédlépésébe) arányának növekedése, ezek a változások is a beszélgetések fesztelenebbé válása felé mutatnak. A vége előtti megszólalás típusa a beszélgetések minden fázisában gyakrabban fordul elő ismerős beszédpartner esetén, a baráti viszonyban lévő beszédpartnerek gyakrabban kezdik el úgy saját beszédlépésüket, hogy a másik beszélő még nem fejezte be a fordulóját, ez a jelenség szintén a beszélgetések fesztelenebb voltára utalhat.

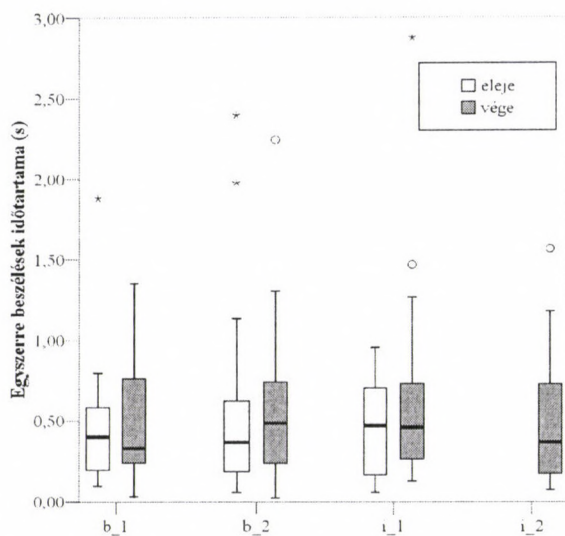
A 4. és 5. ábrán a temporális viszonyok szerinti két nagyobb csoportot, az egyszerre beszéléseket és a hallgatásokat az időtartambeli megoszlásuk szerint ábráztuk. Úgy tűnik, az egyszerre beszélések időtartamainak hossza sem az átlagot, sem a szórást tekintve nem függ az ismeretéstől. A naiv elvárás az lehetne, hogy az idegeneknél kevesebb lesz az egyszerre beszélés, hiszen azt sokan udvariatságnak tekintik. Az eredmények alapján azonban arra következtethetünk, hogy az egyszerre beszélés nem csak versengés, a másik beszédlépésébe beleszólás lehet, hanem együttműködés is, az egyetértés kifejezése, a másik beszédlépéséhez való szoros kapcsolódás jele. A hallgatások időtartama növekszik a beszélgetések végére, mind a szórást, mind a mediánt tekintve, legjelentősebb mértékben a b₂ kóddal jelölt, ismerősök közötti beszélgetés esetében. Az időtartam-növekedést magyarázza az a tény, hogy az utolsó egy-két percben (a beszédpartnerek számára is) érzékelhető, hogy a társalgás a lezáráshoz közeledik: a téma nyugvópontra jut, a megszólalások rövidülnek, a társalgás résztvevői már nem érznek erős késztetést a szó átvételére. Emellett feltételezhető, hogy ha a beszédpartnerek érzékelik a társalgás lezárásának közeledtét, akkor nagyobb a „kritikus hallgatási időtartam”, azaz a hosszabb hallgatások sem válnak kellemetlenné. Valószínűleg a beszélők a nagyobb időtartamú hallgatásokkal azt érzékeltetik, hogy megadják a lehetőséget beszédpartnereiknek a társalgás befejezésére. A beszélgetések lezárását a felvételvezető végzi el, azután, hogy megbizonyosodott róla, senki nem kíván már a témához hozzászólni, ezért ezt a beszédlépést különösen hosszú hallgatás előzi meg. Az említettek alól látszólag kivételt képez az első ismerősök közötti beszélgetés, amelyben éppen ellenkező tendencia figyelhető meg, azaz a társalgás végére a hallgatások időtartama jelentősen

csökken. Ezt a jelenséget az okozza, hogy ebben az esetben a társalgásban nem figyelhető meg lezáró szakasz, a beszélgetés még talán a többi résztvevő szándéka szerint folytathatna tovább, amikor az adatközlő átveszi a felvétellevezetőtől az irányító szerepet, és ő zárja le a beszélgetést, úgy is fogalmazhatunk, hogy a társalgás nem ér véget, hanem félbeszakad. Az eltérő mintázat mélyebb oka tehát a szerepvizonyok átrendeződésében keresendő.



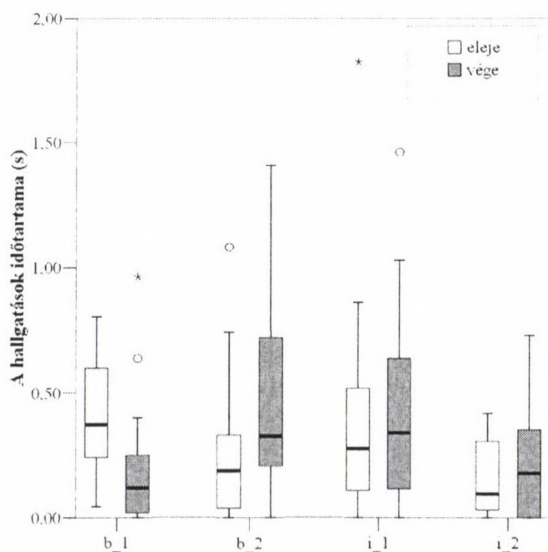
3. ábra

A beszélőváltások formái megoszlása az egyes társalgásokban
(1: párhuzamos forduló, 2: egyszerre kezdés, 3: a vége előtti megszólalás: 4: azonnali váltás, 5: hallgatás)



4. ábra

Az egyszerre beszélések időtartamának eloszlása az egyes társalgásokban (szórás és medián)



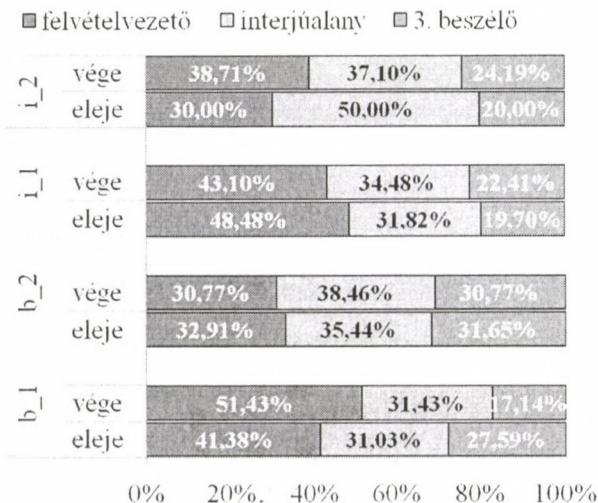
5. ábra

A hallgatások időtartamának eloszlása az egyes társalgásokban

A beszélőlépések **funkció** szerinti megoszlása a következőképpen alakul: a leggyakoribb típus a visszacsatolás, a társalgásban éppen hallgatóként részt vevő fél (felek) által használt típus, amelynek igen fontos funkciója van az együttműködés megvalósításában. Az egyéb együttműködő és versengő funkciójú beszélőváltások megoszlása társalgásonként eltérő, általános tendenciák azonban megállapíthatók: az ismerős beszélőpartnernek részvételekor gazdagabb a mintázat, mindegyik társalgás elején és végén is minden típust adatoltuk, míg az ismeretlen résztvevőknél minden esetben akadt olyan kategória, amely nem fordult elő. Mind a négy társalgás elejére jellemzőbbek a kérdések és a válaszok, a leginkább számottevő változás az egyik ismeretlenek között folyó társalgás esetében jelentkezett, itt a kérdés-válasz miatt létrejövő beszélőváltás a beszélgetés elején a beszélőlépésváltások felét tette ki, míg az utolsó négy percben egyetlen alkalommal sem adatoltunk ilyen funkciójú beszélőváltást. A beszélgetések végére pedig növekszik a versengő beszélőlépésváltások aránya (új téma felvetése), amely típusnál a beszélőnek nem célja a beszélőpartnernek által elmondottakhoz kapcsolódni, hanem új témával, új szemponttal lép be a beszélgetésbe, minden egyes alkalommal önmagát kiválasztva, azaz nem kérdésre válaszolva, illetve a beszélésre felszólítva. Ennél a típusnál a különbség a beszélgetés eleje és vége között az ismeretlen beszélőpartnernek esetében igen jelentős, az egyik társalgásban az előfordulási arány 0%-ról 21%-ra, a másikban pedig 8%-ról 26%-ra nőtt, feltételezhetjük, hogy az ismerőssé válás hatására a beszélőpartnernek felbátorodtak, és szívesebben kezdeményeztek beszélőlépést, szívesebben kapcsolódtak a beszélgetésbe új témával.

A beszélőváltásokat elemeztük az azt kezdeményező személyek szempontjából is, azt találtuk, hogy nagyjából a háromnegyed részük a felvételvezető és az adatközlő között oszlik meg, minden esetben a harmadik személy kezdeményezi a legkevesebb beszélőlépésváltást (6. ábra). Ezek alapján igazolódni látszik az a hipotézisünk, hogy az ismerőssé válás jelentősen befolyásolja a társalgás menetét: az akkorra már egy ideje együtt beszélgető, egymás számára többé-kevésbé ismerős felvételvezető és adatközlő aktívabban vesz részt a társalgásban, mint a harmadik személy, ő kevésbé vonódik be a beszélgetésbe, kevesebb beszélőváltást kezdeményez, és ritkábban is intéznek hozzá kérdéseket a beszélőpartnernek. A társalgásban betöltött szerep is befolyásolja a beszélőváltások típusainak előfordulási arányát: a felvételvezető és az adatközlő között a társalgások elején még érzékelhető a felvételsorozat korábbi interjúrészeire jellemző szereposztás: az adatközlő elmondja a véleményét, a felvételvezető pedig reagál ezekre, megszólalásaira leginkább a visszacsatolás különböző formái (egyetértés, a figyelem jelzése) jellemzőek. A felvételvezető szerepéhez csatlakozik a harmadik személy, az ő beszélőlépései is legnagyobb arányban visszacsatolások. Ez a szereposztás a beszélgetés végére fellazul, az adatközlők egyre többször kerülnek hallgatói szerepbe, amelynek elkerülhetetlen része a visszacsatolások használata, a másik két beszélőpartner (külön-

sen a felvételvezető) pedig többször közli a saját álláspontját, mint a felvétel elején. Azokban a beszélgetésekben, amelyekben baráti viszony fűzi a felvételvezetőt az adatközlőhöz, az adatközlők már az első percekben jelentősen többször jelzik figyelmüket, egyetértésüket, mint az idegenek, azaz az ismeretség hamarabb feloldja az addigra kialakult interjúszerepeket.



6. ábra

A beszélők által kezdeményezett beszélőváltások aránya a társalgásokban

Következtetések

A beszédpartnerek személyes kapcsolata hatással van a társalgás szerkezetére, mind a már meglévő, mind a társalgás során kialakuló kapcsolat tekintetében. Az évek óta tartó barátság és az ismeretlenség kettőssége mellett a beszéd jellemzőire hatott az ismerőssé válás is: a társalgások első perceiben más módon alakultak ki a beszédlépések tekintetében, mint az utolsó percekben.

A beszédlépések temporális viszonyait, valamint funkcióját tekintve is különbségek adódtak a beszédpartnerek személyes kapcsolatának függvényében. A formát tekintve elmondható, hogy valóban gyakrabban fordult elő hallgatás az ismeretlen beszédpartnerek közötti társalgásban, az egyszerre beszélések bizonyos típusa (az előző beszédlépés vége előtti megszólalás) gyakoribb ugyan az ismerősök közötti társalgásban, az egyszerre beszélések időtartam-jellegzetességei azonban nem függenek az ismeretstől. Az ismerőssé válás jelensége is adathozható a beszédlépések jellegzetességeit tekintve, a társalgás végére jellemző a beszélgetések fesztelenebbé válása: a beszédlépé-

sek száma növekszik, a kezdetben jellemzően hosszabb monologikus egységeket tartalmazó társalgást felváltja egy pergőbb diskurzus, valamint gyakoribbá válnak az egyszerre beszélést eredményező beszélőváltások.

A beszédpartnereknek a társalgásban betöltött szerepének változása is befolyásolta a beszédlépésváltások szerkezetét, a társalgások kezdetén az ismeretlen interjúalany a felvételsorozatban addig betöltött szerepének megfelelően kérdés-válasz típusú beszédlépésegységekben vett részt, ezzel szemben a beszélgetések végére ő is tett fel kérdéseket, és gyakrabban helyezkedett hallgatói szerepbe, visszacsatolásokkal (helyeslés, figyelem jelzése) erősítve meg beszédpartnereit.

További célunk a társalgásban megjelenő beszédalkalmazkodási jelenségek vizsgálatának kiterjesztése több paraméter – többek között a beszédpartnerek száma, életkora, neme, státusa; az ismeretség különböző fokozatai – mentén.

Irodalom

- Auer 1996. On the prosody and syntax of turn-continuations In: Couper-Kuhlen, Elizabeth – Selting, Margret (eds.) *Prosody in conversation*. Cambridge University Press, Cambridge, 57–101.
- Bata Sarolta 2009. A társalgás fonetikai jellemzőinek alakulása a beszédpartnerek életkorának függvényében. In: Várad Tamás (szerk.): *III. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. MTA Nyelvtudományi Intézet, Budapest (megjelenőben).
- Bata Sarolta – Grácz 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, 74–83.
- Bell, Alan 1984. Language style as audience design. *Language in Society* 13. 145–204.
- Boronkai Dóra 2008. Konverzációelemzés és anyanyelvtanítás I–II. *Anyanyelv-pedagógia* I. évf. 1. és 3-4. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=115>
- Bourhis, Richard Y. 1978. *Language in ethnic interaction: a social psychological approach*. Előadás a IX. Szociológiai Világkonferencián, Uppsala, Svédország, 1978. augusztus 14–19.
- Beattie, Geoffrey W. – Cutler, Anne – Pearson, Mark 1982. Why is Mrs. Thatcher interrupted so often? *Nature* 300. 744–747.
- Brown, Gillian – Yule, George 1989. *Discourse analysis*. Cambridge University Press, Cambridge–New York–Port Chester–Melbourne–Sydney.
- Cutler, Anne – Pearson, Mark 1986. On the analysis of prosodic turn-taking cues. In John-Levis, Catherine (ed.): *Intonation of discourse*. Croom Helm, London, 139–155.
- De Ruiter, Jan Peter – Mitterer, Holger – Enfield, N. J. 2006. Projecting the end of a speakers turn: A cognitive cornerstone of conversation. *Language* 82. 515–535.
- Drake, G. F. 1979. The social role of slang. In Giles, Howard – Robinson, W. P. (eds.): *Language: Social psychological perspectives: Selected Papers from the*

- First International Conference on Social Psychology and Language*. Pergamon Press, Oxford, 63–70.
- Douglas-Cowie, Ellen 1978. Linguistic code-switching in a Northern Irish village: Social interaction and social ambition. In: Trudgill, Peter (ed.): *Sociolinguistic patterns in British English*. Edward Arnold, London, 37–51.
- Eisikovits, Edina 1998. Sex differences in intergroup and intra-group interaction among adolescents. In Pauwels, A. (ed.): *Women and language in Australian and New Zealand society*. Australian Professional Publications, Sydney, 45–58.
- Evellei Kata Dóra 2009. *A szünetek szerepe a társalgásban. Különös tekintettel a beszélőváltásokra*. Alapszakos záródolgozat, ELTE, Budapest.
- Giles, Howard. 1977. Social psychology and applied linguistics: toward an integrative approach. *J. T. L. Review of Applied Linguistics* 33. 27–42.
- Giles, Howard 1979. Sociolinguistics and social psychology. In Giles, Howard – St. Clair, Robert N. (eds.): *Language and social psychology*. Blackwell, Oxford, 1–20.
- Giles, Howard – Smith, Philip 1979. Accommodation theory: Optimal levels of convergence. In Giles, Howard – St. Clair, Robert N. (eds.): *Language and social psychology*. Blackwell, Oxford, 45–65.
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszéd kutatás* 2003. 19–43.
- Gósy Mária 2008. Magyar spontánbeszéd adatbázis – BEA. *Beszéd kutatás* 2008. 194–207.
- Gráczki Tekla Etelka 2009. Temporális jellemzők a beszédpartnerek ismeretségének függvényében. *Beszéd kutatás* 2009. 121–133.
- Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.
- Kátainé Koós Ildikó 2001. Az anya személyiségjegyeinek hatása a csecsemő preverbalis kommunikációjára. *Beszéd kutatás* 2001. 101–113.
- Labov, William 1972. *Sociolinguistic patterns*. University of Pennsylvania Press, Philadelphia.
- Markó Alexandra 2006. Beszélőváltások a társalgásban. IX. *Pszicholingvisztikai és Alkalmazott Nyelvészeti Nyári Egyetem*, Balatonalmádi, május 21–24. Elérhető: fonetika.nyud.hu/letolt/ma_2.pdf
- Meyerhoff, Miriam 1998. Accomodating your data: the use and misuse of accommodation theory in sociolinguistics. *Language and Communication* 18. 205–225.
- Ogden, Richard 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetics Association* 31. 139–152.
- Ogden, Richard 2004. Non-modal voice quality and turn-taking in Finnish. In Couper-Kuhlen, Elizabeth – Ford, Cecilia (eds.) *Sound patterns in interaction*. Benjamins, Amsterdam, 29–62.
- Réger Zita 1986. *Utak a nyelvhez. Nyelvi szocializáció – nyelvi hátrány*. Akadémiai Kiadó, Budapest.
- Sacks, Harvey – Schegloff, Emanuel A. – Jefferson, Gail 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50. 696–735.
- Stephens, Jane – Beattie, Geoffrey 1986a. On judging the ends of speaker turns in conversation. *Journal of Language and Social Psychology* 5/2. 119–134.

- Stephens, Jane – Beattie, Geoffrey 1986b. Turn-taking on the telephone: textual features which distinguish turn-final and turn-medial utterances. *Journal of Language and Social Psychology* 5/3. 211–222.
- Stivers, Tanya – Enfield, N. J. – Brown, Penelope – Englert, Christina – Hayashi, Makoto – Heinemann, Trine – Hoymann, Gertie – Rossano, Federico – de Ruiter, Jan Peter – Yoon, Kyung-Eun – Levinson, Stephen C. 2009. Universals and cultural variation in turn-taking in conversation. In: *Proceedings of the National Academy of Sciences of the United States of America*, 106 (26). 10587–10592.
- ten Bosch, Louis – Oostdijk, Nelleke – Boves, Lou 2005. On temporal aspects of turn taking in conversational dialogues. *Speech Communication* 47. 80–86.
- Wells, Bill – Peppé, Sue 1996. Ending up in Ulster: prosody and turn-taking in English dialects. In Couper-Kuhlen, Elizabeth – Selting, Margret (eds.): *Prosody in conversation*. Cambridge University Press, Cambridge, 57–101.
- Wilson, Margaret – Wilson, Thomas P. 2005. An oscillator model of the timing of turn-taking. *Psychonomic Bulletin and Review* 12 (6). 957–968.

TEMPORÁLIS JELLEMZŐK A BESZÉDPARTNEREK ISMERETSÉGÉNEK FÜGGVÉNYÉBEN

Gráci Tekla Etelka

Bevezetés

Mindennapi kommunikációnk során számos, a beszédstílusunkat, stílusváltásainkat befolyásoló tényező jelenhet meg. Ilyen tényező lehet a beszédpartnerünk, az elérni kívánt célunk, a megszólalás tárgya, oka stb. A mindezekhez és további tényezőkhez való alkalmazkodás vizsgálatának, elméletének az alapjai a stílus kutatásokban gyökereznek. A beszédstílusok számos paraméterben különbözhetnek. A szakirodalomban eltérést találunk annak tekintetében, hogy melyeket emelik ki meghatározó tényezőkként. A fonetikai paraméterek mellett néhány esetben a lexikális, grammatikai stb. választásokkal is foglalkoznak a szerzők (pl. Zwicky 1972). Több tanulmány a szegmentális tényezőket emeli ki (l. Joos 1968, Labov 1972, Bell 1984), mások a szupraszegmentális, illetve paralingvisztikai jellemzőket vizsgálják (vö. pl. Giles-Smith 1972).

Míg számos kutató a stílusokat eleve adottnak tekinti, és annak nyelven kívüli kiváltó tényezőitől függetlenül elemzi azokat (vö. pl. Joos 1968, Zwicky 1972, Labov 1972), több stíluselmélet a stílusváltozatokat magukat nem független változóként, hanem éppen a számos nyelven kívüli tényező függvényeként értelmezik. Így például a beszédalkalmazkodási modellekben a beszédpartnerek közötti viszonyokat elemzik, és azok függvényében vizsgálják a beszédjellemzőket. Bell 1984-es munkájában a hallgatóság tagjainak szerepét és a beszélő hozzájuk fűződő kapcsolatát tekinti a beszédstílusváltások okaként. Elemzi, hogy annak függvényében, hogy az adott hallgatónak milyen „szerepe” van a társalgásban, mennyire alkalmazkodik hozzá a beszélő. Ladegaard (1995) finomítja Bell elképzelését. Bevezeti a megnevezett és a valódi címzett közötti eltérést. Diákinterjúkat elemez, mely során a diák a tanár kérdéseire válaszol ugyan, így mintegy látszólag felé címzi beszédét, néhány ponton a valódi címzett mégis a teremben ülő többi diák lesz. Ezeket a pontokat többnyire kódváltás kíséri/jelöli. A beszédpartnerek közötti viszonyok és azoknak a beszédre gyakorolt hatásának pszichológiai alapjait Giles és Smith (1972) négy pszichológiai elmélet felhasználásával fogalmazza meg. A szerzők ezen elemzésekben a beszédet szociális interakcióként értelmezik, mely során a beszélő a várt/ható jutalom és az annak elérése érdekében szükséges költségek mérlegelése alapján hoz döntéseket. Ennek megnyeréséhez azonban a beszélőnek saját beszédét úgy kell formálnia, hogy azzal

megnyerje a kívánt figyelmet. A mutatni kívánt arculat (Goffmann 1955, Brown–Lewinson 1987) nagymértékben függ attól, hogy a beszélő a hallgatóság csoportjához milyen viszonyt igyekszik mutatni (vö. pl. Giles–Smith 1972). Amennyiben a beszélő a csoporthoz való tartozást vagy annak vágyát kívánja kifejezni, igyekszik a csoportban megszokott stílushoz közel eső regisztert választani, míg a csoporttól, hallgatótól való elkülönülési szándék kifejezésekor az ő beszédstílusuktól eltérő beszédmodot választ. A beszédet így az eltávolodás, divergencia célja is jellemezheti. Giles és Smith elgondolásában a negyedik pszichológiai elmélet a mások cselekedeteinek való oktatájdönítés. Bármit tesz egy másik személy, annak, amit tesz, és ahogyan teszi, minden esetben keressük az okát, a személy mögöttes szándékát. Egy adott stílus használata vagy éppen a stílusváltás is kiváltja az oktatájdönítés folyamatát, így mind a hallgatót, mind visszahatván a beszélőt, a stílusválasztást és az esetleges stílusváltást is befolyásolja.

A hallgatósághoz való alkalmazkodás lehetőségét több tényező is befolyásolja. Az ismeretség, a partnerről tudott információk – pl. a másik személynek tulajdonított presztízs, befolyás –, az újabb találkozások valószínűsége stb. (vö. pl. Ladegaard 1995) is meghatározó jegy lehet. Az egyes kutatások rámutattak további, a beszédet befolyásoló situációs és a beszélő szándékából fakadó tényezőkre is (összefoglalóan vö. Eskénazi 1993). Ilyen beszélői szándék többek között a közlés tartalma. Coupland (1981) utazási ügynökök beszélgetéseit vizsgálva azt találta, hogy a munkáról folyó társalgások során a presztízsvaltozatot követte a beszéd stíláriis alakulása, míg a személyes témák során a lezserebb stílusok jelentek meg.

Természetesen az egyes jellemzők nemcsak külön-külön, hanem egymás függvényében, összetetten hatnak a beszédre. A megszólalás tartalma például, célja a címzett és a beszélő közötti viszony függvényében (számos egyéb tényezővel együtt) határozza meg a formai szerkezetet (vö. Blum-Kulka et al. 1989 és Szili 2004).

A stílus elemzésében nem elhanyagolható tényező a beszélő és a beszédpartnerek repertoárja sem (Giles–Smith 1972, Eskénazi 1993), hiszen a stílusválasztás és stílusváltás a másik személyhez való akkomodáció vagy a tőle való divergálás kizárólag a beszélő és a hallgató repertoárjának függvényében alakulhat.

Eskénazi (1993) egy háromdimenziós koordinárendszerben összegzi a fonetikai stílus kutatás eredményeit. A tengelyek az érthetőségre, valamint a presztízsvaltozat felé való törekvésnek, illetve a beszédpartnerek összeszokottságának, ismeretségének mértékét tükrözik. E koordináta-rendszerben helyezi el a címzettek függvényében a beszédstílusokat, valamint a fonetikai jegyeket, mint tartam, CV-hangátmenetek aránya a hangidőtartamhoz viszonyítva, szünetezés, vagy a koartikuláció. A familiárisabb társalgásokban, jellemzőbb például a magánhangzók artikulációjában a célalulmúlás (target undershoot – Lindblom 1990, magyarrá: Mády pl. 2008), a mássalhangzók és

a magánhangzók közötti koartikuláció erőteljesebb mértéke, az explozívák felpattanásának elmaradása stb. Összességében a három tengely egymásra vetítésével, összegzésével a lindblomi túl- és alulartikulált beszédtypust megkülönböztető modelljéhez hasonló megoszlást találhatunk. A kevésbé összeszokott beszédpartnerek közötti és/vagy a pontosabb érthetőségre, illetve a presztízsváltozat felé való törekvés a hiperartikuláció irányába hat, míg a kötetlenebb beszélgetési helyzetek a kevésbé feszes artikulációval jellemezhető stílus felé mutatnak.

A magyar, célzottan a beszédpartner hatását elemző tanulmányok a gyermekekhez szóló kommunikációval foglalkoztak (Réger 1990, Kátainé Koós 2001). A későbbi kutatások is elsősorban az életkor mentén elemezték a beszéd alakulását. Vizsgálták a fiatal, azonos korúak közötti és fiatalok idősekkal való beszédinterakcióinak során megjelenő temporális szerkezeteket több beszédmódban (instrukciók, társalgás stb.; Bata–Grácsi 2009), valamint a beszélőváltásokat társalgásban (Bata 2009a). A két tanulmány eredményei különböző mintázatot mutattak a beszédpartnerek függvényében. Az idősebb résztvevő jelenlétében például rövidültek a hallgatások, lassabb volt a beszédtempó. A különbségek hátterében a beszédpartner saját temporális jellemzőihez való alkalmazkodás mellett a vélt vagy valós percepciók, feldolgozási folyamatok jellemzőinek figyelembevétele is állhat. Markó (2005a, 2005b) négy beszélő temporális szerkezeteinek adataiban azt találta, hogy a különböző beszédmodok között jelentős az eltérés mind az artikulációs, mind a beszédtempó, mind a szünetezés tekintetében. A feladat kognitív összetettsége mellett befolyásoló volt a beszédpartnerek jelenléte is. A társalgás során, melyben egyszerre vettek részt mind a négyen, a tempóértékek növekedtek, közeledtek egymáshoz, a szünetek pedig rövidültek. (Hasonló közeledési tendenciákat mutattak az alaphangbeli jellemzők is.)

A jelen vizsgálat tudatos beszédalkalmazkodást követelő interjúfelvételek alapján készült, célja a felvételvezető és a felvételi személy ismeretségének a temporális jellemzőkre gyakorolt befolyását elemezni. Korábbi kutatások kimutatták, hogy a beszéd szegmentális jellemzőire befolyással van a beszédpartner vélt vagy valós ismerete. Labov (1972) egy vizsgálatában különböző társadalmi osztályokat kiszolgáló üzletek eladóitól *r* hangzókat elicitált, és azok képzési módját jellemezte. Az eredmények alapján elmondható volt, hogy az eladók többsége a várható vásárlói körtől megszokottnak tekinthető ejtési formát alkalmazta. (Az eladók szociológiai jellemzői mentén is mutakoztak különbségek.) Douglas-Cowie (1978) egy interjú első és második órájának elejéről vett mintákat hasonlított össze. Az interjú résztvevőinek beszéde az összeszokás során egyre inkább a kevésbé presztízsváltozathoz tartozó megjelenés felé tolódott. Az idő elteltével, ahogyan ismerősebbé vált az interjúkészítő és a résztvevő, beszédük a kevésbé sztenderd felé tolódott. Magyarra a beszédlépés vagy forduló jellegzetességeit Bata (2009b) vizsgálta. A felvétel során ismerősként érkezett, valamint idegenként érkezett adatközlők

felvételeinek társalgási részét elemezte. Ezen felvételrészek a teljes interjúanyag utolsó előtti felvételei voltak, vagyis, eddigre a felvételvezető és az adatközlő már jobban ismerte egymást, míg a harmadikként belépő társalgó fél nem ismerte semmilyen mértékben az adatközlőket, azonban a felvételkészítő kollégája volt mind a négy esetben. Eredményeiben összevetette a társalgás elején és végén található forduló és fordulóváltási jellemzőket. Az adatok eltérő mintázatokat mutattak egyrészt a felvételvezető és az adatközlő ismerettségének függvényében, másrészt a társalgás elején és végén. A beszélőváltások gyakorisága és típusa is egyrészt eltért, másrészt változott a beszélők összeszokása során.

Vizsgálatunk kérdése, hogy a beszédpartnerek személyes kapcsolata (ismeretlenség, illetve barátság) milyen hatással van a beszéd temporális szerkezetére; a beszéd- és artikulációs tempó, valamint a szünetezés jellemzőire egy speciális nyelvi interjúfeladatban: mondatismételtetési helyzetben. Mivel ezen helyzetekben az interjúkészítőnek célja, hogy az adatközlő a hallott mondatokat minél pontosabban meg tudja ismételni, ezért feltételezhető részéről a minél pontosabb artikulációra való törekvés, így akár a presztízsváltozathoz képest túlartikuláltabb beszéd is megjelenhet. Feltételeztük, hogy a felvételvezető mondatfelolvasása lassú és tagolt lesz, amelyet az adatközlőknek több okból kevésbé fognak követni – egyrészt az ismétlési feladat nehezítettsége (pl. memóriagénybevétel) miatt, illetve mivel számukra kisebb mértékben fontos a pontosan érthetőség. Feltételeztük azonban, hogy a beszédpartnerek összehozhatósága hatással lesz mindkét résztvevő temporális jellemzőinek megvalósulására, mivel a régebből jól ismert beszédpartner esetében a felvételvezető ismerheti a percepció és produkció képességeiket, míg a felvételre ismeretlenül érkező adatközlők esetében csak feltevései lehetnek.

Anyag, módszer, kísérleti személyek

A vizsgálatához a BEA-adatbázis (Gósy 2008) négy felvételének mondatismétlési feladatát választottuk. Két felvétel esetében az adatközlő az interjúkészítővel közeli viszonyban állt, valamint a felvételvezető másodállásában kollégák is. A másik két interjú alanya teljesen ismeretlenként érkezett a felvételre, azelőtt nem találkoztak, és várhatóan a későbbiekben sem fognak találkozni a felvételvezetővel. Mind a négy felvételt ugyanazon személy vezette, egy 27 éves nő. Az adatközlők férfiak. Az ismerősök 38 és 45 évesek, foglalkozásuk színész; az idegenek pedig 39 és 44 éves színész, illetve biztosítási ügynök, akik foglalkozása és egyéb szociális jellemzőik legkorábban a mondatfelolvasás utáni interjú során válik ismertté a felvételvezető számára.

A mondatismétlés 25 különböző hosszúságú és összetettségű mondatot tartalmaz. Például: *A nagymama specialitása kétséget kizáróan a barackbefőtt. Az ügyfeleknek kompromisszumot kellett kötniük. Nem kötött biztosítást, ezért kisebb vagyonba került a kórházi ellátás. Szerencsétlenül alakultak az ese-*

mények a tegnapi túsmentő akció során. A mondatokat a felvételvezető olvassa fel, a felvételi személynek pedig meg kell ismételnie.

A temporális viszonyok jellemzéséhez a beszédtempót és a mondatra jellemző átlagos, a szakaszonkénti artikulációs tempót, valamint a szünet-előfordulást és -időtartamokat számítottuk ki beszélőnként.

A felvételek szereplőinek megnevezését a következőképpen fogjuk használni: *barát* = a felvételre már ismerősként érkezett (b_1 , b_2), közeli ismerőse a felvételvezetőnek; *idegen* = a felvételre ismeretlenként érkezett adatközlő (i_1 , i_2).

Az annotálást a Praat 5.0 verziójával (Boersma–Weenink 2008), a statisztikai elemzéseket pedig az SPSS 13.0 verziójával végeztük. Egytényezős varianciaanalízist, Tukey-féle post hoc tesztet, valamint Pearson-féle korrelációelemzést alkalmaztunk.

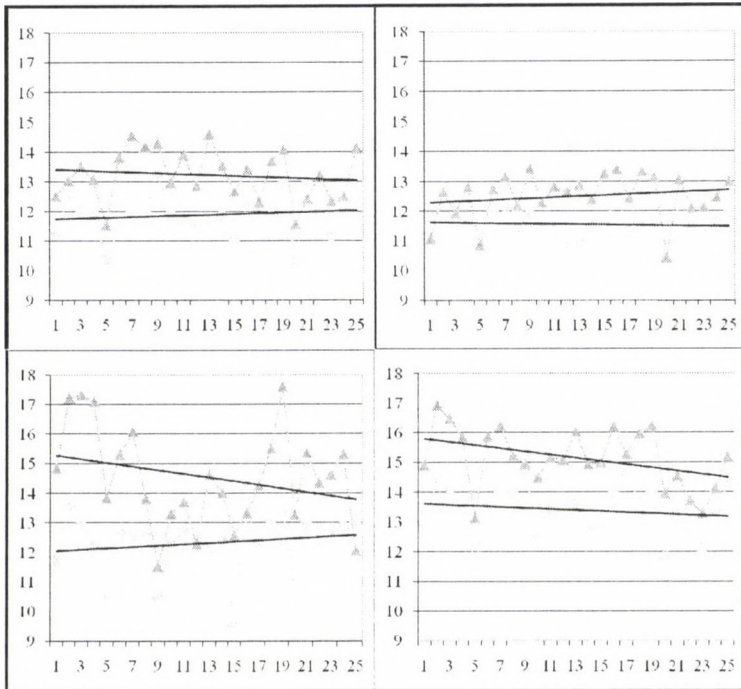
Eredmények

A vizsgált mondatismétlések artikulációs és beszédtempójuk tekintetében is jellegzetes mintázatot mutattak a felvételi helyzet függvényében.

A mondatismétlési feladatok során a felvételvezető **artikulációs tempója** (1. ábra) minden felvétel esetében szignifikánsan lassabb volt az adatközlő artikulációs tempójánál mintegy 5–15%-kal [egytényezős varianciaanalízis minden esetben: $p < 0,001$ – $F(1, 49) = 36,058$; $F(1, 49) = 24,243$; $F(1, 49) = 33,171$; $F(1, 49) = 44,176$]. Az egyes mondatok esetében számított átlagos artikulációs tempók között összesen négy olyan előfordulást adathallhattunk, mely során valamivel (2–8%) gyorsabban artikulált a felvételvezető az adatközlőnél. Ezek egy-egy esetben az ismeretlenként érkezett adatközlők esetében fordultak elő, míg az egyik korábbiól jól ismert adatközlő esetében kétszer, a másik esetében pedig egyszer sem adathallhattunk ilyen előfordulást. (Nem találtunk összefüggést a négy magasabb artikulációs tempó között sem a mondatok típusával, sem egyéb tekintetben.) Az artikulációs tempók szóródása minden esetben tendenciájukban magasabb volt az adatközlők beszédében, mint a felvételvezető olvasásában (3–117%). A felvételvezető átlagos artikulációs tempója a két idegen számára történő felolvasás során valamivel lassabb volt (11,71 és 11,54 hang/s), mint a baráti kísérleti személyek számára történő felolvasás során (12,48 és 13,60 hang/s), valamint az adatok is kisebb szóródást mutattak (átl. eltérés: 1,25 és 1,06, illetve 1,45 és 2,00 hang/s). Ez az eltérés az egytényezős varianciaanalízis alapján szignifikáns [$F(3, 214) = 20,898$; $p < 0,001$]. A Tukey post hoc teszt alapján az idegeneknek történő felolvasás artikulációs tempója lényegében azonosnak tekinthető ($p = 0,923$), míg a baráti adatközlők számára ($p = 0,001$) és azok, valamint a másik két felolvasás között ($p \leq 0,027$) szignifikáns az eltérés.

Az adatközlők artikulációs tempója is eltérő volt a mondatismétlések során. Az egytényezős varianciaanalízis [$F(3, 187) = 17,561$; $p < 0,001$] és a Tukey post hoc teszt alapján a két idegen adatközlő (átl.: 12,76 és 12,12

kivételével a felvételvezető és az adatközlő artikulációs tempója egymáshoz közeledő tendenciát mutatott.



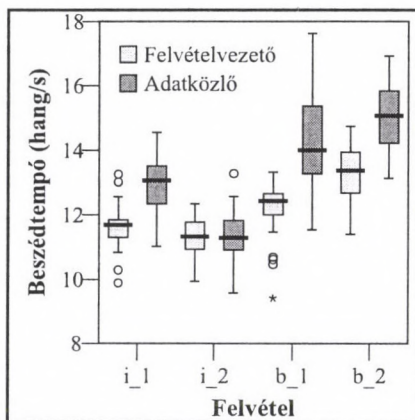
2. ábra

A felvételvezető és az adatközlők artikulációs tempójának (hang/s) alakulása a mondatismétlési feladat során: értékek és trendvonal [vízszintes tengely: a mondatok sorszáma, függőleges tengely: artikulációs tempó (hang/s); bal lent: i_1, jobbra lent: i_2, bal lent: b_1, jobb lent: b_2; sötétszürke háromszögek: adatközlő, világosszürke korongok: felvételvezető]

A felvételek **beszédtempó**beli jellemzői az artikulációs tempó eredménye-
ihez hasonlóan alakultak (3. ábra). A felvételvezető beszédtempója jellemző-
en lassabb volt az idegen számára történő felolvasáskor (11,62 és 11,30
hang/s), valamint kisebb szóródást mutatott (átl. eltérés: 0,75 és 0,59 hang/s),
mint a két baráti személlyel készült felvétel során (12,13 és 13,24 hang/s, ill.
átl. eltérés: 0,94 és 0,89 hang/s). Az egytényezős varianciaanalízis alapján az
eltérés szignifikáns [$F(3, 99) = 28,229$; $p < 0,001$]. Az egyes felvételeket ösz-
szevetve a felvételvezető beszédtempója a két idegen ($p = 0,506$) és az egyik

idegen és az egyik baráti adatközlő számára történő felolvasás során hasonlóan alakult a Tukey post hoc teszt alapján ($p = 0,114$).

Az adatközlők beszédtempója a felvételvezetőtől eltérő mintázatot mutat. A beszélők közötti eltérés szignifikáns [egytényezős varianciaanalízis: $F(3, 99) = 45,400$; $p < 0,001$]. A felvételeket egyesével összevetve a két barát beszédtempójában nincs jelentős eltérés (átl.: 14,39 és 14,98 hang/s, átl. eltérés: 1,72 és 1,06 hang/s; $p = 0,307$), míg az ő eredményeik és a két idegen eredményei között, valamint a két utóbbi adatközlő adatai között (átl.: 12,97 és 11,36 hang/s, átl. eltérés: 0,96 és 0,84 hang/s) az eltérés statisztikailag szignifikáns ($p < 0,001$).



3. ábra

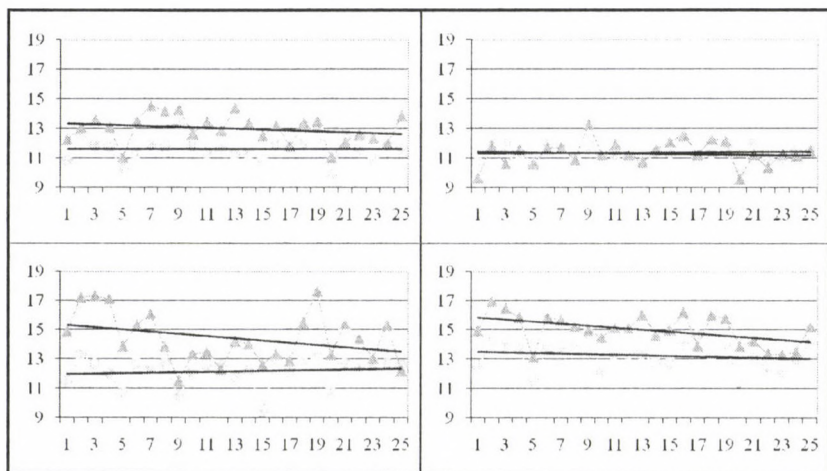
A felvételvezető és az adatközlő beszédtempójának (hang/s) mediánja és szóródása az egyes felvételek során

A felvételvezető és az adatközlő beszédtempóját összevetve a felvételvezető beszéde egy (idegen adatközlő számára történő) felolvasás kivételével minden esetben szignifikánsan lassabb volt 12–19%-kal [egytényezős varianciaanalízisek alapján $p < 0,001$ – $F(1, 49) = 31,164$, $F(1, 49) = 33,125$, $F(1, 49) = 39,068$ –, illetve $F(1, 49) = 0,094$, $p = 0,760$]. Az egyes mondatok beszédtempója azonban már változatosabb eredményeket mutatott, mint az artikulációs tempójuk. Több mondat esetében volt közel hasonló vagy gyorsabb a felvételvezető beszédtempója, mint az adatközlőé. A felvételvezető gyorsabb beszédtempója főként az i_2-es felvételen volt jellemző, amelynél a beszédtempók nagymértékben hasonlóak a post hoc teszt alapján.

A felvételvezető és az adatközlő beszédtempója között alacsonyabb korrelációt találtunk, mint az artikulációs tempók között, azonban ez az összefüggés is jelentős ($r = 0,520$, $p < 0,001$). Az egyes felvételeket összevetve ismét nem találunk egyértelmű összefüggést a felvételvezető és az adatközlő fami-

liaritásának tekintetében. A két idegen esetében a korrelációs együttható 0,520, illetve 0,674; a két régebbiről már jól ismert adatközlő esetében pedig 0,471 és 0,620 ($p < 0,018$ minden esetben).

Az egyes mondatokra kapott beszédtempó-értékek a felvétel végére ismét közeledő tendenciát mutatnak (4. ábra – a mondatok elhangzásának sorrendjében szemlélítve a beszédtempó-értékek alakulását). A két barát esetében nagyobb mértékű a közeledés, míg a két idegen esetében alacsonyabb. Ennek egyik oka, hogy a két idegen beszélő esetében a kezdeti tempóértékek is alapvetően nagyobb hasonlóságot mutattak a felvételvezetőével.



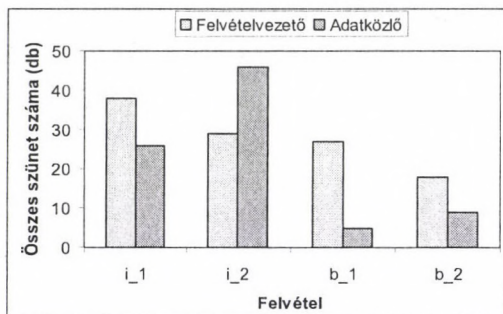
4. ábra

A felvételvezető és az adatközlők beszédtempójának (hang/s) alakulása a mondatismétlési feladat során: értékek és trendvonal

[vízszintes tengely: a mondatok sorszáma, függőleges tengely: artikulációs tempó (hang/s); bal fent: i_1, jobbra fent: i_2, bal lent: b_1, jobb lent: b_2; sötétszürke háromszögek: adatközlő, világosszürke korongok: felvételvezető]

A mondatismétlési feladat **szünetezési** jellemzői az artikulációs és beszédtempó adatok tendenciáitól részben eltérő eredményeket mutatnak (5. ábra). A felvételvezető egy felvétel kivételével minden esetben gyakrabban tartott szünetet a mondatok felolvasása során, mint az adatközlők. Az artikulációs és beszédtempó esetében közepes-erős korrelációt találtunk az beszédpartnerek adatai között, a szünetezés esetében azonban az adatközlők kevésbé követték a felvételvezető felolvasásának tagoltságát, pl. a b_1-es felvétel esetében mintegy ötöde annyi szünetet tartalmaztak az ismételt mondatok, mint a felolvasottak. A felvételvezető összesen kétszer tartott megakadás miatt szünetet, amelyből egy nem szerepel a szünetek között, mivel a mondatot az elejé-

től újraolvasta. Az adatközlők összesen négyszer (három megakadás a b_1-es felvétel során, egy az i_1-es során), melyek mindegyike szerepel a szünetadatok között, mivel a javítási stratégia nem a mondat újbóli teljes ismétlése volt.



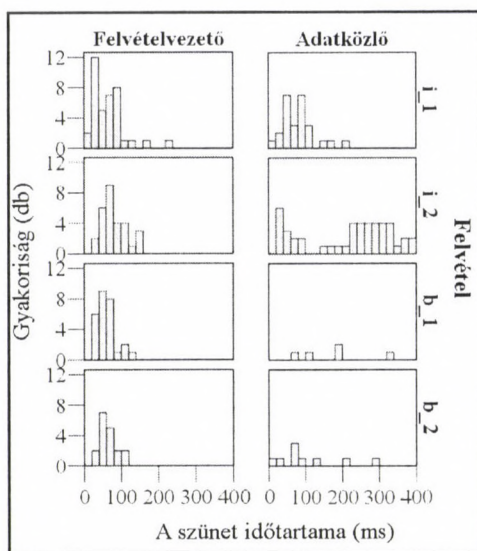
5. ábra

A felvételvezető és az adatközlő összes szünetének száma

A legtöbb és leghosszabb szüneteket az i_2-es felvétel adatközlője tartotta. Az ő artikulációs tempója (és ezen két ok következtében a beszédtempója is) volt a leglassabb az adatközlők között.

A szünetidőtartamok tekintetében (6. ábra) a felvételvezető felolvasásaiban nincs jelentős eltérés az egyes felvételek között [egytényezős varianciaanalízis: $F(3, 111) = 2,621$; $p = 0,054$]. Azt mondhatjuk azonban, hogy tendenciájában az i_2-es felvétel során hosszabb szüneteket tart. Három felvétel alatt 57,88–64,65 ms, addig itt 82,57 ms az átlagos szünethossz. Az adatközlők szüneteinek hossza a felvételvezető szünetidőtartamaival összevetve nem mutat jellegzetes mintázatot sem a felvételvezető és az adatközlő kapcsolata, sem a szünetek száma, sem a szünetek oka alapján.

Az adatközlők és a felvételvezető szünettartása ritkábban esett azonos helyre. Az adatközlők több esetben olyan helyen is tartottak szünetet, ahol a felvételvezető egy-egy koartikulációs hatást nem érvényesített (pl. zöngésségi hasonulás maradt el), vagy érdes zöngét használt.



6. ábra

A felvételvezető és az adatközlő szünetek időtartambeli (ms) megoszlása (db)

Következtetések

A vizsgálat a beszédpartnerek ismeretségének függvényében egy „sajátos” beszédalkalmazkodási helyzetet elemzett. A mondat-visszamondatási feladatok során a felvételvezető a minél pontosabban érthető és ismételhető mondatfelolvasásra törekszik, ezért a beszédében lassú artikulációs és beszédtempót vártunk, valamint hosszabb és gyakori szüneteket, mint az adatközlők ismétlésében, akiknek kevésbé kell törekedni a hiperartikuláltságra és a nagymértékű tagolásra. Feltételeztük azonban, hogy az egyes adatközlők támaszkodni fognak a hallott tagolásra, valamint hogy az ismétlések során követik a hallott mondatok egyéb temporális jellemzőit is, egyrészt a mondatismétlés megkönnyítése érdekében (annak memóriaigénye miatt), másrészt az önkéntelenül is bekövetkező alkalmazkodás miatt.

Az artikulációs és a beszédtempó eredményei azt mutatták, hogy a felvételvezető jelentősen lassabban beszél az adatközlőknél, azonban adataik közepes-erős korrelációt mutattak. Váradi (2009) hasonló eredményeket talált 10 felvételen végzett elemzéseiben. Az adatközlőknél minden esetben szignifikánsan lassabban olvasott fel a felvételvezető. A korreláció mértéke alacsonyabb volt tanulmányában (az együttható értéke 0,354, ill. 0,375 az artikulációs és a beszédtempóra), mint a jelen tanulmányban kapott adatok.

Anyagunkban három beszélő esetében a felvételvezető gyakrabban tartott szünetet, azonban az átlagos szünetidőtartam minden felvétel esetében rövidebb volt az adatközlők értékeinél.

A feladat során az adatközlők artikulációs és beszédtempója közeledett a felvételvezetőjéhez. Ez alól egy, a leglassabb beszédű férfi volt kivétel. Az ő beszédtempója közel azonos volt a felvételvezetőjéhez már a felvétel elején is. A felvételvezető tempója kisebb változást, stagnálást mutatott a felolvasások során. Feltételezhetjük, hogy az adatközlők esetében a felvételvezető önkéntelen követése, illetve a memóriaigénybevétel együttesen okozta a közeledést, míg a felvételvezetőnél a pontos, jól követhető (túlargumentált beszédre való törekvés okozhatta a viszonylag egyenletes, kisebb változásokat mutató temporális szerkezetet.

A jelen felvételek esetében a két régebből ismert beszélő a felvételek elején kevésbé törekedett a felvételvezetőhöz hasonló temporális szerkezetek megtartására, a feladat során azonban nagymértékben közeledtek ahhoz. Ennek az ismeretségen kívül számos oka lehet. Elsősorban a személyiség (pl. extrovertált – introvertált), másrészt például az ismeretség típusa (pl. barát, szülő-gyermek, beosztott-főnök) is meghatározó lehet.

Irodalom

- Bata Sarolta 2009a. A társalgás fonetikai jellemzőinek alakulása a beszédpartnernek életkorának függvényében. In Váradai Tamás (szerk.): *III. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. MTA Nyelvtudományi Intézet, Budapest (megjelenőben).
- Bata Sarolta 2009b. Beszélőváltások a beszédpartnernek személyes kapcsolatának függvényében. *Beszéd kutatás 2009*. 107–120.
- Bata Sarolta – Gráczai 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, 74–83.
- Bell, Alan 1984. Language style as audience design. *Language in Society* 13. 145–204.
- Blum-Kulka, Shoshana – House, Juliana – Kasper, Gabriele 1989. *Cross-cultural pragmatics: request and apologies*. Ablex Publishing Corporation, Norwood.
- Paul Boersma – David Weenink 2008. *Praat: doing phonetics by computer (Version 5.0.20)* [Computer program]. <http://www.praat.org/>
- Brown, Penelope – Levinson, Stephen 1987. *Politeness. Some universals in language use*. Cambridge University Press, Cambridge.
- Coupland, Nicolas 1981. *The social differentiation of functional language use: A sociolinguistic investigation of travel agency talk*. PhD-disszertáció, University of Wales Institute of Science, and Technology, Cardiff.
- Douglas-Cowie, Ellen 1978. Linguistic code-switching in a Northern Irish village: Social interaction and social ambition. In Trudgill, Peter (ed.): *Sociolinguistic patterns in British English*. Edward Arnold, London, 37–51.

- Eskénazi, Maxine 1993. Trends in speaking style research. In *Proceedings of Eurospeech '93*, Berlin, 501–509. http://www.cs.cmu.edu/~max/#_Publications. A letöltés ideje: 2009. január.
- Giles, Howard – Smith, Philip 1979. Accommodation theory: Optimal levels of convergence. In Giles, Howard – St. Clair, Robert (eds.): *Language and social psychology*. Blackwell, Oxford, 45–65.
- Goffman, Erving 1955. On face-work: An analysis of ritual elements in social interactions. *Psychiatry* 18. 213–231.
- Gósy Mária 2008. Magyar spontánbeszéd adatbázis – BEA. *Beszéd kutatás 2008*. 194–207.
- Joos, Martin 1968. The isolation of styles. In Fishmann, Joshua A. (eds.): *Readings in the sociology of language*. Mouton, The Hague, 185–191.
- Kátainé Koós Ildikó 2001. Az anya személyiségjegyeinek hatása a csecsemő preverbális kommunikációjára. *Beszéd kutatás 2001*. 101–113.
- Labov, William 1972. *Sociolinguistic patterns*. University of Pennsylvania Press, Philadelphia.
- Ladegaard, Hans J. 1995. Audience design revisited: Persons, roles and power relations in speech interactions. *Language & Communication* 15/1. 89–101.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H and H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.
- Mády Katalin 2008. Magyar magánhangzók vizsgálata elektromágneses artikulográffal gyors és lassú beszédben. *Beszéd kutatás 2008*. 52–66.
- Markó Alexandra 2005a. *A spontán beszéd néhány szuprasegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. PhD-disszertáció. ELTE, Budapest.
- Markó Alexandra 2005b. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszéd kutatás 2005*. 63–77.
- Réger Zita 1986. *Utak a nyelvhez. Nyelvi szocializáció – nyelvi hátrány*. Akadémiai Kiadó, Budapest.
- Szili Katalin 2004. *Tetté vált szavak*. Tinta Könyvkiadó, Budapest.
- Váradi Viola 2009. Hallásalapú és vizuális alapú közlések. *Beszéd kutatás 2009*. 228–239.
- Zwicky, Arnold 1972. On causal speech. In Peranteau, Paul M. – Levi, Judith N. – Phares, Gloria C. (eds.): *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, 607–615.

A kutatást az OTKA 78315 számú pályázata támogatta.

FORMÁNSMÉRÉSEK AUTOMATIZÁLÁSA, FORMÁNSADATBÁZISOK LÉTREHOZÁSA

Olaszy Gábor – Rácz Zsuzsanna Zsófia – Bartalis Mátyás

Bevezetés

A formánsok vizsgálata a beszédkutatás egyik legrégebbi témaköre. Fonetikai meghatározásuk egyértelmű: a felhangszerkezet bizonyos csoportjai intenzitásban jelentősen kiemelkednek a környezetükből, ezeket a nagy energiájú felhangnyalábokat nevezik formánsnak. Frekvenciaértékükkel, a kiemelkedés fokával és a sáv szélességükkel jellemzik őket. Fizikai vonatkozásban a formánsokat a vonalas spektrum maximumhelyei adják, ezeket matematikai módszerekkel meg lehet határozni. A magyar magánhangzók formánsait első ízben Tarnóczy Tamás számította ki (1941). A formánsfrekvenciák meghatározására számos általános számítógépes eljárást fejlesztettek (Reddy–Swamy 1984, Böhm–Németh 2007), de ha nagy pontosságú adatokra van szükségünk, akkor a gépi eljárások eredményeit sok esetben csak emberi munkával lehet pontosítani. Könnyítheti a formánsmeghatározó algoritmus dolgát, ha ismerjük a hanghullám tartalmának fonetikai átírását, azaz a hangokat, amelyekben a formánsokat meg kell határozni (Lee et al. 1999). Ilyenkor jobban behatárolhatóak a formánsávok. Ha ismerjük a nyelvet is, akkor a jellemző formánsmozgások is meghatározhatók a hangkörnyezet függvényében. Az a gondolat, hogy formánsmérések eredményeit adatbázisba tömörítsék, nem régi. A Microsoft kutatólaboratóriumában (Deng et al. 2006) minősítési és beszédkutató feladatok támogatására hoztak létre formánsadatbázist angol nyelvre, a TIMIT beszédatadátbázis kiválasztott mondataira. Az adatbázis létrehozására kifejlesztettek egy vizuális-grafikus editort, amelynek segítségével szakértők rajzolták be a mondatok minden hangjára a formánsmeneteket. Ezek jelentik a referencia adatokat, vagyis a formánsadatbázist. Más ilyen adatbázisról nincs tudomásunk.

A beszédkutatás tendenciája a statisztikai jellegű vizsgálatok felé tolódik, nagyméretű adatbázisokat hoznak létre, és ezekben végeznek vizsgálatokat. Ehhez nagy mennyiségű adatot kell kinyerni a beszédjelből. Az ilyen munkák támogatására mutatjuk be saját fejlesztésű eljárásunkat, valamint az első magyar formánsadatbázist.

Célkitűzés

A fejlesztési munkának két célja volt. Az első, hogy olyan eljárást dolgozzunk ki, amely lehetővé teszi, hogy a magyar magánhangzókban létrejövő

formánsmozgások jellemzésére nagy pontossággal lehessen formánsadatokat meghatározni bármely formánselemző algoritmus támogatásával. A második, hogy létrehozzunk egy referencia formánsadatbázist a magyar magánhangzókban végbemenő formánsmozgások jellemzésére. Ez utóbbit egyrészt az első célkitűzés támogatására, másrésztől későbbi kutatások kiszolgálására szántuk (például bármely formánselemző algoritmus hatásosságának mérése). A formánsadatbázis az oktatásban is használható (feladatok, mérések kiadása, tendenciák bemutatása stb.), mivel stabil és pontos vizsgálati anyagot biztosít.

A vizsgálati anyag és módszer

A jelen munka első lépése volt a formánsadatbázis létrehozása. Ezt tekintettük kiindulási alapnak. Ennél a munkánál más utat követtünk, mint a hivatkozott, angol nyelvre kidolgozott formánsadatbázis fejlesztői. Egyrésztől nem folyamatos formánsmeneteket rögzítettünk, hanem előre meghatározott pontokra adtuk meg a formánsértékeket a magánhangzókban, másrésztől a formánsadatokhoz való hozzájutás is más módszerrel történt, harmadrészt, nem mondatokra, hanem szavakra épül a vizsgálati anyag. Ugyanakkor követtük Lee és munkatársai (1999) módszerét, vagyis ismertnek tekintettük a mért magánhangzót, valamint a nyelvet.

Az adatbázisokkal való munkánál tudnunk kell, hogy nincs teljesen hibátlan adathalmaz. A célkitűzés az lehet, hogy az adatbázis pontossága minél magasabb százalékkénttel legyen jellemezhető. Ebben a munkában a fonetikai elvárásnak nem megfelelő mérési adatokat tekintjük pontatlannak (például a spektrogramon látható a harmadik formáns vonulata, a program mégsem ad adatot erre a formánsra, ugyanakkor más elemző esetleg jelezne ezt).

A kísérletek nyelvi anyagául egy nyilvános, nagy pontosságú belső adatokkal rendelkező magyar beszédatadbázist használtunk, amely hozzáférhető a <http://fonetika.nytud.hu/cvvc> honlapon (más, nyilvános magyar beszédatadbázisról nem tudunk, amelyen bárki ellenőrző méréseket végezhet). Ez az adatbázis szólista felolvasást tartalmaz. Minden szó két bemondó hangján szerepel, női és férfi hangon (összesen 2912 elem). A szóadatbázis régebbi keletű, már több éve aktívan segíti a kutatást és az oktatást. Készítésekor a hanghatárok kijelölése manuálisan történt, a szóhangsorok fonetikus átírását automatikus átíróval végeztük, majd manuálisan ellenőriztük és javítottuk (Abari et al. 2006).

A jelen munka célkitűzését követve a felolvasott szavak összes magánhangzójában határoztuk meg a formánsértékeket a hang 25, 50, 75%-os pontján. A mérések során figyelembe vettük a magánhangzót megelőző és követő hangot is. Így összesen 9975 olyan hanghármast találtunk, amelyek középső hangja magánhangzó volt (zömmel CVC kapcsolatok). A szókezdő és szózáró magánhangzókat is vizsgáltuk, a # jelet használva a hanghármast definiálásánál. A hangok jelölésére a BME TMIT számítógépes jelölési rendszerét

használtuk (lásd a fenti honlapon). A kiindulási adathalmazban tehát hullámformák, azok hangátírásai és a hanghatárok címkézései álltak rendelkezésre.

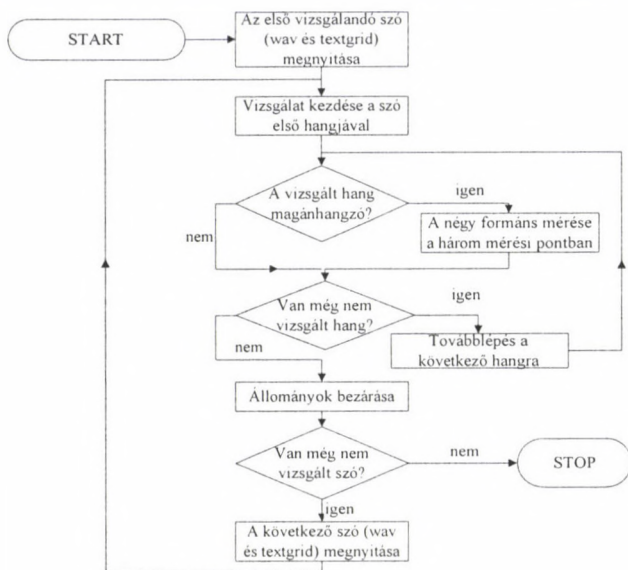
A tömeges formánsmérések végrehajtását a széles körben ismert Praat programra alapoztuk, ehhez készítettünk kiegészítő algoritmust és szkriptet, amely elvégezte a szóállomány hullámformáinak automatikus feldolgozását. A gyári alapbeállításokon nem változtattunk, a mérési sáv maximuma 5500 Hz volt. A 3 mérési ponton való adatgyűjtéssel az volt a célunk, hogy a későbbiekben vizsgálni lehessen a hangkörnyezet hatását is, vagyis a formánsmozgások jellemző tendenciáit a magánhangzón belül. A szókezdő és szózáró magánhangzókban csak két mérési ponton rögzítettük az eredményeket (50, 75%, illetve 25, 50%), mivel a kísérletek során kiderült, hogy az induló, illetve lecsengő hangokban a kezdeti illetve a végső szakaszokban az elemző sokszor hibás értéket mért. Ezeken a helyeken a beszéd spektrális tartalma bizonytalannak tűnt. A mérések során 29926 mérési ponton rögzítettünk adatot. Minden mérési ponton négy formáns adatait mértük meg. Összesítve tehát 119704 formánsadat szerepelt a Praat program által előállított adathalmazban. Ebből készítettük el a referencia formánsadatbázist oly módon, hogy félautomatikus módszerrel feltártuk a vélhetően hibás formánsadatokat, majd azokat manuálisan ellenőriztük, ha kellett, javítottuk (a fenti honlap hangspektrogramjainak tanulmányozása alapján). A véglegesített formánsadatokat SQL adatbázisba töltöttük, és ez képezi a referencia formánsadatbázist. Ebből nyertük az itt bemutatott adatokat. A statisztikai elemzéseket SPSS programmal végeztük.

A mérési rendszer fejlesztése, elemei, jellemzői

A mérési rendszert két pillérre építettük: *a)* automatikus Praat alapú elemzés (bármely más elemző is megfelelt volna), *b)* manuális hibafeltáró és javító eljárás. A két feldolgozási forma együttes felhasználásával terveztük létrehozni a referencia formánsadatbázist.

Automatikus, Praat alapú elemző

Az elemzővel szemben támasztott követelmény volt, hogy kezelni tudjon könyvtárakba rendezett wav fájlokat, valamint a hozzájuk tartozó hanghatárjelölő fájlokat. A mérés során a program kiszámította a magánhangzón belüli három mérési pontot, és az ott végzett mérést rögzítette a következő adatokkal: a beszélő neve, a szó hangsora, a hang sorszáma a szóban, a magánhangzón belüli mérési pontok, a magánhangzót megelőző és követő hang, a formánsok értéke a 25, 50, 75%-os ponton, valamint a hangfájl elérési útvonala. A mérés folyamatát az 1. ábra mutatja. Az algoritmust alkalmazva a Praat létrehozta a formáns adatbázis kiindulási, nyers változatát. Az 1. táblázatban mintát mutatunk be az adatbázisból.



1. ábra
Az automatikus formánsmérő szkript algoritmus

1. táblázat: Az *ácsjelvény* szó három magánhangzójában mért formánsadatok megadása az adatbázisban

Szó	Hang sorszáma	Előző hang	Mért hang	Következő hang	Pont	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
férfi A:CjelvE:N	1	#	A:	C	50%	723	1478	2561	3529
férfi A:CjelvE:N	1	#	A:	C	75%	567	1654	2586	3593
férfi A:CjelvE:N	4	j	e	l	25%	490	1845	2466	3506
férfi A:CjelvE:N	4	j	e	l	50%	576	1683	2452	3566
férfi A:CjelvE:N	4	j	e	l	75%	585	1597	2481	3596
férfi A:CjelvE:N	7	v	E:	N	25%	382	1965	2472	3508
férfi A:CjelvE:N	7	v	E:	N	50%	392	2027	2560	3550
férfi A:CjelvE:N	7	v	E:	N	75%	361	2105	2644	3522

Manuális hibafeltárás és javítás

Az automatikus elemző által produkált nagy mennyiségű formánsadattal kapcsolatosan tudjuk, hogy vannak köztük hibás értékek is. A hibák több forrásból is eredhetnek (rossz feldolgozás, algoritmikai tévesztés, téves emberi jelölés stb.). A kérdés az, hogy milyen módszerrel tudjuk megtalálni a hibás

számértékeket az adathalmazban, valamint, hogy mi alapján javítunk. A hibakeresésre többlépcsős megközelítést dolgoztunk ki. Első lépésként formánsokként emelkedő sorrendbe rendeztük az adatokat. Tájékozódásul felhasználtuk a magyar magánhangzókra vonatkozó korábbi kutatások adatait (Olaszy 1989, Gósy 2004), és ezek tükrében meghatároztuk a várható és a mért érték közötti különbségből azt, hogy elfogadható vagy hibás-e az adat. Az ellenőrzéshez és a javításhoz felhasználtuk a fenti honlap hangspektrogramjait is, vizuális elemzés segítette a döntéshozatalt. Hibás adat esetén azt átírjuk.

Nézzünk meg néhány kiugró hibát az algoritmus által készített nyers adathalmazból. Az F_1 esetében a legalacsonyabb érték 164 Hz volt női hangnál, ami egy [ɔ] hangra vonatkozott. Ez egyértelműen rossz adat. A legmagasabb F_1 érték pedig 1902 Hz volt, ami egy női [i] hanghoz volt jelölve. Az ilyen durva hibák többnyire abból eredtek, hogy a hangdefiníciók elcsúsztak a szón belül a konverziónál, vagy korábbi feldolgozási hiba volt az adatokban (kiindulási adatbázis hiba). Az F_2 és F_3 esetében is hasonló hibákat tartalmazott a skála eleje és vége. Az ilyen hibák kijavítása után a kivonásos elvet alkalmazva olyan adatsorokat képeztünk, amelyekben az F_2-F_1 , F_3-F_2 és F_4-F_3 értékeket szerepeltettük sorba rendezve. A formánsok távolsága ugyanis megadott határokon belül mozog minden magánhangzónál. A különbségképezés legkirívóbb eredményeire is mutatunk példát, ezeket a hibás értékeket már a formánselemző program tévesztései eredményezték. Az F_2-F_1 legkisebb értéke 14 Hz volt, és 100 Hz alatti különbség 25 esetben fordult elő. Az F_3-F_2 különbségnél 7 Hz volt a legkisebb, és 127 esetben volt 100 Hz alatt a két különbség. Az F_4-F_3 mérésből 8 Hz volt a legkisebb, és 58 esetben 100 Hz alatti volt a különbség a két formáns között. Az ilyen eltéréseket már mérési hibának tekintettük. Ez a hibacsoport önmagában is elgondolkodtató, hiszen egy formánselemzőtől elvárható, hogy 10-20 Hz távolságra ne jelöljön két külön formánst. Ezeket a hibás mérési értékeket a vizuális feldolgozási módszerrel javítottuk. A manuális hibafeltárás harmadik lépéséhez felhasználtuk azt a fonetikai ismeretet, hogy egy magánhangzón belül az egyes formánsok mozgástere behatárolható és a hangkörnyezet ismeretében tipizálható is. Így a mért formáns frekvenciákból és a három mérési pontból automatikusan megállapíthatók a gyanús elemek. Ezeket manuálisan ellenőriztük és javítottuk. A hibakeresés utolsó fázisában eloszlásokat rajzoltunk az $x = F_1$ és $y = F_2$, majd az $x = F_3$ és $y = F_4$ síkokra a mért 14 féle magánhangzóra, külön a férfi és külön a női ejtésre. Ezeken a diagramokon a kiugró értékeket ellenőriztük és javítottuk. Mindösszesen 10307 formánsadatot javítottunk, ebből a férfi ejtésű 916 darab volt. Ennek a munkának a végén érkeztünk el ahhoz a ponthoz, amikor azt mondhattuk, hogy a formánsadatbázisunkban csak minimális számú hiba maradt, tehát referencia-adatbázisnak tekinthető.

A mérési, javítási eredmények összefoglalása

Az összes magánhangzó 25,1%-ában kellett kézi javítást végezni (egy vagy több formáns adatát kellett korrigálni), összesen 2669 hangban, ebből a férfi ejtésű magánhangzók száma 298 (11,2%). A hibák zöme tehát a női ejtéshez kapcsolódik. Az automatikus mérés hibáinak összesítését nemekre bontva a 2. táblázat mutatja.

2. táblázat: A hibásnak talált formánsadatok részletezése

	F₁ javítása a V-ben (db)	F₂ javítása a V-ben (db)	F₃ javítása a V-ben (db)	F₄ javítása a V-ben (db)	Összes javítás (db)
Férfi	177	215	235	289	916
Nő	185	1516	3104	4586	9391

A férfi hangban a mért magánhangzók 88,8%-ában voltak pontos adatok, a nőiben 76,4%-ban. A Praat program nemektől független mérési pontossága ebben a mérésben tehát 74,9%. A részletezett adatok azt mutatják, hogy minél magasabb formánsról van szó, annál kevésbé pontos a formáns meghatározása. Az F₁ hibái azt mutatják, hogy ezt a formánst nagy pontossággal meg tudja mérni a Praat elemző, ha mégis téveszt, akkor egyforma gyakorisággal fordul elő a hiba a férfi és a női hangban. A férfi hangban minden formáns tévesztési aránya közel egyenlő. A női hangnál az F₂ esetében a hibás adatok száma jelentősen megugrik, és ez fokozatosan növekszik az F₃ és az F₄ esetében. A hibás formánsmérések hangfüggőek (3. táblázat). A férfi hangnál a leggyakrabban az [ɔ o ɔ:] hangoknál tévesztett az automatikus elemző függetlenül a formáns sorszámától. Ezekben az esetekben a Praat úgynevezett elcsúszási hibát mutatott, ami azt jelentette, hogy egy hamis formánst észlelt az F₂-nek, így a valódi F₂-t az F₃-ra, a valódi F₃-at pedig az F₄-re tette.

3. táblázat: A hibás formánsértékek előfordulása hangonként

Bemondó	Mért hang	F₁-hibák (db)	F₂-hibák (db)	F₃-hibák (db)	F₄-hibák (db)
férfi	A:	2	15	15	46
férfi	a	30	36	69	99
férfi	o	59	83	79	79
férfi	o:	7	33	15	16
férfi	u	8	18	10	16
férfi	u:	3	12	7	6
férfi	i	5	32	44	37
férfi	i:	0	0	5	3
férfi	U	0	0	1	0
férfi	U:	1	5	9	7
férfi	E:	0	24	28	28

Bemondó	Mért hang	F ₁ -hibák (db)	F ₂ -hibák (db)	F ₃ -hibák (db)	F ₄ -hibák (db)
férfi	O	0	0	3	10
férfi	O:	0	0	1	7
férfi	e	3	14	37	49
nő	A:	46	268	694	836
nő	a	32	63	256	628
nő	o	14	34	116	269
nő	o:	0	1	7	60
nő	u	1	7	28	97
nő	u:	1	7	5	28
nő	i	4	406	518	709
nő	i:	0	50	67	75
nő	U	1	4	14	46
nő	U:	0	5	1	18
nő	E:	1	322	469	586
nő	O	1	10	11	57
nő	O:	0	0	9	36
nő	e	34	361	902	1097

Megfigyeltük, hogy ha nazális környezetben van a V, akkor az elemző nehezen tud jó adatokat kinyerni, mert a formánsok jobban elmosódnak (például az [o] hang a *demonstráció* szóban).

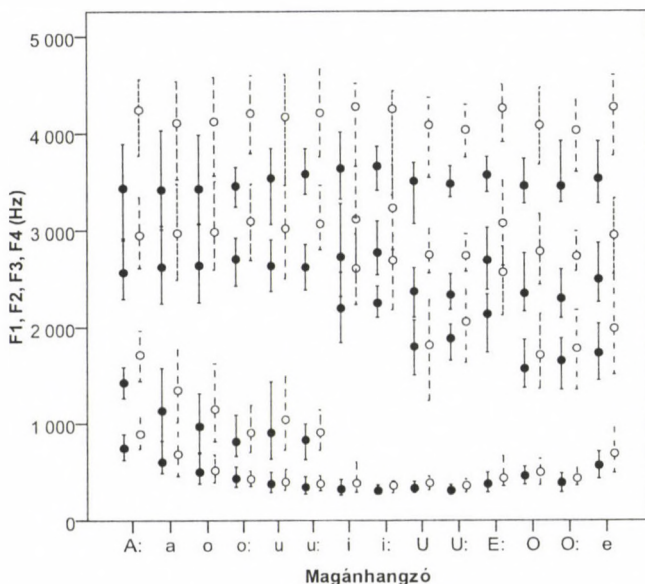
A formánsadatbázis adatainak végzett elemzések

A formánsadatbázisból a különböző változók szerinti lekérdezésekkel válogatásokat hajthatunk végre, és csoportokba rendezhetjük az adatokat. Ilyenek lehetnek a férfi és női ejtés közötti sajátosságok, általános átlagok, szűrt átlagok, V csoportok viselkedése a megelőző és követő hang (képzési helyek) függvényében stb. Az így összegyűjtött adattáblázatokon a továbbiakban végezhetünk bármilyen statisztikai feldolgozást. Ilyen mérésekre mutatunk be néhány példát.

Formánsátlagok és -sávok

Az egyes magánhangzók formánsainak frekvenciasávjai hangonként ábrázolhatók (2. ábra). Bolla (1978) közölt először ilyen mérési eredményeket, magyar beszédre, beszédsszintézis felhasználásával. A most bemutatott eloszlási sávok valós magánhangzókából származnak, és hasonló tendenciákat mutatnak, mint a szintézisből kapott korábbi adatok. A diagramok mutatják, hogy a női hang formánsai átlagosan magasabbak, mint a férfié. A második jellemző, ami kiolvasható az eloszlási sávokból, hogy az egyes formánsok sávjai egymásba csúszhatnak. Bolla ábrázolásában ilyen átfedés nem látható, mivel formánsszintetizátort használt, amelyben csak meghatározott formáns-sávokat lehetett beállítani. A harmadik jellemző, hogy az F₃, F₄ mozgási sáv-

ja lényegesen szűkebb, mint az elő két formánsé. Az egyes magánhangzók formánsaira kapott átlagokat és szórásokat mutatja a 4. táblázat.



2. ábra

A magánhangzók formánssávjai négy formánssra a formánsadatbázis adatai alapján (szaggatott: nő, folytonos vonal: férfi, karika: átlag)

4. táblázat: A formánsadatbázis magánhangzóinak formánsadatai

Bemondó	Mért hang	F ₁ (Hz) átlag/szórás	F ₂ (Hz) átlag/szórás	F ₃ (Hz) átlag/szórás	F ₄ (Hz) átlag/szórás
férfi	A:	746/46	1426/63	2566/119	3449/156
férfi	a	602/48	1132/117	2624/138	3422/118
férfi	o	499/52	970/105	2640/136	3430/117
férfi	o:	433/42	814/77	2704/90	3461/65
férfi	u	379/47	903/135	2633/124	3547/129
férfi	u:	342/40	848/173	2619/141	3581/106
férfi	i	326/26	1796/143	2385/137	3514/102
férfi	i:	307/26	1889/104	2365/166	3507/142
férfi	U	322/29	2193/99	2725/156	3650/117
férfi	U:	304/19	2248/79	2795/165	3682/104
férfi	E:	375/32	2142/110	2699/148	3591/137
férfi	Ö	456/34	1569/111	2352/109	3461/95
férfi	Ó	387/35	1643/144	2301/108	3461/98
férfi	e	563/43	1730/86	2495/93	3542/133

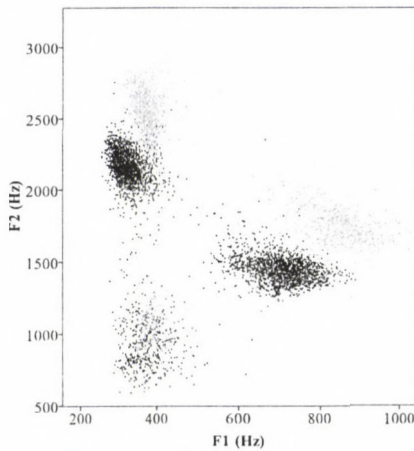
Bemondó	Mért hang	F ₁ (Hz) átlag/szórás	F ₂ (Hz) átlag/szórás	F ₃ (Hz) átlag/szórás	F ₄ (Hz) átlag/szórás
nő	A:	890/64	1713/93	2950/145	4240/138
nő	a	683/60	1346/126	2972/163	4106/145
nő	o	516/50	1148/127	2982/175	4121/184
nő	o:	427/33	904/70	3093/145	4200/173
nő	u	396/38	1057/191	3014/199	4164/234
nő	u:	378/37	908/98	3037/173	4196/203
nő	i	383/28	1815/241	2745/125	4088/170
nő	i:	356/28	2049/157	2739/109	4032/131
nő	U	382/40	2604/127	3110/200	4274/101
nő	U:	356/21	2697/101	3240/181	4261/92
nő	E:	435/49	2553/148	3070/138	4253/99
nő	O	494/44	1709/181	2778/134	4083/143
nő	O:	431/44	1778/175	2727/96	4028/137
nő	e	682/73	1984/138	2944/131	4265/123

F₁-F₂ eloszlások

A magánhangzók F₁-F₂ síkon történő ábrázolása megmutatja az egyes hangokra jellemző halmazokat az első két formáns függvényében. Ilyen ábrák a szakirodalomban szerepelnek a magyar magánhangzókra is (Gósy 2004). Példaként megmutatjuk, hogy milyen halmazképet mutat a három kardinális magánhangzó az adatbázis adatait feldolgozva (3. ábra). Az [a:], [i] hangok esetében jól elkülönül egymástól a férfi és a nő ejtésből keletkezett halmaz, az [u]-nál viszont nem. Míg az [a:], [i] esetében egyértelműen látható, hogy a női ejtés formánsai átlagosan magasabbak, mint a férfié, az [u]-nál ez csak az F₂-re vonatkoztatható némileg.

Szűkített mérések

A mérés terét szűkíthetjük a hangkörnyezet figyelembe vételével. Példaként bemutatjuk annak a lekérdezésnek az eredményét, amelyik arra keres választ, hogy a palatális mássalhangzók milyen hatást gyakorolnak az [o] és [u] magánhangzó formánsaira (változik-e az átlaghoz képest a formáns értéke a hang közepén)? A méréshez kétféle lekérdezést végzünk. Az [o] hangra lekérdeztük azokat az eseteket, amikor palatális mássalhangzó volt az egyik-szomszéd, majd azokat az eseteket, amikor nem volt palatális szomszédja. Az eredményt az 5. táblázatban mutatjuk meg. A másik lekérdezésnél azokat az [u] hangokat válogattuk ki, amelyeknek mindkét szomszédja palatális volt. Az eredményt a 6. táblázat mutatja. Mindkét esetben a palatális környezet jelentősen megemeli az F₂ értékét.



3. ábra

Az [a:], [u], [i] hangok formánsfrekvenciái az F_1 - F_2 koordináta-rendszerben
(fekete: férfi, szürke: nő)

5. táblázat: Az [o] hang első két formánsának átlaga, ha palatális előzi meg,
és ha egyik szomszédja sem palatális

Hang	Hangkörnyezet	Bemondó neme	F_1 (Hz) 50%	F_2 (Hz) 50%	Mért eset- szám (db)
[o]	palatális előzi meg	férfi	512	1050	18
[o]	nem palatális előzi meg	férfi	498	968	526
[o]	palatális előzi meg	nő	494	1258	17
[o]	nem palatális előzi meg	nő	517	1145	527

6. táblázat. Az [u] hang első két formánsának átlaga, ha mindkét szomszédja
palatális, és ha egyik sem az

Hang	Hangkörnyezet	Bemondó neme	F_1 (Hz) 50%	F_2 (Hz) 50%	Mért eset- szám (db)
[u]	palatálisok a szomszédjai	férfi	323	1153	3
[u]	nem palatálisok	férfi	378	890	162
[u]	palatálisok a szomszédjai	nő	381	1256	2
[u]	nem palatálisok	nő	394	1029	158

Lekérdezhetők például az egyes formánssávok adott frekvenciaértékeire jellemző hangkörnyezetek. Vegyük például a férfi ejtésű [a:] hang F_1 sávjának alsó és felső határához közeli értékekhez tartozó hangkörnyezeteket. Az

alsó érték vizsgálatához a 650 Hz alatti hangokat kértük le (7. táblázat), a felsőéhez a 870 Hz felettieket (8. táblázat).

7. táblázat: Az [a:] hangra kapott szavak $F_1 < 650$ Hz esetére

	Szó	A hang sorszáma	Előző hang	Mért hang	Köv. hang	Mérési pont	F_1 (Hz)
férfi	A:dZ:A:moj	3	dZ:	A:	m	50%	629
férfi	jajkiA:ltA:S	6	i	A:	l	50%	642
férfi	pajzA:n	5	z	A:	n	50%	625
férfi	fudZijA:ma	6	j	A:	m	50%	626
férfi	kiutazA:S	7	z	A:	S	50%	644

8. táblázat: Az [a:] hangra kapott szavak $F_1 > 870$ Hz esetére

	Szó	A hang sorszáma	Előző hang	Mért hang	Köv. hang	Mérési pont	F_1 (Hz)
férfi	A:SvA:ny	1	#	A:	S	50%	884
férfi	A:tmeG	1	#	A:	t	50%	873
férfi	A:tnE:z	1	#	A:	t	50%	873
férfi	A:tfordi:t	1	#	A:	t	50%	880
férfi	A:trepUl	1	#	A:	t	50%	890

Első vizsgálati eredményként a táblázatok adataiból azt a következtetést vonhatjuk le, hogy a magas F_1 értékek inkább hangsorkezdő helyzetben jellemzőek, az alacsonyok pedig akkor, ha nazális követi a magánhangzót, illetve ha a képzési hely az artikulációs csatorna középső területeire esik (dentalveoláristól a palatálisig), és a magánhangzó hangsorbelseji pozícióban van.

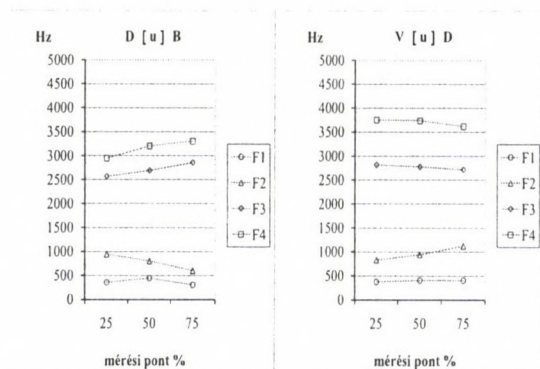
Az adatbázisból lekérdezhetők a jellemző formánstávolságok is az egyes magánhangzókra, a megelőző és követő mássalhangzó függvényében. A képzési helyek rövidítései: B = bilabiális; L = labiodentális; D = dentális; A = alveoláris; P = palatális; V = veláris. Hasonlítsuk össze például a B[a:]B és a D[a:]P hanghármasok magánhangzójára jellemző formánstávolságokat férfi ejtésben. A mérésnél a fenti két mintacsoport összes reprezentánsát kigyűjtöttük az adatbázisból, az 50%-os mérési pontokra (9. táblázat).

9. táblázat: A jellemző formánstávolságok az [a:] magánhangzó közepén két hangkörnyezeti konfigurációra, férfi ejtésben

	F_1 és F_2 átlagos távolsága (Hz)	F_2 és F_3 átlagos távolsága (Hz)	F_3 és F_4 átlagos távolsága (Hz)
B[a:]B	585	1335	813
D[a:]P	784	1098	863

A példában látható, hogy az F_1 – F_2 , illetve az F_2 – F_3 távolságában lényeges az eltérés, az F_3 – F_4 közel azonos. Az így meghatározott formánstávolságok jellemzőek lehetnek a hangkörnyezetre, tehát ezekből további következtetéseket vonhatunk le. Az ilyen jellegű adatok felhasználhatók a formánskinyerő algoritmusok finomítására is.

Célzott lekérdezéssel olyan adatokhoz is juthatunk, amelyekből felrajzolhatók, hogy a mássalhangzó képzési helye hogyan befolyásolja a hozzá csatlakozó magánhangzó formánsainak a mozgását. Ez a mozgás két tényezőtől függ. Egyrésztől a mássalhangzó képzési helyének akusztikai vetületétől (Olaszy 2007), másrésztől magától a magánhangzótól, annak formánsaitól. A két hatás együttesen jelenik meg a V formánsmozgásaiban. Mivel az adatbázis minden magánhangzóra három belső ponton tárolja a formánsadatokat, kirajzoltathatók és rendszerezhetőek a hangon belüli jellemző formánsmozgások (vonalas közelítéssel). Példaként bemutatjuk, hogy a férfi ejtésű [u] hang formánsmozgásai hogyan alakulnak a D[u]B, illetve az V[u]D esetekben (4. ábra). A kétféle hangkörnyezeti konfiguráció hatása egyértelműen tükröződik a formánsmenetekben. Az F_2 mozgásiránya egyrésztől ellentétes, másrésztől más frekvenciatartományba esik. Az F_3 , F_4 esetében is hasonló a tendencia. Vizsgálhatók továbbá, hogy a formánsmozgások hogyan alakulnak a magánhangzó függvényében.

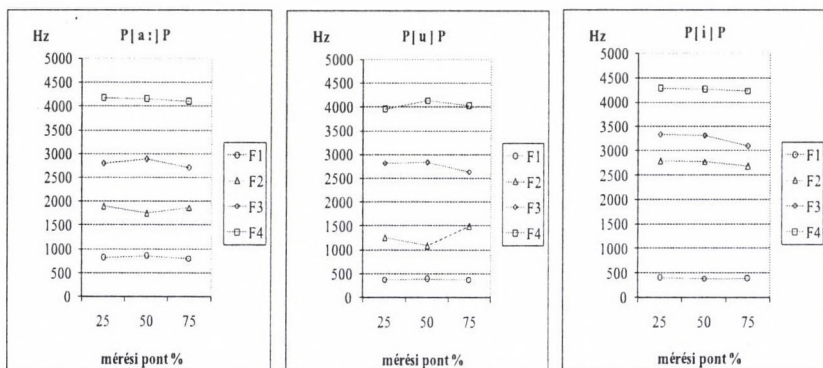


4. ábra

Jellemző formánsmozgások az [u] hangban a két hangkörnyezeti konfigurációban

Készítettünk ábrákat a női ejtésű három kardinális magánhangzóra, CVC hanghelyzetre, ha palatális (5. ábra), illetve ha veláris mássalhangzók veszik közre őket (6. ábra). A palatális képzési helyhez tartozó akusztikai vetület jellemző értéke a 2000–2500 Hz-es sávba esik, továbbá ismert, hogy a palatálisok frekvenciailleszkedésre nem hajlamosak, magukhoz idomítják a magán-

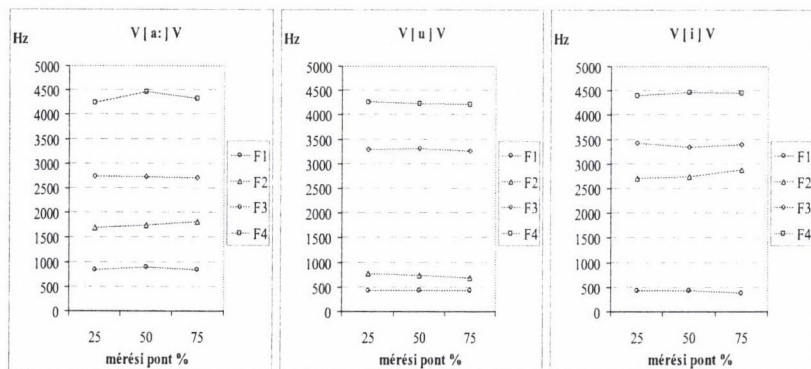
hangzók formánsait. Ez az illeszkedés olvasható le az 5. ábrából, különösen az F_2 és F_3 vonatkozásában. Ha a magánhangzóra jellemző F_2 alacsonyabb, mint a palatális mássalhangzó akusztikai vetülete, akkor a középponttól kiindulva az F_2 a 2000–2500 Hz-es tartomány felé mozdul el (emelkedik), ha magasabb, akkor csökken. Az F_3 -ra is ugyanez vonatkozik.



5. ábra

A palatális mássalhangzó-környezet hatása a közrefogott magánhangzó formánsaira a kardinális magánhangzókban

A veláris mássalhangzóról azt tudjuk, hogy a legnagyobb az illeszkedési hajlandósága a hangkörnyezetéhez, tehát nem kényszeríti a hozzá kapcsolódó magánhangzó formánsait illeszkedő mozgásra. A 6. ábrán látható formánsvonalatok ezt maximálisan alátámasztják.



6. ábra

A veláris mássalhangzó-környezet hatása a kardinális magánhangzók formánsaira

Összefoglalás

Az ismertetett új eljárás közelebb hoz ahhoz, hogy a jövőben nagyméretű beszédatadtbázisokon formánsméréseket lehessen végezni, az adatokat hatékonyan lehessen javítani, majd adatbázisba rendezni. A bemutatott eljárást egy nyilvános szóadatbázisra alkalmaztuk, így az adatok bármikor újra ellenőrizhetők. A munka konkrét eredménye a létrehozott formánsadatbázis, amely referenciának tekinthető. Az elkészített formánsadatbázis lehetőséget ad összehasonlító mérésekre, valamint bármely formánselemző algoritmus hibaszázalékának kimutatására is. Ezt a Praat program jelen méréseire meg is határoztuk. A formánsadatbázis számos további kutatás alapját képezheti. Az adathalmaz jól használható az oktatásban is, használati lehetőségeinek érzékeltetésére bemutattunk néhány lekérdezést. További terveink között szerepel, hogy a referencia-adatbázisra támaszkodva olyan adatjavítási gépi eljárást dolgozzunk ki, amely kiküszöböli a manuális javítási fázist. Ezzel elérhető, hogy más nagyméretű magyar beszédatadtbázisokban is megbízható formánsméréseket lehessen automatikusan végezni, ha azok megfelelő belső címkékkel vannak ellátva. Tervezzük az adatbázis internetes közzétételét is.

Irodalom

- Abari Kálmán – Olasz Gábor 2007. A magyar beszéd hangkapcsolódásainak bemutatása az interneten. *Beszédkutatás* 2007. 178–186.
- Bolla Kálmán 1978. A magyar magánhangzók analízise és szintézise. *Magyar Fonetikai Füzetek* 1. 53–68.
- Böhm Tamás – Németh Géza 2007. Algorithm for formant tracking, modification and synthesis. *Híradástechnika* 2007/1. 15–20.
- Deng, Li – Cui, Xiaodong – Pruvencok, Robert – Huang, Jonathan – Momen, Safiyy – Chen, Yanyi – Alwan, Abeer 2006. A database of vocal tract resonance trajectories for research in speech processing. *Proceedings of the ICASSP 2006*. 369–372.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Lee, Minkyu – van Santen, Jan – Möbius, Bernd – Olive, Joseph 1999. Formant tracking using segmental phonemic information. *Proceedings of Eurospeech 1999. Vol. 6*. 2789–2792.
- Olasz Gábor 1989. *Elektronikus beszédelőállítás*. Műszaki Könyvkiadó, Budapest.
- Olasz Gábor 2007. *Mássalhangzó-kapcsolódások a magyar beszédben*. Tinta Könyvkiadó, Budapest.
- Tarnóczy Tamás 1941. *A magyar magánhangzók akusztikai szerkezete*. Kir. Magy. Pázmány Péter Tudományegyetem Általános Nyelvészeti és Fonetikai Intézete, Budapest.

A SVÁVARIÁCIÓK AUTOMATIKUS FELISMERÉSE MAGYAR NYELVŰ SPONTÁN BESZÉDBEN

Beke András – Szaszák György

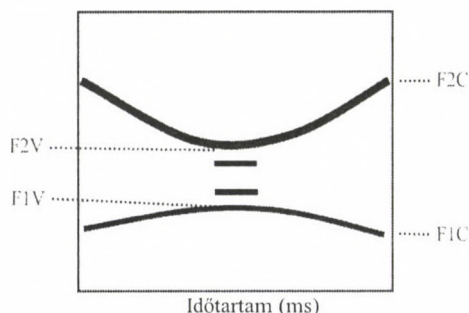
Bevezetés

A magánhangzó redukciójáról akkor beszélünk, amikor annak képzésekor az artikulációs konfiguráció a centrális irányba tolódik el, megváltoztatva ezzel a magánhangzó minőségét. A magánhangzó redukálódását a magyar szakirodalomban szokás neutralizálódott, semleges magánhangzónak és svának nevezni. A jelenség számos tényezővel mutat összefüggést: megváltozhatnak a szótag hangsúlyviszonyai, rövidülhet az időtartam, csökkenhet az intenzitás vagy az artikulációs erőfeszítés.

A felolvasásban a magánhangzók általában redukálódás nélkül valósulnak meg (természetesen a nyelvspecifikus sajátosságok figyelembevételével), hiszen a beszéd tervezésére kisebb energiát kell fordítani, így a szöveg artikulációs megvalósítása pontosabb, mint a spontán beszédé. Bergem (1994) kimutatta, hogy a svá F_2 -értéke egyes hangkörnyezetekben majdnem egyenes irányban mozog a magánhangzó kezdetétől a végéig, amely minimális artikulációs mozgásra utal a magánhangzó képzésében. Másrészt, a svá bizonyos esetekben teljesen asszimilálódik a fonetikai hangkörnyezetéhez. A folyamatos beszédben megfigyelhető egyfajta artikulációs gazdaságosság, amely többek között a svá megjelenésében is tetten érhető. A magyar nyelvben a neutrális magánhangzók képzésekor a vokális csatornára a semleges konfiguráció a jellemző, amelynek megfelelően az ajkak nem kerekítettek (Gósy 2006). A H&H elméletre (Lindblom 1990) is figyelemmel, az artikuláció során a beszédben kettős cél húzódik meg: a beszélő egyrészt törekszik a minél érthetőbb kivitelezésre – ami olykor a hiperartikulált beszéd létrejöttéhez vezet –, másrészt igyekszik minimalizálni az artikulációra fordított energiát, ami pedig a magánhangzók redukciójához vezethet (ez a célalulmúlás: target under-shoot; Lindblom 1990). A célalulmúlás akkor következhet be, ha a magánhangzó olyan hangkörnyezetben realizálódik, amelynek célkonfigurációja távol áll a szándékolt magánhangzó célkonfigurációjától, illetve akkor, ha a magánhangzó időtartama jelentősen csökken (Lindblom 1990, Gósy 2006, Koopmans 1994). A jelenség tetten érhető a formánsértékek mozgásában is (1. ábra).

A hazai és a nemzetközi szakirodalom az első két formánst rendszerint elégséges akusztikai jegynek tartja a magánhangzó meghatározására, mivel az első formáns értéke jól korrelál a nyelv függőleges mozgásával, a másodi-

ké pedig a nyelv vízszintes mozgásával (Gósy 2006, Patterson et al. 2003, Slifka 2005, Ivan Yuen et al. 2007).



1. ábra

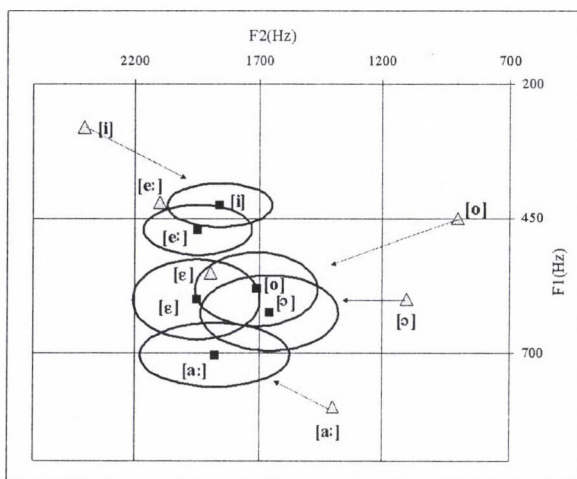
A svá magánhangzó formánsstruktúrájának sematikus képe
(Flemming et al. 2006 nyomán)

A semleges magánhangzó realizációi akusztikailag nem azonosak. A nemzetközi és a hazai kutatások szerint a redukált magánhangzó akusztikai realizációja függ a fonetikai kontextustól (Bunne–Lilley 2008), a fonetikai pozíciótól (Flemming et al. 2007), a szótag helyzetétől a szóban (hangsúlyos-hangsúlytalan – Lindblom 1963; Delattre 1969; Gay 1978) és a beszéd típusától (Masanobu 2006).

A magánhangzók különféle funkcióban redukálódhatnak: hezitáció hatására, koartikulációs következményként (*r* + svá) stb. A semleges magánhangzó a spontán beszédben gyakran egy másik magánhangzó helyett realizálódhat – ezt a jelenséget **helyettesítő funkciónak** nevezik. A franciában 7,8%-ban, az angolban 22,9%-ban, a németben 30%-ban, utóbbi nyelvben is előfordulhat a svá helyettesítő funkcióban (Gósy 2006, Gósy 2008, Beke 2008, Beke–Grácsi 2009). A spontán beszédben a magánhangzók akár 50%-a is realizálódhat sváként. E magas előfordulás oka, hogy a spontán beszédre, a felolvasással ellentétben, a laza artikuláció a jellemző, ezért a magánhangzó minősége eltolódik a semleges artikulációs konfiguráció felé (2. ábra).

Az felolvasott beszédre és a hasonló beszéd típusokra (például újságfelolvasás, hírbemondás, időjárás-jelentés) már léteznek olyan felismerő rendszerek (speech-to-text rendszerek), amelyek sokszor 90%-ot meghaladó pontossággal fordítják át a beszédet szöveggé. A beszéd felismerő rendszerek pontossága azonban a spontán beszédén drasztikusan romlik (Furui 2007). Az eredmények romlását az okozza, hogy az akusztikai és nyelvi modelleket általában írott nyelvtan szabályaiból és a felolvasott szövegek nyelvéből építik ki, a spontán beszéd és az olvasott beszéd pedig jelentősen különbözik mind

akusztikailag, mind nyelvtanilag (Furui 2005). Az akusztikai és a nyelvtani jelenségek sokkal változatosabban jelennek meg a spontán beszéd közben, mint az előre megtervezett beszédben. A spontán beszédben a magánhangzók akusztikai realizációjukban jelentősen eltérhetnek a normáltól, így bonyolítják a felismerésben használt modell nyelvtanát: a magánhangzó realizálódhat teljes spektrális minőségében, ahogyan a felolvasásban, vagy realizálódhat redukált, céllalumult konfigurációval, amelynek számos akusztikai megvalósulása is lehetséges a spontán beszédben. A spontán beszéd felismerésében ezért olyan fonetikai variációkkal és a spontán beszéd jellegéből adódó egyéb sajátosságokkal is számolni kell a nyelvi modellben, amelyek mint elkülönült egységek vesznek részt a felismerésben. Ilyen jelenség lehet a svá magánhangzó is.



2. ábra

A magánhangzók F_1/F_2 értékei olvasott (szürke háromszög), illetve spontán beszédben (fekete négyzet) (Beke–Grácsi 2009 alapján)

A magánhangzók általában a hangsúlytalan szótagban redukálódhatnak, míg a hangsúlyos szótagokban általában nem. Ezt a szabályosságot felhasználva a hangsúlyos-hangsúlytalan szótagok elkülönítésében más akusztikai jegek mellett (pl: alaphangmagasság, intenzitás stb.) fontos szerepe lehet a magánhangzó minőség modellezésének is. A hangsúlyos-hangsúlytalan szótagok és a magánhangzó minősége közötti összefüggés felhasználható a szavak kezdetének automatikus felismerésében a kötött hangsúlyú nyelvben, mivel ezekben a nyelvekben a szó első szótagja hangsúlyos, ami valószínűsíti a magánhangzó tiszta realizációját, míg hangsúlytalan szótagban a redukált variáns megjelenését (Xie et al. 2004, Halpern 2006, Lindblom, 1963; Delattre, 1969;

Gay, 1978; Beinum, 1989; Engstrand–Nordstrand 1984, Cruttenden 1997): jöllehet ez az összefüggés nem teljes, hiszen a magánhangzó hangsúlytalan szótagon megjelenhet eredeti realizációjában és semleges realizációjában is (Ladefoged 1993, Jeans et al. 2000).

Az akusztikai kutatások nagy része a magánhangzók akusztikai szerkezetét a formánsaikkal jellemzi, jöllehet a formánsok mérése időigényes és mindig kézi javításra szorul. A redukálódott magánhangzók formánsainak meghatározása igen nehéz, hiszen a svá időtartama sokszor nagyon rövid, ami miatt a formánskövető algoritmusok pontatlan eredményt adnak. A hagyományos akusztikai jellemző alternánsaként a jelen tanulmányban MFC együtthatókat használtunk (Mel-Frequency Cepstral Coefficients), amely bementi jellemző vektorként szolgál a HMM (rejtett Markov-modell) betanításához és teszteléséhez. A formánsfrekvencia meghatározásával szemben az MFC együtthatók számítása gyors, és nem igényel kézi ellenőrzést (Bunnell–Lilley 2008).

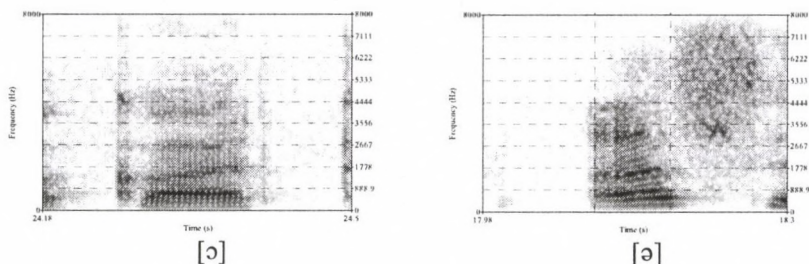
A jelen tanulmány célkitűzése, hogy a helyettesítő funkcióban álló svárealizációk spektrális paramétereit HMM-ekkel modellezzük magyar nyelvű, spontán beszédatadabázison, amely az elemeknek és lexikális egységeknek nagyszámú elfordulását, és változatos fonetikai kontextust tartalmazza.

A vizsgálat során feltételeztük, (i) hogy a svá magánhangzók spektrális paramétereikben eltérnek a többi magánhangzótól, illetve hogy (ii) a svá realizációja függ az eredeti magánhangzó célkonfigurációjától, továbbá, hogy (iii) a svá realizációja függ a fonetikai kontextustól, illetve a szótag pozíciójától a szóban (hangsúlyos-hangsúlytalan), a magyar nyelv esetében.

Anyag, kísérleti személyek, módszer

A jelen kutatásban a BEA (BESzélt nyelvi Adatbázis, Gósy 2008) korpuszból 19 magyar beszélő spontán beszédét dolgoztuk fel (8 férfi és 11 nő). A BEA felvételeit csendesített szobában készítették el Goldwave hangrögzítő programmal. A felvételeket 44,1 kHz-en mintavételezeten és 16-bites felbontással tárolták. Minden hangfájlnak elkészítettük a fonetikus átíratát. Az annotáció során a beszédhangokat kézzel szegmentáltuk a hangszínképük alapján a PRAAT beszédelemző szoftverben. A jelen kutatásban a következő magánhangzókat elemeztük: [ɔ], [a:], [ɛ], [e:], [i], [i:] [o], [o:], [u], [u:]. A beszédhangot akkor jelöltük az annotáció során svá magánhangzónak, (i) ha a beszédhang a Beke–Grácsi (2009) tanulmányban meghatározott formánssztruktúrával rendelkezett¹ (3. ábra), illetve (ii) ha a beszédhang a lehallgatás során svának hatott.

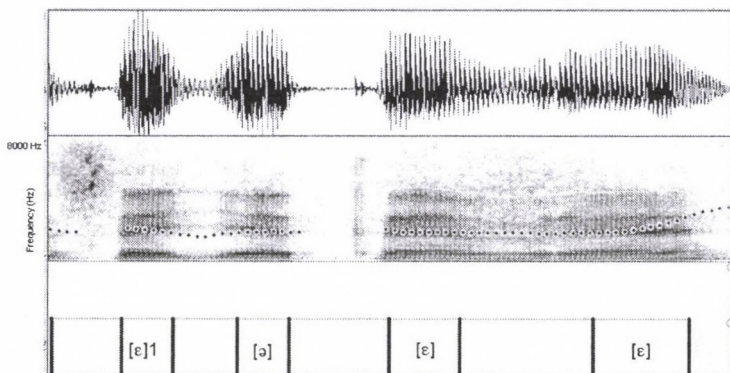
1. A spontán beszédben a leggyakoribb magánhangzók és az azok helyett realizálódott svák formánsértékeit határozták meg K-közép algoritmussal, amelyet auditive ellenőriztek.



3. ábra

Az [ɔ] magánhangzó és a svá [ə] magánhangzó spektrogramja (ugyanazon beszélőtől)

A szegmentálás során az annotációban azt is jelöltük, hogy a magánhangzó hangsúlyos vagy hangsúlytalan szótagon realizálódott. A magyar nyelvben mindig a szó első szótagjára esik a hangsúly (Siptár–Nádasdy 2000, Kálmán–Nádasdy 2001; 4. ábra). Azt, hogy egy szótag valóban hangsúlyos vagy sem, a szótag F_0 - és intenzitásértékével ellenőriztük (vö. Szaszák 2009).



4. ábra

A hangsúlyos szótagon realizálódott [ɛ]1 magánhangzó és a hangsúlytalan szótagon realizálódott [ɛ] és az [ɛ]-t helyettesítő svá magánhangzó spektrogramja, alaphangmagassága és intenzitása

A svát akkor jelöltük [ə]-val, ha az egységes, osztatlan beszédhangként szerepelt a hangfelismerésben. Nagy karakterrel jelöltük a svávariációkat: [A], [E], [O], ha a svát mint az eredeti magánhangzótól függő realizációt szerepeltettük a modellben. A kutatásban 4000 magánhangzót annotáltunk. A magánhangzókat és a svávariációkat HMM-ekkel (rejtett Markov-modell) modelleztük. A HMM tanításhoz a beszéd-adatbázisból származtatott nagy meny-

nyiségű paraméter szükséges, ezért a magánhangzókából kinyertük a spektrális tartalomra utaló MFC jellemzőket, majd kiszámítottuk ezek első két deriváltját (delták, delta-delták) is. A svárealizációk felismerését statisztikai elven, rejtett Markov-modell alapján végeztük el, amelyre a HTK fejlesztői rendszert (Young 2005) alkalmaztunk. A HMM-ek tanításához 2500, a teszteléséhez 1500 magánhangzót használtunk. A csoportosításra használt HMM modellek működésének kiértékelésére és összehasonlítására meghatároztuk az osztályozás pontosságát. A pontosság azt jellemzi, hogy az osztályozó algoritmus milyen mértékben azonosítja helyesen a beszédhangokat:

$$p = tp / (tp + fp),$$

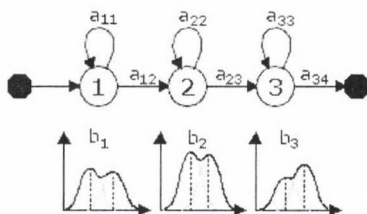
ahol a tp (true positive) a helyesen azonosított hangok száma, a fp (false positive) a tévesen osztályozott hangok száma.

A rejtett Markov-modell

Az automatikus beszédfelismerésben elterjedten használják a rejtett Markov-modelleket az egyes beszédhangok modellezésére. A beszédhangok modelljein kívül a rejtett Markov-modelles beszédfelismerő fontos tudásbázisa még a szótár – amely megadja, hogy mely szavak milyen beszédhangsorozatból épülnek fel – és az ún. nyelvi modell, amely azt adja meg, hogy adott szókörnyezetet feltételezve mely egyéb szavak előfordulása megengedett, illetve mely előfordulás mennyire valószínű. Visszatérve a beszédhangok modelljeihez, ezek szerepe az akusztikai beszédjel megfeleltetése, leképezése az egyes beszédhangoknak (ra). A jó megfeleltethetőséghez az akusztikai beszédjelet előzetes feldolgozásnak vetik alá (előfeldolgozás), például igen elterjedten használják az emberi hallást is modellező Mel-frekvenciás kepsztrális együtthatókra való átalakítást, de más eljárások is ismeretesek. A korábban már említett Mel-frekvenciás kepsztrális együtthatókat (angolul: Mel Frequency Cepstral (Coefficients) az alábbi módon számíthatjuk ki: elsőként gyors Fourier-transzformáció történik (Fast Fourier Transform, vagy röviden FFT), ennek során a beszédből egy rövid részt (pl. 25 ms hosszú darabot) kivágunk, súlyozzuk egy ún. ablakfüggvénnyel, és Fourier-transzformációval meghatározzuk a spektrumát. Ezután 10 ms-ot továbblépünk, és ugyanezt ismétljük mindaddig, amíg el nem érünk a feldolgozandó szakasz végéig. Ily módon 10 ms-onként megkaptuk a beszédjel spektrumát. Második lépésként ezeken a spektrumokon ugyanolyan elemzést végzünk, amelyet az emberi fül is, azaz a spektrumot az ún. kritikus frekvenciasávok szerint bontjuk fel. Ezt a műveletet szűrősoros elemzéssel hajtjuk végre, amely a kritikus sávoknak megfelelő frekvenciatartományokba bontja a jelet. A szűrősor általában 20 sáváteresztő szűrőből áll, amelyek kimenetén egy a sávba eső intenzitással arányos számszerű érték jelenik meg, azaz tulajdonképpen 10 ms-onként egy 20 dimenziós vektort kapunk. Mivel ezek az értékek egymással korrelálnak, ezért a dimenziószám csökkenthető, mégpedig egészen 12-re az ún. diszkrét koszinusz transzformáció segítségével (Discrete Cosine Transform, DCT).

Ezután általában hozzáveszik a vektorhoz a teljes beszédjel-szelet átlagos energiáját, majd az így összesen 13 érték első és második deriváltjait, így összesen egy 39 dimenziós vektort, ún. jellemzővektort kapunk.

A beszédhangok rejtett Markov-modelljei lényegében a hangra jellemző vektorok eloszlását adják meg. Figyelembe véve a jellemzővektorok spektrális származtatását ez tehát frekvenciatartománybeli modellezést jelent. A Markov-modellek leggyakrabban 3 állapotú, balról-jobbra felépítésű modellek – utóbbi azt jelenti, hogy a Markov-modell egyes állapotai között átmenet csak balról jobbra lehetséges. Minden állapothoz tartozik tehát egy valószínűségi eloszlást megadó függvény, amelyet statisztikai eszköztárral, leggyakrabban normális eloszlások szuperponálásával becslünk a modell ún. betanítása során.



5. ábra

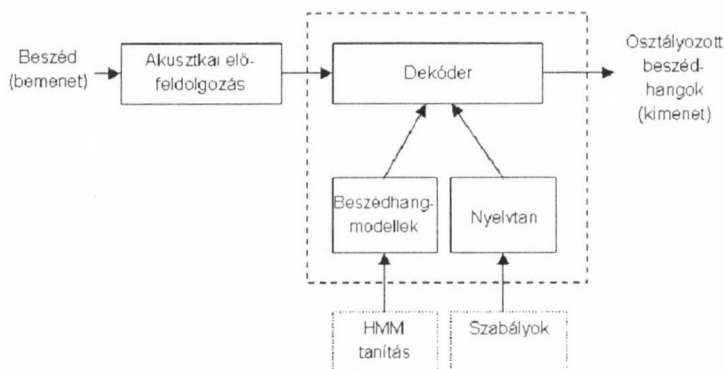
3 állapotú balról-jobbra felépítésű Markov-modell

[A modell tanítás során becsülendő paraméterei az állapotátmeneti valószínűségek (a_{ij}), valamint a normális eloszlások súlya, várható érték- és szórásvektora, amelyek együttesen megadják az eloszlást ($b_j(t)$).]

A betanítás során ezeknek a függvényeknek a paramétereit becsüljük meg. A beszédfelismerés során pedig a beszédből előállított jellemzővektorokat hasonlítjuk az egyes beszédhangok állapotainak megfelelő eloszlásokhoz. Minél inkább illeszkedik a jellemzővektor egy adott állapot eloszlásához, annál nagyobb súlyt rendel a hozzá kapcsolódó útvonalhoz a dekóder, azaz a tulajdonképpen beszéd-szöveg átalakító.

A Markov-modellek a beszédfelismerőben valójában kettős feladatot látnak el, a jellemzővektorok osztályozása mellett a beszédjelet illesztik is a neki megfelelő beszédhangsorozatra, azaz meghatározzák az egyes beszédhangok kezdő- és végidőpontjait. A Markov-modellek úgy is használhatók, hogy előre elkülönített osztályokra betanított modellek alapján osztályozzanak. Ilyen esetben a beszédfelismerőben használatos szótár szerepét az egyes osztályok listája, a nyelvtan szerepét pedig az osztályozás szabályai veszik át. A beszédhangok osztályozása esetén a lista az osztályozni kívánt beszédhangokból áll, az osztályozás szabályai pedig megadják, milyen beszédhangokat milyen sorrendben lehet illeszteni.

A szerzők a magánhangzók osztályozását (svá vagy teljes magánhangzó) valósították meg rejtett Markov-modellekkel. A HTK környezetben megvalósított osztályozó felépítése a 6. ábrán látható.



6. ábra
Gépi beszédfelismerő rendszer

Eredmények

1. A magánhangzó és a redukálódott magánhangzó osztályozása

A semleges magánhangzók akusztikai realizációi jóval változatosabbak, mint a magánhangzókéi (Browman–Goldstein 1992). A szegmentális és szuprasegmentális modellekben fontos szerephez juthat a svá automatikus felismerése, hiszen a folyamatos szófelismerésben a magánhangzó nem redukálódott, teljes realizációja jelezheti a szó kezdetét a beszédben (Dressler 1984, Madelska–Dressler 1996). Kopecký (2008) a beszédfelismerő rendszerébe beépítette a svá fonémát, amely a rendszer felismerési pontosságának javulását eredményezte.

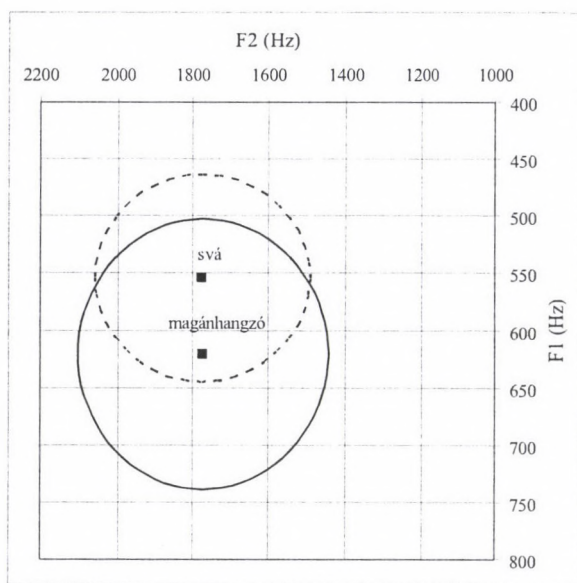
A magánhangzókat és a svárealizációkat 3 állapotú HMM-mel modelleztük. A tanítás során „V” szimbólummal jelöltük a magánhangzókat, és „S” szimbólummal a svárealizációkat. Mind a két modellt rendre 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvényvel tanítottuk. A nyelvtanban mindkét hangmodellt („V”, „S”) egyenlő súllyal rögzítettük (azaz egyenlő valószínűség mellett). A legjobb felismerési eredményt a 4 gaussos modell adta (1. táblázat).

A semleges magánhangzókra tanított HMM-modell független az eredeti magánhangzótól, illetve a szótag hangsúlyosságától. Az eredmények azt mutatják, hogy a spontán beszéden tanított HMM modellel a svárealizációk 71,97%-át osztályozta helyesen a rendszer. Az eredményből arra következtethetünk, hogy a semleges magánhangzó rendelkezik egy jól meghatározható spektrális karakterrel, amely megkülönbözteti a többi magánhangzótól. A svá

akusztikai realizációi között azonban további kisebb csoportok vannak, amelyek lehetséges sváalcsoportokra utalnak. Ilyen csoport lehet az, amelyik átfedésben lehet a magánhangzókkal is. Az átfedés szemléltetéséhez kimértük a beszédhangok első két formánsát. A 7. ábrán látható, hogy az F_1/F_2 terében a semleges magánhangzó és a magánhangzó között nagy az átfedés. Az átfedést a magánhangzó és a semleges magánhangzók között a realizációk széles skálája okozza (Bondarko et al. 2003).

1. táblázat: A magánhangzók és a svák felismerési eredményei (4 Gauss)

	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Töröltek száma ²
„V”	706	79,46	79,46	0
„S”	157	71,97	71,97	0



7. ábra

A magánhangzók (V) és a semleges magánhangzók (S) F_2/F_1

2. A magánhangzók és az egységes svámodell

Flemming (2007) kimutatta, hogy a svá fonéma realizációinak lehetnek különböző alcsoportjai: közép-centrális svá és kontextusfüggő svá. A svá reali-

2. Törlésnek nevezik a beszédfelismerésben a valós tesztanyagban megjelenő, azonban a felismerő által kihagyott – így tévesen fel nem ismert – elemet.

zációk variációinak egy része a kontextus hatására megváltozik, és egy sajátos kontextusfüggő hangminőséget hoz létre, amely a svának egy akusztikai alcsoportja lehet. A nemzetközi és a hazai szakirodalom sem egységes abban, hogy milyen tényleges okai vannak a svá variáltságának, illetve melyek a lehetséges sváalcsoportok. Flemming szerint nyilvánvaló, hogy a semleges magánhangzónak két típusa létezik, azonban a levonható következtetések nem egyértelműek. A magánhangzó redukciója jelezheti a hangsúlytalan és hangsúlyos szótag közötti szembenállást is, ami az angol nyelvben szabályszerűnek tekinthető. A közép-centrális svá a hangsúlytalan alacsony nyelvtárlású magánhangzóból jön létre kismértékű redukció során, éppen ezért nem minden magánhangzó minőségből keletkezhet. A svá nem közép-centrális variánsai a magas nyelvtárlású magánhangzókól jönnek létre a redukció során. Például, a katalán nyelv gironai dialektusában hat magánhangzó minőség létezik hangsúlyos szótagban [i, e, ɛ, a, o, u], amelyekből hangsúlytalan szótagon csak három magánhangzó redukálódik [i, ə, u], ahol a svá közép-centrális minőségű magánhangzó (Herrick 2003). Az /e, ɛ, a/ magánhangzó [ə]-vá redukálódik hangsúlytalan szótagban, míg az /o, a/ [u]-nak megfelelő minőségű magánhangzóra redukálódik a katalán nyelvben.

A svárealizációk egységességének vizsgálatára, illetve hogy meghatározzuk, melyik magánhangzó minőséghez esik a legközelebb, négy HMM modellt építettünk. Három modellt készítettünk a három leggyakrabban előforduló magánhangzóra [ɔ, ɛ, o] és egy egységes modellt a semleges magánhangzóra „S”. A három magánhangzó-minőséget és a svárealizációkat 3 állapotú HMM-el modelleztük. A tanítás során [ɔ], [ɛ], [o] szimbólummal jelöltük a magánhangzókat, és „S” szimbólummal a svárealizációkat. Mind a négy modellt 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvényrel tanítottuk. A legjobb felismerési eredményt a 4 gaussos modell adta [2. a) táblázat].

2. a) táblázat: Az [ɔ], [ɛ] és [o] magánhangzók és az egységes svá („S”)

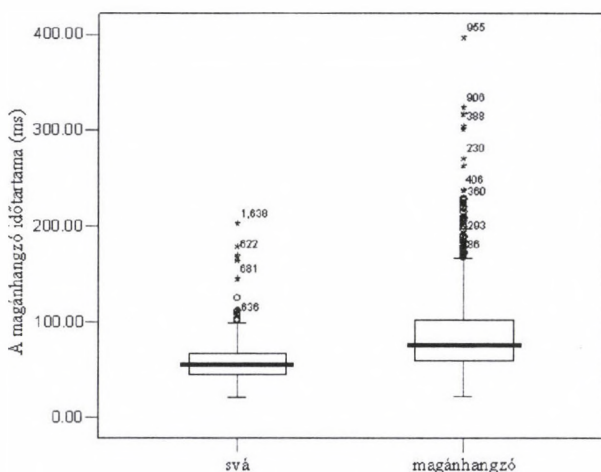
	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Töröltek száma
S	140	65,00	65,30	8
[ɔ]	167	70,65	70,95	14
[ɛ]	225	75,11	75,36	8
[o]	115	73,04	73,51	8

Az eredmények azt mutatják, hogy a svá magánhangzók helyes osztályozásának eredménye 7%-kal romlott ezzel az eljárással. A 2. b) táblázatban a négy modell tévesztési mátrixa mutatja, hogy a svá magánhangzó az [o] modellhez van a legközelebb, mivel az [o] hangok 18%-át téveszti össze a rendszer a svá magánhangzóval.

2. b) táblázat: Az [ɔ], [ɛ] és [o] magánhangzók, és az egységes svá „S” tévesztési mátrixa

Magánhangzók	[ɔ]	[ə]	[ɛ]	[o]
[ɔ]	118	9	10	16
[ə]	15	91	8	18
[ɛ]	16	13	169	19
[o]	12	7	4	84

A nagyobb tévesztési arány oka az lehet, hogy az [o] vokális artikulációs konfigurációja közel esik a semleges magánhangzóéhoz, illetve az [o] időtartama alacsonyabb, mint az [ɔ] és [ɛ] időtartama (Gósy 2004, 2006). Ennek igazolására kimértük a spontán beszédben előforduló [ɔ], [ɛ], [o] és a redukálódott magánhangzók időtartamát. A svá időtartama szignifikánsan rövidebb, mint a magánhangzóké. A svá magánhangzó időtartama átlagosan 53 ms, míg a magánhangzóké 84 ms [ANOVA: $F(1, 2917) = 252,757$; $p = 0,000^{**}$ – 8. ábra]. Ez a tendencia megegyezik a nemzetközi és hazai szakirodalomban leírtakkal (Bondarko et al. 2003, Gósy 2006, Flemming 2007, Swerts et al. 2007).

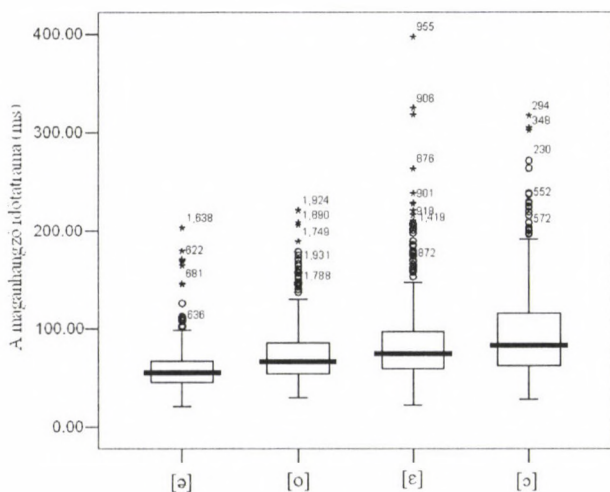


8. ábra

A redukálódott magánhangzók és az alap minőségű magánhangzók

Az adatok szerint a három magánhangzó időtartama közül az [o] magánhangzóé áll a legközelebb a svá időtartamához. Az [o] magánhangzó időtartama (77 ms) szignifikánsan rövidebb, mint az [ɔ] (83 ms) vagy az [ə] (90 ms)

időtartama [ANOVA: $F(2, 2313) = 19,86$; $p = 0,000^{**}$; csoportok közötti különbség (post hoc test) $p > 0,000^{**}$ – 9. ábra].

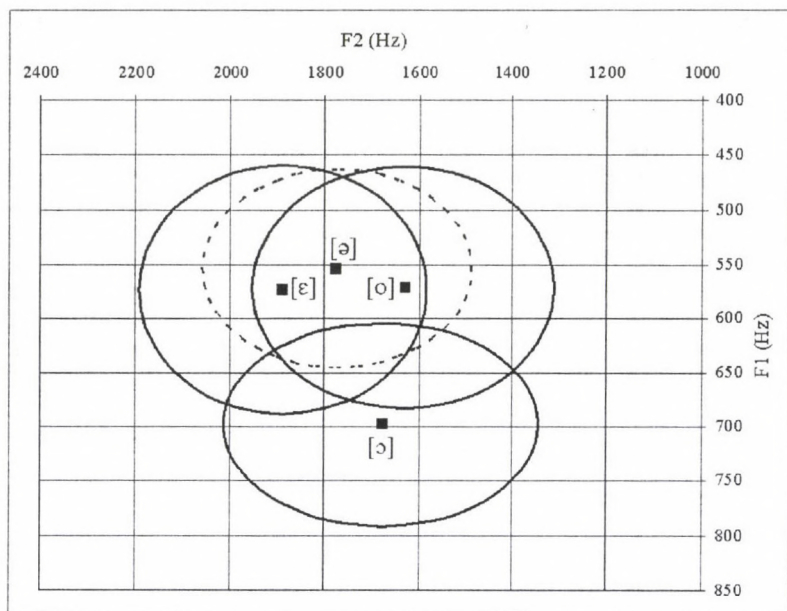


9. ábra

Az [ə], [ɛ] és [o] magánhangzók időtartama

A magánhangzók első két formánsának dimenziójában látszik (10. ábra), hogy a magánhangzók realizációi átfedésben vannak egymással és a redukálódott magánhangzókkal is.

Bondarko et al. (2003) kimutatta, hogy a magánhangzó átmeneti része minden magánhangzó esetében meghatározható mind az olvasott, mind a spontán beszédben, azonban a magánhangzó tiszta fázisa a spontán beszédben sokszor eltűnik a magánhangzók redukálódása miatt. A jelen kutatás adatai szerint az [o] magánhangzó időtartama jelentősen rövidebb, mint az [ə] és [ɛ] magánhangzóé, ami utal a magánhangzó ejtésekör bekövetkezett célalulmúlásra, ez pedig a magánhangzó tiszta fázisának redukációjához vezethet: így az [o] magánhangzó redukálódása olykor erősebb lehet. Eredményeinket alátámasztja Padget (2005) kutatása is, amelyben 9 beszélő beszédében a magánhangzók redukálódását vizsgálta. Azt találta, hogy az [ə] és az [o] magánhangzót nehezebben lehet elkülöníteni a többi magánhangzótól mind felolvasásban, mind spontán beszédben.



10. ábra

A magánhangzók [ɔ], [ɛ], [o] és a redukálódott magánhangzók F₁/F₂ értékei

3. A magánhangzók és a magánhangzófüggő svá

A helyettesítő funkcióban realizálódott svá akusztikai képe feltételezésünk szerint függ az eredetileg kiejteni kívánt magánhangzó artikulációs konfigurációjától is, amely helyett megjelenik a beszéd során. Ha a svá realizációi függenek a helyettesített magánhangzó minőségétől, akkor a svárealizációk modellezhetők a helyettesített magánhangzó minősége mentén. A svárealizációnak a következő alcsoportjai léteznek: az [ɔ] magánhangzót helyettesítő svá [A], az [ɛ] magánhangzót helyettesítő svá [E], az [o] magánhangzót helyettesítő svá [O]. A helyettesített magánhangzó minőségétől függő svárealizációkat és az eredeti minőségben realizálódott magánhangzókat ún. garbage modellel³ (G) és 3 állapotú HMM-el modelleztük. A tanítás során [ɛ], [o], [ɔ]-val jelöltük az eredeti minőségben realizálódott magánhangzókat, míg [A], [E], [O]-val a helyettesített magánhangzó minőségétől függő svárealizációkat. Mind a hat modellt 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvénnel tanítottuk. A nyelvtanban mind a hat hangmodellt egyenlő súllyal szerepeltettük (az-

3. Azokra a hangokat, amelyeket a tanítás során nem szeretnénk külön-külön betanítani, egyetlen közös, ún. garbage modell készül.

az osztályozáskor egyformán valószínűek voltak). A legjobb felismerési eredményt a 4 gaussos modell adta (3. táblázat).

3. táblázat Az [ɛ], [o], [ɔ]-és az [A], [E], [O] osztályozásának eredményei

	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Törölték száma
[ɔ]	169	65,08	69,62	11
[A]	47	68,08	72,72	3
[ɛ]	227	69,60	71,49	6
[E]	65	63,07	68,33	5
[o]	116	61,20	61,73	1
[O]	29	62,06	64,28	1

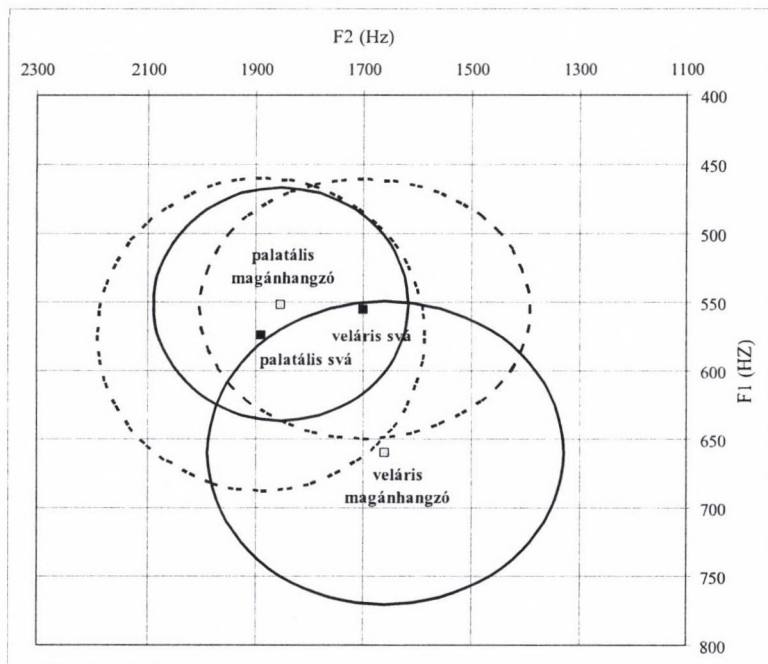
Az eredmények azt mutatják, hogy az osztályozó az [ɔ] magánhangzó helyett realizálódó [A] svát osztályozta a legjobb arányban. Az [o] magánhangzót és az [o] magánhangzó helyett realizálódott svát az algoritmus nem tudta elválasztani olyan pontosan a többi modelltől. Az osztályozás legnagyobb nehézsége ebben az esetben az [o] magánhangzó és a helyette realizálódott svá minőségének variabilitása és az időtartamának csökkenése. A véletlen találgatásnál sokszorosan jobb eredmények mindenesetre alátámasztják, hogy a svá realizációja helyettesítő funkcióban függ a helyettesített magánhangzó eredeti célkonfigurációjától.

4. Veláris-palatális magánhangzó és veláris-palatális svá

A magánhangzófüggő svá jobb osztályozhatóságának vizsgálata érdekében megpróbáltunk modelleket összevonni. Korábban megjegyeztük, hogy számos nemzetközi tanulmány foglalkozik a svárealizációk csoportosítási lehetőségével. A tanulmányok többsége a magánhangzó F_2 dimenziójának és időtartamának módosulását tartja a magánhangzó redukálódás akusztikai paraméterének, ezért a modelleket az F_2 dimenzióban vontuk össze. Ha a svá veláris magánhangzó helyett realizálódik, akkor a redukálódott magánhangzóra a veláris magánhangzó-minőség lesz jellemző a svárealizációkon belül. Ha a svá palatális magánhangzó helyett realizálódik, akkor a redukálódott magánhangzóra a palatális magánhangzó-minőség lesz jellemző a svárealizációkon belül. A jelen tanulmányban elsősorban a svárealizációk palatális és veláris alcsoportjainak elkülöníthetőségét teszteltük. Jason (2008) a svá lehetséges palatális-veláris alcsoportját monofón HMM modellel tanította és tesztelte fonetikailag variábilis, angol nyelvű, egy beszélőtől származó korpuszon. Eredményei alátámasztották, hogy a svárealizációknak létezik egy veláris és egy palatális alcsoportja. Azt is kimutatta, hogy a svá magánhangzók kezdeti fázisukban különíthetők el egymástól, míg a végső fázisukban nem.

A 11. ábrán látható, hogy a redukálódott magánhangzók realizációiban ugyanúgy létezik veláris-palatális különbség, ahogy a magánhangzók realizá-

cióiban. Az adatokból jól látható, hogy a palatális svárealizációk az F_1/F_2 térben közelebb vannak a palatális magánhangzókhoz: magasabb F_2 értékkel realizálódnak.



11. ábra

A veláris-palatális magánhangzók és veláris-palatális svák F_1/F_2 értékei

A négy magánhangzó-minőséget 3 állapotú HMM-el modelleztük. A tanítás során VV-vel jelöltük a veláris magánhangzókat, PV-vel a palatális magánhangzókat, VS-sel a veláris svákat és PS-sel a palatális svákat. Mind a négy modellt 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvénnyel tanítottuk. A legjobb felismerési eredményt a 4 gaussos modell adta [4. a) táblázat].

Az eredmények azt mutatják, hogy a svárealizációk felbonthatóak veláris és palatális svárealizációkra, amely alátámasztja a formánsok alapján leírt megállapításokat. Az osztályozásban a veláris svámodell adta a legjobb eredményt (63,15%). A palatális svák viszonylag alacsony osztályozási képessége azzal magyarázható, hogy a palatális magánhangzó realizációinak artikulációs tere jóval nagyobb, mint a veláris magánhangzóké, ezért jóval magasabb a realizációk variációinak a száma is (azaz a modell nagyobb szórást enged meg, és

emiatt relative pontatlanabb modellezést tesz csak lehetővé). Az eredményeinket alátámasztják Bunnel (2008) eredményei is: megállapította, hogy a palatális svák felismerési eredménye jobb, mint a veláris sváké. Az osztályozási feladatot a veláris és a palatális svák elkülönítésére egyszerűsítve jól látható, hogy a veláris svák felismerése sokkal biztosabb [4. b) táblázat].

4. a) táblázat: A osztályozás eredményei a VV, PV, VS és PS modellekre

	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Töröltek száma
Veláris mgh.	318	56,91	71,54	65
Palatális mgh.	375	53,86	68,70	81
Veláris svák	76	63,15	73,84	11
Palatális svák	66	40,90	54	16

4. b) táblázat: Az osztályozás eredménye a VS és PS modellekre trifónos (fonetikaikontextus-érzékeny) HMM-mel

	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Töröltek száma
Veláris svák	89	79,77	79,77	0
Palatális svák	69	66,66	66,66	0

Ez visszavezethető arra, hogy a veláris svák sokkal egységesebb kategóriát képez, ami meggyezik a nemzetközi szakirodalomban leírtakkal (Flemming–Johnson 2007). Harmegnies–Poch–Olivé (1992) kimutatták, hogy a redukálódás markánsabban jelenik meg a palatális magánhangzók esetében, mint a veláris magánhangzók esetében.

A nemzetközi eredmények és a jelen kutatás eredményei azt mutatják, hogy a svának helyettesítő funkcióban két alcsoportja különíthető el: a palatális és a veláris svák.

A modellek kiértékelése

A jelen tanulmányban használt HMM modellek közül az egységes svák és az egységes magánhangzót modellező 3 állapotú HMM-ek pontossága volt a legjobb (78%), 4 Gauss kibocsátási valószínűséget leíró függvényvel, ami azt jelenti, hogy ezekkel a modellekkel osztályozta helyesen a legtöbb hangot az algoritmus (5. táblázat). A legkevesebb helyes találatot az eredeti minőségű magánhangzót és a helyettesített magánhangzó-minőségtől függő svák modellező 3 állapotú HMM-ek adták (69,46%). A veláris és palatális svákat modellező 3 állapotú HMM-ek pontossága 74%.

A pontosságértékekből látszik, hogy a svárealizációk csoportosításai közül az egységes svámodell és a veláris – palatális svámodell adta a legjobb osztályozási arányt.

5. táblázat: A tanított modellek pontossága

A tanított modellek	Pontosság (%)
Eredeti magánhangzó-minőség és a helyettesített magánhangzó minőségtől függő svá	69,46
Veláris palatális magánhangzó és veláris palatális svá (monofon)	70,35
Veláris palatális magánhangzó és veláris palatális svá (trifon)	74,05
Egységes svá és magánhangzók [ɛ], [o], [ɔ]	75,86
Egységes magánhangzó és egységes svá	78,09

5. Hangsúlyos-hangsúlytalan szótagok osztályozása a magánhangzó minőségének segítségével

A beszédfolyam automatikus, szavaknak vagy néhány szóból álló szócsoportoknak megfelelő szintaktikai egységekre való tagolásában bizonyítottan fontos szerepe van a szupraszegmentális jegyeknek, az alaphangnak és az intenzitásnak (Szaszák 2009). A szupraszegmentális jegyek mellett a magánhangzó minősége is alkalmazható lehet, elsősorban a hangsúlyos-hangsúlytalan szótagok osztályozására, másodsorban pedig a szóhatár meghatározására is. Ha ugyanis a magánhangzó az eredeti minőségében áll, akkor a hangsúlyos szótag megjelenésének esélye növekszik, míg ha a magánhangzó redukálódott formában realizálódik, akkor a hangsúlyos szótag megjelenésének esélye csökken (vö. Swerts et al. 2007, Heuvel et al. 2003).

A hangsúlyos és hangsúlytalan szótagok osztályozásához 3 állapotú HMM-eket használtunk. A hangsúlyos szótagokat „XA”-val, a hangsúlytalan szótagokat „XT”-vel jelöltük. Mind a két modellt 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvénnel tanítottuk. A nyelvtenban mind a két hangmodellt egyenlő súllyal szerepeltettük (azaz egyenlő valószínűség mellett). A legjobb felismerési eredményt a 8 gaussos modell adta [6. a) táblázat].

6.a táblázat: A XA és a XT osztályozásának eredményei

Szótagok	Összesen	A helyes csoportosítási aránya (%)	Törlés nélkül (%)	Töröltek száma
XA	309	82,80	95,93	39
XT	855	70,72	91,87	195

A modellek átlagos pontossága a törlések figyelmen kívül hagyása mellett mintegy 93%, ami azt jelenti, hogy a hangsúlyos és hangsúlytalan szótagok 93%-át helyesen csoportosította az algoritmus. Ebből a hangsúlytalan szótagok osztályozása közel 71%-os, ami arra utal, hogy ez az osztály nem egységes a modellezett szótagok hangminősége szempontjából. Ebben a csoportban magasabb a törlések relatív gyakorisága is.

A hangsúlytalan szótag modellezésében ronthatja az osztályozási eredményeket, hogy a hangsúlytalan szótagban a magánhangzó minősége nem egyértelmű. A hangsúlytalan szótagban a magánhangzó megjelenhet redukálódott magánhangzóként, illetve az eredeti magánhangzó artikulációs konfigurációnak megfelelő minőségben is. Ennek igazolására három HMM-et építettünk. A kísérlet során modelleztük a hangsúlyos szótagban realizálódott magánhangzót (XA), a hangsúlytalan szótagban realizálódott eredeti magánhangzó minőséghez közeli magánhangzót (XT), és a hangsúlytalan szótagon megjelenő redukálódott magánhangzót (XS). Ezeket a magánhangzó-minőségeket 3 állapotú monofón HMM-ekkel modelleztük. Mind a három modellt ismét rendre 2, 4, 8, 16 Gauss kibocsátási valószínűséget leíró függvénnyel tanítottuk. Az osztályozás nyelvtanában mind a három hangmodellt egyenlő súllyal rögzítettük (azaz egyenlő valószínűség mellett). A legjobb felismerési eredményt a 8 gaussos modell adta [6. b) táblázat].

6. b) táblázat: A három magánhangzó-minőség (XA, XT, XS) osztályozási eredménye

Magánhangzó-minőségek	Összesen	A helyes osztályozás aránya (%)	Törlés nélkül (%)	Töröltek száma
XA	299	80,70	93.11	38
XT	646	68,80	79.52	86
XS	218	73,20	86.44	32

A modellek átlagos osztályozási pontossága 84%-os (törléssel), ami azt jelenti, hogy a hangsúlyos és hangsúlytalan szótagok 84%-át helyesen osztályozta az algoritmus. A hangsúlytalan szótagok helyes osztályozása (közös az XT és az XS modell) 81%-os. A hangsúlytalan modell kettéválasztása az eredeti hangminőséghez közeli magánhangzóra és redukálódott magánhangzóra közel 10% javulást eredményezett a hangsúlytalan szótagok osztályozásában. Az eredmények azt mutatják, hogy bár a hangsúlyos szótagok osztályozása 2%-kal pontatlanabb, eközben a hangsúlytalan szótagok osztályozásának eredménye 10%-kal javult. Ez azt jelenti, hogy (i) a svá függ a szótag helyétől a szóban (hangsúlyos-hangsúlytalan), illetve (ii) hogy a svámodell megkülönböztetése a felismerésben javítja az osztályozás pontosságát.

Összefoglalás

A jelen tanulmány célja az volt, hogy a helyettesítő funkcióban lévő svárealizációkat spektrális jellemzőik alapján modellezze HMM-ekkel magyar nyelvű spontán beszédben.

Az elemzések során bemutattuk, hogy (i) a [ə] és a svárealizációk MFC együtthatók alapján előfeldolgozva HMM-ekkel modellezhetőek a magyar nyelvű spontán beszédben; (ii) a svávariációk realizációi függenek az általuk

helyettesített magánhangzó artikulációs konfigurációjától. Ezen megállapítások igazolására hat különböző modellt építettünk, amelyek reprezentálták a „svá” és a lehetséges sváalcsoportok realizációit. Jóllehet az osztályozás során globális pontosság szempontjából a legjobb eredményt az osztatlan svá, az eredeti minőségben realizálódott magánhangzó modellhalmaz adta, a svá-realizációk közötti leghatékonyabb osztályozást a veláris és palatális svá alcsoportra épített HMM-ek adták (79%).

A vizsgálat során összehasonlítottuk az eredeti minőségben realizálódott magánhangzók és a redukálizálódott magánhangzók akusztikai szerkezetét (időtartam és formánsszerkezet). Az eredmények azt mutatták, hogy az [o] magánhangzó artikulációs konfigurációjában közelebb áll a svá artikulációs konfigurációjához a spontán beszédben, mint a vizsgálatban szereplő többi magánhangzó. Ennek oka az, hogy az [o] artikulációs konfigurációja és időtartama jóval nagyobb variációt mutat, mint a vizsgálatban szereplő többi magánhangzóé.

A svárealizációk lehetséges alcsoportjait HMM-ekkel modelleztük. A hipotézisünk és a nemzetközi szakirodalom szerint a svárealizációk alapvetően két csoportra bonthatók, méghozzá palatális és veláris svávariációkra. Az eredmények azt mutatták, hogy a svárealizációknak ez a két alcsoportja létezik: veláris és palatális svá. Ennek oka az, hogy a helyettesítő funkcióban lévő svá függ az általa helyettesített magánhangzó minőségétől: a veláris magánhangzó redukálizációja közben megőrzi az alapvető veláris spektrális jegeket, ahogy a palatális svá is megőrzi a palatális magánhangzó spektrális jellemzőit. Ez természetesen nem zárja ki azt, hogy a svárealizációk esetleg más dimenziókban is elválaszthatók egymástól.

A hangsúlyos és hangsúlytalan szótagokban realizálódott különböző magánhangzó-minőségeket HMM-ekkel modelleztük MFC előfeldolgozás alapján. A hangsúlyos és a hangsúlytalan szótagok osztályozásának eredménye 73,76% volt. A hangsúlyos szótagok felismerése ezen belül 82,8%-kal a leghatékonyabb volt. A svámodell beépítése összességében javított a hangsúlyos és hangsúlytalan szótagok csoportosításában. A magánhangzó-minőséggel modellezett hangsúlyos és hangsúlytalan szótagok felismerésének eredménye jobbnak bizonyult a hasonló nemzetközi kutatások eredményeihez képest (vö. Jenkin–Scordilis 1996, Kuijk–Boves 1999).

Irodalom

- Beke András 2009. A veláris magánhangzók stabilitása a spontán beszédben. In Gecső Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai*. Kodolányi János Főiskola–Tinta Kiadó, Székesfehérvár–Budapest (megjelenőben).
- Beke András – Gráczy Tekla Etelka 2009. A magánhangzók semleges realizációja a spontán beszédben. Konferencia-előadás. XI. Pszicholingvisztikai Nyári Egyetem, Balatonalmádi.

- van Bergem, Dick R. 1994. A model of coarticulatory effects on the schwa. *Speech Communication* 14. 143–62.
- Bondarko, Liya V. – Volskaya, Nina B. – Tananaiko, Svetlana O. – Vasilieva, Ludmila A. 2003. Phonetic properties of Russian spontaneous speech. In: *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, 3–9 August 2003. 2973–2976.
- Browman, Catherine P. – Goldstein, Louis 1992. Articulatory phonology: An overview. *Phonetica* 49. 155–180.
- Bunel, H. Timothy – Lilley, Jason 2008. Schwa variants in American English. In: *Proceedings of the InterSpeech 2008*. Brisbane, Australia, 1159–1162.
- Cruttenden, Alan 1997. *Intonation* (2nd ed.). Cambridge University Press, New York.
- Delattre, Pierre 1969. An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics* 7. 295–325.
- Dressler, Wolfgang U. 1984. Explaining natural phonology. *Phonological Yearbook* 1. 29–50.
- Engstrand, Olle – Nordstrand, Lennart 1983. Acoustic features correlating with tenseness, laxness and stress in Swedish: Preliminary observations. In Elert, Claes-Christian – Johansson, Irène – Strangert, Eva (eds): *Nordic Prosody III*. Acta Universitatis Umensis, Umeå Studies in the Humanities 59. 51–66. University of Umeå. [Also in: Reports from Uppsala University, Department of Linguistics (RUUL) 11: 8–22.]
- Flemming, Edward 2006. Contrast and schwa vowels in English. Paper read at the web: <http://web.mit.edu/flemming/www/paper/schwa.ppt>
- Flemming, Edward 2007. The phonetics of schwa vowels. Manuscript. MIT.
- Flemming, Edward – Johnson, Stephanie 2007. Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association* 37. 83–96.
- Furui, Sadaoki 2005a. Recent progress in corpus-based spontaneous speech recognition. In: *IEICE-Transactions on Information and Systems*, Volume E88-D, Number 3. 366–375.
- Furui, Sadaoki 2005b. Toward robust speech recognition and understanding. *The Journal of VLSI Signal Processing* 41/3. 245–254.
- Furui, Sadaoki 2007. Recent advances in automatic speech summarization. In: *Proceedings of the IEEE/ACL Workshop on Spoken Language Technology*, IEEE, Los Alamitos (2006), 115–122.
- Gay, Thomas 1978. Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America* 63. 223–230.
- Gósy Mária 2004a. Fonetika, a beszéd tudománya. Osiris Kiadó, Budapest.
- Gósy, Mária 2004b. The manifold function of schwa. *Grazer Linguistische Studien* 62. 15–26.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás* 2008. 194–208.
- Halpern, Jack 2006. The contribution of lexical resources to natural language processing of CJK languages. In: *International Symposium on Chinese Spoken Language*, ISCSLP, Singapore, 768–780.
- Harmegnies, Bernard – Poch-Olivé, Dolores 1992. A study of style-induced vowel variability: laboratory versus spontaneous speech in Spanish. *Speech Communication* 11. 429–437.

- van den Heuvel, Henk – van Kuijk, David – Boves, Lou 2003. Modeling lexical stress in continuous speech recognition for Dutch. *Speech Communication* 40. 335–350.
- Herrick, Dylan 2003. *An acoustic analysis of phonological vowel reduction in six varieties of Catalan*. PhD dissertation. University of California, Santa Cruz.
- Janse, Esther – Sennema, Anke – Slis, Anneke 2000. Fast speech timing in Dutch: the durational correlates of lexical stress and pitch accent. In: *Proceedings of the VIth International Conference on Spoken Language Processing*, Beijing, October 2000, vol. III, 251–254.
- Jenkin, Karen L. – Scordilis, Michael S. 1996. Development and comparison of three syllable stress classifiers. In: *International Symposium on Chinese Spoken Language*, ICSLP, 733–736.
- Kálmán László – Nádasdy Ádám 1994. A hangsúly. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest, 393–467.
- Koopmans-van Beinum, Florian J. 1994. What's in a schwa? Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica* 51. 68–80.
- Kopecký, Jiří – Glembek, Ondrej – Karafiat, Martin 2008. Advances in acoustic modeling for the recognition of Czech. Paper read at the International Conference on Text, Speech and Dialogue; TSD 2008.
- van Kuijk, David – Boves, Loe 1999. Acoustic characteristics of lexical stress in continuous telephone speech. *Speech Communication* 27/2. 95–111.
- Ladefoged, Peter 1993. *A course in phonetics*. (Third edition.) Harcourt Brace Jovanovich, New York.
- Lilley, Jason 2008. Data-driven investigation of subphonemic variation: “Front” schwa vs. “back” schwa. Paper read at the Cognitive Science Graduate Student Conference 2008. Delawer, Friday April 18th, 2008.
- Lindblom, Björn 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35. 1773–1781.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, William J. – Marchal, Alain (eds): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.
- Madelska, Liliana – Dressler, Wolfgang U. 1996. Postlexical stress processes and their segmental consequences illustrated in Polish and Czech. In: Hurch, Bernhard – Rhodes, Richard A. (eds.): *Natural phonology: The state of the art*. Mouton de Gruyter, Berlin–New York, 189–200.
- Masanobu, Nakamura – Furui, Sadaoki – Koji, Iwano 2006. Acoustic and linguistic characterization of spontaneous speech. In: *Speech Recognition and Intrinsic Variation Workshop*, SRIV. Toulouse, France, May 20. 3–8.
- Onishi, Masao (ed.) 1981. *A grand dictionary of phonetics*. The Phonetic Society of Japan, Tokio.
- Padgett, Jaye – Tabain, Marija 2005. Adaptive dispersion Theory and phonological vowel reduction in Russian. *Phonetica* 62. 14–54.
- Patterson, David – LoCasto, Paul C. – Connine, Cynthia M. 2003. Corpora analyses of frequency of schwa deletion in conversational American English. *Phonetica* 60. 45–69.
- Pickett, James M. 1999. *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*. Allyn & Bacon, Needham Heights MA.

- Siptár, Péter – Törkenczy, Miklós 2000. *The phonology of Hungarian*. Oxford University Press, Oxford.
- Slifka, Janet 2005. Acoustic cues to vowel-schwa sequences for high front vowels. *Journal of the Acoustical Society of America* 118. 2037.
- Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3–45.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. The MIT Press, Cambridge, MA.
- Swerts, Marc – Kloots, Hanne – Gillis, Steven – de Schutter, Georges 2007. Vowel reduction in spontaneous spoken Dutch. In: *Proceedings of the ISCA & IEEE Workshop on Spontaneous Speech Processing and Recognition*, Tokyo, 31–34.
- Szaszák György 2009. *A szuprasegmentális jellemzők szerepe és felhasználása a gépi beszédfelismerésben*. PhD-értekezés. BME, Budapest.
- Wright, Helen – Taylor, Paul A. 1997. Modelling intonational structure using hidden Markov models. Paper read at the ESCA workshop on Intonation: Theory Models and Applications, Athens, Greece.
- Xie, Huayang – Andrae, Peter – Zhang, Mengjie – Warren, Paul 2004. Detecting stress in spoken English using decision trees and support vector machines. *ACSW Frontiers 2004*. 145–150.
- Young, Steve – Evermann, Gunnar – Gales, Mark – Hain, Thomas – Kershaw, Dan – Moore, Gareth – Odell, Julian – Ollason, Dave – Povey, Dan – Valtchev, Valtcho – Woodland, Phil 2005. *The HTK Book (for HTK Version 3.3)*. Cambridge University Engineering Department, Cambridge.

A BESZÉD TERMÉSZETESSÉGÉRŐL ALKALMAZOTT FONETIKAI SZEMPONTBÓL

Gósy Mária – Gyarmathy Dorottya – Horváth Viktória

Bevezetés

A beszédet az emberi kommunikáció legtermészetesebb formájaként tartjuk számon. Az embereknek általában fel sem tűnik, hogy milyen könnyen beszélnek – szemben például az olvasással vagy az írással. Ugyanakkor a szokatlan beszédhelyzet, illetve a váratlan beszédtema sokszor problémát okozhat; a beszélő nehezebben fejezi ki magát, elégedetlen azzal, amit mondott, avagy ahogyan azt megfogalmazta; a saját spontán beszédét nem a „megszokottnak” ítéli. A felnőttek egy részének a beszéde ilyenkor – az adott kommunikációs helyzetből függően – szorongóvá, feszélyezettebbé válik. Mások a beszédhelyzettől függetlenül, minden felkészülés nélkül, akár váratlan helyzetekben és témák esetén is képesek spontán módon jól beszélni, azaz jól követhető kiejtésben, megfelelő tagolással, egyértelműen elmondani azt, amit akarnak. A spontán beszédet sokan összekeverik a természetes beszéddel. Ez többféle okra vezethető vissza, például a definíciók sokfélesége, a megközelítések különbözőségeire, az elemzési célok eltéréseire. A jelen tanulmányban egyrészt annak a kérdésnek az elméleti körüljárását tűztük ki célul, hogy a beszéd spontaneitása és természetessége milyen összefüggéseket mutat, másrészt ezekkel a tényezőkkel kapcsolatos saját gyakorlati tapasztalatainkról számolunk be.

A spontán beszédről

A beszédtudomány szakirodalma azt mutatja, hogy a különféle célú kutatók sok hosszú időn keresztül főként felolvasott vagy előre betanult szövegek vizsgálatán (szósorok, színészi, előadóművészi interpretációk, klisémondatok stb. alkalmazásán) alapultak. A spontán beszéd elemzésének igénye viszonylag új keletű a külföldi és a hazai kutatásokban egyaránt. Mikor spontán a beszéd? Úgy fogalmazhatunk, hogy a beszélő spontán módon beszél akkor, ha nem tudja (nem akarja, vagy nem áll módjában) közlendőjét részletesen átgondolni, a meghangosítást tartalmi és formai szempontból pontosan megtervezni (szemantikai relációk, grammatikai szerkezetek, szövegszerűség, lexikális válogatás, szegmentális és szupraszegmentális artikulációs kontroll stb.). A Levelt-féle beszédproduktions modell (1989) értelmében ez azt jelenti, hogy a beszédszándékot nem követi semmiféle felkészülés; a makro- és a mikrotervezés folyamatai előzetes tervezés nélkül működnek az adott beszéd-

helyzetben. A spontán beszéd fogalma egyértelműen definiálható ugyan, a szakirodalom mégis elsősorban azt emeli ki vele kapcsolatban, hogy megakadásjelenségek tarkítják, valamint hogy a beszélők gyakran nem jól formált mondatokat használnak benne (pl. Ward 1989, Lustgarten et al. 2002).

A beszéd spontaneitása nem függ attól, hogy több szereplő közötti társalgásban történik-e a verbális kommunikáció, vagy valaki egy történetet (narratívát) mond-e el, de attól sem, hogy milyen a beszélő aktuális lelkiállapota (különféle spontán beszéd zajlik például az orvossal történő beszélgetés során, munkahelyi megbeszélésen, értekezlet alkalmával, családi, baráti életménybeszámolókor, vita vagy veszekedés esetén). A spontaneitásnak elméletileg akár fokozatai is lehetnének (mint például „nagyon spontán”, „alig spontán”, „részben spontán”, „túlnyomórészt spontán”...). Ennek a feltételezett skálának a fokozatai azonban nehezen (talán egyáltalán nem) meghatározhatók, még nehezebben általánosíthatók, és az alkalmazott fonetika területein szükségtelenek.

A beszéd természetességéről

Wolfson szerint a *természetes beszéd* nem más, mint megfelelő beszéd, azaz szemben áll a beszélő szándékának nem teljesen megfelelő beszéddel (1976). Úgy is fogalmazhatunk, hogy a beszéd akkor tekinthető természetesnek, ha a beszélőnek nem okoz feltűnő, észlelhető nehézséget a folyamat működtetése (például a dadogó beszédet vagy a rekedt hangú beszédet ezért állítják szembe a természetes beszéddel). (A nem patológiás beszéd esetén azonban probléma, hogy minek alapján állítjuk, hogy egy adott beszélőnek nincs különösebb nehézsége a verbális kommunikációban, vagy hogy a megnyilatkozás nem igényelt tőle többletenergiát.)

Az emberi agy kiválóan alkalmazkodott a verbális kommunikáció mellett számos új kommunikációs forma megtanulásához, újabb kommunikációs stratégiák kialakításához (ilyen például az olvasás: Csépe 2006). Képessé váltunk ezek „természetes” alkalmazására. Mit értünk tehát a természetességen a beszéd vonatkozásában? Nusbaum és munkatársai úgy fogalmazzak, hogy a természetesség a beszéd multidimenzionális, szubjektív minősége (1995). Ennek a gondolatnak az alapján elméletileg azt mondhatjuk, hogy a beszéd akkor természetes, ha megfelel a beszélő személyiségének, egyéniségének, szokásos beszédének. Nem tudunk azonban egzakt fonetikai paramétereket kapcsolni a természetes beszéd egyes jellemzőihez. Az sem segítene sokat, ha definiálnánk ilyen „kapcsolatokat”, hiszen nem volna etalon, amelyhez viszonyíthatnánk őket egy ismeretlen beszélő esetében. Egy adott történetet például különbözőképpen (más-más stílusban, szóhasználat, hangképzéssel, prozódiaival stb.) mesélünk el a szüleinknek, idősebb vagy fiatalabb idegennek, rokonunknak, a „haverunknak”, a közeli barátunknak. Ennek ellenére senki sem mondaná azt, hogy a felsorolt helyzetek bármelyike ne feljelne meg a természetesség kritériumainak.

Hogyan közelíthető meg tehát a természetes beszéd fogalma? A beszédprodukciónak startégiák alkalmazását igénylő folyamat (Levelt 1989). A beszédstratégia voltaképpen egy általános séma, amely tartalmazza a beszélő szándékát, a beszédtervet és azt a módot, ahogyan a közlendőjét megformálja az adott hallgató(k)nak. Feltételezhetjük, hogy a természetességnek több tényezője van. Ilyen lehet a beszélő személyisége: van, akit jobban zavar egy adott kommunikációs helyzet, van, akit kevésbé vagy egyáltalán nem. A „közös” beszédtema, a kapcsolódó területek ismerete, és az átadás módja is hatással lehetnek a beszéd természetességre. Valószínűleg nem természetes például az a beszélgetés, amelynek során a beszélő folyamatos meghatározásokat, részletes elemzéseket kell, hogy adjon a hallgatónak a megértés érdekében. A beszélő – többnyire tudat alatt – alkalmazkodik a beszédpartnerhez, ez befolyásolhatja beszédének megformálását (vö. Bata-Grácsi 2009), s ez összefügghet a természetességgel is.

A nem feszélyezett (nem szorongó) viselkedés nyilvánvalóan kapcsolatos a természetes beszéddel; a szorongás mértéke és a beszéd természetessége közötti összefüggés azonban számos további kérdést vet fel (hiszen a mindennapi verbális kommunikáció során is adódhatnak olyan körülmények, amelyek között a beszélő nincs könnyű lélektani helyzetben). A Grice (1975) szerinti kooperációs elv szabályozza a társalgást, és valószínűleg biztosítja vagy lehetővé teszi a beszéd természetességét. Ez az elv négy maximából áll: a mennyiségi, a minőségi, a relevancia-, valamint a módmaximából. Grice nem tartotta rendszerét lezártnak, ő maga is kiegészítette ezeket további elvekkel, ilyenek például az esztétikai, a társadalmi vagy a morális maximák. Leech (1983) megalkotta az udvariassági elvet, amely újabb maximákat tartalmaz (mint például tapintatmaxima, nagylelkűség-maxima, jóváhagyás-maxima, szerénységmaxima, megegyezésmaxima, rokonszenvedésmaxima).

A beszéd természetessége az alkalmazott fonetikai területeken

A beszéd természetességének az előadó-művészi, illetve a retorikai gyakorlatban nagy jelentősége van, mivel a beszélő (itt: a színész, a szónok) verbális viselkedése alapozza meg a megformált személy hitelességét (ehhez egyéb viselkedésformák is társulnak, pl. mimika, gesztusok). A színész azonban előre megtanult szöveget hangosít meg, a szónok pedig rendszerint felkészül a mondandójára, ezért a beszéd természetességét másként ítélik meg náluk, mint az átlagos, ún. naiv beszélő esetében. A beszéd természetessége az alkalmazott fonetikában az utóbbi évtizedekben elsősorban a következő kutatások kapcsán merült fel: *a)* a beszédszintézis (azaz a beszéd mesterséges előállítás), *b)* a mesterséges beszédfelismerés, *c)* a törvényszéki fonetika (a beszélő személy felismerésének kérdése) és *d)* a beszédkorpuszok fejlesztése.

a) A természetesség a beszédszintézisben. Az első szintetizáló rendszerek (nyelvtől függetlenül) erősen gépies hangzásúak voltak; a kezdeti technológia (főként a korlátozott tárhelykapacitás) nem tette lehetővé, hogy a fejlesztők a

szegmentális szerkezet meghangosításán túl egyéb paraméterekkel is hatásosan foglalkozhassanak (Olaszy 1989). A formánsalapú szintetizátorok beszéde jól érthető, hangzásuk azonban erősen gépies (Tóth–Németh 2008). A technológia fejlődésével a hangzás egyértelműen javult, a hangzás gépiessége csökkent (van Santen et al. 1997, Olaszy et al. 2000). A beszédszintetizátorokkal előállított hangot az érthetőség, illetve a természetesség kívánalmait mentén szokták minősíteni. A beszédszintetizálási módszerek változása, valamint a tárkapacitás növekedése mindenekelőtt a prozódiai jellemzők beépítését tette lehetővé, és ezáltal a mesterséges beszéd már jobban megközelítette az emberit (Fék et al. 2005, Németh et al. 2006). Noha ez minőségi változást idézett elő a beszédszintézis hangzásában, a „természetesség” problémáját nem oldotta meg. Az újabb kutatások az érzelmi komponens beépítésével, illetőleg annak prozódiai leképezésével igyekeznek a beszéd gépies hangzását tovább csökkenteni, mind jobban az emberi beszéd hangzását szimulálni (Németh et al. 2006, Zainkó–Fék 2006).

A természetesség megítélését az észlelésben – számos egyéb tényező mellett – a zöngéképzés vagy a szótag prozódiai szerkezete befolyásolta a kísérleti adatok szerint (Nusbaum et al. 1995). Úgy tűnik, viszonylag egyszerű azt megmondani, hogy a mesterségesen előállított beszéd mennyire vagy milyen tekintetben természetes, hiszen az adatközlők a humán beszédprodukciónhoz viszonyíthatnak. Az ember-gép kommunikáció fokozódó igényei következtében napjainkban még jobban előtérbe került ez a kérdés. A legjobban érthető és a legtermészetesebbnek ítélt szintetizált beszédet korpuszalapú elemkiválasztásos módszerrel állítják elő (Tóth–Németh 2008).

b) A természetesség a mesterséges beszédfelismerésben. A beszédfelismerők jó hatásfokkal működnek izolált szavak, felolvasott szövegek, illetve egy adott személy beszédére betanított rendszerek esetében (pl. Becchetti–Ricotti 1999, Mihajlik et al. 2006, Vicsi et al. 2007, Tüske et al. 2007). A spontán beszéd létrehozásának következtében létrejövő sajátosságok (mint pl. a megakadásjelenségek) azonban problémát jelentenek a fejlesztőknek (Schaaf–Kemp 1997, Seneff 2004). A folyamatos spontán közlések mesterséges felismerése ugyanis rugalmas akusztikai modellálást, mondathatár-detektálást, rugalmas kiejtésmodellálást, megfelelő akusztikai és nyelvi modell alkalmazását és mindezek automatikus összegzését teszi szükségessé (vö. Furui 2005 áttekintését).

c) A természetesség a beszélő személy felismerésében. Az emberi hang magában rejtje a beszélő személy egyedi jellemzőit is, ennek kutatásával az utóbbi évtizedekben kezdtek behatóan foglalkozni elsősorban a törvényszéki fonetika területén (Braun–Köster 1995, Dankovičová–Nolan 1999, Butcher 2002, Nikléczy 1996). A kiinduló kérdés az volt, hogy a beszéd szegmentális vagy szupraszegmentális szintjén keresendők-e az egyéni beszédre utaló összetevők, avagy mindkettő tartalmaz ilyeneket. Ha ismerjük a beszélőt, akkor rendszerint meg tudjuk mondani, hogy természetesen beszél-e vagy

nem; nem biztos azonban, hogy tényeket is fel tudunk sorolni véleményünk alátámasztására. Egy nem ismert (idegen) személy esetében ráadásul nincs is mihez viszonyítanunk. Valamennyi segítséget jelent ebben az ún. alkati harmónia, amely a beszélő anatómiai felépítésének megfelelő hangadást jelent (Gósy 1999).

A beszélőre jellemző sajátosságok részben anatómiailag meghatározottak, részben a beszélőnek az adott helyzetben létrehozott artikulációs mozgásaival kapcsolatosak. Ez azt jelenti, hogy a létrehozott beszéd artikulációs, illetve akusztikai fonetikai jellemzői egyfelől szervi eredetű, másfelől tanult különbözőségekre vezethetők vissza. A szakirodalom prototípusmodellje egy prototipikus hangzást tételez fel a hallgató memóriájában, amelyhez a hallgató minden új beszélő esetében viszonyít (Lavner et al. 1999). A természetesség kérdése azonban ez a teória nem ad választ. Nem lehet ugyanis tapasztalunk minden ember beszédéről ún. természetes és nem természetes környezetben. A kutatási eredmények szerint az agy bizonyos kérgi területe a különböző emberi zönghangok megkülönböztetéséért felelős (Belin et al. 2000). A kompenzációs adaptációs elmélet (Kock 2007) elsősorban a kommunikációs helyzet (és nem a beszéd) természetességével foglalkozik. Azt állítja, hogy a kommunikáció természetességének fokozatai attól függenek, hogy egy adott médiumban beszélők milyen mértékben alkalmazzák a beszédükben azokat az elemeket, amelyeket a szemtől álló beszélgetőpartnerek szoktak alkalmazni. Ezeknek az elemeknek a megnevezésével azonban adós marad.

d) A természetesség a korpuszokban. Beszédkorpuszok fejlesztésekor felmerül, hogy a rögzített anyagok mennyire tekinthetők természetesnek, vagyis olyanoknak, amilyenek a mindennapi életben előfordulnak. Minthogy a beszéd-felvételek valamilyen mértékben mindig mesterséges környezetben zajlanak, az egyik legfontosabb kérdés, hogy a beszélő mennyire képes ekkor természetesen, vagyis a megszokott módon beszélni. Vajon milyen mértékben kontrollálja a közléseit, mennyire képes függetlenedni a mikrofontól, illetve a számára többé-kevésbé ismeretlen beszélgetőpartner(ek)től? Itt említjük meg Labovnak a szociolingvisztikai interjúk rögzítése kapcsán bevezetett „megfigyelői paradoxon” fogalmát (1979).

A beszédkorpuszok fejlesztésének általános célja annak feltérképezése, hogyan beszélnek az emberek a mindennapokban. A kérdés az, hogy a korpuszok rögzített anyagai valóban a mindennapok verbális kommunikációját tükrözik-e. Ennek megítélésében sokszor keveredni látszik a spontán beszéd és a természetes beszéd fogalma. Először a természetesség kérdését közelítjük meg. Ha az adatközlők tudják, hogy éppen a beszédüket figyelik, akkor nyilvánvalóan jobban ügyelnek annak tartalmi és formai sajátosságaira. Ez pedig gátolhatja a beszélőt abban, hogy úgy beszéljen, mint a mindennapokban. Ha a mindennapok verbális kommunikációja természetes, akkor vajon ebben az értelemben természetes-e a beszédkorpuszok rögzített anyaga? „Minél jobban

biztosítjuk [...] az adat-, illetőleg a szövegrögzítés technikai föltételeit [...] a kísérleti effektus miatt annál inkább eltávolodnak az adatközlők a megszokott nyelvhasználatuktól, a kutatók pedig egyre távolabb kerülnek eredeti céljuk elérésétől” – írja Kiss Jenő (1995: 37). A megfigyelői paradoxon áthidalásának egyik lehetséges módja az, ha az adatközlőtől többféle típusú beszédet, többféle módon gyűjtünk (Kontra 1990, Wardhaugh 1995). Labov megoldásnak tekinti továbbá azt, hogy csoportinterjúkat készítsünk, mivel ezt tekinti a kommunikáció legtermészetesebb módjának. Véleménye szerint a megfigyelési hatás itt érvényesül a legkevésbé. Az interjút készítő személynek alárendeltként kell „vezetnie” a beszélgetést, és éreztetnie kell, hogy az ő tekintélye kisebb, mint az adatközlőké (Labov 1984).

A beszédkorpuszok felvételi módszereinek sokfélesége tehát azt a célt is szolgálja, hogy minél „természetesebb” helyzetbe kerüljön az adatközlő. A játékok, társasjátékok alkalmazása (pl. Szabó 2004) azért jó, mert valamennyire eltereli a figyelmet a beszédfelvétel tényéről. A feladat jellegű beszéd-rögzítések (pl. a térképmódszer: Horváth 2004) előnye, hogy az adatközlők hosszabb-rövidebb idő elteltével a feladatra és nem a saját verbális viselkedésükre koncentrálnak. A dialógusok, a narratívák, a társalgások, a viták, az értekezletek vagy a hallott történet, illetve látott film reprodukálásának vagy megbeszélésének természetessége sok tényezőtől függ [pl. az adatközlő(k) és az adott téma kapcsolata, mennyire zavarja a beszélőt a mesterséges helyzet, mennyire képes feloldódni, milyenek a beszédpartnerek stb.]. Nemegyszer a felvételi technika rovására rögzítenek beszédanyagokat éppen a természetesség mind jobb megközelítése érdekében (pl. otthonokban, mindennapi helyszíneken, utcán, iskolában). Kérdés azonban, hogy az ismert vagy szokásos helyszín, avagy a nem személyes kapcsolat (telefon) valóban növeli-e a beszéd természetességét (pl. Hutchinson–Pereira 2001, Keszler 1983, repülőjegy-rendeléssel kapcsolatos beszélgetések).

A beszédkorpuszok fejlesztése kapcsán felmerülő másik kérdés a spontaneitás. (A jelen tanulmányban nem térünk ki az olvasásalapú beszédkorpuszokra.) Láttuk, hogy különféle módjai vannak a nem olvasott beszédanyag-felvételeknek. Az ilyen módon rögzített anyag spontán beszéd, hiszen a beszélőnek nincs módja az adott témára felkészülni. A gondolatait abban a helyzetben, azonnal kell verbalizálnia. A mesterséges szituáció, illetve a beszéd-rögzítés ténye hatással lehet (általában negatívan) a beszéd folyamatosságára (pl. hezitálások arányának növekedése) vagy a megszokott légzéstechnikára. (Ezek a mindennapi kommunikációban tapasztalható jelenségek, mint amilyen a hezitálás is, úgy tűnik, egyszersmind a természetesség benyomását növelik, vö. Csapó et al. 2009.) A beszéd-rögzítéses helyzet befolyásolhatja továbbá a lexikai válogatást (ilyen helyzetben nyilvánvalóan jobban odafigyel az adatközlő arra, hogy milyen szavakat mond ki), valamint valamilyen mértékben a kiejtést is (vö. Lindblom 1990: hiperejtés és hipoejtés). Szende szerint (1973) egy ilyen helyzetben a beszélők törekszenek a nyelvi norma

megközelítésére, sokkal inkább, mint a mindennapi megnyilatkozásokban. Mindez rendszerint jogos feltételezés, de nehéz a megállapítást általánosítani és objektíven igazolni.

Spontaneitás és természetesség a BEA tapasztalatai alapján

Láttuk, hogy a spontán beszédet tartalmazó korpuszok esetében merül fel újra és újra a beszéd természetességének kérdése, amely azonban (mint említettük) gyakran keveredik a spontán beszéd fogalmával. A természetességet nemegyszer a spontaneitással rokon értelmű fogalomként kezelik, így összekeveredik a beszédtervezési folyamat jellege és a verbális nyelvhasználat minősége.

A BEA (BESzélt nyelvi Adatbázis, vö. Gósy 2008) célja különféle spontán beszédek rögzítése stúdiókörülmények között, amely alkalmassá teszi a beszédmintákat akusztikai fonetikai elemzésekre is. Egy felvétel közel 60 percig tart, különböző részekből áll, az adatközlőnek van ideje „hozzászokni” a felvételi körülményekhez, az interjúkészítőhöz, így beszéde és viselkedése egy bizonyos idő elteltével (közel) természetessé válik. A felvételi protokoll kialakításánál törekedtünk arra, hogy az egyes részek egymásra épülésének sorrendje már önmagában is segítse az adatközlőt az esetleges kezdeti szorongásának leküzdésében.

A **spontaneitás** nem sérül egy mesterségesen létrehozott kommunikációs helyzetben, például egy interjú során, hiszen a beszélő az adott pillanatban válogat a gondolatok között, és rendeli hozzájuk a nyelvi formát. Az adatközlő ekkor sokkal inkább a közlés tartalmára fordít figyelmet, kevésbé koncentrálna a nyelvi átalakításra és legkevésbé az artikulációra. A rögzített beszéd tehát egyértelműen spontán akkor, ha a beszélőnek a feltett kérdés(ek)re felkészülés nélkül azonnal reagálnia kell, illetve a társalgásban minden előzetes megbeszélés és tervezés nélkül vesz részt (a társalgás során felmerülő témák váratlanok a számára). A beszéd spontaneitásának mértéke természetesen változhat az adott téma vagy az adott feladat függvényében, de ez adatközlő-specifikus.

A BEA spontánbeszéd-anyagai is értelemszerűen különböző fokozatúak lehetnek egy „spontaneitási skálán”. A beszédprodukció feltehetően a legnagyobb mértékben akkor spontán, amikor az interjút készítő egy váratlan problémáról kérdezi az adatközlő véleményét. Némelykor az adatközlő erre reagál is, és ez a beszédének egyre természetesebb formájához vezet. Például: *jaa igen ez igaz na látod milyen praktikus kérdés tényleg hát nem tudom...* Valamivel kevésbé lehet spontán a narratíva típusú beszéd, amikor az adatközlő a családjáról, a munkájáról, a hobbijáról beszél, hiszen ilyenkor óhatatlanul aktiválja a témával kapcsolatos tartalmi és szerkezeti kliséket. A beszélő ilyenkor is „válogat”, átgondolja, hogy mit mondjon, mit hagyjon ki, miről beszéljen és mennyit. (Nincs tudományos kutatási eredmény arra vonatkozó-

an, hogy a beszédtervezési folyamatok milyen különbségeket mutatnak a mindennapi kommunikációban és egy beszédrögzítéses helyzetben.)

A BEA-felvétel egyik szakaszában két elhangzó szöveget kell az adatközlőnek saját szavaival összefoglalva elmondania. Ezek a narratívák egyértelműen spontánok ugyan, de a mondandó tartalma adott (hiszen azt a beszélő megelőzően hallotta), így a beszédtervezési folyamatok valamilyen mértékben másként működnek, mint az előzőekben említett narratívák vagy a társalgások során. Nyilvánvalóan spontán közléseket hallhatunk a társalgásokban, hiszen az adott témára nincs mód előzetesen felkészülni, ugyanakkor a társalgás sajátosságai mégis több időt hagynak az adatközlőnek a beszédtervezésre (pl. Markó 2007), mint a narratívák esetében. További kutatás adhat számot arról, hogy mindennek vannak-e nyelvi és/vagy kiejtési következményei, és ha igen, milyenek.

A spontánbeszéd-adatbázis interjúinak készítésekor a beszédszándék nem a beszélőtől indul ki, hanem mesterségesen hozzák létre oly módon, hogy felkérjük a személyt a részvételre. Megjegyezzük, hogy az adatközlők nemegyszer átveszik az irányítást, ami növeli a beszéd spontaneitását, s egyúttal feltételezhetően a természetességét is. Például az egyik adatközlő, amikor a felvételvezető elővezette a beszédtemát, a következőképpen reagált: *na jó akkor kezdjék maguk*. Egy másik példa: *akkor viszont győri megkérdezem magától hogy a győri történetek közül tudja-e azt hogy a Széchenyi téren a bencés templommal szemben van egy Mária-szobor és ez a Mária-szobor arról nevezetes, hogy nagypénteken délben amikor harangoznak a Mária meghajtja a fejét miért*.

A szorongás hatással van a spontán beszéd folyamatosságára és a beszéd természetességére, de nem a spontán voltára. A mikrofon zavaró tényező (Szende 1973), hiszen azt jelzi, hogy az elhangzott közlés visszavonhatatlan, korrigálhatatlan, megőrződik. Ez (hosszú távon) nem jellemző a természetes beszédhelyzetekre, ezért erre az adatközlők nincsenek felkészülve, nincsenek stratégiáik az ebből fakadó zavar ellensúlyozására. A legtöbb emberben szorongást kelt, hogy felveszik és később különböző szempontokból elemzik majd a beszédét. Úgy érzik, hogy bizonyos elvárásoknak kell megfelelniük, bár a felvételeket megelőzően a kísérletvezető informálja az adatközlőt (biztosítja, hogy anonim marad a felvétel stb.), és ez általában csökkenti is a beszélő szorongását. A beszédrögzítés tényéből adódó zavar azonban nem egyformán érint mindenkit; akadnak, akikre ez nagyobb mértékben hat, míg mások (láthatólag) egyáltalán nem foglalkoznak vele, látszólag nem frusztrálja őket az adott helyzet. A mikrofon okozta esetleges beszédmódosulások az ő esetükben nehezen vagy egyáltalán nem megítélhetők.

A BEA-felvétel előtt a beszélők egy része bevallja, hogy nagyon izgul (sokszor a testtartáson, a mozdulatokon is látható a feszélyezettség). Ezek a külsődleges jelek viszonylag gyorsan megszűnnek, az adatközlők hamar feloldódnak. Ezt utólag (jóleső érzéssel) el is mondják. Nemegyszer kapunk

olyan visszajelzést, hogy mennyire jó volt beszélgetni, a beszélőnek – saját maga számára is – meglepően gyorsan sikerült feloldódnia, belehelyezkednie a beszélgetésbe, a témába; a kérdések, a társalgás, a beszédpartnerek és az egész légkör olyan volt, hogy nagyon gyorsan elterelte a figyelmét a hangrögzítés és az idegen helyzet tényéről. Egy beszédfelvétel hasonlóan bizalmi szituáció, mint például egy orvossal vagy egy pszichológussal történő beszélgetés. A közös munka csak akkor hatékony, ha a kommunikációs partnerek rövid idő alatt egymásra tudnak hangolódni, és kialakul közöttük ez a bizalmi viszony. Ez esetünkben azt is jelenti, hogy az adatközlő nem mérlegeli minden hangos megnyilatkozását, hanem a mindennapokhoz hasonlóan beszél. Ez pedig a beszéd természetességét növeli.

Mindebben a felvételvezetőnek óriási szerepe van. Egyfajta intuitív, empátikus készséggel kell rendelkeznie ahhoz, hogy igen rövid idő alatt elnyerje a beszélő bizalmát, és mintegy „régis ismerősként” beszélgessen vele. Már a beszélgetés kezdetén fel kell tudnia mérni a beszélgetőpartner „szorongási szintjét”, és ennek megfelelően igyekeznie kell oldani azt. Ennek érdekében lehetősége van például felcserélni az egyes részek sorrendjét. Többször előfordult, hogy az ismeretlenként odaérkező adatközlő a felvétel után megkérdezte, hogy vajon voltak-e róla esetleg előzetes információk, mert a felvételvezető olyan témákat hozott fel, amik őt nagyon foglalkoztatják mostanában. Egy másik esetben a beszélgetés során az adatközlő azt mondta, hogy bár tudja, hogy ezt a felvételt kutatók meg fogják hallgatni és elemzik, de akkor is elmondja, hogy... És még egy példa. Megtörtént, hogy a felvetett kérdésre az adatközlőnek az eredeti véleménye a beszélgetés végére az ellenkezőjére változott. Ez pedig csak úgy lehetséges, hogy közben nem a felvétel tényére figyelt, hanem teljesen lekötötte az adott téma, a véleménykülönbség, illetve a saját hozzáállásának a módosítása. Ezek a „vallomások”, vélemények, megjegyzések és tények megerősítik azt, hogy a felvételek legalábbis közel természetes beszédet tartalmaznak. Minden felvétel előtt van továbbá egy úgynevezett „próbafelvétel”, ami egyfajta „bemelegítésként” szolgál az adatközlő számára, ezalatt van ideje hozzászokni a mikrofon jelenlétéhez.

A témaválasztás a mindenkor adatközlő valószínűsíthető személyes érdeklődéséhez igazodik. A felvételvezető legtöbbször közfelháborodást kiváltó, avagy a közvéleményt foglalkoztató aktuális (nem politikai jellegű) eseményekről kérdezi az interjúalany véleményét; figyelembe véve a labovi elvet (1984), végig „alárendelt” szerepben maradva. A témafelvetéskor is tekintettel vagyunk az adatközlőre; ha a választott téma nem tetszik neki, láthatólag nem akar beszélni róla, akkor az interjúkészítő azonnal vált. A témák a beszélőből gyakran érzelmeket is kiváltanak, ami nagyban hozzájárul ahhoz, hogy megfedekezzenek a felvétel rögzítésének tényéről. Nemegyszer előfordult például, hogy a véleménykifejtés hevében trágár kifejezés „csúszott ki” az adatközlő száján, és csak utána döbrent rá, hogy ez egy rögzített felvételtör történt: *és akkor mondtam hogy hát mondom akkor most elmennek a bűdös*

*pi**ba* így széttéptem a bocsánat a felvétel jaj jaj jaj jaj kiszaladt na mindegy. Előfordul, hogy a beszélő a vélemény kifejtés folyamán elakad, megszakad a gondolatmenete, illetőleg a felvételvezető úgy ítéli meg, hogy túl szűkszavúan nyilatkozott az adott témában. Ilyenkor – az adatközlő személyiségétől függően – szándékosan ellenkező nézőpontú kérdésekkel és véleményformálással igyekszik továbblendíteni a társalgást. Mindez megint csak eltereli a figyelmet a mikrofon jelenlétéről.

Amikor azt fejtegetjük, hogy egy spontánbeszéd-felvétel során mennyire beszél természetesen az adatközlő, voltaképpen arra keressük a választ, hogy mennyire hasonlít a beszéde ahhoz, mint amikor informális helyzetben, a mindennapokban, például családi körben beszél. Ha lenne ilyen informális helyzetben gyűjtött beszédanyagunk, akkor lehetne válaszolni erre a kérdésre. Ilyen beszédanyag hiányában csupán a felvételvezetői tapasztalatra támaszkodhatunk.

Összegző megállapítások

A spontán beszéd mibenléte egyértelműen meghatározható, szemben a beszéd természetességének fogalmával, ez utóbbi – mint láttuk – nehezen definiálható, sok bizonytalanságot rejt magában. A kettő nemegyszer keveredik egymással, ami a beszéd kutatásban számos félreértésre adhat okot. A felkészülés nélküli beszéd mindig spontán, jól jellemezhető sajátosságokkal. A beszéd természetességének megítéléséhez számos tényező figyelembevétele szükséges, és jelenleg nem rendelkezünk ehhez objektív kritériumrendszerrel. Az emberek eltérően beszélhetnek mind a spontaneitás, mind a természetesség tekintetében, noha ezen eltérések jellemzése valószínűsíthetően bonyolult feladat. Az ilyen irányú kutatások megalapozásához igyekeztünk a jelen tanulmánnyal hozzájárulni.

Irodalom

- Bata Sarolta – Gráczy 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.
- Becchetti, Claudio – Ricotti, Lucio Prina 1999. *Speech recognition and C++ implementation*. John Wiley and Sons, Rome.
- Belin, Pascal – Zatorre, Robert J. – Lafaille, Philip – Ahad, Pierre – Pike, Bruce 2000. Voice-selective areas in human auditory cortex. *Nature* 403. 309–312.
- Braun, Angelika – Köster, Jens-Peter (eds.) 1995. *Studies in forensic phonetics*. Wissenschaftlicher Verlag, Trier.
- Butcher, Andrew R. 2002. Forensic Phonetics: Issues in speaker identification evidence. Paper at the Inaugural International Conference of the Institute of Forensic Studies: "Forensic Evidence: Proof and Presentation". Prato, Italy 3–5 July (CD-ROM).

- Csépe Valéria 2006. *Az olvasó agy*. Akadémiai Kiadó, Budapest.
- Dankovičová, Jana – Nolan, Francis 1999. Some acoustic effects of speaking style on utterances for automatic speaker verification. *Journal of the International Phonetic Association* 29. 115–129.
- Csapó Tamás – Zainkó Csaba – Németh Géza 2009. Szintetizált beszéd prozódiai változatosságának növelése spontán beszéd alapján. Konferencia-előadás. Beszédkutatás 2009, MTA Nyelvtudományi Intézet, Budapest.
- Fék Márk – Olasz Gábor – Szabó János – Németh Géza – Gordos Géza 2005. Érzelem kifejezése gépi beszéddel. *Beszédkutatás 2005*. 134–144.
- Furui, Sadaoki 2005. Recent progress in corpus-based spontaneous speech recognition. *IEICE Transactions on Information and Systems*. E88-D(3): 366–375.
- Gósy Mária 1999. Az egyéni hangszínezet és a beszélő felismerésének kísérleti-fonetikai megközelítése. *Magyar Nyelvőr* 123. 424–438.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008*. 194–207.
- Grice, H. Paul 1975. Logic and Conversation. In Cole, Peter – Morgan, Jerry L. (eds.): *Syntax and Semantics: Speech Acts*. Academic Press, San Diego, 41–58.
- Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszédkutatás 2004*. 187–199.
- Hutchison, Ben – Pereira, Cécile 2001. *Um, one large pizza*. A preliminary study of disfluency modelling for improving ASR. In Lickley, Robin – Shriberg, Elisabeth (eds.): *Disfluency in spontaneous speech. Proceedings*. Edinburgh, 77–81.
- Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből*. Tankönyvkiadó, Budapest, 164–187.
- Kiss Jenő 1995. *Társadalom és nyelvhasználat*. Nemzeti Tankönyvkiadó, Budapest.
- Kock, Ned 2007. Media naturalness and compensatory encoding: the burden of electronic media obstacles is on senders. *Decision Support Systems* 44. 175–187.
- Kontra Miklós 1990. Az élőnyelvi vizsgálatok tárgyáról és módszereiről. In Szabó Géza (szerk.): *II. DialSzimp*. Veszprém, 131–138.
- Labov, William 1979. A nyelv vizsgálata társadalmi összefüggésben. In Pléh Csaba – Terestyéni Tamás (szerk.): *Beszédaktus – kommunikáció – interakció*. Tömegkommunikációs Kutatóközpont, Budapest, 365–398.
- Labov, William 1984. Field methods of the project on linguistic change and variations. In Baugh, John – Sherzer, Joel (eds.): *Language in use. Readings in sociolinguistics*. Prentice-Hall, Englewood Cliffs, N. KJ., 30–53.
- Lavner, Yizhar – Rosenhouse, Judith – Gath, Isah 1999. The prototype model in speaker verification. In Olasz, Gábor – Németh, Géza – Erdőhegyi, Krisztina (eds.): *Proceedings of Eurospeech '99*. Budapesti Műszaki Egyetem, Budapest, 771–774.
- Leech, Geoffrey 1983. *Principles of pragmatics*. Longman, London–New York.
- Levelt, Willem J. M. 1989. *Speaking. From intention to articulation*. A Bradford Book. Cambridge, Massachusetts.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.

- Lustgarten, Paul C. – Juang, Biing-Hwang 2002. Naturalness in speech communications. <http://users.ece.gatech.edu/~juang/Publications/Naturalness> (Az utolsó letöltés ideje: 2009. augusztus 26.)
- Markó Alexandra 2007. A társalgás hangtana. *Magyar Tudomány* 167. 54–57.
- Mihajlik Péter – Fegyő Tibor – Tatai Péter 2006. Új eljárás gépi beszédfelismerés környezetfüggő beszédhangmodelljeinek kialakítására. *Beszéd kutatás* 2006. 218–231.
- Németh Géza – Olasz Gábor – Fék Márk 2006. Új rendszerű, korpuszalapú gépi szövegfelolvasó fejlesztése és kísérleti eredményei. *Beszéd kutatás* 2006. 183–196.
- Nikléczy Péter 1996. Beszélő személy azonosítása szűkfrekvenciás szavak alapján. *Beszéd kutatás* '96. 20–31.
- Nusbaum, Howard C. – Francis, Alexander L. – Henly, Anne S. 1995. Measuring the naturalness of synthetic speech. *International Journal of Speech Technology* 1. 7–19.
- Olasz Gábor 1989. *Elektornikus beszédelőállítás*. Műszaki Kiadó, Budapest.
- Olasz Gábor – Németh Géza – Olasz Péter – Kiss Géza 2000. Profivox – a legkorszerűbb hazai beszéd szintetizátor és szövegfelolvasó. *Beszéd kutatás* 2000. 167–179.
- van Santen, Jan H. P. – Sproat, Richard W. – Olive, Joseph P. – Hirschberg, Julia (eds.) 1997. *Progress in speech synthesis*. Springer, New York–Berlin.
- Schaaf, Thomas – Kemp, Thomas 1997. Confidence measures for spontaneous speech recognition. *International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*. Vol 2. 875–878.
- Seneff, Stephanie 2004. The use of subword linguistic modeling for multiple tasks in speech recognition. *Speech Communication* 42/3–4. 373–390.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszéd kutatás* 2004. 200–208.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Tóth Bálint – Németh Géza 2008. Rejtett Markov-modell alkalmazása magyar nyelvű gépi szövegfelolvasáshoz. *Beszéd kutatás* 2008. 182–193.
- Tüske Zoltán – Simon Márta – Mihajlik Péter – Fegyő Tibor 2007. Érzelmek automatikus felismerése a beszéd akusztikus jellemzői alapján. *Beszéd kutatás* 2007. 151–162.
- Vicsi Klára – Szaszák György – Németh Zsolt 2007. Folyamatos magyar beszéd mondatfajtáinak automatikus felismerése. *Beszéd kutatás* 2007. 162–172.
- Ward, Wayne 1989. Understanding spontaneous speech. In: *Proceedings of the Workshop on Speech and Natural Language*. Association for Computational Linguistics, Morristown, New Jersey, 137–141.
- Wardhaugh, Ronald 1995. *Szociolingvisztika*. Osiris–Századvég. Budapest.
- Wolfson, Nessa 1976. *Speech events and natural speech: Some implications for sociolinguistic methodology*. Cambridge University Press. Cambridge.
- Zainkó Csaba – Fék Márk 2006. Beszédatadabázis prózadiájának szerepe a gép beszéd hangzásában. *Beszéd kutatás* 2006. 208–217.

A SPONTÁN BESZÉD LEJEGYZÉSE – A BEA ADATBÁZIS TAPASZTALATAI ALAPJÁN

Neuberger Tilda

Bevezetés

Megnyilatkozásaink többféle okból jöhetnek létre, így a beszédnek is több fajtája létezik. Egy elhangzott szövegrészletről hamar el tudjuk dönteni, hogy milyen típusú szöveget hallottunk. Ebben segít a szöveg tartalma, felépítése, stílusa, mondatszerkezete, hangzásvilága stb. A tartalomnak, a grammatikai szerkezeteknek és az egyéni beszéd-sajátosságoknak megfelelően változik/változhat a beszéd tempója, ritmusa, kiszélesedhet a hangterjedelem, változhat a beiktatott szünetek hossza és minősége (Imre 2005).

Wacha Imre az elhangzó beszéd vizsgálata során az akusztikus jellemzőket is figyelembe véve, négy fő stíluskategóriát különböztetett meg (Wacha 1974). E stíluskategóriák a következők:

1. spontán, kötetlen beszéd,
2. felolvasás,
3. reprodukív vagy interpretatív beszéd,
4. félreprodukív beszéd.

Spontán beszéd esetén a gondolkodás, a szöveg megszerkesztése és annak elmondása egy időben zajlik, tehát szinkron tevékenység. Felolvasás során már előre megfogalmazott és papírra vetett szöveget szólatatunk meg. Ebben az esetben érzékelhető, hogy a szövegalkotás és a szövegmondás nem szimultán történt. A reprodukív vagy interpretatív beszéd jellemzője, hogy a beszélő ugyancsak egy előre megírt szöveget hangosít meg, de azt az illúziót akarja kelteni, hogy az elhangzó szöveg spontán beszéd terméke. Ebbe a kategóriába tartoznak a különböző előadó-művészeti beszédtevékenységek. A félreprodukív beszéd átmeneti kategória az előző három között, rendszerint előre megfogalmazott szöveg alapján történik, de a beszélő nem feltétlenül ragaszkodik a leírtakhoz, spontán módon kiegészíti azokat. Ez esetben a szöveg megalkotása hol megelőzi annak elmondását, hol pedig a spontán beszédnek megfelelően szimultán módon zajlik le a két folyamat.

A spontán beszéd tehát a beszédprodukciós folyamat működésének eredménye azokban az esetekben, amikor nem előre megfogalmazott szöveg kiejtése történik, amikor különösebb felkészülés nélkül mondjuk el a mondanókat (Gósy 2005). Andó Éva tanulmányában a spontán beszélt nyelv formai jellegzetességeit így foglalja össze: valós idejű, rögtönzött; a hallgató és a beszélő rövid távú memóriájának kapacitása befolyásolja; rendszerint meg-

határozott kontextusban, szemtől szembe kommunikációban jön létre; akusztikai sajátosságokkal jellemezhető; gesztusok, arckifejezés, testbeszéd kíséri (Andó 2003).

Mivel a spontán közlések során a mondanivaló tervezése és kivitelezése csaknem egyidejűleg zajlik, ez gyakran megnehezíti a beszélő (és olykor a hallgató) dolgát. Míg a közölni kívánt gondolatok, képek aktiválódnak, már folyik a nyelvi tervezés is. Ezt nevezzük a beszédtervezés és -kivitelezés paradoxonának, hiszen olyan folyamatokról van szó, amelyek éppen egyidejűségük miatt korlátozzák egymás működését (Gósy 1998). Ahogy Lux Gyula fogalmaz: „a beszélésnek a tulajdonképpeni nehézsége nem maga a beszéd, a kiejtés, hanem a kiejtés előtti, a tudatunk alatt végbemenő, bizonyára érzelmeken nyugvó munka” (Lux 1926 után: 91). A tervezés és a kivitelezés nincs mindig összhangban, ebből következik, hogy a spontán beszédben gyakran előfordulnak a következő jelenségek: redundancia, hiány, közbeekelés, váratlan bővítmények, alanyváltások, időrendi és/vagy logikai sorrendiség be nem tartása, tévesen kiejtett névelő, egyeztetési tévedések, ismétlés, egyéb megakadásjelenségek. A megakadások a spontán beszéd artikulációs, illetve percepciósi folyamatosságát bizonyos értelemben megakasztó, a beszélő saját nyelvi normájának ellentmondó jelenségek. Funkcionális szempontból ezek lehetnek a beszélő bizonytalanságából adódó jelenségek és téves kivitelezések, mint például: néma és kitöltött szünetek, nyújtások, újrakezdések, ismétlések, szón belüli szünetek és töltelékszavak, illetve freudi elszólások, nyelvbtlások, grammatikai hibák, téves szótalálások, változtatások és a „nyelvem hegyén van” jelenség (Gósy 2005).

Ahhoz, hogy a hallgató hosszabb közléseket megértsen és értelmezni tudjon, fontos tényező a szöveg megfelelő tagolása. A beszélő – mintegy tudat alatt – bizonyos egységekre tagolja a spontán beszédét; a hallgató ezeket az egységeket, ún. virtuális mondatokat elsősorban a beszédben előforduló szünetek, másfelől a szintaktikai-szemantikai sajátosságok mentén dekódolja (Gósy 2003a, Váradi 2008). A tagolás artikulációs megvalósítása túlnyomórészt ösztönös, a beszélő nincs feltétlenül tudatában annak, mely szegmentális vagy szupraszegmentális tényezőt alkalmazza tagoló funkcióban, illetőleg meglehetősen nagyok az egyéni különbségek. A spontán beszéd tagolása jelentősen eltér az interpretatív megnyilatkozások, avagy a felolvasás kiejtési sajátosságaitól; hiszen ez utóbbi esetekben a beszédprodukciós mechanizmus kezdeti folyamatainak tervezésére nincsen szükség, a beszélő „készen kapott” szöveget hangosít meg. Ezeknek a szövegeknek az eredeti központosítása befolyással van a meghangosításukra.

A tagolás egyik eszköze a szünettartás. A spontán beszédben előforduló szünetek funkciójuk szerint a következők lehetnek: az artikulációhoz szükséges légzési szünet (oxigén biztosítása), akaratlagos szünet (értelmi tagolás céljából vagy hatásszünet), gondolkodási szünet, bizonytalanság miatt bekövetkező szünet stb. A beszédszüneteknek kettős funkciója van a résztvevők

szempontjából: *a)* a beszélő felől a tervezés és a kivitelezés összehangolatlanságát semlegesítő tényező, *b)* a hallgató felől pedig biztosítják az elhangzott közlések megértését (Gósy 2000).

A beszélt nyelv vizsgálata sajátos metodológiát igényel. A módszerek a kezdeti önfigyelésektől a mások megfigyelésén át a különféle kísérletekig igen sokrétűek. Tágabb körű elemzésre adnak lehetőséget a gyűjtésen alapuló korpuszok, illetve a fejlesztett beszédatadtbázisok. Ezek olyan gyűjtemények, melyek meghatározott szempontok szerint összeválogatott, egységesen kódolt – írott vagy hangzó – szövegeket tartalmaznak. Azokon a területeken lehetnek igazán hasznosak, ahol megfelelő mennyiségű és minőségű felgyűjtött adat nélkül a kitűzött cél nem érhető el, vagy a jelenség nem vizsgálható.

A gyűjtés kétféle módon történhet: „on line” vagy „off line” módszerrel. Az „on line” módszer lényege, hogy a gyűjtő előre megállapított szempontok alapján feljegyzí (majd rendszerezi) a mindennapi élet folyamán észlelt nyelvi adatokat (pl. nyelvbottlásokat, egyéb megakadásjelenségeket). Ez az egyszerűbb eljárás, hiszen csak papír és írószerszám kell hozzá, és segítségével nagyméretű anyag válik elemezhetővé. „Off line” módszer esetén a kutató hangfelvételeket és/vagy ezek fonetikailag hű leírását vizsgálja. Az utóbbi módszerrel gyűjtött adatállomány fő erőssége, hogy nem kell számolni a gyűjtő előzetes elvárásai és a figyelmének ingadozásai miatt bekövetkező torzításokkal, több független személy is meghallgathatja a hangfelvételt, egymás után többször is. A lejegyzések tehát nagyon megbízhatók (Huszár 2005: 22–23). A kinyert adatokból lehetőség nyílik gyakorisági elemzésekre, valamint statisztikai jellegű megállapítások tehetők. Ma már többféle célú és szempontú fonetikai, pszicholingvisztikai és beszédtechnológiai kutatások számára állnak rendelkezésre adatbázisok. Egyre bővül azoknak a nyelveknek a száma, amelyek rendelkeznek ilyen jellegű korpuszal (Gósy–Horváth 2007). Az amerikai, angol, német, ausztrál, japán, lengyel stb. adatbázisok némelyike olvasott beszéd alapú, mások spontán közlésekre, monológokra, dialógusokra épülnek, köznyelvet vagy nyelvjárási beszédet is tartalmaznak. A magyar nyelvű korpuszok, illetve adatbázisok között szerepel az ELTE-n kiadott Magyar nyelvjárási hangfelvételtár, a MTA Nyelvtudományi Intézet Hegedűs-archívuma, a BUSZI (Budapesti Szociolingvisztikai Interjú) vagy a MTBA (Magyar Telefonbeszéd-adatbázis).

A beszédatadtbázisok a hangzó anyag mellett általában azok valamilyen szintű írásos változatát is tartalmazzák. Minden lejegyzés előre meghatározott szempontok, követelmények szerint történik, amelyeket a lejegyzők legjobb tudásuk szerint betartanak. Az átírás során azonban felmerülhetnek olyan (formai) problémák, amelyekre az előírások nem terjednek ki, ezért előfordulhat az, hogy a különböző lejegyzők máshogyan rögzítik az adott jelenséget. A lejegyzés nagyban függ a mindenkor lejegyző percepciók működésétől is, hiszen egy adott hanganyag az egyén percepciók mechanizmusának

szűrőjén megy keresztül: egyes jelenségekre jobban figyel, másokra kevésbé. Az objektív szabályok mellett tehát az írásos változatok bizonyos mértékben szubjektívek.

Különbféle adatbázisok feldolgozási lehetőségeiről már többen beszámoltak. Markó Alexandra és Bóna Judit tanulmánya (2008) a spontán beszéd lejegyzésének módszertani elveit és gyakorlati problémáit mutatja be két felhasználási terület szerint, a szupraszegmentális vizsgálatok és a beszédtechnológiai alkalmazások számára. A Tóth–Kocsor szerzőpáros a fonetikai szintű szegmentálás nehézségeit és érdekességeit osztja meg az olvasóval a Magyar Telefonbeszéd-adatbázis tapasztalatai alapján (Tóth–Kocsor 2003).

A jelen tanulmány célja bemutatni a lejegyzés folyamatát, nehézségeit egy most épülő magyar „off line” adatbázis, a BEA (Beszélt nyelvi Adatbázis) alapján.

BEA: Beszélt nyelvi adatbázis

A fonetikai elemzések igénye és a szélesebb értelemben vett beszéd kutatás új feladatai szükségessé tették egy modern technológián alapuló, a minőségi hangrögzítés minden kritériumát teljesítő többfunkciós korpusz létrehozását. Az MTA Nyelvtudományi Intézet Fonetikai Osztályán 2007 őszén indult meg a BEA elnevezésű adatbázis fejlesztése (Gósy 2008). A beszélt nyelvi adatbázis egynyelvű budapesti adatközlők beszédét rögzíti, nagy mennyiségű és jó minőségű anyagot biztosít különféle kutatásokhoz és gyakorlati alkalmazásokhoz. Többféle típusú spontán beszédet, felolvasást és ismétléseket is tartalmaz. A felvételek öt állandó része: 1. mondatvisszamondás, 2. spontán beszéd, interjú (narratíva személyes kérdésekről és vélemény kifejtés közeleti kérdésekről), 3. társalgás (adatközlő és két társalgó partner), 4. interpretált beszéd (két szöveg tartalmi összefoglalása), 5. szöveg- és mondatfelolvasás. A tervezett adatbázis 500 különböző életkorú (20–70 év) és iskolázottságú magyar anyanyelvű nő és férfi hangfelvételeit fogja tartalmazni a megfelelő átírással együtt. A lejegyzési szabályok megbeszélések sorozatának eredményeként jöttek létre. Egy fonetikusokból álló szakmai bizottság határozta meg az útmutatót korábbi tapasztalatok, illetőleg a BEA felhasználási céljait tekintve véve. A próbalejegyzések után az útmutatóban néhány finomítás történt. Különösen fontos, hogy a mindenkor lejegyző következetesen betartsa a lejegyzési szempontokat (a lejegyzési útmutatót vö. Gósy 2008).

A spontán beszéd lejegyzésének sajátosságai

Az elmúlt másfél évben – amióta az adatbázis fejlesztőivel és a többi lejegyzővel együtt kialakítottuk a lejegyzés végleges szabályait – összesen tizenkét BEA-hanganyag írásos változatát készítettem el. Ez mintegy 450 percnyi hanganyag átírását jelenti, ami több mint 260 oldal terjedelmű szöveges anyagot tesz ki. A következőkben saját tapasztalataim alapján kívánom bemutatni a lejegyzés folyamatát és nehézségeit.

A lejegyzés időigényes munka. Egy percnyi elhangzó szöveg lejegyzéséhez – a szöveg típusától, az adatközlő beszéd-sajátosságaitól és a lejegyző gyakorlottságától függően – átlagosan tíz percre van szükség. Ebbe az időtartamba beletartozik a kijelölt szövegrész első meghallgatása, az ismételt meghallgatások, az adott (hangzó) szakasz írásos változatának elkészítése, majd annak újbóli meghallgatás utáni ellenőrzése és esetleges javítása.

A BEA-hanganyagok részeit tekintve legkönnyebben és leggyorsabban lejegyezhetők a mondatvisszamondást, illetve a felolvasást tartalmazó részek, a legtöbb időt és figyelmet kívánó rész pedig a társalgás. A visszamondásban és a felolvasásban szereplő mondatokat előzetesen írott formában (egy Word dokumentumban) megkapjuk, ezért ezeket a részeket bemásolhatjuk a lejegyzésünk dokumentumába, majd az adatközlő által produkált eltéréseket bejelelhetjük rajtuk (ha vannak ilyenek). Az interpretált beszéd alapját adó szövegeket ugyanilyen formában megkapjuk. Azok előzetes elolvasása után – a szövegek ismeretében – lehetnek bizonyos elvárásaink az interpretációt illetően, amely megkönnyítheti annak lejegyzését. Például: az egyik ilyen szöveg egy történelmi anekdota, amely Varkocs György haláláról szól; az adatközlő nagy valószínűséggel említeni fogja a történet főhősének nevét, valamint a történet főbb mozzanatait (helyszín, török ellenség, csata, visszavonulás, polgárok árulása). Az interjút és a társalgást tartalmazó részekben nincsenek ilyenféle megjósolható megnyilatkozások, ezek számunkra sok új információt tartalmaznak, ezért nehezebb és hosszadalmasabb őket lejegyezni. Az interjúban az adatközlő (A) hosszabb-rövidebb narratívákat hoz létre, s ezeket csak a beszédpartner (T1) kérdései szakítják meg olykor. A felváltva történő megszólalás szerkezete itt kötöttebb, mint a mindennapi életben előforduló beszélgetések során: a kérdező türelmesen végighallgatja az alany mondani-valóját, ha az elakad, akkor segítő kérdésekkel ösztönzi a további beszédre. Az interjút tartalmazó rész lejegyzésének nehézsége az adatközlő beszéd-sajátosságainak függvénye (lásd később). Ezzel szemben a társalgás szerkezete bonyolultabb, lejegyzése a társalgás sajátosságaiból fakadóan több figyelmet igényel: gyakoriak a gyors beszélőváltások, a közbevágás, a félbeszakítás, az együttbeszélés, a hiányos és tagolatlan, valamint a különféle modalitású mondatok stb. Az 1. ábra a gyors beszélőváltást szemlélteti. Egy BEA-hanganyag 15 másodperces részletét mutatja, amelyen jól látszik, hogy a három személy egymásnak adogatja át a szót (1. ábra). Közbevágás, együttbeszélés esetén a lejegyzőnek többször vissza kell hallgatni az adott részt, és még akkor sem biztos, hogy megérti az elhangzott szöveget (2. ábra). Nem ritka, hogy a három személy egyszerre beszél. Ezekben az esetekben igen nehéz mindhárom személy közléseit megérteni, a megértés sikere nagyban függ az együtt beszélő személyek hangerejének eltérő szintjétől. Általában az adatközlő (A) hangjának intenzitása a legnagyobb (ő ül a legközelebb a mikrofonhoz).

A: átjáznak egymáshoz a gyerekek meg hát a szülőknek is valahogy
 T1: persze
 A: ki kell alakítani a kapcsolatot és ti
 T1: ! mesélj
 T2: hát ööö ☐ semmi különöset nem tervezek csak dolgozni
 A: dolgozni !
 T1: húsvétkor ugyan már

1. ábra

Lejegyzésrészlet egy BEA-felvételről: gyors beszélőváltás

T1: amiben van esély hogy meg fognak locsolni
 (T1: minket merhogy [merthogy])
 (A: hát)
 (T1: kábé hét fiú)
 (A: mindig találni olyat !)
 T1: kábé hét fiú lesz úgyhogy

2. ábra

Lejegyzésrészlet egy BEA-felvételről: közbevágás, együttbeszélés

Az adatközlő beszédsajátosságai közül kiemelhetjük a beszéd-, illetve artikulációs tempóját. A beszédtempó az időegységre jutó nyelvi jelek számát jelenti, függetlenül attól, hogy a közlésben volt-e szünet vagy megakadásjelenség. Az artikulációs tempó pedig az artikuláció tiszta idejére eső nyelvi jelek száma, vagyis szünetek nélkül a hangok képzésének sebessége. Nyilvánvaló, hogy minél gyorsabb egy személy beszéd-, illetve artikulációs tempója, annál nehezebb beszédének lejegyzése. Ilyen esetekben a lejegyzést megnehezíti, ha az illető rövid szüneteket tart, és gyors beszédtempójából adódóan általában relative sok megakadásjelenséget produkál. Nemegyszer nehézséget jelent a beszédhangok szegmentálása, hiszen a szavakat alkotó beszédhangok szerialitása felbomolhat, egyes beszédhangok torlódhatnak, olykor hangok és szótagok ki is eshetnek. Szemléltetésül egy-egy 5 másodperces rész két adatközlő beszédéből:

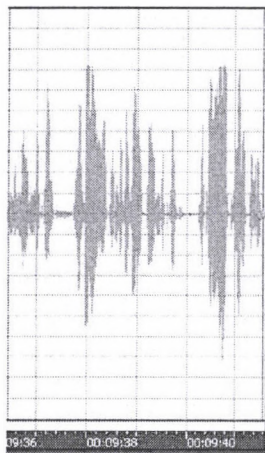
1. példa: *hogy mentünk a kocsival a feleségemmel és mondom vigyázz má [már] mondom mer [mert] izé má [már] mom [mondom] itt egy néni itten bandukol*

2. példa: *sose szoktam ☐ úgy beleülni a kocsiba hogy előtte iszok vagy ☐ akár órákkal előtte mondjuk*

Az első példában egy bizonyos időegységre több hang jut, mint a másodikonál, továbbá több a hangelhagyás, tehát az első példa gyorsabb beszéd-, illetve artikulációs tempójú adatközlőtől származik. Tekintetbe véve egyes nyelvi elemek kiesését, a gyors artikulációs tempójú adatközlők ugyanannyi idő alatt rendszerint több információt közölnek, mint a lassú artikulációs tempójú beszélők. A hangelhagyások miatt töredékesen maradt szavak felis-

merése néha igen nehézkes, ezért a gyors beszélők szövegének átírása több figyelmet kíván a lejegyzőktől.

A lejegyzést megkönnyíti, ha a hanganyagot oscillogrammot mutató hang-elemző programmal vizsgáljuk. Ilyen program például a Praat vagy a GoldWave. Az oscillogram tartalmazza az egyes időpillanatokban megjelenő rezgések amplitúdóját az idő függvényében, így támpontot nyújt az egyes beszédhangok sajátosságairól, képzési kezdetéről, végéről, valamennyire az átmenetektől is. A legszembeütőbb, amit láthatunk rajta, hogy a beszélők mikor közölnek valamit, illetve mikor tartanak szünetet. A szegmentálás ekkor a szöveges szakaszok (beszédszakaszok) és a szünetek, hallgatások határának kijelölését jelenti. A program segítségével kijelölhetünk a hanganyagban szakaszokat, azokat egymás után többször lejátszhatjuk (meghallgathatjuk), így elkerülhetjük a folyamatos visszatekerést, ha újból meg szeretnénk hallgatni az adott beszédszakaszt. A 3. ábrán egy kijelölt szakasz rezgésképe láthatjuk a GoldWave programban, az adatközlő szövege: *sose szoktam □ úgy beleülni a kocsiba hogy előtte iszok vagy □ akár órákkal előtte mondjuk.*



3. ábra

Rezgéskép a GoldWave programban

A lejegyzés egyéni menetét – a lejegyzési útmutató szigorú figyelembevételével – minden lejegyző saját magának alakítja ki. Tapasztalataim alapján azt figyeltem meg, hogy érdemes először a szöveg tartalmára koncentrálni, így hosszabb szósorokat tudunk megjegyezni és lejegyezni. Majd újra meghallgatjuk a szövegrészletet, ekkor a lejegyzésben bejelöljük a (kitöltött és a néma) szüneteket és a megakadásokat. Ha együttbeszélés történik, akkor felbontjuk a szöveget: a fent említett jelöléssel egymás alatt zárójelben közöljük a különböző beszélőktől származó, azonos időben elhangzó szövegrészleteket.

Az, hogy ki milyen terjedelmű szöveget képes megjegyezni egyszerre, három tényezőtől függ: a lejegyzőtől, az adatközlőtől és magától a szövegtől. Egyrészt befolyásolja a lejegyző munkamemória-kapacitása, az aktuális figyelmi állapota, fáradtsága, másrészt a szöveg számos tényezője, mint a mondat szerkezet, a tartalom bonyolultsága, valamint az adatközlő beszéd-sajátosságai, főként a beszédtempója. Pár percnyi anyag folyamatos lejegyzése során a lejegyző hozzászokik az adatközlő beszédstílusához, beszéd-sajátosságaihoz, ebből adódik, hogy a lejegyzés kezdetén még nagyobb a tévesztések megjelenésének aránya, mint később. Saját tapasztalat alapján, egy húszpercnyi anyag lejegyzése után, ami megközelítőleg három órás munkát jelent, a figyelem lankad, ezért a lejegyzés sok pontatlanságot, tévesztést tartalmazhat. A gyakorlott lejegyzők nem javasolják, hogy egyúttal három óránál tovább folytassuk a lejegyzést.

A gyakorlottság a lejegyzés stratégiáinak elsajátítását és automatikus alkalmazását jelenti a lejegyzések során, amelyet körülbelül két-három teljes BEA-felvétel lejegyzése után érhetünk el. A stratégiák automatikus működtetésekor a lejegyzőnek már nem szükséges folyton az útmutatóból kikeresnie, hogy mit mivel jelöljön, hanem a szöveg tartalmára és formájára koncentrálhat, s ez pontosabb lejegyzéshez vezet. Nehezen korrigálható és sok felesleges munkával jár, ha a lejegyző tévesen sajátítja el a stratégiákat, és az átiratokban következetesen a rossz jelölést használja. Például a szón belüli szünetet következetesen nem tapadó négyzettel jelöli tapadó négyzet helyett. További hibás jelölések adódhatnak a következőkből is: kötőjelet (-) újraindításkor tesszünk a töredék után, téves szókezdés és hiány esetén nem, pl. *imá- imádtam, jönnek újr [újra] újak, akk [akkor] menjete*k. Téves szókezdés és hiány esetén tehát nincs kötőjel (-), viszont a feltételezett szót szögletes zárójelben ([...]) feltüntetjük. Mivel ezek a töredékek megakadásjelenségek, vastagon szedjük őket. Lejegyzési hiba, ha a kívánt esetekben hol tesznek kötőjelet, hol nem. Gyakran előfordul az is, hogy a megakadásjelenségeket nem jelölik félkövérrel, vagy hogy a feltételezett szót nem tüntetik fel szögletes zárójelben.

A lejegyzési útmutató szerint a néma beszédhez tartozó, de a beszélő által produkált hangokat (pl. köhögés, belégzés, nevetés) egységesen felkiáltójellel (!) kell jelölni, a néma szünetet pedig négyzettel (□). Problémát okozhatnak a lejegyzőnek azok az esetek, amikor a beszélő szünetet tart, de az nem teljesen „néma”, mert lehet hallani a be- vagy kilégzést. Ilyenkor el kell döntenie, hogy felkiáltójellel vagy négyzettel jelölje az adott jelenséget. Az egyértelműség érdekében a levegővételhez szükséges szünetet, még ha a légzés hallatszik is, a néma szünetet mutató négyzettel jelöljük. Ezt azzal indokoljuk, hogy sokszor nem hallani a légzést vagy csak nagyon halkán, így nem lehet pontosan bejelelni minden esetben, ezért ettől eltekintünk, és csak az egyértelműen biztos tényt jelöljük, jelen esetben azt, hogy a beszélő szünetet tartott.

A lejegyzés során gyakorlottságtól függetlenül is előfordulhatnak tévesztések. Ezek nem a jelölésekre vonatkoznak, hanem legtöbbször a szavak azo-

nosításában és leírásában fordulnak elő. Saját lejegyzéseimben a következő tévesztéstípusok fordultak elő, amelyeket az ellenőrzés során vettem észre és utólag korrigáltam:

1. Egy adott egység (szó/kifejezés vagy szünet) más helyre került a lejegyzésben, mint ahol a hanganyagban eredetileg volt (sorrendi tévesztés) pl. *azon fogok megtanulni igazán vezetni* helyett *azon fogok igazán megtanulni vezetni*.

2. Egy elhangzott szó/kifejezés helyett más (általában hasonló hangzású és rövidebb) szó/kifejezés került a lejegyzésbe (csere) pl. *igazság szerint* helyett *igazából*, *gyökereket eresztenek* helyett *gyökereket növesztenek*, *ez is azért [azért] annyira durva* helyett *ez is azért [azért] olyan durva*.

3. A hanganyagban el nem hangzó egység került a lejegyzésbe (belehallás, kiegészítés) pl. *akkor, mondjuk, tehát*.

4. Egy adott egység kimaradt a lejegyzésből (kihagyás) pl. *és mondjuk ebbe a kategóriába beletartozik* helyett *és ebbe a kategóriába beletartozik*.

Ezek a tévesztések abból adódhatnak, hogy olykor a lejegyzők is (ahogyan a hallgatók vagy a társalgás szereplői) a beszélgetés tartalmára jobban figyelnek, mint annak megformáltságára. Lejegyzéskor természetesen működik a humán beszédészlelési és beszédmegértési mechanizmus, amely a feldolgozársal összefüggő, azaz az elhangzottak tartalmára, információira. Azért nem tűnik fel első hallásra minden megakadási jelenség sem, mert – ahogyan a társalgás szereplői is – a morfológiai, szintaktikai, sőt a tartalmi-logikai tévesztéseket figyelmen kívül hagyják a globális szövegtartalom követése érdekében (Fábricz 1988). A megakadási jelenségek ugyan a beszéd tartalmához nem járulnak hozzá, de a kutatások bizonyították, hogy bizonyos értelemben szükségesek mind a beszélő, mind a hallgató számára. Egyes megakadási típusok a produkcióban a tervezés és kivitelezés diszharmoniaját oldják fel, a beszédmegértésben pedig hozzájárulnak a feldolgozási folyamat sikeréhez (Gósy 2003b). A beszédhelyzet résztvevőit a megakadások tehát általában nem zavarják, a lejegyző munkáját azonban hátráltatják, hiszen pontos felismerésük olykor nehézkes, jelölésük sok időt vesz igénybe. A megakadási jelenségek lejegyzésénél nagyobb az esélye az átírási tévesztések előfordulásának, mint a valódi szavak esetén.

Eddigi munkám során azt figyeltem meg, hogy lejegyzésbeli tévesztések legnagyobb számban a társalgást tartalmazó részben fordulnak elő, a mondatviisszmondást és a felolvasást tartalmazó részekben szinte soha. Ezeket a tévesztéseket az ellenőrzés során korrigálhatjuk. Az önellenőrzésnek két típusa van. Az egyik, amikor kisebb szakaszokat hallgatunk újra a lejegyzés közben (pl. néhány mondatnyi rész), a másik pedig a végső ellenőrzés, vagyis a teljes szöveg újrahallgatása. A minél pontosabb lejegyzés érdekében mindkettőt célszerű alkalmazni.

Azért, hogy még többet megtudjunk a lejegyzés folyamatáról és nehézségeiről, összehasonlítottuk egyazon hanganyagnak két különböző lejegyzőtől

származó átiratát. Véletlenszerűen kiválasztottunk két lejegyzőt, és mindkettőjüket megbíztunk egy 39'10" terjedelmű hangfelvétel lejegyzésével, majd az átiratokat összehasonlítottuk formai és tartalmi szempontok szerint.

A lejegyzés formai hibáit tekintve a legszembevetőbb az volt, hogy a lejegyzési követelményeket, stratégiákat nem sajátították el helyesen, egyes jelöléseket néhol hibásan, néhol nem következetesen alkalmaztak. Az egyik lejegyző az újraindítást, a téves szókezdést és a hiányt (ti. hiányzik egy szó eleje vagy vége) is kötőjellel (-) látta el. Egy másik hiba, hogy az idegen szavak lejegyzésénél a helyesírást követte, nem a kiejtést, holott a lejegyzési útmutatóban az áll, hogy azt először kiejtés szerint jegyezzük le, majd szögletes zárójelben [] megadjuk a szót az adott nyelvnek megfelelő írásmódban. A másik lejegyző szögletes zárójel [] helyett mindenhol kerek zárójelet () használt.

A lejegyzéseket tartalmi szempontból vizsgálva, további különbségeket találtunk: olyan jelenségek jelölését, amelyek a lejegyzők saját percepciók érzékenységtől függenek, ilyenek például a beszédhangok átlagostól eltérőnek ítélt időtartamai, egyes beszédhangok minőségének pontos észlelése, a beszédpszünetek, az idegen szavak, a megakadásjelenségek, illetve a nehezen azonosítható hangsorok észlelése. Ezekben az esetekben az átiratok ellenőrzése (több kutató percepciók döntésén alapuló konszenzus) alapján alakulnak ki a végleges változatok.

A két lejegyző különbözőképpen ítélte meg némely hangok időtartamát (1. táblázat). A jelen tanulmány írója az 1-es lejegyző észleletét ítéli korrektnek ezekben az esetekben.

1. táblázat: A beszédhangok átlagostól eltérőnek ítélt időtartamai a két lejegyzőnél

Lejegyző 1	Lejegyző 2
raboltak eel	raboltak el
galériabóól	galériából
kisebb	kissebb

A beszédpszünetek jelölése több más beszédatadabázisnál is problémát okozott. A problémák formai (a szünetek jelölése) és percepciók (az észlelés pontossága) tényezőkből születnek. Általában mindenhol jelölik valahogyan, de a jelölésben nincs egységes módszer: egyesek fokozatokat különböztetnek meg a szünetek hosszát illetően, mások hosszuktól függetlenül egyetlen jellel jelölik. Az utóbbi fő oka az, hogy félő, hogy a füllel való megkülönböztetésük sok esetben önkényes volna (Kontra 1988). A BEA-ban minden néma szünetet egyetlen jellel (□) jelölünk, ennek magyarázata a szünethossz észlelésének szubjektív volta. Nincs meghatározva, hogy milyen hosszú jelkimaradás számít szünetnek, ez teljesen a lejegyző percepciójára van bízva. Nyilvánvalóan ebből (is) ered, hogy a két lejegyző a szüneteket másképpen észlelte és

jelölte (2. táblázat). A jelen tanulmány írója a 2-es lejegyző ítéletével ért egyet az alábbi esetekben. Az átiratok jó támpontot adnak a további kutatásokhoz, de a pontos méréshez a hanganyag percepció, illetve hangfeldolgozó programmal történő elemzése szükséges.

2. táblázat: A szünetek jelölése a két lejegyzőnél

Lejegyző 1	Lejegyző 2
<i>áldozatainak ☐ az állapota</i>	<i>áldozatainak az állapota</i>
<i>jó ☐ köszönöm</i>	<i>jó köszönöm</i>
<i>évvel</i>	<i>évvel</i>

A megakadásjelenségek észlelése és reprodukálása gyakran nehézséget okoz a lejegyzőknek. Ismeretes, hogy a beszédprodukciós megakadások döntő többségét a hallgatók észre sem veszik, illetőleg percepció mechanizmusuk működése során (akár tudattalanul is) korrigálják azokat (Gósy-Bóna 2006). A lejegyző feladata a megakadásjelenség észlelésén túl annak reprodukálása is, s mivel a megakadások nem járulnak hozzá a szöveg információ-tartalmához, így még több figyelmet igényelnek. A két lejegyző különféleképpen jegyezte le ugyanazt a megakadást, pl. belehallottak vagy elhagytak hangokat/szavakat, másképp hallotta az egyik, mint a másik (3. táblázat). A jelen tanulmány írója az 1-es lejegyző észleletét ítéli korrektnek az első két esetben, és a 2-es lejegyzőjét a következő két esetben.

3. táblázat: A megakadásjelenségek eltérő észlelése és jelölése a két lejegyzőnél

Lejegyző 1	Lejegyző 2
<i>kellett megkép ☐ megküzdenniük</i>	<i>kellett megké- küzdenniük</i>
<i>a várat a túel a túlerő ! ellen</i>	<i>a várat a türe- túlerő ☐ ellen</i>
<i>szerencsétlen alak ☐ szerencsétlenül alakultak</i>	<i>szerencsétlenül ala- szerencsétlenül alakultak</i>

A beszédészlelési és -megértési folyamat támaszkodik a predikciókra is, az előfeltevések nagyrészt segítik, de olykor nehezítik is a percepció működést. Utóbbi eset eredményei a félrehallások. Ez magyarázhatja azokat az eseteket, amelyekben a két lejegyző átirata teljesen mást tartalmazott egy adott helyen. Például eltérő szavakat, illetve kifejezéseket észleltek és jegyeztek le (4. táblázat). A jelen tanulmány írója az 1-es lejegyző átiratát tartja megfelelőnek.

Arra is találhatunk példát, hogy az egyik lejegyző nem érti, hogy mit mondanak egy bizonyos beszédszakaszban, ezt kérdőjellel vagy csillaggal jelöli, a másik viszont megérti az elhangzott közlést (5. táblázat). A jelen tanulmány írója a hanganyag meghallgatása után az 1-es lejegyző átiratát korrektnek találta az alábbi esetekben.

4. táblázat: Eltérő szavak és kifejezések egyazon szövegrészben a két lejegyzőnél

Lejegyző 1	Lejegyző 2
<i>nálunk a nálunk úgy működik</i>	<i>nálunk ugye úgy működik</i>
<i>ők Budán laknak mi meg Dunakeszin úgyhogy</i>	<i>ők Budán laknak mindketten úgyhogy</i>
<i>úgy is volt</i>	<i>úgy is mondta</i>
<i>gratulálok ! és így</i>	<i>a!ha és így</i>

5. táblázat: Megértési nehézségek a két lejegyzőnél

Lejegyző 1	Lejegyző 2
<i>gondolná tovább tehát hogy mi mit kellene</i>	<i>gondolja tehát ???</i>
<i>aztán a Fel nagy örömrre ma született</i>	<i>aztán a ??? meg</i>
<i>akkor lehet hogy még apácázárda volt</i>	<i>akkor még lehet hogy ??? volt</i>

Már két lejegyző munkájának összehasonlításakor is láthatjuk, hogy milyen mértékű eltéréseket okoz a percepciók mechanizmus szubjektivitása. Eppen ezért döntöttek úgy a BEA fejlesztői, hogy az írásos változatok ne fonetikus átírásban készüljenek, hiszen az esetleges téves észlelet félrevezetné a későbbi kutatót. Ezzel szemben a lejegyzés mégsem (teljesen) ortografikus elvek szerint épül fel, hiszen például központozást egyáltalán nem tartalmaz. Ez arra az okra vezethető vissza, hogy ez a fajta jelölésrendszer az írott nyelvben jellemző, a beszélt nyelvi lejegyzésekben félrevezető lehet, ráadásul a virtuális mondatok meghatározása a hibázás lehetőségét növeli.

A lejegyzés tehát nem tartalmaz sem központozást, sem intonációs jelöléseket, nincsenek bejelölve a (virtuális) mondatok, ennek hátránya, hogy a modalitás sem mindig következtethető ki a kontextusból. A beszélő viszonya a mondat valóságtartalmához csak a hanganyag meghallgatásával válik egyértelművé, a hanglejtés ismeretében. A már korábban idézett szövegrész jó példa az átíratban tapasztalható többértelműség szemléltetésére (4. ábra).

A „mesélj” felszólító jellege egyértelmű az ige toldalékolásának köszönhetően. Az utána következő mondat kijelentő módban áll, a társalgó partner (T2) egy tényt közöl. Az ezt követő „dolgozni” szóról azonban pusztán a leírásból nem tudjuk eldönteni, hogy az adatközlő (A) meglepettségében viselkedik vagy felkiált, esetleg csak megismétli a tényt. (A felkiáltó jel a lejegyzésben – mint láttuk már – nem mondat végi írásjelként szerepel, hanem a különböző, nem a beszédhez tartozó, de a beszélő által produkált hangokat jelöljük vele.) Ugyanígy az 1-es számú társalgó partner (T1) következő megnyilatkozása sem egyértelmű a modalitást tekintve, valószínűleg úgy értelmeznénk, hogy a „húsvétkor” egy kérdés, az „ugyan már” pedig egy felkiáltás. A különböző vizsgálatokhoz tehát elengedhetetlen a hanganyagot az átíratlal együtt kezelni.

T1: ! mesélj
 T2: hát ööö □ semmi különöset nem tervezek csak dolgozni
 A: dolgozni!
 T1: húsvétkor ugyan már

4. ábra

Lejegyzésrészlet egy BEA-felvételről: modalitás

Összegzés

A lejegyzett korpuszok bizonyos kérdések megválaszolására jó és könnyen kezelhető adatokat nyújthatnak, más problémák vizsgálatára önmagukban esetleg teljesen alkalmatlanok (Kontra 1988). A BEA lejegyzett szövegei a további kutatásokhoz jól használható kiindulást nyújthatnak. Természetesen az átiratok mellett elengedhetetlen a hanganyag alapján történő ellenőrzés, a keresést azonban nagymértékben megkönnyíti a rendelkezésre álló leírt szöveg. Ha valaki például a szüneteket kívánja vizsgálni, több munkát kell befektetnie, mert az átiratok tartalmazzák ugyan a szünetek előfordulásának helyét, de ez a bejelölés szubjektív, a lejegyzőknek a lejegyzés pillanatában működő percepciójától függ, továbbá a szöveg nem tartalmaz a szünet hosszára vonatkozó jelölést.

A tanulmányban bemutatott lejegyzések célja nem az, hogy az átiratból az elhangzott beszéd minden tekintetben rekonstruálható legyen, hanem hogy segítséget, kiindulást nyújtson a további kutatásokhoz.

A BEA beszédadatbázis egy adott időszakban a nyelv hangzó változatának megismerését teszi lehetővé. A hangtani sajátosságokon túl jól kutatható anyagot nyújt a nyelvészet más területei számára is.

Irodalom

- Andó Éva 2003. Beszélt nyelvi narratívumok szerkezeti összetevőinek és beszédtempójának összefüggése. In Tóth Szergej (szerk.): *Nyelvek és kultúrák találkozása*. A XII. Magyar Alkalmazott Nyelvészeti Kongresszus kiadványai III. Szeged, 98–103.
- Fábricz Károly 1988. A beszélt nyelvi szövegalkotás kérdéséhez. In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. MTA Nyelvtudományi Intézet, Budapest, 78–89.
- Gósy Mária 1998. A beszédtervezés és a beszéd kivitelezés paradoxona. *Magyar Nyelvőr* 122. 3–15.
- Gósy Mária 2000. A beszéd szünetek kettős funkciója. *Beszéd kutatás* 2000. 1–15.
- Gósy Mária 2003a. Virtuális mondatok a spontán beszédben. *Beszéd kutatás* 2003. 19–44.
- Gósy Mária 2003b. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.

- Gósy Mária – Bóna Judit 2006. A megakadásjelenségek javítása a beszédmegértésben. *Magyar Nyelvőr* 130. 33–49.
- Gósy Mária – Horváth Viktória 2007. Fonetikai elemzések a spontán beszédben: alapok, kihívások. *Beszédkutató* 2007. 7–18.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutató* 2008. 194–207.
- Huszár Ágnes 2005. *A gondolatól a szóig*. Tinta Könyvkiadó, Budapest.
- Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- Kontra Miklós (szerk.) 1988. *Beszélt nyelvi tanulmányok*. MTA Nyelvtudományi Intézet, Budapest.
- Lux Gyula 1926 után. *A nyelv*. (Nyelvlélektani tanulmány). Athenaeum, Budapest.
- Markó Alexandra – Bóna Judit 2006. A spontán beszéd lejegyzésének néhány módszertani kérdése. *Beszédkutató* 2006. 124–133.
- Tóth László – Kocsor András 2003. A Magyar Telefonbeszéd-adatbázis (MTBA) kézi feldolgozásának tapasztalatai. *Beszédkutató* 2003. 134–146.
- Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszédkutató* 2008. 134–147.
- Wacha Imre 1974. Az elhangzó beszéd főbb akusztikus stiluskategóriáiról. *Általános Nyelvészeti Tanulmányok* X. 203–216.

A BESZÉLŐ BIZONYTALANSÁGÁNAK JELZÉSEI: ISMÉTLÉSEK ÉS ÚJRAINDÍTÁSOK

Gyarmathy Dorottya

Bevezetés

A beszédprodukciónak folyamata bonyolult, összetett művelet, mely során gondolataink beszéddé alakulnak át. A beszélést megelőzi az átadásra szánt gondolatok megtervezése, illetőleg az ennek megfelelő nyelvi forma hozzárendelése. A beszéd szándékától a kiejtésig tartó tervezési folyamat két fő részből tevődik össze, a makro-, illetve a mikrotervezésből. Az előbbi szakaszban a tervezés döntően képi formában zajlik, a beszélő behatárolja a közölni kívánt tartalmat és a közlés célját, majd a mikrotervezés során grammatikailag és pragmatikailag releváns nyelvi formába önti az eredeti beszéd-szándékot, és az elhangzásra vonatkozólag is megtervez egy bizonyos időrendet (Levelt 1989, Gósy 2005). Az artikuláció következtében létrejön a beszéd akusztikai hullámformája. Az akusztikai szerkezetet több tényező befolyásolja. Az emberek beszédszervei kisebb-nagyobb mértékben különböznek egymástól, továbbá a működtetésük még az ugyanazon nyelvet beszélők esetében sem egyforma. Ennek következtében az egyes beszélők által létrehozott beszéd akusztikailag különböző (Gósy 2004a). Nem csupán az egyes beszélők között tapasztalható az akusztikum nagymértékű variabilitása; ugyanazon beszélő ugyanazon beszédhangjai sem azonosak (vö. Gósy 2009).

A beszélő artikulációját nagymértékben befolyásolja a beszédhelyzet, illetve a beszédpartner is (Bata-Grácsi 2009). Lindblom (1990) „hipo- és hiperartikulációs modelljében” az artikuláció egyfajta gazdaságosságra törekvő motoros tevékenységként jelenik meg; a beszélő önkéntelenül is a hallgató igényeihez igazítja az artikulációját. Ha a hallgatónak a sikeres percepcióhoz maximális akusztikai információkra van szüksége, a beszélő jobban odafigyel az artikulációra, ezt nevezzük hiperartikulációnak. Ha azonban a hallgató beszédfeldolgozását az akusztikai információkon kívül más információk is segítik (pl. előrejelezhető kontextusok), a beszélő egy redukáltabb artikulációs módra vált, a hipoartikulációra. Az artikuláció ilyen mértékű különbözősége a beszédfelismerés számára is problémát jelent. A felolvasott beszéddel, avagy izolált szavas ejtéssel ellentétben ugyanis a spontán beszéd nagy variabilitást mutat, ami az automatikus felismerést különösen bonyolulttá teszi (vö. Beke-Szaszák 2010, Xie et al. 2004, Halpern 2006).

A spontán beszéd egyik jellegzetessége, hogy nem minden esetben folyamatos, az emberek többé-kevésbé akadozva, töredezetten beszélnek. A be-

széd folyamatosságát (nyelvtől függetlenül) szünetek, hezitálások, nyelvbotlások és más hibák akasztják meg, melyek a beszédproduktions folyamatok bonyolultságára vezethetők vissza (Gósy 2004b, 2005; Horváth 2004; Markó 2004; Szabó 2004; Dabóczi 2008). A tervezés és a kivitelezés gyakorlatilag egyszerre zajlik, ezért gyakran összehangolatlanná válhat. Ez a diszharmonia a felszíni szerkezetben megakadásjelenségekként realizálódik. A magyar nyelvre jellemző megakadásjelenségek alapvetően két nagy csoportba, a beszélő bizonytalanságából adódó jelenségek, illetőleg a hiba típusú jelenségek közé sorolhatók (Gósy 2002). Ezek különös jelentőséggel bírnak a beszéd kutatásban, hiszen azokra a folyamatokra és működéseikre utalnak, amikor a hibátlan közlések elemzésekor nem kaphatunk közvetlen információt. Ugyanazon produktions szabályok hozzák létre ugyanis mind az ép, mind a hibás alakokat (Pouplier–Hardcastle 2005).

Kutatásokkal igazolt tény, hogy az utóbbi évtizedekben megnőtt a megakadásjelenségek előfordulási gyakorisága. A múlt század hatvanas éveinek spontánbeszéd-elemzései alapján 1000 szavanként mindössze egy-két hezitálás volt adatolható (Garnham et al. 1981), csaknem húsz évvel később ez az arány 100 szavankénti 2–26 előfordulásra nőtt (Lutz–Mallard 1986). Napjaink adatai szerint a spontán közlésekben átlagosan 6 szavanként fordul elő megakadás (néma szünetek nélkül): narratívákban 100 szavanként átlagosan 3,6 szó (Fox Tree 1995), míg dialógusokban 100 szavanként 5,5–8,83 szó (Bortfeld et al. 2001). A magyar nyelvre vonatkozó vizsgálatok hasonló előfordulási gyakoriságot mutatnak (vö. Gósy 2003). Gósy és Gyarmathy (2008) kutatásukban mintegy hat évtized elteltével hasonlították össze a spontán beszéd megakadásjelenségeit. Eredményeik igazolták, hogy a mai beszélők megnyilatkozásait jóval gyakrabban tarkítják „nyelvbtlások”, mint a régieket. A hezitálások, az ismétlések és a hiba típusú jelenségek jelentős növekedését adatolták.

A beszélő nemegyszer észreveszi hibázásait, és lehetőség szerint javítja is azokat (Markó 2006). Mindezt a beszédproduktio közben folyamatosan működő önellenőrző, önmonitorozó folyamatok teszik lehetővé. A produktio során ugyanúgy kontrolláljuk saját beszédünket, mint hallgatóként beszédpartnerünkét (Levelt 1989). Valószínűsíthetően egyetlen, percepció alapokon működő önmonitorozó mechanizmus létezik, amely mind a hibadetektálásért, mind pedig a javítás megtervezéséért és végrehajtásáért felelős (vö. Gósy 2008a). Levelt (1983) az önjavítás folyamatát három fő részre osztja. Az első rész a hiba megvalósulása: amikor az önmonitorozó mechanizmus a tervezés során hibát észlel, a beszélő félbeszakítja az artikulációt. Ezt követi a szerkesztési szakasz, ami alatt végbemegy a hibajavítás megtervezése. Ezt az időt a beszélők rendszerint néma szünettel hezitálással, töltelékszavakkal, egyéb nyelvi (*jaj, bocs*) vagy nem nyelvi (torokköszörülés) jelekkel töltik ki. Végül a harmadik részben végbemegy a hiba javítása. A hiba felismerése és javítása végbemehet a saját beszédünk hallás alapú megértése révén (felszíni önellen-

örzés), illetőleg a beszédtervezési folyamat részeként (rejtett önellenőrzés) (Gósy 2008a).

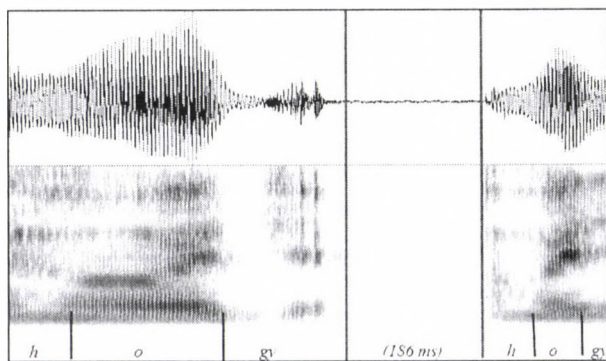
A szerkesztési szakaszok hossza utalhat az önmonitorozás rejtett vagy felszíni voltára. Marslen-Wilson (1990) igazolta, hogy minden szó esetében létezik egy ún. „felismerési pont”, amelytől az adott szó már biztonsággal beazonosítható. Az angol szavakat a hallgatók átlagosan 200 ms-mal az elhangzás kezdete után felismerték, (ennyi időre van tehát szüksége a hallás alapú percepciónak a biztos azonosításhoz), ezért feltételezik, hogy amennyiben a hiba detektálása és esetleges javítása 200 ms-nál rövidebb idő alatt megy végbe, az a rejtett, míg ha ennél hosszabb időt vesz igénybe, az a felszíni önmonitorozás eredménye (Nooteboom 2005). A hosszabb magyar szavak esetében a felismerési pont meghatározása nehezebb. Egy még nem publikált kísérlet eredményei azt mutatják, hogy a spontán beszédből kivágot szavak kezdő hangjától számított 200 ms-nyi időtartam csupán 50%-ban tette lehetővé a szófelismerést. Az a kérdés tehát, hogy a magyar nyelv esetében hány ms-tól beszélhetünk biztosan felszíni monitorozásról, egyelőre még megválaszolásra vár, de az eredmények alapján az kijelenthető, hogy a 200 ms-nál rövidebb lexémarészletek felszíni monitorozással nem azonosíthatók. Ha tehát a hiba detektálása és korrekciója 200 ms-nál rövidebb idő alatt megy végbe, rejtett önmonitorozásról beszélhetünk. A javítandó hiba sok esetben nem jelenik meg a felszínen, de egyéb (nem hiba típusú) jelenségek jelezhetik a beszédtervezés során bekövetkezett hibákat (Horváth 2009). A beszélőnek a tervezés közben fellépő bizonytalansága érhető tetten az ismétlés és az újraindítás jelenségében is. Az előbbi esetében a teljes lexéma kiejtése után a beszélő megismétli azt (*ez a versenyszellem **nem** [szünet: 152 ms] **nem** olyan erős*), míg az utóbbinál az aktivált és félig kiejtett szót a teljes szó kiejtése követi (*ezeket **i- időnként** körbe kell járni*). Felmerül a kérdés, hogy milyen rejtett folyamatok vezetnek a felszínen az ismétlések és az újraindítások megjelenéséhez, illetve hogy a kétféle megakadásjelenség azonos vagy különböző folyamatok eredményeképpen keletkezik-e. Mivel mind az ismétlés, mind az újraindítás egyfajta tervezési bizonytalanságot jelez, feltételezzük, hogy különbség csak az elbizonytalanodás időpontjában van közöttük. A már aktivált célszó az artikulációs tárban várakozik a kiejtésre, amikor a beszélő a folyamatos önmonitorozás eredményeképpen bizonytalanná válik vagy a tervezett, vagy a már elhangzott nyelvi elem megfelelőségében.

Az újraindításoknál az elbizonytalanodás a célszó kiejtése közben következik be. Az egyik lehetséges ok, hogy a beszélő nem biztos abban, hogy a közlési szándékának megfelelő szót aktiválta, ez a lexikális előhívás problémájára utal. A másik ok az lehet, hogy a folyamatosan működő önellenőrző mechanizmus valahol hibát észlel. Mindezek miatt az artikuláció megszakad, majd a zavar elhárultával megtörténik a teljes szó kimondása.

Az ismétlések esetében – az újraindításokkal ellentétben – még a célszó artikulációjának megkezdése előtt következik be az elbizonytalanodás. A be-

szelőnek időre van szüksége, hogy a produkciós folyamat valamely szintjén bekövetkezett hibát kijavítsa. A tervezési problémát a felszínen úgy próbálja megoldani, hogy az utoljára kiejtett szó megismétlésével időt nyer. Egy példával szemléltetjük, illetve magyarázzuk az ismétlés és az újraindítás háttérében végbemenő feltételezett folyamatokat. A „*megettük az ebédet majd hazamentünk*” megnyilatkozás kiejtése közben az önmonitorozó mechanizmus zavart jelez, ami miatt a beszélő elbizonytalanodik a folytatásban. Ha az elbizonytalanodás a „*majd*” szó kiejtése közben vagy után következik be, a beszélő megismétli azt (*megettük az ebédet majd majd hazamentünk*), ha pedig már a „*hazamentünk*” – mint az artikulációs tárban várakozó célszó kiejtése közben –, akkor az artikuláció leáll, majd megtörténik a teljes szó kimondása (*megettük az ebédet majd hazame- hazamentünk*). Mindkét jelenség esetében megjelenhet a szerkesztési szakasz, melynek hosszából következtethetünk a rejtett és a felszíni önmonitorozás működéseire. Clark és Wasow (1998) Levelt önjavítási modelljét az ismételt szavak visszaállító modelljeként kiterjeszti az ismétlés jelenségére. Az első rész a „kiinduló beszédszándék”, ezt követi a modell második részében a beszéd megszakítása, amely a megnyilatkozás bármely részén bekövetkezhet attól függően, hogy a beszélő mikor veszi észre a hibát. A harmadik rész a közlés megszakítása és újratekzdése közötti „hiátus”, ami lehet néma szünet, de tartalmazhat hezitálást és töltelékszavakat is. Az utolsó részben a beszélő folytatja az aktuális megnyilatkozást.

Az ismétlések és az újraindítások folyamán a percepció számára azonos fonémásorok hangzanak el. Mindkét elemzett megakadás esetében a beszélő ugyanazt a lexikai egységet ismétli meg nagyon rövid időn belül, az első és a második kimondás akusztikai vetülete mégis különböző lehet (Benkenstein–Simpson 2003). Jól szemlélteti ezt az 1. ábra rezgés- és hangszinképe, ahol a beszélő a *hog*y szót 186 ms-os időkülönbséggel ismételte meg. Látható, hogy a két kimondás akusztikai szerkezete különbözik egymástól.



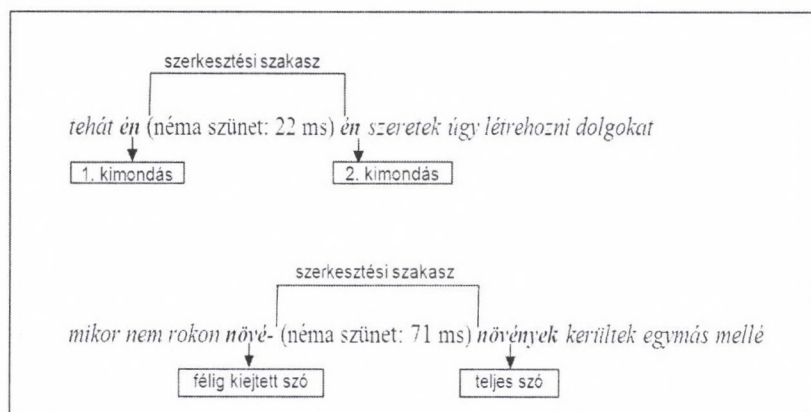
1. ábra

A *hog*y (186 ms) *hog*y rezgés- és hangszinképe

Választ kerestünk arra, hogy a spontán beszédben a két kimondás közti különbség milyen akusztikai paraméterek változásában jelenik meg. Feltételeztük, hogy az alaphangmagasság, az időtartamok és a magánhangzók artikulációja lesz elsősorban érintve. A szerkesztési szakaszok hosszának elemzésével választ kerestünk arra, hogy a két megakadásjelenség esetében a javítás valószínűsíthetően a rejtett vagy a felszíni önmonitorozás működésének eredményeképpen ment-e végbe.

Anyag, módszer, kísérleti személyek

A kutatáshoz a BEA spontánbeszéd-adatbázis 16 beszélőjének (9 férfi, 7 nő, átlagéletkoruk 42,7 év) felvételéből a spontán beszédet rögzítő részeket (narratíva, véleménykifejtés, tartalomösszegzés, társalgás) használtuk fel (vö. Gósy 2008b). Az összesen 8 óra 15 perces anyagban az általunk elemzett két megakadásjelenségből 858 előfordulást adatoltunk. Az ismétléseket és az újraindításokat a Praat programmal (Boersma-Weenink 2009) annotáltuk. Adatoltuk az első és a második ejtés közötti szakasz (szerkesztési szakasz) időtartamát, ami az ismétlés esetében a megismételt szó két kimondása között eltelt időt, míg az újraindításnál a félig kiejtett és az újraindított teljes szó közti időt (az első kimondás utolsó hangjának végétől a második kimondás első hangjának kezdetéig tartó szakasz) jelentette (2. ábra).



2. ábra

Szerkesztési szakasz az ismétlésnél és az újraindításnál

Elemeztük, hogy a szerkesztési szakaszok milyen arányban realizálódnak néma szünetként, hezitációként, illetve a kettő kombinációjaként. Az oszcillogram és a spektrogram figyelembevételével, folyamatos auditív ellenőrzés mellett kimértük az első és a második kimondás időtartamát. Az ismétléseknél a részt vevő két szó teljes hosszát vetettük össze, míg az újraindításoknál

a második kimondásból is csak az első kimondásban már elhangzott szórészlet időtartamát vettük figyelembe (a fenti példában tehát mindkét esetben a *nővé-* szórészlet időtartamát adatoltuk). Az alaphangmagasság és a magánhangzók formánsértékeinek meghatározásához automatikus mérést alkalmaztunk a Praat program 5.0.03-as verziójában. Az automatikus mérés minden 10 ms-ban rögzítette az F_0 és az első, a második, valamint a harmadik formáns átlagát. A kapott adatokat ellenőriztük és szükség esetén manuálisan javítottuk. Az adatok statisztikai elemzéséhez az SPSS szoftver 8.0-s verzióját használtuk. Az eredményeket egytényezős ANOVA-val (95%-os szignifikanciasinten), párosított t -próbával, illetőleg többváltozós lineáris regresszióval, valamint binominális nemparaméteres teszttel elemeztük.

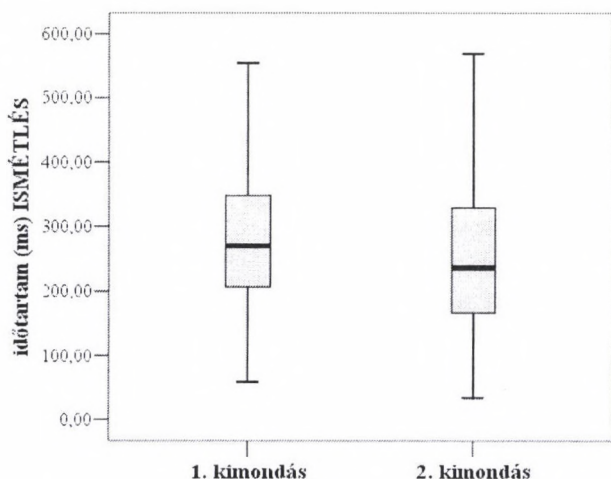
Eredmények

Az elemzett spontán beszédben az ismétlésre 705, míg az újraindításra 153 példát találtunk. A 16 beszélőtől egyénenként átlagosan 30 perc 45 másodperces beszédrészleteket elemeztünk; a legrövidebb felvétel 14 perc 36 másodperc, a leghosszabb 50 perc 50 másodperc hosszúságú volt (átlagos eltérés: 10 perc 13 másodperc). A beszédanyagban az adatközlők közel 4,5-ször annyi ismétlést produkáltak (82,17%), mint újraindítást (17,83%). Az ismétlések átlagosan 42 másodpercenként, míg az újraindítások átlagosan 3 perc 14 másodpercenként fordultak elő. Ez összhangban áll azzal a már számos magyar és külföldi kutatás által igazolt ténnyel, miszerint a spontán beszédben az egyik legtöbbször előforduló megakadásjelenség az ismétlés (vö. Clark–Wasow 1998; Gósy 2003; Gyarmathy 2006, 2007; Horváth 2007; Bóna 2006). Arra, hogy az újraindítások az ismétlésekhez képest ilyen keveszer fordulnak elő, lehet az magyarázat, hogy az önellenőrzési folyamat gyakran még az artikuláció megkezdése előtt jelzi a hibát, és mód van a rejtett korrekcióra. Az is feltehető továbbá, hogy az ismétlések különféle hibák felszíni megjelenései, míg az újraindítások elbizonytalanodási következmények.

Az ismétlések eredményei

Elemeztük, hogy az első, avagy a második kimondás igényelt-e nagyobb időráfordítást a beszélőktől. Amennyiben a hibadetektálási és javítási folyamatok a tervezés során megfelelően működnek, a beszélő a szerkesztési szakasz alatt képes korrigálni a hibát, ezért a megismételt szó kiejtése feltehetően kevesebb időt vesz igénybe. Az esetek 71,95%-ában a második, míg 28,05%-ában az első kimondás időtartama volt a rövidebb. Az ismétlésben részt vevő szavak az első elhangzáskor átlagosan 289,22 ms (59–901 ms) hosszúságúak voltak (átlagos eltérés: 171,03 ms), ez az időtartam a második elhangzáskor 261,32 ms-ra (34–1043 ms) rövidült (átlagos eltérés: 134,65 ms). Mivel az ismétlés egy hangból (pl. határozott névelő, egyes szám harmadik személyű személyes névmás), illetőleg több hangból álló szavakat (pl. *akkor*, *valami*) egyaránt érinthet, a szóródás mindkét esetben nagy. Az időtartamértékek statisztikai elemzése (egytényezős ANOVA) alapján megállapítható, hogy az

ismétléseknél a második kimondás szignifikánsan rövidebb [$F(2, 1372) = 16,293$; $p = 0,000$] az elsőnél (3. ábra). Az ismétlés a tervezési folyamat valamely részén bekövetkezett hiba javításához biztosított megfelelő időt a beszélő számára. Ekkor tehát nem maga az ismétlés tárgyát képező szó okozza a zavart, az csak egyfajta időnyerő stratégiaként funkcionál. A beszélő fenn kívánja tartani a folyamatos beszéd látszatát, ezért ismételi. Az esetek többségében a szerkesztési szakasz alatt sikeresen végbemegy a javítás, ennek következtében a második kimondás időtartama lerövidül.



3. ábra

Az első és a második kimondás időtartamértékei az ismétléseknél (szóródás és medián)

Választ kerestünk arra a kérdésre, hogy mi történik azokban az esetekben (az ismétlések 28,05%-a), amikor a második kimondás időtartama – a fenti tendenciával ellentétben – hosszabb, mint az első. Az esetelemzésből kiderült, hogy a beszélő ilyenkor még túlnyomórészt a közlés folytatásán gondolkodik, esetleg a mentális lexikonban való kereséssel vannak problémái. Az első és a második példa a lexikális előhívás problémáit szemlélteti, a harmadik pedig a fogalmi tervezés nehézségeit:

1. *nem tudhatjuk hogy mi lesz tíz húsz év múlva mit hoz az élet hiszen (néma szünet: 717 ms) **az** (171 ms) **az** (300 ms) *életben egy ööö (1031 ms) (néma szünet: 1055 ms) foly-* (néma szünet: 176 ms) *ööö (237 ms) folyamatos dolog van vagy egy fix pont van az pedig a változás**

A beszélőnek ebben a példában az okozhatta a problémát, hogy nem tudta idejében aktiválni a *fix pont* szavakat (esetleg nem tudta, hogyan fogalmazza

meg nyelvillel a gondolatot), ezért időre volt szüksége. A szóaktiválási nehézséggel már a *hiszen* szó kimondása közben szembesült, ezt jelzi a 717 ms hosszúságú néma szünet. A diszharmóniát az *az* névelő megismételt kimondása alatt is igyekszik feloldani – sikertelenül. Ezt nem csupán az jelzi, hogy a második kimondás időtartama az elsőhöz képest csaknem a duplájára növekszik, hanem a két szóval később megjelenő hosszú hezitálás és néma szünet is. A beszélő bizonytalansága a folytatásban sem szűnik meg: a *folyamatos* szó artikulációját megszakítja, mivel nem ez az általa kimondani szándékozott célzó. Mivel a *folyamatos dolog* kifejezés beleillik a kontextusba, nem okoz értelmi zavart a hallgatónak (a döntés meghozatalára a megfelelő időt a 176 ms-os néma szünet és a 237 ms-os hezitálás biztosítja), a beszélő kimondja azokat, és ezalatt sikeresen aktiválja a *fix pont* kifejezést is.

2. *neki is kellett hogy legyen valami vízszintes vagy* (233 ms) *ööö* (169 ms) *vagy* (327 ms) *egyenés vonal ami párhuzamos a fallal és akkor valamihez tud viszonyítani*

A második példában a beszélő tévesen aktiválta és ejtette ki a *vízszintes* szót, amit azonnal észrevesz, a hiba javítása azonban nem fejeződik be a 233 ms-nyi *vagy* és a 169 ms-nyi hezitálás alatt, ezért megtörténik a *vagy* ismételt kimondása, és az ideje megnövekszik. A teljes, mintegy 730 ms-nyi időtartam már elegendő a célzó (*egyenés*) aktivációjához, grammatikai, fonológiai és artikulációs kódolásához.

A beszélő nemegyszer hangot is ad annak, hogy az adott pillanatban még a közlés folytatásán gondolkodik. Ezt szemlélteti a következő példa:

3. *én nem vagyok ez* (229 ms) (néma szünet: 50 ms) *ez* (554 ms) *hát mmm* (298 ms) (néma szünet: 279 ms) *hogy mondjam* (néma szünet: 1809 ms) *ööö* (720 ms) *tehát nyilván olyan értelemben toleránsnak kell lennem*

A beszélőnek egy váratlanul feltett kérdésre kell válaszolnia, az artikuláció megkezdésekor azonban még láthatólag nincs kialakult gondolata arra vonatkozólag, hogy mit fog mondani. A tervezésben bekövetkezett bizonytalanság feloldására sem az *ez* szó ismétlése, sem az utána következő töltelékszó, hezitálás és néma szünet ideje nem elegendő. A *hogy mondjam* szavakkal a hallgatót egyfelől informálja az aktuális tervezési zavaráról, másfelől biztosítja arról, hogy szándékában áll a folytatás, a jelen esetben a kérdés megválaszolása.

Vizsgáltuk a szerkesztési szakaszok időtartamát és felépítését. Az ismétlések szerkesztési szakasza átlagosan 277,41 ms volt, az adatok azonban 0 ms és 2909 ms között szóródnak (átlagos eltérés: 465,20 ms). Előfordul, hogy a szerkesztési szakasz teljesen hiányzik. Ezekben az esetekben biztosra vehető a rejtett önmonitorozás hibátlán működése; a hiba felismerése és sikeres korrekciója még a tervezési folyamat részeként végbement, ezért a hiba nem jelenik meg a felszínen, csak az ismétlés ténye utal a beszélő tervezési problémára.

májára (pl.: *ez nem lehet egy jó érv hiszen hiszen egy gyereknek szüksége van arra*). Az ismétlések 33,96%-ára volt jellemző a zéró ms időtartamú szerkesztési szakasz; ekkor a megakadás valójában nem is szakítja meg a közlés folyamatosságát. Nem tudunk pontos határt megállapítani arra vonatkozólag, hogy hány ms felett beszélhetünk hallás alapú önmonitorozásról. Valószínűsíthetjük, hogy amennyiben a szerkesztési szakasz 200 ms-nál rövidebb, úgy a javítás a rejtett monitorozás eredményeként ment végbe (vö. Bevezetés). Az ismétlések 63,54%-ánál a szerkesztési szakasz 200 ms-nál rövidebb, tehát az elbizonytalanodást okozó hiba detektálása és javítása a rejtett önellenőrzés során történhetett (a hiba a felszínen nem is jelent meg).

A szerkesztési szakasz felépítése különbözően alakulhat; állhat csak néma szünetből [*bármilyen olyan dolog is felmerülhet a felvétel kapcsán ami* (néma szünet: 667 ms) *ami nem teszi lehetővé az azonosítást*], kitöltött szünetből [*azt gondolom hogy ööö* (50 ms) *hogy nem jó hogyha a társadalom ennyire elfogadó*], és lehet e kettő kombinációja [*lehet találni támpontokat ami ööö* (215 ms) (néma szünet: 551 ms) *ami közelebb vihet*]. Az ismétlések szerkesztési szakasza legnagyobb arányban néma szünetként valósult meg (80,04%-ban); 7,76%-ban hezitálásként, 12,19%-ban néma szünet és hezitálás különböző kombinációiként (1. táblázat).

1. táblázat: Az ismétlések szerkesztési szakaszának hossza

Szerkesztési szakasz	Átlagos időtartam	Minimum	Maximum
Néma szünet	354,14 ms	7 ms	2909 ms
Hezitálás	180,11 ms	42 ms	909 ms
Néma szünet és hezitálás	1005,85 ms	90 ms	3537 ms

A tartalmas szavak, illetve a funkciószók a spontán beszédben eltérő arányban fordulnak elő. Szende Tamás 1973-as adatai alapján a spontán beszéd 39,25%-a tartalmas szó, 60,75%-a pedig funkciószó, míg egy tíz évvel később elvégzett, kötetlen beszélgetéseket vizsgáló kutatás eredményei szerint a tartalmas szavak előfordulási gyakorisága 41,21%, míg a funkciószavaké 58,79% (Keszler 1983). A BEA-n végzett, nem publikált adatok szerint a tartalmas szavak előfordulása 37,92%, a funkciószavaké 62,08%. Az angol nyelvű Switchboard-korpusz elemzése hasonló tendenciát mutat (eltekintve a nyelvspecifikus különbségektől); a funkciószavak 37,6%-a több mint 1000-szer fordult elő, ellentétben a tartalmas szavakkal, melyeknek csupán 0,5%-áról mondható el. Kevesebb mint 10 előfordulás jellemezte azonban a tartalmas szavak 75%-át, a funkciószavaknak pedig az 5,7%-át (Clark–Wasow 1998). A tartalmas, illetve a funkciószavak aránya az ismétlések esetében eltérően alakul; a spontán beszédbeli gyakoriságnak megfelelően az ismétlések nagyobb mértékben érintettek funkciószavakat, mint tartalmas szavakat. Az ismétlések 92,91%-a funkciószavakon fordult elő, csupán 7,09%-uk volt tar-

talmas szó. Annak igazolására, hogy ez az előfordulási arány véletlenszerű-e, binominális nemparaméteres tesztet készítettünk. Nullhipotézisünk (H_0) szerint a funkciósavak és a tartalmas szavak egyenlő arányban jelennek meg, ezért a H_0 szerinti gyakoriságot 0,5-re, 50%-ra állítottuk. A teszt lefuttatása után az eredmények azt mutatják, hogy a H_0 szerinti gyakoriság (0,5) mellett, a tapasztalati gyakoriság $91/152 = 0,60$, ami egyoldali szignifikanciát eredményez ($p = 0,0175$). Nullhipotézisünket ezek alapján elvethetjük, ami azt jelenti, hogy az ismétléseknél a funkciósavak túlsúlya nem véletlenszerű. Ez egybevág az angol nyelvre kapott kutatási eredményekkel, amelyekben megállapították, hogy a beszélők a funkciósavakat sokkal gyakrabban ismétlik, mint a tartalmas szavakat (Fox–Jasperson 1995). Ez részben a gyakorisági tényezőkkel van összefüggésben, részben pedig az ismétlés elsődleges funkciójával, az időnyeréssel. Úgy tűnik, fontos, hogy a beszélő olyan szót válaszson, amely tartalmilag nem változtatja meg a közlését, nem okoz értelmi zavart a hallgató számára. Kiváltja ugyanakkor a (további) hezitálást, tehát fenntartja a folyamatos beszéd látszatát. A másik lehetséges magyarázat a funkciósavak klisészerű tárolásával van összefüggésben (Gósy 2005); ismétlésük valószínűleg egyszerűbben megoldható.

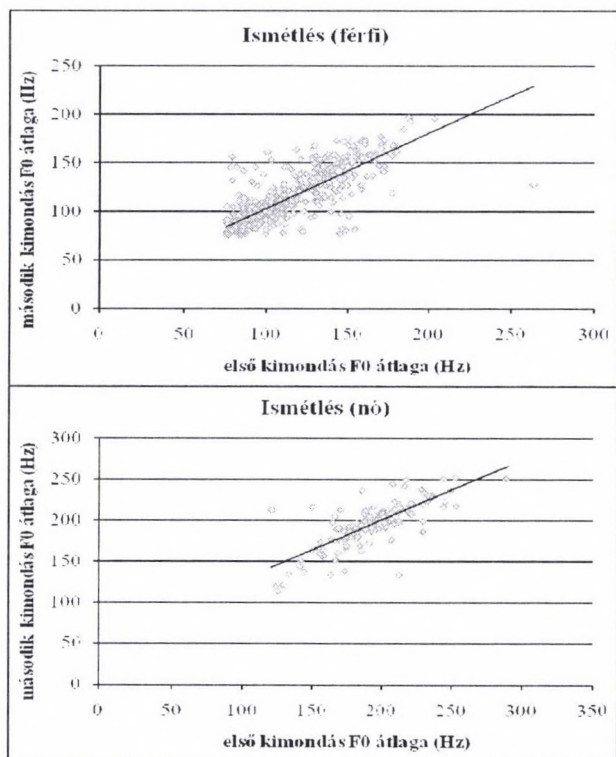
Elemeztük a beszélők alaphangmagasságának változását mindkét nemnél a két kimondás során. A nőknél a második kimondáskor magasabb átlagos alaphangmagasságot mértünk, mint az elsőnél, de a különbség rendkívül csekély. A férfiaknál a második kiejtés F_0 -átlaga alacsonyabb volt az elsőnél (2. táblázat), ami az érdes zöngé (Böhm–Ujváry 2008) második kiejtésben való megjelenésével magyarázható. A különbség itt is kicsi. Az ismétlések közel 10%-ában adatoltunk érdes zöngét. Ebből az esetek 8,96%-ában mindkét, 25,37%-ában az első, míg 65,67%-ában a második kimondás valósult meg érdes zöngével. A beszélők az ismételt ejtésre kevesebb energiát fordítanak.

2. táblázat: A férfiak és a nők alaphangmagassága az ismétlésekben

	1. kimondás		2. kimondás	
	F_0 -átlag	Átlagos eltérés	F_0 -átlag	Átlagos eltérés
Férfi	106,72 Hz	35,39 Hz	102,58 Hz	34,79 Hz
Nő	190,14 Hz	41,32 Hz	192,94 Hz	33,44 Hz

Az alaphang hangterjedelme a nőknél a második kimondás során szűkül; de az adatok statisztikai elemzése nem mutatott szignifikáns különbséget. A férfiaknál azonban a megismételt szó F_0 -hangterjedelme szignifikánsan szélesebb [$t(494) = -2,460$; $p = 0,014$]. Curl és munkatársai (2006) angol nyelvű társalgásokból származó ismétlésekben elemezték az alaphangmagasságot, és a kapott eredmények alapján megállapították, hogy az ismételt részben az F_0 -hangterjedelem általában szűkebb.

Választ kerestünk arra a kérdésre, hogy a szerkesztési szakasz hosszától függően módosul-e a második kimondásban az alaphang átlagértéke, a két változó között azonban nincsen statisztikailag igazolható összefüggés. Annak megválaszolására, hogy milyen összefüggés mutatható ki az első kimondás F_0 -átlaga és hangterjedelme, valamint a második kimondás F_0 -átlaga és hangterjedelme között, többváltozós lineáris regressziót alkalmaztunk. Mindkét nemnél lineáris pozitív kölcsönhatás van a két kimondás F_0 -átlaga és hangterjedelme között. A férfiaknál a regresszió értéke $r = 0,624$, a nőknél $r = 0,685$, (4. ábra). Ez erős lineáris kapcsolatot jelent a két változó között. Az alkalmazott modell segítségével tehát az első kimondás F_0 -átlagából és hangterjedelméből következtethetünk a második kimondás értékeire.



4. ábra

Lineáris kölcsönhatás az ismétlések első és második kimondásának F_0 -átlagértékei között

A magánhangzók első, második és harmadik formánsának elemzése választ ad arra, hogy módosul-e az artikulációs konfiguráció a két kimondás folyamán. Sem a férfiak, sem a nők ejtésében nem volt statisztikailag igazolható különbség az egyes magánhangzók között a szavak első és második kimondásában (3. táblázat), ez megegyezik az angol nyelvre kapott eredményekkel. A produkciós folyamat során az első és a második kimondás tervezése nem különül jelentősen el, az ismétlés feltehetően ugyanazon fonetikai terven belül valósul meg (Curl et. al. 2006).

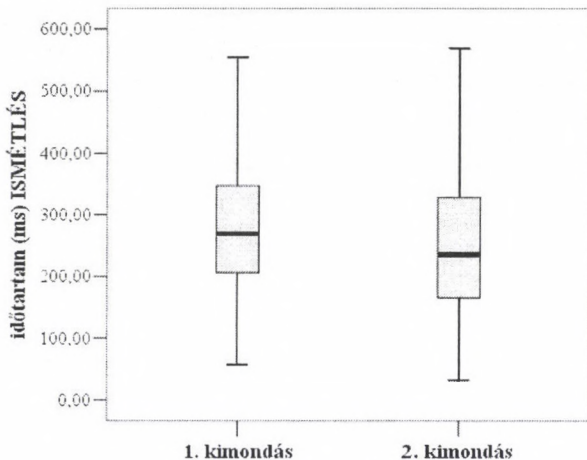
3. táblázat: Az ismétlésekben leggyakrabban előforduló magánhangzók formánsainak átlagértékei

		F ₁ (Hz) 1. ejtés	F ₁ (Hz) 2. ejtés	F ₂ (Hz) 1. ejtés	F ₂ (Hz) 2. ejtés	F ₃ (Hz) 1. ejtés	F ₃ (Hz) 2. ejtés
férfi	ɔ	580,87	562,74	1130,69	1123,06	2502,79	2520,09
	o	458,22	433,75	1119,89	1157,97	2412,97	2404,31
	ε	565,10	565,57	1611,43	1631,94	2509,59	2513,45
	i	382,17	387,56	2141,94	2104,39	2682,50	2681,83
nő	ɔ	652,87	628,62	1463,06	1458,78	2749,13	2711,59
	ε	654,80	606,63	1977,73	1966,57	2913,03	2900,37

Az újraindítások eredményei

Az újraindítás arra utal, hogy a beszélő a lexikális előhívás közben elbizonytalanodik. A szó kiejtése először csak részleges, majd rövid szünet után végbemegy a teljes szó kiejtése. Elemeztük, hogy a beszélő az első vagy a második kimondást valósítja meg hosszabban. Az ismétlésekéhez hasonló eredményeket kaptunk: az újraindítások 80,29%-ban a második, míg 19,71%-uknál az első kimondás volt a rövidebb. Az első kimondás átlagosan 259,76 ms (96–636 ms) volt (átlagos eltérés: 117,18 ms), a második ennél 23%-kal rövidebb: 199,82 ms (51–579 ms) (átlagos eltérés: 109,38 ms). A második kimondás időtartama szignifikánsan rövidebb [$F(2, 272) = 19,152$; $p = 0,000$], mint az elsőé (5. ábra). Ezek alapján feltételezhetjük, hogy a beszélők hibadejektálási és javítási folyamatai a tervezés során megfelelően működtek, a hibát a szerkesztési szakasz alatt korrigálták, ezért a második kiejtés már kevesebb időt vett igénybe.

Azokban az esetekben, amikor a második kimondás időtartama hosszabb volt (a példák 19,71%-a), mint az elsőé, az okozta az időtartam növekedését, hogy a beszélő közvetlenül az újraindítás előtt téves kezdési vagy téves szótaglálási hibát ejt. Emiatt nagyobb mértékben bizonytalanodik el, hogy ezúttal valóban a célszót aktiválta-e (4. példa). A második kimondás időtartamnövekedését okozhatja továbbá az is, hogy a beszélő még mindig bizonytalan a közlés tartalmában (5. példa).



5. ábra

Az első és a második kimondás időtartamértékei az újraindításoknál (szóródás és medián)

4. és mindegyik valami módon vigye- **vé-** (239 ms) (néma szünet: 143 ms) **védekezett** (vé = 271 ms)

A példában a beszélő elbizonytalanodását egy megelőző téves szókezdés okozza. Mivel a téves kezdés után hiányzik a szerkesztési szakasz, valószínűsíthető, hogy annak felismerése és javítása a rejtett önmonitorozás eredményeképpen megtörtént. A fonetikai hasonlóság miatt azonban a beszélő a javított alak helyességében kételkedik.

5. *unatkozó* (néma szünet: 290 ms) *kanadai tudósok* (néma szünet: 525 ms) *megállapították* (néma szünet: 357 ms) *aa* (néma szünet: 117 ms) *párhuzamot* (néma szünet: 571 ms) *az emberek bizonyos negatív tulajdonságai és* (néma szünet: 378 ms) *a növények* (néma szünet: 933 ms) *életviteli körülményei között* (néma szünet: 146 ms) *tehát ho-* (155 ms) (néma szünet: 130 ms) **hoggy** (ho = 177 ms) *fogalmazzam meg*

Ebben az esetben a beszélő azt a feladatot kapta, hogy minél részletesebben próbálja meg összefoglalni egy általa először hallott, ismeretlen szöveg tartalmát. A két-három szavanként elforduló néma szünetek egyértelműen jelzik a beszélő fogalmi tervezéssel kapcsolatos nehézségeit. Az adatközlő sem a közlés tartalmában, sem annak formájában nem biztos. Ez okozza a példában szereplő újraindítás második kimondásában tapasztalható időtartam-növekedést is. A beszélőnek ezek után még további összesen 2625 ms-os szünetre (egy 714 ms-os néma szünet után 537 ms-os hezitálás, majd 1373 ms-os néma

szünet következett) volt szüksége ahhoz, hogy gondolatait összerendezze, és megfelelő nyelvi formába öntse.

Elemeztük a szerkesztési szakaszok időtartamát és felépítését. Az újraindítások szerkesztési szakaszának időtartama átlagosan 152,41 ms volt. A legrövidebb szerkesztési szakasz 0 ms, a leghosszabb 1155 ms hosszúságú (átlagos eltérés: 249,03 ms). Az újraindítások 34,5%-ánál a szerkesztési szakasz hiányzott. Ezekben az esetekben biztosra vehető, hogy a beszélőnek még a tervezési folyamat részeként sikerült feloldania az elbizonytalanodás okául szolgáló diszharmoniót. Az újraindítások 78,17%-ában a szerkesztési szakasz 200 ms-nál rövidebb idő alatt realizálódott. Ebből arra következtethetünk, hogy az esetek többségében a beszélő a lexikális előhívással kapcsolatos bizonytalanságát nem a saját beszédének hallás alapú megértése alapján javítja, hanem a rejtetten működő folyamatok részeként.

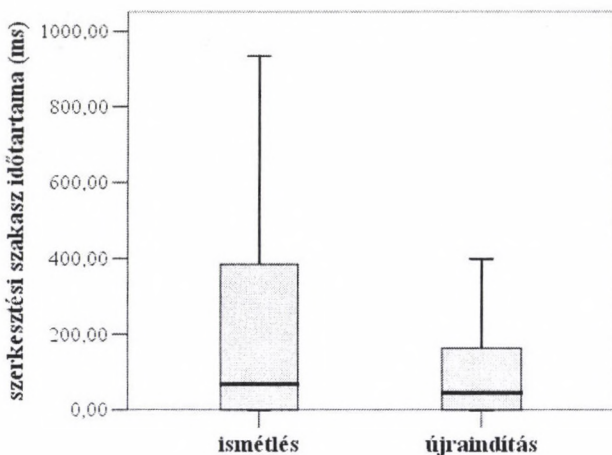
A szerkesztési szakasz felépítése az ismétlésekhez hasonlóan változó volt. Az újraindításoknál a részben kiejtett szót az adatok 84,95%-ában követte néma szünet [*akkor amikor nem rokon növé-* (néma szünet: 71 ms) *növények kerültek egymás mellé*]. Az újraindított szót 8,60%-ban előzte meg hezitálás [*mivel én mé-* öö (170 ms) *mérnök végzettségű vagyok*]; a hezitálás és a néma szünet kombinációja az adatok 6,45%-ára volt jellemző [*védekezett az ellen hogy a másik elvonja előle a tá-* (néma szünet: 118 ms) *öö* (285 ms) (néma szünet: 72 ms) *tüpléléket*] (4. táblázat).

4. táblázat: Az újraindítások szerkesztési szakaszának hossza

Szerkesztési szakasz	Átlagos hosszúság	Minimum	Maximum
Néma szünet	211,39 ms	12 ms	1155 ms
Hezitálás	185,12 ms	69 ms	351 ms
Néma szünet és hezitálás	577,00 ms	78 ms	1123 ms

A két megakadásjelenség (ismétlés és újraindítás) szerkesztési szakaszának összehasonlító elemzéséből kiderült, hogy a megakadásjelenség típusa meghatározó a szerkesztési szakasz hosszát illetően [egytényezős ANOVA: $F = (2, 824) = 9,669$; $p = 0,002$]. Az újraindítások során a beszélőnek csaknem feleannyi időre van szüksége a diszharmonia feloldásához, mint az ismétléseknél (6. ábra). Az előbbi esetben ugyanis az önmonitorozás magára az újraindítás tárgyát képező hangsorra irányul, a kiejteni kívánt szó tehát már nem csupán aktiválódott a mentális lexikonban, hanem az artikuláció közben szakadt meg. A kiejtett szótöredékről kell tehát eldöntenie a beszélőnek, hogy valóban a kívánt hangsor artikulációját kezdte-e meg. Az ismétlések során bonyolultabb hibajavítási folyamatokat kell feltételeznünk, amelyeknek következtében a szerkesztési szakaszok időtartama hosszabb lesz. A két megakadásjelenség között nem csak a szerkesztési szakasz időtartamának átlagában figyelhető meg nagy különbség, de az adatok szóródásában is. Az ismét-

léseknél az időtartamértékek nagyon tág tartományban realizálódnak. Ekkor – mint korábban írtuk – a beszélő bizonytalanságát kiváltó zavar a tervezési folyamat bármely szintjén bekövetkezhetett.



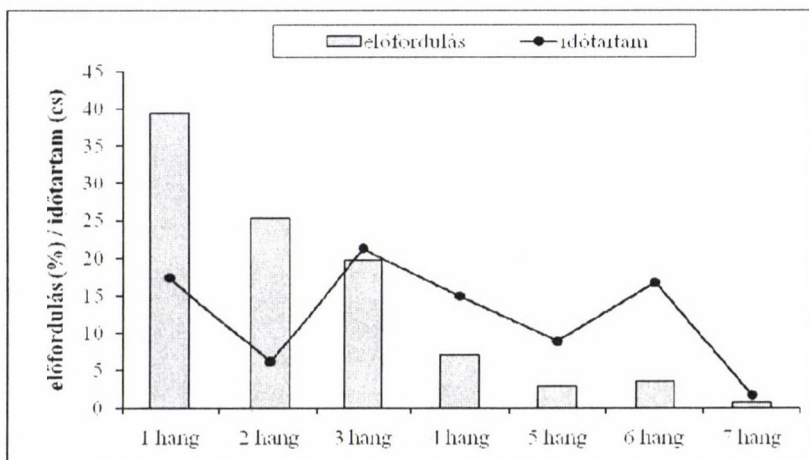
6. ábra

Az elemzett megakadásjelenségek követő szerkesztési szakasz tartama (medián és szóródás)

A beszélők önmonitorozási stratégiáira jellemző, hogy az újraindításkor hány beszédhang megvalósulása után állítják le, majd indítják újra az adott szó kiejtését. Az artikuláció leggyakrabban 1–3 beszédhang után szakad meg; 4, 5, 6 és annál több hang kiejtése után a beszélők már ritkán állítják le és indítják újra az adott hangsor artikulációját (7. ábra). Minél több hang valósul tehát meg a célszóból, a beszélő annál biztosabb benne, hogy a megfelelő szót aktiválta. Az adatok statisztikai elemzése azt igazolta, hogy az első kiejtésben megvalósult hangok száma összefügg a szerkesztési szakasz időtartamával [egytényezős ANOVA: $F(8, 141) = 2,292$; $p = 0,025$]. Egyetlen hang megvalósulása esetén a szerkesztési szakasz átlagos időtartama 174,30 ms (átlagos eltérés: 279,32 ms), két hang esetén pedig 62,14 ms (átlagos eltérés: 98,47 ms). A szerkesztési szakasz időtartama még több kiejtett beszédhang esetén nem rövidül.

A funkciósók és a tartalmas szavak aránya az újraindításoknál másként alakul, mint az ismétléseknél. Funkciósókat 59,87%-ban, tartalmas szavakat 40,13%-ban indítanak újra a beszélők. Ez az arány gyakorlatilag megfelel a két kategória spontán beszédbeli gyakoriságnak (Szende 1973, Keszler 1983). Azért, hogy igazoljuk, hogy ez az előfordulási arány nem véletlenszerű, ezúttal is binominális nemparaméteres tesztet készítettünk. Nullhipotézi-

sünk (H_0) szerint a funkciószavak és a tartalmas szavak egyenlő arányban jelennek meg az újraindításokor. A teszt a H_0 szerinti gyakoriság (0,5) mellett, $655/705 = 0,93$ tapasztalati gyakoriságot igazolt, ami egyoldali szignifikanciát eredményezett ($p = 0,0175$). Megállapítható tehát, hogy csakúgy, mint az ismétléseknél, itt sem véletlenszerű az, hogy az újraindítások nagyobb mértékben a funkciószavakat érintik. Ahhoz ugyanis, hogy a beszélő megfelelően tudjon dönteni az adott funkciószó helyességéről, már rendelkeznie kell a teljes közlés(rész) grammatikai szerkezetével.



7. ábra

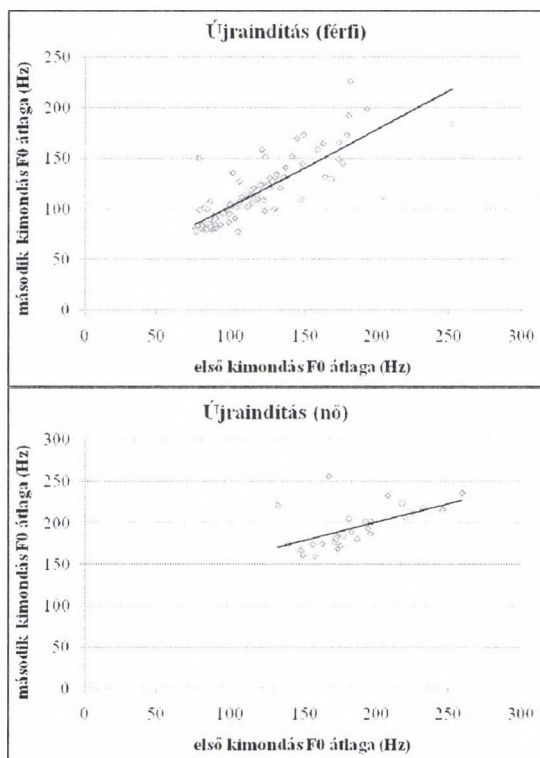
A kiejtett beszédhangok száma újraindításokor és a szerkesztési szakasz időtartama

Az alaphangmagasság a nőknél az ismétlésekhez hasonlóan az újraindításoknál is csekély mértékben magasabb volt a második kimondáskor, a férfiaknál nem volt kimutatható alaphangmagasság-változás (5. táblázat). Az újraindítások 8,45%-ánál találtunk érdes zöngét, de csak a férfiak ejtésében. 41,6%-ban mindkét kimondás, 33,3%-ban a második, 25%-ban pedig az első esetben vált aperiodikusá a zöngékepzés.

5. táblázat: A férfiak és a nők alaphangmagassága az újraindításokban

	1. kimondás		2. kimondás	
	F ₀ -átlag	Átlagos eltérés	F ₀ -átlag	Átlagos eltérés
Férfi	105,03 Hz	43,14 Hz	105,66 Hz	41,42 Hz
Nő	185,28 Hz	31,89 Hz	194,76 Hz	24,72 Hz

Az alaphang hangterjedelmének változása a két nemnél eltérően alakult. A nők hangterjedelme a második kimondáskor szűkülte, a férfiaké nem változott. Az újraindítások első kimondásának F_0 -átlaga és hangterjedelme, valamint a második kimondásának F_0 -átlaga és hangterjedelme közötti összefüggés igazolására többváltozós lineáris regressziót alkalmazunk. Az eredmények arra engednek következtetni, hogy az újraindítások esetében mindkét nemnél lineáris pozitív kölcsönhatás van a két kimondás F_0 -átlaga és sávszélessége között. A férfiak újraindításainál a regresszió értéke $r = 0,624$, míg a nőknél $r = 0,638$ (8. ábra). Ez azt jelenti, hogy a két változó között erős lineáris kapcsolat áll fenn. Az általunk alkalmazott modell segítségével az első kimondás F_0 -átlagából és hangterjedelméből következtethetünk a második kimondás értékeire.



8. ábra

Lineáris kölcsönhatás az újraindítások első és második kimondásának F_0 -átlagértékei között

A magánhangzók első, második és harmadik formánsának elemzésével kapott adatok szerint egyik magánhangzó esetében sem találtunk szignifikáns különbséget az újraindításokban (6. táblázat). Arra következtethetünk ennek alapján, hogy az újraindítás során ugyanaz a fonetikai terv valósul meg az első és a második kimondáskor. Az artikuláció megszakadásakor a beszélőnek egy egyszerű döntést kell hoznia, hogy a megfelelő szó kiejtését kezdte-e meg. Ha a kérdésre a válasz igen, megtörténik a szó teljes kimondása, és ehhez az artikulációs tárban már meglévő hangsor lehívása történik meg. A célszó fonetikai terve az ellenőrzés ideje alatt ugyanis valószínűsíthetően „vára-kozik” az artikulációs tárban. Amennyiben az ellenőrző kérdésre az önmonitorozó rendszer nemleges választ ad, a tervezési folyamat újraindul, és végbemegy a célszó aktivációja és kimondása, ez azonban már a téves kezdés esete.

6. táblázat: A leggyakoribb magánhangzók formánsainak átlagértékei az újraindításokban

		F ₁ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₃ (Hz)
		1. ejtés	2. ejtés	1. ejtés	2. ejtés	1. ejtés	2. ejtés
férfi	ɔ	598,64	567,54	1134,50	1160,18	2533,32	2524,04
	ε	550,67	542,89	1626,89	1657,78	2483,33	2503,56
nő	ɔ	713,21	655,71	1394,50	1430,57	2806,71	2869,01

Következtetések

Kutatásunkban a magyar spontán beszédben előforduló – a beszélő bizonytalanságát jelző – ismétléseket és újraindításokat elemeztük. Az eredmények alapján választ kaptunk arra, hogy mely akusztikai fonetikai paraméterek változnak meg a második kimondás során, továbbá a temporális eredményekből következtethetünk az önmonitorozó mechanizmus működésére. A két megakadásjelenség időviszonyainak elemzése igazolta azt, hogy a második kimondás rövidebb időt vesz igénybe. A szerkesztési szakasz időtartama tehát vélhetően elegendő időt biztosít a beszélőnek a zavar elhárítására, ennek következtében a második kimondás időtartama lerövidül. Ez jelzi az önjavítási stratégiák sikerességét. A két kimondás közötti szerkesztési szakasz időtartama az önellenőrzés rejtett vagy felszíni voltára utal. Az ismétlések szerkesztési szakasza átlagosan csaknem kétszer olyan hosszú, mint az újraindításoké. Ez a tény, illetőleg az, hogy az újraindításoknál nagyobb a 200 ms-nál rövidebb szakaszok előfordulási aránya, arra enged következtetni, hogy az ismétlés ideje alatt a beszélőnek a tervezési folyamat valamely szintjén előállt hibát kell korrigálnia. Az újraindításoknál az első kimondásban a célszóból csupán néhány hang valósul meg. Minél hosszabb a szórészlet, a beszélő annál biztosabb abban, hogy a kívánt szót aktiválta.

Mindkét elemzett megakadásjelenség túlnyomórészt funkciószókat érintett, ez a jelenség nem véletlenszerű. Az ismétlések esetében ez töltelékelem-funkcióra utal. Az újraindításoknál a célszó megfelelőségét ellenőrzi a beszélő; a tartalmas szavaknál elsősorban szemantikai tekintetben, a funkciószóknál a közlésrészlet teljes grammatikai szerkezetének ellenőrzése is szükséges.

Az ismétlések és az újraindítások első és második kimondásában nem változott a beszélők alaphangmagassága. A férfiak ismétléseiben a második kimondás F_0 -ja nemegyszer érdes zöngével valósult meg. Többváltozós lineáris regresszióval sikerült igazolnunk, hogy az első kimondás F_0 -átlagából és hangterjedelméből mindkét vizsgált megakadásjelenség esetében következteni lehet a második kimondás értékeire.

A formánsértékek sem az ismétléseknél, sem az újraindításoknál nem változtak jelentős mértékben. Ez azt bizonyítja, hogy az újraindításoknál a már kész és félig kiejtett szóra vonatkozó fonetikai terv az ellenőrzési és javítási folyamat végéig az artikulációs tárban „várakozik”; az ismétlések esetében pedig a tervezési folyamatban nem választható külön a két rész tervezése, hanem az egyetlen fonetikai terv részeként jön létre.

Irodalom

- Bata Sarolta – Gráczy 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 András – Szaszák György 2010. Automatic recognition of schwa variants and stress syllables in spontaneous Hungarian speech. *Acta Linguistica Hungarica* Vol. 57. (megjelenésre elfogadva).
- Benkenstein, Ramona – Simpson, Adrian P. 2003. Phonetic correlates of self-repair involving word repetition in German spontaneous speech. In Eklund, Robert (ed.): *Gothenburg Papers in Theoretical Linguistics 90. Proceedings of DiSS '03. Disfluency in Spontaneous Speech Workshop, 5-8 September 2003*. Göteborg University, Göteborg, 81–84.
- Boersma, Paul – Weenink, David 2009. Praat: doing phonetics by computer (Version 5.1.07) [Computer program]. <http://www.praat.org> (A letöltés ideje: 2009. július. 20.)
- Bóna Judit 2006. A megakadásjelenségek akusztikai és percepcióis sajátosságai. *Beszédkutató* 2006. 101–114.
- 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.
- Bóhm Tamás – Ujváry István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutató* 2008. 108–121.
- Clark, Herbert H. – Wasow, Thomas 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.

- Curl, Traci – Local, John – Walker, Gareth 2006. Repetition and the prosody–pragmatics interface. *Science Direct Journal of Pragmatics* 38. 1721–1751.
- Dabóczy Viktória 2008. Korrektur der Versprecher in der deutschen und ungarischen gesprochenen Sprache. In Boszák, Gizella (Hg.). *Wissenschaften im Dialog. Studien aus dem Bereich der Germanistik. Band 3.* Partium Verlag, Klausenburg, Großwardein, 53–68.
- Fox, Barbara - Jaspersen, Robert 1995. A syntactic exploration of repair in English conversation. In Davis, Philip W. (ed.): *Alternative linguistics: Descriptive and theoretical modes.* Current Issues in Linguistic Theory. Vol. 102. Benjamins, Amsterdam, 77–134.
- Fox Tree, Jean E. 1995. The effect of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34. 709–738.
- Gósy Mária 2002. A megakadásjelenségek eredete a beszédprodukción tervezési folyamatában. *Magyar Nyelvőr* 126. 192–204.
- Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2004a. *Fonetika, a beszéd tudománya.* Osiris Kiadó, Budapest.
- Gósy Mária 2004b. A magyar spontán beszéd megakadásainak hallás alapú gyűjteménye. *Beszéd kutatás 2004.* 6–19.
- Gósy Mária 2005. *Pszicholingvisztika.* Osiris Kiadó, Budapest.
- Gósy Mária 2008a. Önellenőrzési folyamatok a beszédben. *Magyar Nyelv* 104/4. 402–426.
- Gósy Mária 2008b. Magyar spontánbeszéd-adatbázis – BEA. *Beszéd kutatás 2008.* 194–208.
- Gósy Mária – Gyarmathy Dorottya 2008. A nyelvhasználati változás egy jelensége. *Magyar Nyelvőr* 132. 206–222.
- Gósy Mária 2009. Szójelentés és szóészlelés: változatosság és adaptálódás. *Beszéd kutatás 2009.* 46–75.
- Garnham, Alan – Shillcock, Richard C. – Brown, Gordon D. A. – Mill, Andrew I. D. – Cutler, Ann 1981. Slips of the tongue in the London-Lund corpus of spontaneous conversation. *Linguistics* 19. 805–817.
- Gyarmathy Dorottya 2006. A beszédpercepció és beszédprodukción folyamat összefüggései a megakadásjelenségek tükrében. In Heltai Pál (szerk.): *MANYE XVI. Tanulmánykötet.* Vol. 3/2. 449–455.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukción. *Beszéd kutatás 2007.* 108–121.
- Halpern, Jack 2006. The contribution of lexical resources to natural language processing of CJK languages. In *ISCSLP 2006.* 768–780.
- Horváth Viktória 2004. Megakadásjelenségek a párbeszédokban. *Beszéd kutatás 2004.* 187–200.
- Horváth Viktória 2007. Vannak-e „női” és „férfi” megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131. 315–323.
- Horváth Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben.* Doktori disszertáció. ELTE, Budapest.

- Keszler Borbála 1983. Kötetlen beszélgetések mondat és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szöveg-tana köréből*. Tankönyvkiadó, Budapest, 164–202.
- 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*. A Bradford Book. Cambridge, Massachusetts.
- 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 Academic Publishers, Dordrecht, 403–439.
- Lutz, K. C. – Mallard, A. R. 1986. Disfluencies and rate of speech in young adult nonstutterers. *Journal of Fluency Disorders* 11. 307–316.
- Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszéd kutatás 2004*. 209–223.
- Markó Alexandra 2006. Megakadásjelenségek a beszédprodukción és a beszédészlelésben. In Heltai Pál (szerk.): *MANYE XVI. Tanulmánykötet Vol. 3/2*. 96–102.
- Marslen-Wilson, William 1990. Activation, competition, and frequency in lexical access. In Altman, Gerry (ed.): *Cognitive models of speech processing*. MIT Press, Cambridge, 148–172.
- Nooteboom, Sieb 2005. Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication* 47. 43–48.
- Poupplier, Marianne – Hardcastle, William J. 2005. A re-evaluation of the nature of speech errors in normal and disordered speakers. *Phonetica* 62. 227–244.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszéd kutatás 2004*. 200–209.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Xie, Huayang – Andreae, Peter – Zhang, Mengjie – Warren, Paul 2004. Detecting stress in spoken English using decision trees and Support Vector Machines. *ACSW Frontiers 2004*. 145–150.

A kutatást az OTKA 78315 számú pályázata támogatja.

MEGAKADÁSOK A SZINKRONTOLMÁCSOK SPONTÁN BESZÉDÉBEN

Bakti Mária

Bevezetés

A spontán beszéd produkciója az a folyamat, amely a megszólalás szándékától a kiejtésig tart. A spontán beszédnek számos változata létezik, ezeket két csoportba lehet osztani, a félspontán (az angol szakirodalomban *extemporaneous speech*) beszéd során a beszélő felkészül a beszédre, míg a valódi spontán beszéd (*impromptu speech*) során „semmilyen beszédtervezés nem előzi meg az aktuális közlést” (Gósy 2005: 72).

A spontán beszéd definiálásának egy másik lehetséges megközelítése a beszélt és az írott nyelv közötti különbségek megállapítása. Erre a feladatra az 1960-as, 1970-es években több kutató is vállalkozott. Heeschén az alábbi lényeges különbségeket sorolja fel a területen végzett kutatások eredményei alapján: a beszélt nyelv gyakrabban használja a mellérendelő szerkezeteket, és gyakoribbak benne a deiktikus utalások, mint a határozott névelővel bevezetett főneves szerkezetek. A beszélt nyelvben gyakoribbak az aktív szerkezetek és a kevésbé bonyolult mondatnapi szerkezetek (Heeschén 1993).

Heeschén azt is megjegyzi, hogy a későbbiekben paradigmaváltás ment végbe a kutatásban, és a korábbi merev írott nyelv – beszélt nyelv elhatárolást felváltotta az a megközelítés, mely szerint a szövegeket egy kontinuum mentén lehet elhelyezni, melynek végpontjai az írott és a beszélt nyelv (1993).

A spontán beszéd tervezési folyamatainak jobb megértésében segít a megakadásjelenségek kutatása, mert a beszédprodukció olyan rejtett folyamatokra utalnak, amelyek a hibátlan beszéd során nem vizsgálhatóak. A megakadások definíciója Gósy szerint a következő: „a spontán beszéd artikulációs illetve percepciósi folyamatosságát megakasztó, különféle jelenségek” (2004: 7). A megakadásjelenségeket a bizonytalanságból adódó jelenségekre és téves kivitelezés jelenségeire lehet felosztani. A bizonytalanságból adódó megakadások a következők: néma szünetek, hezitálások, nyújtások, újramezések, ismétlések és a töltelékszavak. A téves kivitelezés jelenségei: freudi elszólás, morfológiai, szintaktikai hibák, kontamináció, téves szótalálás, „nyelvem hegyén van” jelenség, változtatás, újraindítás, sorrendiségi problémák, egyszerű nyelvbottlások (Gósy 2005). A hiba típusú megakadások közé tartoznak a nyelvbottlások, ezek egyik első definícióját adja Boomer és Laver: a nyelvbottlás nem szándékos eltérés a beszélő folyamatban lévő fonológiai,

nyelvtani vagy lexikai szándékától (1973: 123). Huszár szerint a „nyelvbottlásnak az adott nyelvet első nyelvként beszélő, mentálisan egészséges, általános intelligenciájú [...] felnőtt ember – szándéka ellenére létrejövő, saját fonológiai, morfológiai, szemantikai stb. normatudatának meg nem felelő – hibás beszédprodukcióját nevezzük (2005: 15). A spontán magyar beszédben előforduló megakadásokat vizsgálta többek között Gósy (2003), Horváth (2004), Markó (2004) és Szabó (2004).

A tolmácsolás kutatói közül többen a szinkrontolmácsok beszédprodukcióját is spontán beszédnek tekintik, mivel az „itt és most” születik (Enkvist 1982, Pöchlhacker 1995). A szinkrontolmácsolás és a spontán beszéd kapcsolatát számos kutató vizsgálta, különösen az 1960-as, 1970-es években, ekkor elsősorban pszichológusok és pszicholingvisták kutatták a szinkrontolmácsolás folyamatát. Henri C. Barik (1972, 1973) kutatásaiban hangsúlyozta, hogy a szinkrontolmácsolás során fontos változók a bejövő szöveg típusa és előadásmódja. Kísérletében hat tolmács (2 profi, 2 tanuló, 2 tapasztalatlan) négy típusú szöveget tolmácsolt: spontán beszédet, félíg előkészített beszédet, előre megírt beszédet, és egy újságcikk felolvasott változatát. A célnyelvi szövegek elemzésekor Barik arra az eredményre jutott, hogy a szinkrontolmácsolt szövegek temporális jellemzői alapján véve azonosak a forrásnyelvi szövegek vagy a természetes beszéd temporális jellemzőivel. Érdekes módon azonban a tolmácsok több időt töltenek beszéddel, mint a forrásnyelvi szövegek előadói, amit Barik a tolmácsok lassabb artikulációs tempójának tulajdonít. A tolmácsoknál azért lassúbb az artikuláció, mivel a tolmácsok többet hezitálnak, mert az előadóra várnak, hogy kifejtse a forrásnyelvi szöveg tartalmát. Összességében Barik a szinkrontolmácsolást *spontaneous mode of expression*, vagyis spontán kifejezésmódnak tekinti (Barik 1972).

Goldman-Eisler a szinkrontolmácsolást a spontán beszéd egy típusának tekintette (Goldman Eisler 1968). A *Segmentation of input in simultaneous translation* című cikkének bevezetőjében a spontán beszéd és a szinkrontolmácsolás kapcsolatáról szólva megjegyzi, hogy a szinkrontolmácsolás a spontán beszédhez hasonló, mivel a tolmácsnak saját szavaival kell elmondania a forrásnyelvi üzenetet a célnyelven. Bizonyos értelemben a szinkrontolmács feladata jobban hasonlít az átvitelre vagy a parafrázisra, mivel a tolmács feladatának része az, hogy „saját szavaival mondja el a forrásnyelvi üzenetet” (Goldman-Eisler 1972: 128). Déjean-ra utalva Seleskovich (1982) a spontán beszéd és a szinkrontolmácsolás kapcsolatáról szólva megjegyzi, hogy a szóbeli fordítás, azaz a tolmácsolás mindig spontán (252).

Kopczynski a szinkrontolmácsok célnyelvi beszédprodukcióját félspontán beszédnek tekinti, a fogalmat így definiálja: „olyan beszéd, amely a helyszínen születik, egy korábban ismeretlen szöveg alapján” (1982: 257). A szerző azt is hangsúlyozza, hogy a szinkrontolmácsolás, mivel szóban történik, rendelkezik a valódi spontán beszéd jellegzetességeivel is. Ezeket a jellemzőket Kopczynski három csoportba sorolja. Az első csoportba azok a jelenségek

tartoznak, amelyek eltérést jelentenek a forrásnyelvi szöveg folyamatos visszaadásától: hezitálás, újraindítás, ismétlés, téves szótalálás. Másodszorban a tolmácsolt szöveg szerkezete „lazább”, több benne a mellérendelés, kevesebb a beágyazott szerkezet, a tolmácsok gyakran elhagyják az alárendelő mellékmondatokat. Harmadszor pedig a tolmácsolt szövegek általában kevésbé formálisak, mint a forrásnyelvi megfelelőjük.

Összegzőképpen Kopczynski (1982) elmondja, hogy a mivel a szinkrontolmácsolás szóbeli tevékenység, itt és most történik, ezért rendelkezik a spontán beszéd több jellemzőjével is. Ennek kihatása lehet a tolmácsolás minőségének megítélésére is, különösen akkor, ha a forrásnyelvi szöveg nem spontán, hanem szerkesztett beszéd.

Shlesinger MA-dolgozatában azt vizsgálta, hogy a szinkrontolmácsolás során hogyan változik a szövegek szerkesztettségi foka. A dolgozat címe: *Simultaneous interpretation as a factor in effecting shifts in the position of texts on the oral-literate continuum*. A dolgozat héber nyelven született, angol nyelvű összefoglalását Pym (2007) adja.

A szövegek szerkesztettségét Shlesinger a tervezettség foka, a hallgató és az előadó közös tudása, a szókincs és a hallgatóság bevonásának mértéke alapján vizsgálta. A fenti paraméterek alapján került meghatározásra a forrásnyelvi szövegek szerkesztettségi foka, vagyis elhelyezkedésük a szóbeli/spontán–szerkesztett/formális skálán.

Shlesinger hipotézise az volt, hogy a szinkrontolmácsolás során változik a szövegek szerkesztettségi foka: a szerkesztett forrásnyelvi szövegekből spontánabb célnyelvi szövegek születnek, a spontán forrásnyelvi szövegekből pedig formálisabb célnyelvi szövegek születnek majd.

Az 1. táblázat Shlesinger és Pym alapján mutatja, hogy milyen jellemzőket tulajdonít Shlesinger a skála két végpontján elhelyezkedő szövegeknek.

Shlesinger hipotézise kézenfekvőnek látszik, hiszen a szinkrontolmácsolási gyakorlat alapján logikus, hogy a forrásnyelvi szövegben előforduló hezitálásokat a tolmács kihagyja, így időt takarít meg, és az ilyen kihagyások a szerkesztett szövegek felé vezetnek. Amikor viszont a tolmácsoknak szerkesztett szöveget kell tolmácsolniuk, például egy előre megírt beszédet vagy jogi szövegeket, akkor a célnyelvi szöveg az eredetinel spontánabb lesz, mivel ez az írott szöveg szóbeli változata.

Shlesinger eredményei az összes szerkesztett forrásnyelvi szöveg esetében azt mutatták, hogy célnyelvi megfelelőjük sokkal spontánabb lett, a skálán a spontán beszéd jellemzői felé mozdultak el. A spontán forrásnyelvi szövegek esetében azonban nem ilyen egyszerű a helyzet, itt az eredmények vegyesebb képet mutatnak. Úgy tűnik, hogy a spontán forrásnyelvi szövegek a fordítás során még spontánabbak lettek. Összegzőképpen elmondható, hogy Shlesinger eredményei alátámasztják a szinkrontolmácsolási beszédprodukción alapvetően beszélt, spontán természetét (Pym 2007).

1. táblázat: A formális és a spontán szövegek jellemzői Shlesinger és Pym nyomán (Pym 2007)

	Formális szövegek	Spontán szövegek
1. Tervezettség foka		
1.a) Lexikai sűrűség	nagyobb lexikális sűrűség; nominalizáció; alárendelések	töredezett szintakszis; mellérendelés
1.b) Szövegkohézió	erős szövegkohézió	gyenge szövegkohézió; az előadó feltételezi a kontextus ismeretét
1.c) Megakadások	–	több a redundancia, szünet és ismétlés; az előadó figyel arra, hogy a hallgató megértse az üzenetet
2. Közös tartalom és tudás	explicittebb a szemantikai utalások területén	feltételezik, hogy a címzett ismeri a kontextust
3. Szókincs	több kohéziós eszközt használ	mindennapi nyelvhasználat; innovatív
4. A hallgatóság bevonásának mértéke	–	minél jobban hangsúlyozza a kommunikáció résztvevői közötti kapcsolatot, annál spontánabb; nonverbális jellemzők; töredezett prozódia

Korábbi kutatásaink alapján (Bakti 2007a, Bakti 2007b, Kusztor és Bakti 2007) feltételezzük egy szinkrontolmácsolás-specifikus megakadásjelenség-mintázat meglétét. A szinkrontolmácsolás során előforduló megakadásokat vizsgáló tanulmányában Pöchlhammer (1995) azonban bizonyos megakadások előfordulását azzal magyarázza, hogy a szinkrontolmácsok célnyelvi beszédprodukciója a spontán beszéd egy formája. A jelen dolgozat célja, hogy pontosítsa a tolmácsolás kutatói által kialakított képet, és összevesse a szinkrontolmácsok spontán és célnyelvi beszédprodukciójában előforduló megakadásokat, és az eredmények segítségével alátámassza a szinkrontolmácsolás-specifikus megakadás-mintázat létezését. A hipotézisünk az volt, hogy különbségek lesznek a két szituációban előforduló megakadások között, mivel a szinkrontolmácsolás során párhuzamosan zajlik a beszédprodukció, beszédpercepció és a fordítás folyamata. A szinkrontolmácsolás esetében a beszédprodukció tervezésének különös esetével van dolgunk, mivel a szinkrontolmácsok kívülről megadott program alapján dolgoznak; a forrásnyelvi tartalom és az illokúciós szándékot a forrásnyelvi beszélő határozza meg és ő ren-

dezi sorba a közölni szándékozott információkat is, a szinkrontolmácsok ennek alapján tervezik meg a célnyelvi beszédprodukción.

Anyag, módszer, kísérleti személyek

A vizsgálatban négy profi szinkrontolmács vett részt (T1, T2, T3, T4), akik több éve dolgoznak szinkrontolmácsként. Közülük ketten az Európai Unió intézményeinek szabadúszó tolmácsai, ketten pedig Magyarországon dolgoznak szinkrontolmácsként, A nyelvük magyar, B nyelvük angol. A kísérletben egy nő (33) és három férfi (31, 33, 40) vett részt.

A tolmácsok egy előre rögzített nyolc perc hosszúságú angol beszédet tolmácsoltak angolról magyarra. Az előre rögzített szöveg a brit belügyminiszter beszéde volt, az egyházak és vallási csoportok társadalmi szerepéről szólt. A beszédet egy magyar anyanyelvű személy olvasta fel, a felolvasott szöveget rögzítettük, majd a szinkrontolmácsok ezt a szöveget tolmácsolták angolról magyarra. A felvételek 2007. januárjában készültek Szegeden, a Szegedi Tudományegyetem Irinyi épületében, az I. Nyelvi Laborban, TANDBERG EDUCATIONAL Media Centre IS-10 MM felszereléssel, Sony HF 60 típusú kazettákra. A rögzített anyag hossza 36 perc. A szinkrontolmácsolási helyzet során a szinkrontolmácsok nem látták az előadót, és egyedül voltak a tolmácsfülkében, nem számíthattak a tolmácskolléga segítségére. Erre azért volt szükség, hogy minden résztvevő számára egyenlő feltételeket tudjunk biztosítani.

Később a tolmácsokat arra kértük meg, hogy egy fordítási feladat után értékeljék teljesítményüket, majd köznapi témákban feltett kérdésekre válaszoltak. A tolmácsok spontán beszédprodukción rögzítettük, a hangfelvételeket ártírtuk, majd az előforduló megakadásokat Gósy (2004) alapján csoportosítottuk. Az eredményeket összehasonlítottuk a tolmácsok célnyelvi beszédprodukción előforduló megakadásokkal.

Eredmények

A 2. táblázat a megakadások százalékos eloszlását mutatja a szinkrontolmácsok spontán magyar beszédében. A táblázat adatai jól mutatják az egyéni eltéréseket a szinkrontolmácsok spontán beszédprodukción. T1 tolmács esetében nagyon kevés megakadás fordult elő, a beszédprodukción végig szabatos, ellenőrzött volt. T2 tolmács esetében volt a legmagasabb a megakadások előfordulási aránya, kiemelkedő az újraindítások és az ismétlések gyakori előfordulása. T3 tolmács esetében viszonylag kevés megakadás fordult elő, leggyakoribbak az ismétlések voltak. T4 tolmács esetében viszonylag magas volt a megakadások előfordulási gyakorisága, esetében a kiesések, az újraindítások és az ismétlések fordultak elő a leggyakrabban.

Összességében elmondhatjuk, hogy a vizsgálatban szereplő profi szinkrontolmácsok spontán beszédében előforduló leggyakoribb megakadások az ismétlések és az újraindítások.

2. táblázat: A megakadások százalékos eloszlása a szinkrontolmácsok spontán magyar beszédében

Megakadás	T1	T2	T3	T4
Morfológiai, szintaktikai hiba	–	7%	–	–
Változtatás	40%	–	–	12%
Újraindítás	20%	26%	17%	22%
Betoldás	–	–	–	11%
Kiesés	–	–	–	22%
Ismétlés	40%	60%	66%	22%
Többféleképpen osztályozható jelenség	–	7%	17%	11%
Megakadás/100 szó	1,37	5,22	2,37	2,7

Az ismétlés bizonytalanságra visszavezethető jelenség, a tervezés korai szintjéhez kapcsolódó megakadás. Habár az ismétlések minden tolmács spontán beszédében előfordultak, mégis nagyon élesek az egyéni különbségek a beszélők között, az ismétlések aránya kiemelkedően magas T2 tolmács esetében. Az első példa egy ismétlést mutat T2 tolmács spontán beszédéből.

(1) *De végül is ezek ezek indiánok voltak, ezek az emberek...* (T2)

Általánosságban véve az ismétlések a kötőszókat vagy a többi töltelékszót érintik (Gósy 2005), ez jellemezte a tolmácsok spontán beszédprodukciónak is; a (2) példa is ezt mutatja T3 tolmács beszédprodukciónak:

(2) *hogya hogya van szöveg és van legalább öt perc az embernek arra, hogy átolvassa* (T3)

A tolmácsok spontán beszédprodukciónak előfordultak újraindítások is, azonban számuk viszonylag csekély, kivéve T2 tolmácsot, akinél ez az egyik leggyakrabban előforduló megakadás. A (3) példában a beszélő egy egész kifejezést indít újra, míg a (4) példában a beszélő az első szótág indítja újra a szót.

(3) *Ööö de úgy é _ úgy éreztem, hogy ő hogy bár a tartalom sérült* (T2)

(4) *am _ amikor ezekkel a problémákkal szembetalálkoztam, ő általában nem éreztem mégsem,* (T2)

Hasonló az (5) példa is, T4 tolmács beszédprodukciónak:

(5) *mert csak azt tudtam, hogy min _ minden nap bemondják, hogy* (T4)

A tolmácsok spontán beszédprodukciónak egy beszélő esetében (T2) fordult elő morfológiai szintaktikai hiba.

A (6) példában a mondatból hiányzik az ige, azonban mind a beszélő, mind a hallgató számára érthető az üzenet:

- (6) *ő általában nem éreztem mégsem, hogy a javításnak ő fontos szerepét*
(T2)

A változtatás, vagyis a kiejtett szó szándékos javítása kontextusba illő szavakkal, viszonylag ritkán fordult elő a gyűjteményben. Változtatást csak T1 és T4 tolmácsok beszédprodukciónak találtunk, a (7) példa T4 tolmács spontán beszédéből mutat változtatást:

- (7) *vannak ilyen visszatérő öö motívumok, ilyen ismétlődő struktúrák* (T4)

Tanulságos a szinkrontolmácsok spontán beszédében és célnyelvi beszédprodukciónak, angolról magyarra tolmácsoló szövegekben előforduló megakadások összehasonlítása. A 3. táblázat a kísérletben szereplő szinkrontolmácsok célnyelvi beszédprodukciónak előforduló megakadások gyakoriságát mutatja.

3. táblázat: Megakadások százalékos eloszlása a szinkrontolmácsok célnyelvi beszédprodukciónak, Bakti (2007b) alapján

Megakadás	T1	T2	T3	T4
Morfológiai, szintaktikai hiba	36,5%	17,5%	20,0%	20,0%
Kontamináció	–	–	7,0%	–
Téves szókezdés	–	6,0%	7,0%	–
Változtatás	45,5%	11,7%	26,0%	10,0%
Újraindítás	–	41,3%	33,0%	50,0%
Sorrendiségi hibák	9,0%	6,0%	–	10,0%
Csere	–	–	–	10,0%
Kiesés	–	–	7,0%	–
Ismétlés	9,0%	17,5%	–	–
Megakadás/100 szó	1,6	2,4	1,9	1,5

Az eredmények nagy egyéni eltéréseket mutatnak, azonban bizonyos tendenciák meghatározhatóak az adatok alapján. T1 tolmács kivételével a tolmácsok célnyelvi beszédprodukciónak kevesebb megakadás fordul elő. Ennek feltételezett magyarázata az lehet, hogy a szinkrontolmácsolás során a tolmácsok kívülről megadott program szerint dolgoznak, ahol a makrotervezés nagy részét a célnyelvi beszélő végzi.

A leggyakoribb hibatípusok a grammatikai, morfológiai hibák, a változtatások és az újraindítások. A grammatikai, morfológiai hibák a spontán beszédben előforduló morfológiai, szintaktikai hibáktól eltérően a leggyakrabban az egyeztetési tévedések, amelyeket a fül-száj ívhossz nagysága, vagyis a forrásnyelvi és a célnyelvi szöveg közötti időbeli különbség és a rendelkezésre álló mentális energia mennyisége vagy a hiánya magyaráz. Míg a spontán beszédben a grammatikai, morfológiai hibákat főleg a hiányos szerezetek jellemezték, például az ige hiánya a mondatban, a tolmácsok esetében sokszor

maradt le vagy nem volt megfelelő a toldalék. Ennek oka a forrásnyelvi és a célnyelvi szöveg között eltelt idő hosszában kereshető. A (8) példa egy több szót érintő grammatikai hibát mutat, ahol a tolmács, valószínűleg a fül-száj ívhossz nagysága miatt, kénytelen menet közben változtatni a már kiejtett szerkezetén.

(8) *a nagy kihívás tehát azt jelen, az jelenti, hogy*

A (9) példában a tolmács nem megfelelő végződést használ:

(9) *a hit szerepe fontos az emberek életébe*

A változtatások mind a spontán, mint a tolmácsolt szövegekben előfordulnak, a tolmácsolt szövegekben azonban gyakoribbak voltak, minden tolmács beszédprodukciónak megtalálhatóak. A változtatás a szinkrontolmácsolt beszédprodukciónak jellemző eleme. Ennek számos oka lehet. A változtatást tekinthetjük időnyerési stratégiának, amikor a tolmácsok átfogalmaznak a már kiejtett nyelvi elemeket, de az is előfordulhat, hogy a tolmácsok a mentális lexikonból elsősre nem a kontextusnak megfelelő szót választják ki, így szükség van a változtatásra. A szinkrontolmácsolás során olyan esetek is előfordulhatnak, amikor a célnyelvi normáknak való megfelelés miatt dönt a változtatás mellett a tolmács.

A (10) példában a tolmács valószínűleg nem tudta aktiválni a 'mecset' szót a mentális lexikonból, először körülírta (muzulmán templom), majd sikerült megtalálnia a megfelelő szót:

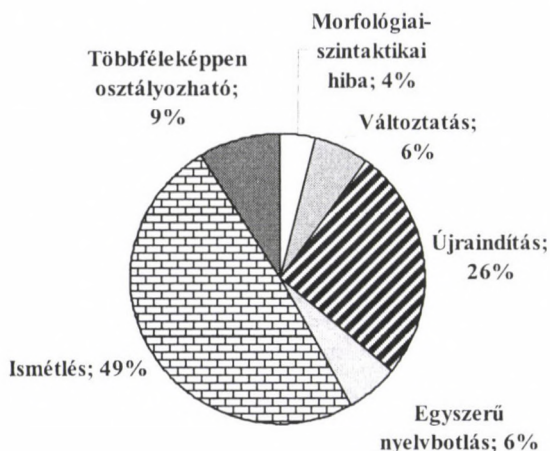
(10) *vannak ő olyan muzulmán templomok is olyan mecsetek*

A (11) példában a tolmács az eredeti *focal point* kifejezést fordítja le szó szerint magyarra, majd ezután talál egy valamivel magyarosabb megfelelőt:

(11) *ezenkívül fókuszpontot biztosítanak, központi pontot biztosítanak a közösség számára*

Az újraindítások három tolmács beszédprodukciónak jellemzők; előfordulásuk jóval gyakoribb, mint a tolmácsok spontán beszédében. Az újraindítás a lexikális hozzáférés és az artikulációs tervezés összehangolatlanságából ered. Gyakori előfordulására magyarázat lehet az, hogy a tolmácsok beszédzajban dolgoznak; Gósy kutatásai szerint a beszédzajban beszélők beszédprodukciónak is kiemelten jellemzők az újraindítások (Gósy 2008).

Az 1. ábra a tolmácsok spontán beszédprodukciónak előforduló megakadásokat mutatja. A tolmácsok spontán beszédében a leggyakoribb megakadások az ismétlések és az újraindítások, míg a tolmácsolt szövegekben a leggyakrabban az újraindítások, a morfológiai-szintaktikai hibák és a változtatások. A beszédprodukción ezen formájában az ismétlések aránya viszonylag alacsony.



1. ábra

A tolmácsok spontán beszédében előforduló megakadások

Összegzés

Eredményeink különbségeket mutatnak a szinkrontolmácsok spontán beszédében és célnyelvi beszédprodukciójában előforduló megakadások között. A különbségeket a beszédprodukció két típusa között a beszédprodukció körülményei magyarázhatják. A szinkrontolmácsolt szövegekben alacsonyabb volt a megakadások előfordulási aránya, mivel a szinkrontolmácsok előre megadott program alapján dolgoznak, ahol a beszédtervezés során adott, hogy a forrásnyelvi beszélő milyen információt szeretne közölni a hallgatónak, és a forrásnyelvi beszélő meghatározza a kifejezni szándékozott információ sorrendjét is. A tolmács a gondolatok sorrendjét átrendezheti és az ő feladata a megfelelő pragmatikai forma megtalálása a célnyelven.

A szinkrontolmácsolt szövegekben a leggyakrabban a morfológia-szintaktikai hibák, a változtatások és az újraindítások fordulnak elő, míg a spontán beszédben az ismétlések és az újraindítások voltak a leggyakoribbak. Az újraindítások a zajban történő beszédprodukció jellemzői, azaz a szinkrontolmácsolás során a forrásnyelvi szöveg zajhatása okozhatja az újraindítások magas arányát. A tolmácsok spontán beszéde lényegesen nagyobb egyéni eltéréseket mutatott, mint a célnyelvi beszédprodukciójuk. Az eredmények alátámasztják azt a korábbi megállapítást (Bakti 2008), hogy a szinkrontolmácsolási beszédhelyzet a célnyelvi beszédprodukcióban tolmácsolás-specifikus megakadás-mintázat kialakulásához vezet.

A szinkrontolmácsolás és a spontán beszéd közötti különbségek jobb megértését segíthetné elő a szinkrontolmácsok felpontán beszédének elemzése,

ahol a tervezési mechanizmusok jobban hasonlítanak a szinkrontolmácsolási beszédhelyzetre.

Irodalom

- Bakti Mária 2007a. Interferenciális nyelvbtlások a szinkrontolmácsolásban. In Heltai Pál (szerk.): *Nyelvi modernizáció. Szaknyelv, fordítás, terminológia*. MANYE Vol. 3. 625–629.
- Bakti Mária 2007b. Retrospektív interjúk a szinkrontolmácsolás kutatásában. In Váradi Tamás (szerk.): *I. Alkalmazott Nyelvészeti Doktorandusz Konferencia Előadásai*. MTA Nyelvtudományi Intézet, Budapest 12–23.
- Bakti Mária 2008. Megakadásjelenségek a szinkrontolmácsolásban. *Fordítástudomány* X/2. 22–38.
- Barik, Henri C. 1972. Interpreters talk a lot, among other things. *Babel*, 18/1. 3–10.
- Barik, Henri C. 1973. Simultaneous interpretation: Qualitative and linguistic data. *Language and Speech* 16/3. 237–270.
- Boomer, Donald S. – Laver, John 1973. Slips of the tongue. In Fromkin, Victoria A. (ed.): *Speech errors as linguistic evidence*. Mouton, The Hague, 120–131.
- Enkvist, Nils E. 1982. (ed.) *Impromptu speech: A symposium*. Abo Akademi, Abo.
- Goldman-Eisler, Frieda 1968. *Psycholinguistics: experiments in spontaneous speech*. Academic Press, London.
- Goldman-Eisler, Frieda 1972. Segmentation of input in simultaneous translation. *Journal of Psycholinguistic Research* 1/2. 127–140.
- Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar nyelvőr* 127. 257–277.
- Gósy Mária 2004. A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye. *Beszéd kutatás 2004*. 6–18.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2008. A zaj hatása a beszédre. *Beszéd kutatás 2008*. 5–21.
- Heeschen, Claus 1993. Morphosyntactic characteristics of spoken language. In Blanken, Gerhard et al. (eds.): *Linguistic disorders and pathologies. An international handbook*. Walter de Gruyter, Berlin–New York, 16–34.
- Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszéd kutatás 2004*. 187–199.
- Huszár Ágnes 2005. *A gondolatól a szóig. A beszéd folyamata a nyelvbtlások tükrében*. Tinta Kiadó, Budapest.
- Kusztor Mónika – Bakti Mária 2007. Megakadásjelenségek németre és angolra szinkrontolmácsolott szövegekben. *Beszéd kutatás 2007*. 121–133.
- Kopczynski, Andrzej 1982. Effects of some characteristics of impromptu speech on conference interpreting. In Enkvist, Nils E. (ed.): *Impromptu speech: A symposium*. Abo Akademi, Abo, 255–266.
- Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszéd kutatás 2004*. 209–222.
- Pöschhacker, Franz 1995. Slips and shifts in simultaneous interpreting. In Tommola, Jorma (ed.): *Topics in interpreting research*. The University of Turku, Centre for Translation and Interpreting, Turku, 73–90.

- Pym, Anthony 2007. On Schlesinger's proposed equalizing universal. In: Pöhhacker, Franz – Jakobsen, Arnt Lykke – Mees, Inger (eds.): *Interpreting studies and beyond: A tribute to Miriam Schlesinger*. Samfundslitteratur, Fredriksberg.
- Seleskovitch, Danica. 1982. Impromptu speech and oral translation. In Enkvist, Nils E. 1982. (ed.): *Impromptu speech: A symposium*. Abo Akademi, Abo, 241–253.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszédkutatás* 2004. 200–208.

HALLÁSALAPÚ ÉS VIZUÁLIS ALAPÚ KÖZLÉSEK

Váradi Viola

Bevezetés

A beszédpercepció folyamatát két nagy szakaszra oszthatjuk: először megértjük, majd értelmezzük a hangjelenségeket. Az első nagy szakaszban megtörténik a beszédhangok és a hangkapcsolatok észlelése, felismerése, a másik nagy szakaszban a szavak, a mondatok, a szöveg megértése zajlik. Mindkettő több szintből épül fel, melyek együttműködése biztosítja a hallott beszédjelenségek megértését. A beszédmegértés tehát komplex folyamat, amelyben a beszéd nem válik szét értelmes és értelem nélküli egységekre, illetve szegmentális és szupraszegmentális részre. A folyamat működésének jelentős részéről azonban ismereteink még hiányosak. A megértési folyamatokat modellekben próbálják ábrázolni, amelyekben egyszerűsítve jelennek meg az egyébként rendkívül összetett és bonyolult működéssorozatok. A ma ismert beszédpercepció modellek legjelentősebb része a teljes megértési folyamatot próbálja ábrázolni, de vannak olyanok is, amelyek külön foglalkoznak a beszéd észlelésével, illetve megértésével. Meglehetősen sok modell ismeretes – motoros teória, analízis szintézissel vagy aktív-passzív modell, beszédmegértés globális modellje, Bondarko-féle elmélet, hierarchikus megértési modell, kognitív modellek –, amelyek kisebb-nagyobb mértékben különböznek egymástól (Gósy 2005).

Hagyományos megfogalmazás szerint az olvasás a leírtak által vezetett gondolkodás, vagyis vizuális alapú közlés (Imre 2006). Nagyon sok olvasási modell létezik, melyek mind a kognitív pszichológiából erednek. A modelleket három csoportba oszthatjuk: 1. az agy megismerő folyamatainak alapuló modellek (bottom-up, top-down, interaktív modellek); 2. kettős kódolású modell és tranzakciós elmélet; 3. a kognitív folyamatokat szociális kontextusban vizsgáló modellek (A. Jászó 2006). Abban minden modell egyetért, hogy az olvasás egyfajta dekódolási készség: a leírt szavakat kell beszélt, kiejtett szavakká transzformálni. Ez a folyamat rendkívül összetett, és két fő részre oszthatjuk: dekódolás és megértés. A dekódolási folyamat során megfejtjük a betűsorok alapján a vizuális élményt, a betűket megfeleltetjük a megfelelő fonémákkal, illetve elkülönítjük, szegmentáljuk a szavakat. A megértés során felismerjük a szegmentált szó morfológiai struktúráját, és azonosítjuk a jelentését. Az agy bizonyos folyamatait vizuálisan megjelenítő képalkotó eljárások azt mutatják, hogy más és más területek aktiválódnak a bal agyféltekében attól függően, hogy az olvasás hangos (felolvasás) vagy néma. Hangos olva-

sáskor az aktiválódás a beszédfeldolgozó területen, míg néma olvasásnál inkább a vizuális feldolgozó területen jelentkezik (Gósy 2005).

A nemzetközi és a hazai szakirodalomban mind a beszédpercepció, mind az olvasás folyamatával számos kutatás foglalkozott és foglalkozik napjainkban is, de a kétféle működéssorozat összehasonlítására vonatkozó kutatások magyar nyelven nem ismertek. A jelen kísérletben a hallásalapú és vizuális alapuló közlések néhány akusztikai sajátosságát – időtartam, szünetezés, tempó – hasonlítjuk össze. Feltételezésünk szerint a kétféle ingeren alapuló közlések produkciója során alapvetően más-más folyamatok játszódnak le az agyban. Ezekről a folyamatokról közvetlenül nem, de közvetetten, a beszédprodukció vizsgálatával kaphatunk információt. Hipotézisünk szerint a beszélőknek nagyobb nehézséget okoz a hallás utáni visszamondás, mint a felolvasás, mivel egy időben, szimultán több folyamat játszódik le ekkor az agyban. Felolvasáskor csak a meghangosítást kell megtervezni, a mondatokat nem kell előhívni a memóriából. Ekkor abból fakadhatnak a nehézségek, hogy rosszul hangosítja meg a mondatot a beszélő, javítja a szupraszegmentális szerkezetet, esetleg félreolvassa a mondatot. Feltételezésünk szerint a bonyolultabb működéssorozatnak köszönhetően a beszélőknek hosszabb időre van szüksége a mondatok meghangosításához, ezért lassabb lehet az artikulációs és beszédtempójuk a hallásalapú közlésekben.

Számtalan tényező befolyásolhatja beszédünk szupraszegmentális szerkezetét, például a kor, a beszédképző szervek állapota, a fizikai-fiziológiás állapot, a beszéd típusa. A nemzetközi és a magyar szakirodalomban is egyre több kutatás foglalkozik a beszédpartnerhez való alkalmazkodással, mint befolyásoló tényezővel (Bata-Grácsi 2009). Hipotézisünk szerint a mondatok utánmondása során hatással van egymás beszédprodukciójára a mondatokat felolvasó beszélő, és aki visszamondja a mondatokat. Feltételezésünk szerint a felolvasó lassabban, érthetőbben próbál artikulálni, megfelelően, értelem-szerűen tagolja a mondatokat, hogy ezzel segítse az utánmondó percepcióját. A mondatai ezért hosszabbak lesznek, és lassabb lesz az artikulációs és beszédtempója.

Kísérleti személyek, anyag és módszer

Az MTA Nyelvtudományi Intézet Fonetikai Osztályán 2007 őszén indult meg a beszélt nyelvi adatbázis (BEA) fejlesztése, majd 2008-ban megkezdődtek a BEA-t kialakító beszédfelvételek. Az összehasonlíthatóság érdekében nemcsak spontán beszédanyagok felvételére került sor, hanem felolvasásokat és ismétléseket is tartalmaz az adatbázis (Gósy 2008).

A jelen kísérletben a mondatfelolvasásokat – vizuális alapú közlések – és a visszamondásokat – hallásalapú közlések – elemeztük: az adatközlőknek 20 egyszerű, illetve 5 összetett mondatot (pl. *A nagymama specialitása kétséget kizáróan a barackbefőtt. Nem kötött biztosítást, ezért kisebb vagyona került a kórházi ellátás.*) kell először, első feladatként a felvételvezető után megis-

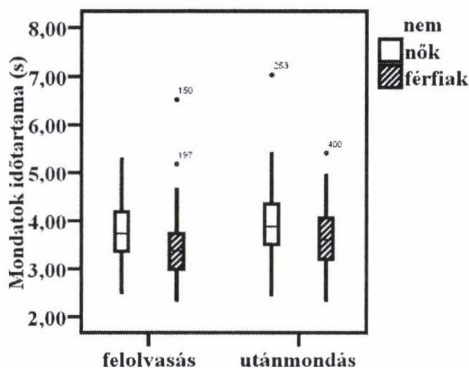
mételniük, majd legvégül, a további feladatok után fel kell olvasniuk ezeket a mondatokat. A 25 mondat átlagosan 8 szóból (szóródás: 5–11), 48 hangból (szóródás: 38–58), 52 betűből (szóródás: 41–62) áll. A kísérlethez 5-5 fiatal férfi és nő hanganyagát választottuk ki véletlenszerűen. Az adatközlők átlagos életkora 25 év (a nők 26 év, szóródás: 22–29 év; a férfiaké 24 év, szóródás: 20–30 év).

A hanganyagok feldolgozásához a Praat 4.3 szoftvert használtuk, ezzel végeztük el az annotálást és a méréseket. Az adatközlők és a felvételvezető felolvasásában, valamint az adatközlők mondatutánmondásában megmértük a mondatok hosszát, a szünetek időtartamát és előfordulását, az artikulációs és beszédtempót, illetve elemeztük a megakadásjelenségeket. A kapott adatokat a nemek szempontjából is értékeltük. A statisztikai elemzéseket 95%-os szignifikanciaszinten az SPSS 15.0 szoftverrel végeztük el.

Eredmények

A mondatok időtartama

Elemeztük a mondatok időtartamát a felolvasásban és az utánmondásban (1. ábra). A hallásalapú közléseknél hosszabbak a mondatok, mint a vizuális alapú közléseknél. Az egytényezős varianciaanalízis szerint a különbség szignifikáns [$F(1, 498) = 9,765; p < 0,001$].



1. ábra

A mondatok időtartama felolvasáskor és utánmondáskor nemenként (szóródás és medián)

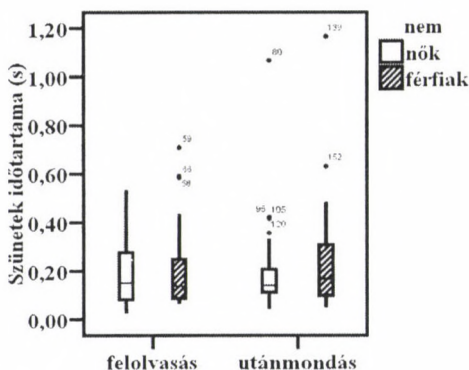
Amikor hallás után kell visszamondani a mondatokat, akkor meg kell jegyezni, memorizálni kell őket, majd az adatközlőnek a megszólalással egy időben fel kell idéznie a hallott mondatot. Ehhez a gondolkodási folyamathoz több időre lehet szüksége, ebből fakadhat a mondatok időtartama között mért

szignifikáns különbség a két feladattípus között. Megvizsgáltuk a mondatok időtartamát nemek szerint csoportosítva is arra a kérdésre keresve a választ, hogy a nem befolyásolja-e az időtartamviszonyokat. Az 1. ábrán láthatjuk, hogy mind a mondatok felolvasása, mind hallás utáni megismétlése során a nők mondatai hosszabbak. A különbségek a statisztikai elemzések (egytényezős varianciaanalízis) szerint szignifikánsak [felolvasás: $F(1, 248) = 19,703$; $p < 0,001$; utánmondás: $F(1, 248) = 17,010$; $p < 0,001$].

A mondatok időtartamát több tényező is befolyásolja: a szünetek gyakorisága, időtartama, az artikulációs és beszédtempó; a megakadásjelenségek. A nemek közti különbségek okainak feltárásához szükséges, hogy a fenti befolyásoló tényezők elemzésénél összehasonlítsuk a férfiak és a nők produkcióját.

Szünetek, megakadásjelenségek

Elemeztük a felolvasott és az utánmondott mondatokban a szünetek előfordulásainak gyakoriságát és időtartamát (2. ábra). A két feladattípus között nincs jelentős különbség a szünetek előfordulásának tekintetében. A nők mindkét feladattípusban gyakrabban tartanak szünetet, mint a férfiak (felolvasás: nők átlagosan 11 db, férfiak 4 db; utánmondás: nők átlagosan 12 db, férfiak 4 db).



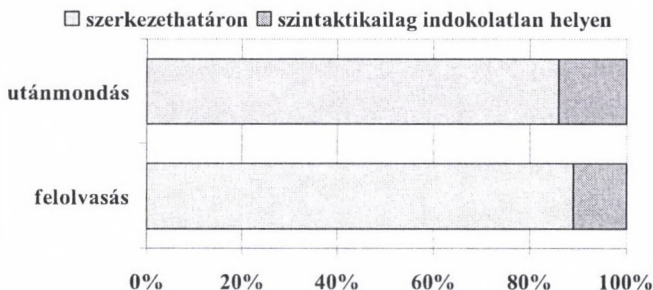
2. ábra

A szünetek időtartama felolvasáskor és utánmondáskor nemenként (szóródás és medián)

A felolvasott és az utánmondott mondatok szünetidőtartamai közötti különbség a statisztikai elemzés szerint nem szignifikáns [egytényezős varianciaanalízis; $F(1, 154) = 0,031$; $p = 0,861$]. Elemeztük a szünetek időtartamainak eltéréseit mindkét feladattípusban a nemek szempontjából is, de a statisztikai elemzések szerint egyik esetben sem szignifikáns az eltérés a nők és a

férfiak által tartott szünetek időtartamaiban [egytényezős varianciaanalízis; felolvasás: $F(1, 74) = 0,832$; $p = 0,832$; utánmondás: $F(1, 56) = 1,539$; $p = 0,220$]. A szünetek időtartamát és gyakoriságát nem befolyásolta szignifikánsan, hogy az adatközlőnek leírt vagy hallott mondatokat kellett-e meghangsúlyoznia, illetve hogy az adatközlő nő vagy férfi.

Elemeztük mindkét feladattípusban, hogy hol tartanak szünetet az adatközlők (3. ábra): szerkezet határán vagy szintaktikailag indokolatlan helyen (pl. a jelző és a jelzett szó között). Utánmondáskor a szünetek 89%-a, míg felolvasáskor 86%-a szerkezethatáron fordul elő. A szintaktikailag indokolatlan helyen előforduló szünetek feltehetően bizonytalanságból fakadó megakadásjelenségek, az adatközlőnek időre van szüksége ahhoz, hogy felidézzé, amit hallott. A szerkezethatáron előforduló szünetek nem sorolhatóak egyértelműen a megakadásjelenségek közé, ezeknek inkább tagoló funkciója van.



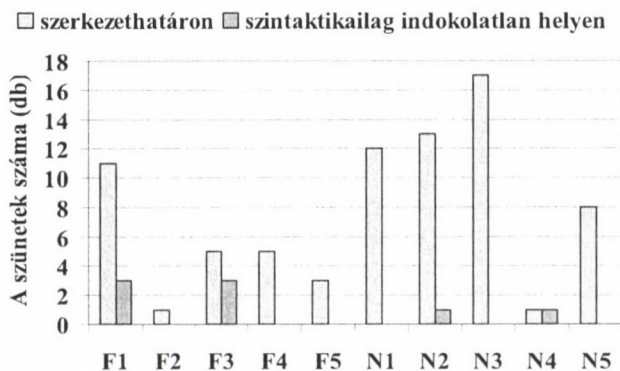
3. ábra

A szünetek előfordulása felolvasáskor és utánmondáskor

A 4. ábrán láthatjuk, hogy 8 adatközlőnél felolvasáskor a szerkezetek határán fordul elő leggyakrabban szünet. Szintaktikailag indokolatlan helyen egy adatközlő többször tart szünetet, mint szerkezethatáron, és egy adatközlőnél egyszer szerkezet határán, egyszer szintaktikailag indokolatlan helyen fordul elő szünetet. Szintaktikailag indokolatlan helyen csak 4 adatközlő tart szünetet, viszont ebből két férfi a 25 mondatból háromban is. Egy férfi adatközlőnél a szünetek háromnegyede, az egyik női adatközlőnél a szünetek fele szintaktikailag indokolatlan helyen fordul elő. A mondatok egyötöde összetett, ezért megvizsgáltuk, hogy a tagmondathatároknál milyen gyakran fordul elő szünet. A tízből csak három adatközlő tart szünetet minden tagmondathatáron felolvasáskor.

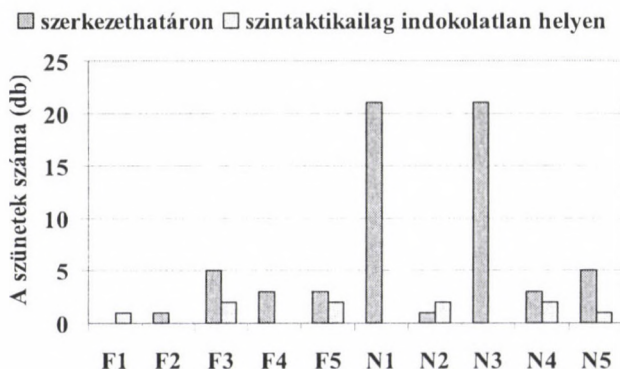
Az adatközlők közül nyolcan szerkezethatáron, ketten szintaktikailag indokolatlan helyen tartanak leggyakrabban szünetet utánmondáskor (5. ábra). A legtöbb szünetet tartó női beszélő (N3) utánmondásában fordul csak elő minden összetett mondat tagmondathatárán szünet, de 4 adatközlőnél egyszer

sem. Az adatközlőknek több mint a fele (6 fő) tart legalább egyszer szünetet szintaktikailag indokolatlan helyen, ami azt mutatja, hogy nehézségeik voltak a produkció során, mivel emlékezetből kellett meghangosítaniuk a mondatot. Feltételezésünk szerint azért gyakoribb a szünettartás felolvasáskor a tagmondatok határán, mert ez írásban vesszővel jelölve van az írott szövegben.



4. ábra

A szünetek előfordulása felolvasáskor adatközlőnként



5. ábra

A szünetek előfordulása utánmondáskor adatközlőnként

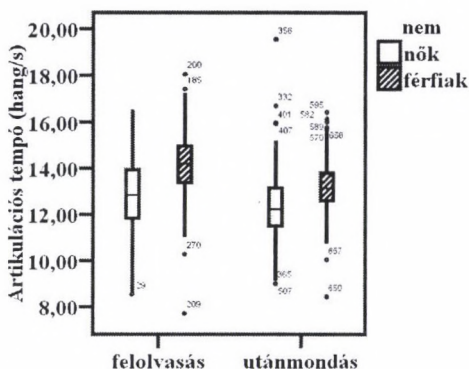
Mind a felolvasásokban, mind az utánmondásokban előfordulnak megakadásjelenségek, amelyek mindig valamilyen nehézségre utalnak a beszédtervezés, illetve a beszédprodukció során (Gósy 2004a). A 10 adatközlő utánmondáskor 12-szer, felolvasáskor 8-szor akadt meg. A vizuális alapú közlé-

sekben 4, a hallásalapú közlésekben 5 adatközlőnél fordul elő olyan egyéb megakadásjelenség, amely nem szünet. Az előfordulások alacsony száma miatt messzemenő következtetéseket nem vonhatunk le, de megfigyelhető az a tendencia, hogy utánmondáskor gyakrabban akadnak meg a beszélők.

Tempóértékek

A beszéddel foglalkozó tudományokban kétféle tempóértéket különböztetünk meg, a kettő között az alapvető különbség az, hogy a beszédtempó mérésénél figyelembe vesszük a szüneteket és a megakadásjelenségeket is a jelidő részeként, míg az artikulációs tempó mérésénél nem. Az artikulációs tempó a beszélő artikulációjának sebességét mutatja meg, és mindig gyorsabb, mint a beszédtempó (Gósy 2004b).

A felolvasás és az utánmondás artikulációs tempói (6. ábra) között mért különbség a statisztikai elemzés szerint szignifikáns [egytényezős varianciaanalízis; $F(1, 664) = 38,48$; $p < 0,001$]. Adatközlőink 80%-a átlagosan gyorsabban artikulál felolvasás során, mint amikor hallás után kell megismételnie a mondatokat. Az adatközlőknek a mondatok visszamondásakor folyamatosan koncentrálniuk kell – főleg a bonyolultabb szerkezetű, ritkább szavakat tartalmazó mondatok esetén – arra, hogy helyesen ismételjék meg a mondatokat. A munkamemória állandó igénybevételének, illetve az eltérő dekódolási folyamatoknak köszönhető ez az eltérés.



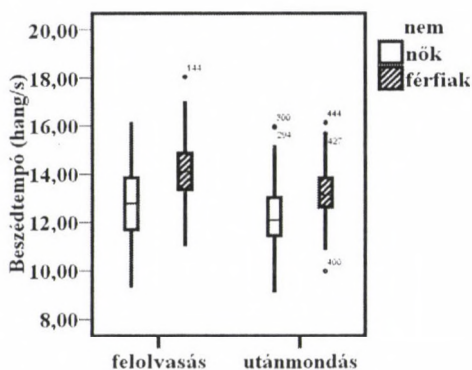
6. ábra

Artikulációs tempóértékek a mondatok felolvasásában és utánmondásában nemenként (szóródás és medián)

Összehasonlítottuk az artikulációs tempóértékeket mindkét feladattípusban nemenként is. A 6. ábrán láthatjuk, hogy mind felolvasáskor, mind visszamondáskor a férfiak artikulációs tempója gyorsabb. Az egytényezős varianciaanalízis szerint mindkét feladattípusban szignifikáns a különbség [felolva-

sás: $F(1, 324) = 66,680$; $p < 0,001$; utánmondás: $F(1, 335) = 35,514$; $p < 0,001$].

Adatközlőink 90%-ának nemcsak az artikulációs tempója, hanem a beszédtempója is gyorsabb akkor, amikor felolvassa a mondatokat, mint amikor hallás után ismétli meg (7. ábra). Az egytényezős varianciaanalízis alapján a különbség szignifikáns [$F(1, 498) = 31,200$; $p < 0,001$]. Összehasonlítottuk mindkét feladattípusban a férfiak és a nők beszédének tempóját is. A férfiaknak nemcsak az artikulációs tempója, hanem a beszédtempója is gyorsabb, mint a nőknek. A statisztikai elemzések szerint a különbségek szignifikánsak [egytényezős varianciaanalízis; felolvasás: $F(1, 248) = 69,159$; $p < 0,001$; utánmondás: $F(1, 248) = 39,029$; $p < 0,001$].



7. ábra

Beszédtempóértékek felolvasáskor és utánmondáskor nemenként (szóródás és medián)

Az adatközlőknek az artikulációs és a beszédtempója is gyorsabb a vizuális alapú közlésekben, mint a hallásalapúakban. A mondatok utánmondása során nemcsak a meghangosítás minőségére kell koncentrálnia az adatközlőknek, hanem fel kell idézniük, majd pontosan meg kell ismételniük az egyszerű, illetve összetett mondatokat. A tartalom felidézése egy időben, szimultán zajlik a szupraszegmentális szerkezet tervezésével, ami nehézséget okozhat az adatközlőknek. Feltételezésünk szerint emiatt lassabban artikulálnak.

A hazai szakirodalomban Gocsál (2001) férfiak és nők spontán beszédében vizsgálta, hogy befolyásolja-e a nem a beszédünk tempóját. 30 fővel végzett kutatásában arra a megállapításra jutott, hogy a női beszélők nem beszélnek gyorsabban a férfiaknál. A szünettartás értékeinek összehasonlításakor talált szignifikáns eltérést a két nem között. A jelen kutatás eredményei alapján azonban úgy tűnik, hogy a férfiak és a nők artikulációjának és beszédének

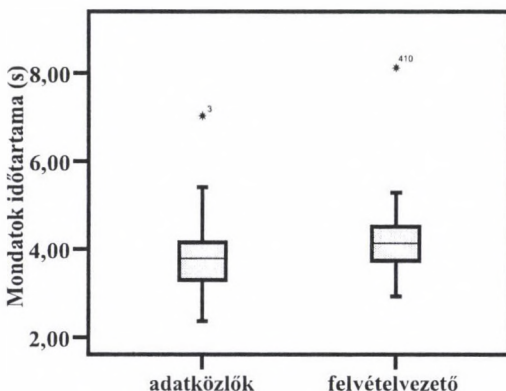
tempója között vannak eltérések – nem spontán beszédben. Néhányan úgy vélik, hogy a nők gyorsabban beszélnek, mint a férfiak, a jelen kutatás – mely nem spontán beszédben elemezte a tempóértékeket – eredményei azonban cáfolják ezt a hétköznapi megfigyelést.

A felvételvezető és az adatközlők beszédprodukcójának viszonya

A BEA-felvételek során utánmondáskor a felvételvezetője olvassa fel az adatközlőknek a mondatokat, amelyeket meg kell ismételniük. Ez a felvételi eljárás lehetőséget ad arra, hogy megvizsgáljuk, hatással vannak-e egymásra a felvételvezető és az adatközlők beszédének egyéni sajátosságai.

Összehasonlítottuk a felvételvezető és az adatközlők mondatainak időtartamát, illetve az artikulációs és a beszédtempó-értékeket. A Pearson-féle korrelációs számítás szerint a felvételvezető és az adatközlők mondatainak hosszúsága, artikulációs tempója és beszédtempója közepesen és szignifikánsan korrelál egymással (mondathossz: $r = 0,656$; $p < 0,001$; artikulációs tempó: $r = 0,354$; $p < 0,001$; beszédtempó: $r = 0,373$; $p < 0,001$). Az eredmények azt mutatják, hogy hasonlítanak egymásra a felvételvezető és az adatközlők beszédprodukcójának artikulációs sajátosságai. Feltételezésünk szerint a felvételvezető a felolvasása során figyelembe veszi, hogy mekkora nehézséget okozhat a feladat teljesítése az adatközlőknek, akik a hallott mondatoknak nemcsak a szegmentális, hanem a szupraszegmentális jellemzőit is memorizálják.

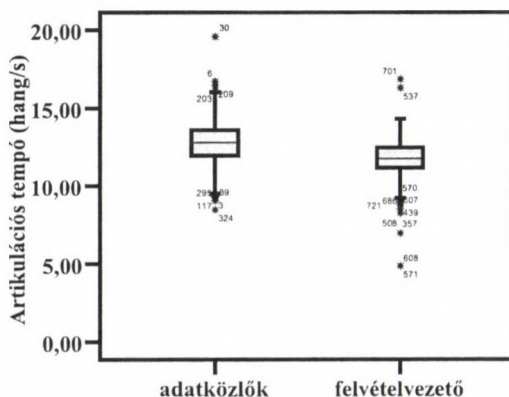
Összehasonlítottuk a felvételvezető és az adatközlők mondatainak időtartamát (8. ábra), és az egytényezős varianciaanalízis szerint a különbség szignifikáns [$F(1, 198) = 40,246$; $p < 0,001$]. A felvételvezető mondatai szignifikánsan hosszabbak, mint az adatközlőké.



8. ábra

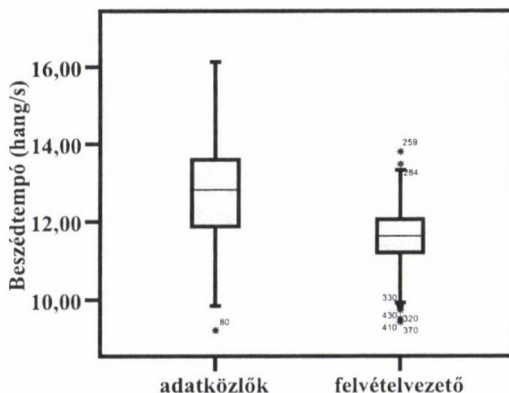
Az adatközlők és a felvételvezető mondatainak időtartama
(szóródás és medián)

Az adatközlők artikulációs tempója és beszédtempója (9. és 10. ábra) az egytényezős varianciaanalízis szerint szignifikánsan gyorsabb, mint a felvételvezetőé [artikulációs tempó: $F(1, 730) = 118,661$; $p < 0,001$; beszédtempó: $F(1, 498) = 153,190$; $p < 0,001$].



9. ábra

Az adatközlők és a felvételvezető artikulációs tempója (szóródás és medián)



10. ábra

Az adatközlők és a felvételvezető beszédtempója (szóródás és medián)

A mondatok felolvasása során a felvételvezetőnek az a célja, hogy az adatközlőkkel visszamondassa a mondatokat. A felolvasónak – a jelen kísérletben

a felvétel rögzítőjének – érdekében áll, hogy segítse a hallgató beszédfeldolgozását. A lassabb artikulációs és beszédtempó, a mondatok felolvasására szánt több idő ennek köszönhető.

Következtetések

A dekódolásbeli különbségek a vizuális és a hallásalapú közlések között megmutatkoztak az adatközlők beszédében. A jelen kísérletben négy beszéd-jellemzőt vizsgáltunk és hasonlítottunk össze a két feladattípusban: a mondatok időtartamát, a szünetek előfordulását és időtartamát, az artikulációs tempót és a beszédtempót. A szünetek előfordulását és időtartamát nem befolyásolja, hogy felolvassa-e a mondatokat az adatközlő, vagy hallás után mondja-e vissza. A mondatok időtartama azonban hosszabb utánmondás esetén, és lassabban beszélnek ekkor az adatközlők. A dekódolásbeli különbségek az artikulálás tempójában jelennek meg, és ennek köszönhető, hogy a beszédtempó és a mondatok időtartama is változik. A hallásalapú közlések során a memóriakapacitás terhelésének hatására lelassul az adatközlők artikulálása, aminek következtében a beszédtempó is lassabb lesz, és megnő a mondatok időtartama. A felvételvezető és az adatközlők mondatainak időtartama, az artikulációs és a beszédtempója közepesen és szignifikánsan korrelál egymással. Feltételezésünk szerint a felvételvezető próbál alkalmazkodni az adatközlők percepciójának feltételezett tempójához, illetve az adatközlők a hallott szupraszegmentális sajátosságokat is memorizálhatják a mondatok szemantikai, szintaktikai és szegmentális sajátosságaival együtt.

Az adatközlők kiválasztása során az egyetlen szempont, amit figyelembe vettünk az volt, hogy egyenlő arányban legyenek a férfiak és nők. Az eredmények elemzése során feladattípusonként külön-külön összehasonlítottuk a produkciójukat. A férfiak mondatai rövidebbek, gyorsabb az artikulációs és beszédtempójuk, mint a nőknek. Ez az eltérő szünetezési szokásokra lenne visszavezethető (vö. Gocsál 2001), ha a szünetek előfordulását, esetleg hosszúságát befolyásolta volna, hogy milyen nemű az adatközlő. A férfiak és a nők beszédprodukciójának összehasonlításakor érdemes lenne összehasonlítani a spontán és a nem spontán beszédet, illetve többféle beszédsszituációt, beszédtypust bevonni a vizsgálatok körébe.

A vizsgálat eredményei rámutattak, hogy a hallásalapú és vizuális alapú közlések akusztikuma több szempontból is eltér egymástól. Ennek oka a kétféle ingeren alapuló dekódolási folyamatokban kereshető, melyekről azonban csak közvetetten, a beszédprodukció vizsgálatával kaphatunk információt. A probléma részletes feltárásához további kutatások szükségesek, amelyek a beszédprodukció más összetevőit vizsgálják.

Irodalom

- Bata Sarolta – Gráczki Tekla Etelka 2009. A beszédpartner életkorának hatása a beszéd szupraszegmentális jellegzetességeire. In: Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Könyvkiadó, Budapest, 74–82.
- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszéd kutatás 2001*. 61–72.
- Gósy Mária 2004a. A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye. *Beszéd kutatás 2004*. 6–18.
- Gósy Mária 2004b. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszéd kutatás 2008*. 194–207.
- Imre Angéla 2006. Az olvasáaszavar és a beszédfeldolgozási folyamatok összefüggései. *Beszéd kutatás 2006*. 16–171.
- A. Jászó Anna 2006. *Az olvasás múltja és jelene*. Trezor Kiadó, Budapest.

A DISZLEXIA TÜNETEI A SPONTÁN BESZÉDBEN

Horváth Viktória – Imre Angéla

Bevezetés

A gyermekek iskolai sikertelenségeinek okai között mind a szakmai fórumokon, mind a mindennapi életben gyakran említik a diszlexiát mint az egyik fő tanulási nehézséget. A gyakorlatban a pedagógus és a szülő azt észleli, hogy a gyermek nem jól olvas; kortársaihoz képest úgy az olvasás technikája, mint az olvasott szöveg megértése kisebb-nagyobb elmaradást mutat. Az olvasási nehézség és a diszlexia nem azonos; elkülönítésük azonban nem egyszerű feladat (Kibby 1995, Frith 1999, Csabay 1999, Gósy 2005, Csépe 2006). A valódi diszlexia diagnosztizálása komplex vizsgálatokat igényel: amelyek között az olvasás szintjének feltérképezése mellett rendszerint szerepel a memória, az intelligencia, az írás, a beszédészlelés és beszédmegértés állapotának felmérése is (vö. Mandel Glaser–Searloff 1992, Kibby 1995, Meixner 1995, Schwark ed. 1997).

A diszlexia kifejezés meghatározása mára már nagyrészt egységes a szakirodalomban (Csépe 2006). A diszlexia egyfajta viszonyfogalomnak is tekinthető, amely esetén ellentét van a gyermekkel szembeni jogos elvárások (képességei, ill. a gyakorlási idő alapján) és az olvasás-írás elsajátításában nyújtott teljesítménye között (vö. Mandel Glaser–Searloff 1992, Meixner 1995). A diszlexiát mint jelenséget kritériumrendszerrel közelítik meg, a legáltalánosabban elfogadott kritériumpontok a következők. A diszlexia *a)* a felnőttkorig elhúzódik; *b)* az olvasásban és az írásban/helyesírásban tapasztalható sajátos hibákkal jellemezhető; *c)* örökletes jelenség; *d)* gyakrabban fordul elő fiúknál, mint lányoknál; *e)* normál, illetőleg magas intelligenciahányadossal jár együtt; *f)* kizárható a szociológiai tényező, mint például a verbálisan ingerszegény környezet (Gósy 2005, Blomert et al. 2004, Vellutino et al. 2004).

A diszlexia (általában) akkor állapítható meg, ha

a) az olvasási teljesítmény az olvasás pontosságát vagy a megértést egyénileg, standardizált tesztekkel vizsgálva, lényegesen alatta marad a személy biológiai kora, mért intelligenciája vagy a képzettség alapján elvárhatónak;

b) az előbbi zavar jelentősen kihat az iskolai teljesítményre vagy az olvasási jártasságot igénylő mindennapi élettevékenységekre;

c) ha érzékelési deficit van jelen, és az olvasási nehézségek meghaladják az ahhoz rendszerint társuló zavar mértékét (Csépe 2005: 227).

A kutatási eredmények alapján több kutató azt feltételezi, hogy a diszlexia nem csupán az olvasás zavara, hanem komplex nyelvfejlődési zavar (Rudel 1985, Catts 1989a, Csépe 2005), amely az írott és a beszélt nyelv használatát egyaránt érinti. A diszlexiásnak diagnosztizált fiatal felnőtt kortársaikhoz képest elmaradást mutatnak a fonetikai és a fonológiai feldolgozás, a munkamemória, a szókincs és az általános ismeretek terén is (Lieberman et al. 1985, Ransby–Swanson 2003). Empirikus kutatási eredmények szerint a diszlexiás tanulók beszédészlelési és beszédmegértési teljesítménye szintén elmarad tipikus beszédfejlődésű kortársaikéhoz képest (Ransby–Swanson 2003, Bloemert et al. 2004, Grácz 2007, Csiszár 2009). Feltételezhető továbbá, hogy a diszlexiásoknál a beszédritmust meghatározó tényezők feldolgozása is atipikus fejlődést mutat (Csépe 2007).

A diszlexia tehát nemcsak az írott nyelvet, hanem a beszédfeldolgozást is érinti, felmerül tehát a kérdés, hogy hatással van-e a beszéd másik nagy mechanizmusára, a beszédprodukcióra is. Annál is inkább, mivel a komplex nyelvfejlődési zavar feltételezése megkérdőjelezheti a beszédprodukciós folyamatok hibátlan működését. Ha a spontán beszéd tervezése és kivitelezése is érintett a diszlexiában, akkor a tanulók szóbeli közlés esetén is hátrányba kerülhetnek a tipikus fejlődésű társaikkal szemben.

Milyen működések lehetnek érintettek a diszlexiások beszédprodukciójában? A spontán beszéd létrehozása ugyanis komplex folyamat, a beszédtervezés és kivitelezés párhuzamosan zajlik: a beszélő egy gondolat nyelvi megvalósítása közben már a következő gondolat megformálását végzi. Az egyidejű folyamatműködések következményei a spontán beszéd megakadásjelenségei, amelyek egyúttal időt is biztosítanak a tervezéshez. A megakadások vizsgálatának legfontosabb célja a rejtetten működő tervezési és önellenőrzési folyamatok minél pontosabb feltérképezése. A felszíni szerkezet hibáit ugyanazok a mechanizmusok állítják elő, mint a hibátlan közléseket, ezért a megakadások egyfajta „ablakok” a rejtetten működő folyamatokra (Fromkin 1973). A hibák elemzésével az egyes tervezési szintek működéséről kaphatunk képet. A megakadások vizsgálata információt nyújt továbbá arról a folyamatról is, hogy a beszéd során miként ellenőrizzük saját közléseinket a rejtett és a felszíni önmonitorozás segítségével, és hogyan megy végbe egy esetleges hiba korrekciója. A beszéd „hibás” jelenségeinek tanulmányozása a patológiás beszéd esetén különösen fontos, mert arról ad információt, hogy mely jelenségek következnek a nem tipikus működésekből.

A jelen kutatás arra a kérdésre keresi a választ, hogy a diszlexiás középiskolások beszédprodukciójában kimutatható-e bármilyen eltérés a tipikus fejlődésű kortársaikéhoz képest. Az a feltételezésünk, hogy a diszlexia hatással van a beszédprodukciós folyamatokra is. Ennek megfelelően a diszlexiás tanulók spontán beszéde a morfológiai és szintaktikai szerkesztettségében, a szavak számában, valamint a megakadások gyakoriságában el fog térni a kortársaik sajátosságaihoz képest. Hipotézisünk szerint továbbá a diszlexiás kö-

zépiskolások kevésbé komplex közléseket hoznak létre, ezért beszédükben kevesebb lesz a megakadásjelenség. A közlés hossza és összetettsége ugyanis fontos tényezői annak, hogy a közlés folyamatosságát megszakítja-e valamilyen megakadásjelenség (McLaughlin–Cullinan 1989, Yaruss et al 1999). Korábbi kutatási eredmények (Meixner 1995, Csiszár 2007) alapján feltételeztük, hogy a diszlexia hatással van a szavak előhívási folyamatára, vagyis a diszlexia érinti a mentális lexikon aktiválását is.

Anyag és módszer, kísérleti személyek

A spontán beszéd gyűjtésére kétféle módszert alkalmaztunk. Az adatközlőknek elsőként egy 12 képből álló képregény alapján kellett mesét alkotniuk. A képeket lefelé fordítva kapták meg, és azokat csak egyenként fordíthatták meg – csak akkor haladhattak tovább a történettel, ha úgy érezték, az adott képről már mindent elmondtak. A képsorozat kiválasztásakor arra törekedtünk, hogy a képek sok részletet tartalmazzanak, és váratlan fordulatok kövessék egymást. A mesealkotást követően interjúkat készítettünk az adatközlőkkel a családról, a nyári terveikről, a továbbtanulásról. A felvételek SONY MZ R-900 típusú minidisc-hangfelvevővel készültek. Minden felvételről lejegyzést készítettünk.

A spontán beszéd szintaktikai komplexitásának megítéléséhez az ún. KFM-módszert használtuk (Gerebenné–Gósy–Laczkó 1992). Az amerikai Lee és Canter által a hetvenes években kidolgozott eljárás azt vizsgálja, hogy milyen a gyermek beszédének morfológiai és szintaktikai szerkesztettsége; a mondatok struktúrája és hosszúsága; a névmások és egyéb szófajok, valamint az igeidők, a kérdő és a tagadó formák használatának szintje. A magyar nyelvre Gerebenné Várbíró Katalin, Gósy Mária és Laczkó Mária adaptálta a teszteléses eljárást, amelyet a közlésegségek fejlődési mutatójaként (KFM) neveztek meg. A spontán beszéd szavainak, szerkezeteinek és mondatainak meghatározott, nyelvspecifikus pontértéke van. A szavak pontozásának alapelve a magyar nyelvtan sajátosságait, illetve a magyar gyermek anyanyelv-elsajátítási szakaszait veszi figyelembe. A gyermek rögzített spontán beszédének lejegyzését követi annak megállapítása, hogy a beszédben milyen pontértékű szavak és szerkezetek vannak. A szavak és a szerkezetek értékelése után az egyes mondatok is pontot kapnak. A KFM-érték kiszámítása a következőképpen történik: a szavak és a szerkezetek értékeléséből összejött pontszámot hozzá kell adni a helyes mondatok számához, majd ezt a teljes összeget elosztani az összes mondat számával. Kutatásunkban ennek a módszernek a felhasználásával elemeztük minden tanuló beszédének grammatikai komplexitását. (A teszt sztenderdizálása magyar nyelvre még nem történt meg. Az adaptálás során a magyar kutatók azonban minden korosztályra megadtak egy tapasztalati értéket, amely az adott életkorban elvárható, ez az érték 6 éves korban például 11.) A módszert elsősorban óvodáskorú gyermekek beszédének elemzéséhez dolgozták ki, de jól alkalmazható idősebbeknél is.

Meghatároztuk minden beszélőnél a különböző szavak számát, vagyis hogy hány különböző tartalmas szó fordult elő a beszédben (nem számoltuk bele tehát a funkciószavakat).

A spontán beszéd vizsgálatának másik szempontja a megakadásjelenségek elemzése volt. Az egyes jelenségek típusba sorolásakor Gósy (2004) kategorizációját vettük alapul.

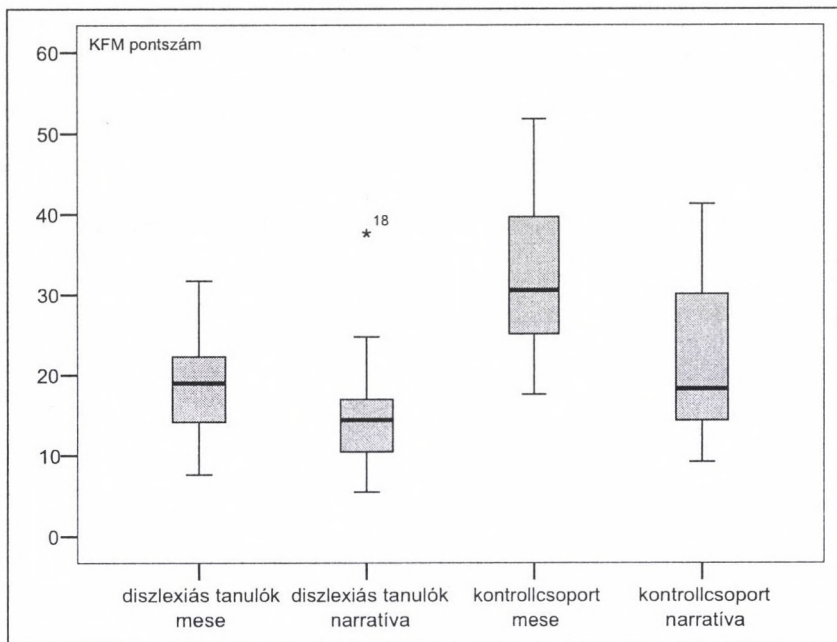
Az összes adat statisztikai elemzését az SPSS 13.0 statisztikai programmal végeztük. Egyváltozós varianciaanalízist, egytényezős varianciaanalízist, az összefüggések feltárására pedig a Pearson-féle korrelációelemzést alkalmaztuk, 95%-os szignifikanciaszinten.

A kutatásban 20, a Beszédvizsgáló Országos Szakértői és Rehabilitációs Bizottság és Gyógypedagógiai Szolgáltató Központ rutin eljárásának alapján diszlexiásnak diagnosztizált első osztályos középiskolás vett részt (valamennyiük intelligenciahányadosa a Raven-teszt alapján a normál övezetben volt, IQ: 93–112). A tanulók ugyanazon budapesti gimnázium diákjai. A kontrollcsoportot életkorban illesztett, szintén 20 fő első osztályos, nem diszlexiás gimnazista alkotta.

Eredmények

A lejegyzett spontán beszédben – a mesében és a narratívában egyaránt – meghatároztuk a közlésegségek fejlődési mutatóját. A statisztikai elemzés eredményei szerint a KFM-értékek szignifikáns különbséget mutatnak a diszlexiások és a nem diszlexiások beszédében [egyváltozós varianciaanalízis: $F(1, 79) = 31,370, p < 0,001$, parciális együttható: 0,292], valamint a különböző típusú szövegekben [$F(1, 79) = 15,607, p < 0,001$, parciális együttható: 0,170]. A két tényező együttes hatása a változóra nem szignifikáns [diszkontr×mese-narr: $F(1, 79) = 3,181, p = 0,078$, parciális együttható: 0,040, R -érték: 0,398]. Ez azt jelenti, hogy a két tényező együtt nem, csak külön-külön befolyásolja a KFM-értéket.

A hipotézisnek megfelelően a diszlexiás tanulók mindkét típusú spontán beszédben szignifikánsan kevesebb KFM-pontszámot értek el, mint a nem diszlexiások (1. ábra). A diszlexiás diákok a képről alkotott mesében átlagosan 18,78 pontot értek el (átl. elt: 6,36); a narratívában kevesebbet, átlagosan 14,89 pontot (átl. elt: 7,15). A mesénél a diszlexiások csoportjában a legalcsonyabb pontszám 7,63 volt, a narratívában 5,48. Az adatok szerint a diszlexiás elsős gimnazisták csoportjában a tanulók egyharmada nem éri el egy iskolába kerülő gyermek beszédének komplexitását (a nem sztenderdizált érték 11 pont). A kontroll csoportot alkotó tanulók meséiben számított átlagos KFM-érték majdnem kétszerese a diszlexiás diákok átlagának (átl.: 32,07; átl. elt: 8,77); ők a narratívában is jóval magasabb átlagos pontszámot értek el (21,75; átl. elt: 9,49).



1. ábra

A KFM-pontszámok a dislexiás tanuló és a kontrollcsoport különböző típusú spontán beszédében

A hipotézisnek megfelelően tehát az eredmények azt mutatták, hogy a diszlexia nemcsak az olvasásban mutat jellegzetes tüneteket, hanem – valószínűsíthetően a teljes nyelvi rendszer érintettsége miatt – negatívan befolyásolja a spontán beszéd komplexitását is. A dislexiás tanulók kevésbé összetett közléseket hoznak létre, beszédükben szegényesebb az igeidők és a bővítmények használata. Az alábbiakban példa következik egy dislexiás és egy nem dislexiás tanuló képek alapján alkotott meséjéből:

Itt a kisfiú beszélget a királlyal...első képnél. Az emberek tanácskoznak a királlyal. Utána meg a királykisasszony és ez fiú összeháza...összeölelkeznek. A király-királylány elmegy lovon. Aztán találkoznak egy medvével. A királylány ismeri a medvét. (dislexiás tanuló)

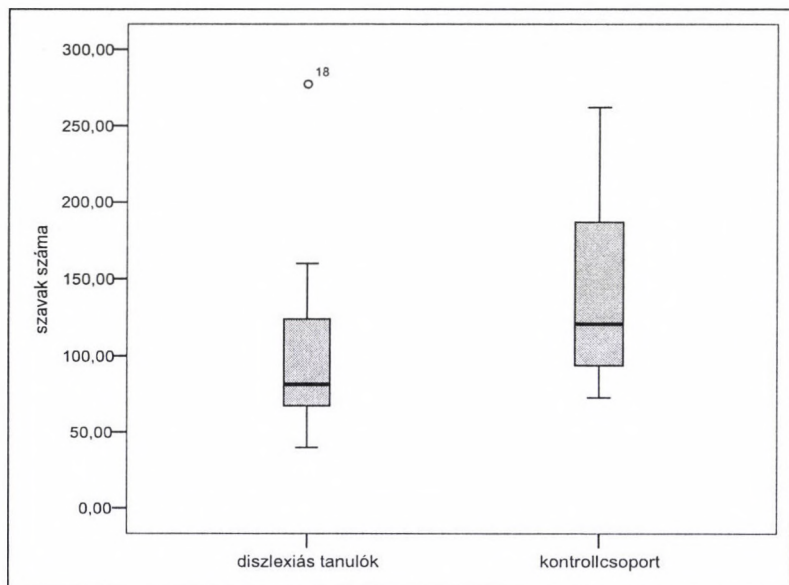
Egyszer volt egy ő szegény pásztor. Egész nap a földjén dolgozott, ő viszont a fia megbetegedett, ezért elindult a királyhoz, hogy pénzt gyűjtsön. Útközben a királyhoz mene...királyhoz elin...menet közben találkozott egy nagy medvével, aki segítséget kért tőle. (nem dislexiás tanuló)

Az első részletben szinte kizárólag egyszerű mondatok – illetve a spontán beszéd műfaji sajátosságaiból adódóan virtuális mondatok (Gósy 2003a) – szerepelnek, nincsenek jelzős szerkezetek vagy több bővítményből álló szintaktikai egységek. A nem diszlexiás tanuló ezzel szemben többszörösen összetett mondatokat alkotott, a közlés jelzős szerkezeteket és különféle alárendeléseket tartalmazott.

A két vizsgált csoport KFM-értékeiben a különbség a mesénél a nagyobb, ebben a beszédműfajban a nem diszlexiások átlaga 14 ponttal meghaladja a diszlexiásokét. A diszlexiás tanulóknál a legmagasabb érték 31,75, a kontrollcsoportban 51,83. A diszlexiás és a nem diszlexiás tanulók egyaránt a mesénél produkáltak magasabb KFM-értékeket. A narratívákban csökkenek az átlagos pontszámok, a diszlexiás tanulóknál a legalacsonyabb pontszám 5,48, a nem diszlexiásoknál 9,2. A maximum értékek 37,62 a diszlexiás tanulóknál és 41,32 a kontrollcsoportban. A két műfajban mért eltérő KFM-pontszámok a műfaji sajátosságokkal magyarázhatók. A beszéd műfaja ugyanis nagymértékben meghatározza a tervezési és megvalósítási sajátosságokat (Markó 2005). A kép alapján történő mesealkotás során a nyelvi formát ugyan az adott pillanatban hozza létre a beszélő, de a mese műfajából adódóan egyrészt klisék is rendelkezésre állnak; másfelől a képek, az adott tartalom miatt körülhatároltabb a keresés a mentális lexikonban. A mese műfajában emiatt mindkét csoport beszélői összetettebb szerkezeteket, bonyolultabb igeidő- és bővítmény-rendszert alkalmaztak.

A szavak számának elemzése felvilágosítást nyújt a mentális lexikonhoz való hozzáférésről, illetve arról, hogy az aktivált szókincs mutat-e eltérést diszlexia esetén a kontrollcsoporthoz képest. A diszlexiás tanulókkal készített interjúkban átlagosan 102 különböző tartalmas szót adatoltunk (átl. elt: 54,27), a kontrollcsoportban 143 szót (átl. elt: 64,09). A diszlexiás diákok spontán beszédét szignifikánsan kevesebb különféle szó jellemzi [egytényezős varianciaanalízis: $F(1, 39) = 4,954, p = 0,032$]. Ez a mentális lexikonhoz történő hozzáférési stratégia sajátosságaira utal. Mindkét csoportban nagyok az egyéni különbségek; a kontrollcsoportban nagyobb az adatok szóródása (2. ábra).

A diszlexiás tanulók egyikének spontán narratívájában mindössze 40 különböző tartalmas szót adatoltunk, a kontrollcsoportban ennek majdnem kétszerese volt a legkevesebb (73 szó). A diszlexiás tanulók csoportjában a maximális érték 277, a kontrollcsoportban 262 lexéma volt. A két csoport teljesítménye között természetesen volt átfedés, de a diszlexiás középiskolások egyharmadának beszédében a kontrollcsoportban tapasztalt minimális szószámnál is kevesebb szó fordult elő. A diszlexiások közül egyetlen diáknál adatoltunk kiugróan sok különböző szót (277 darab). Ezt az értéket kivéve a diszlexiások csoportjában 160 szó volt a maximális teljesítmény – a kontrollcsoport tanulóinak majdnem fele meghaladta ezt a szószámot.



2. ábra

A különböző tartalma szavak száma a két csoport spontán narratíváiban

Az alábbiakban példák olvashatók a diszlexiás tanulók spontán narratíváiból. A beszédszakaszok általában rövidek és egyszerűbben szerkesztettek, sok ismétlést tartalmaznak.

– *Mesélj a testvéreidről!*

– *Hát, egy van, négy éves, kislány, hát nem tudom. Most megy iskolába idén, mármint hogy öt éves, nem tudom, hány éves, nem tudom. Most megy iskolába, azt tudom.*

– *Mesélj, mit csináltál a tavaszi szünetben!*

– *Hát tavaszi szünet...hát az az annak nagy részét lent töltöttem a nagyszüleimnél Sopronban. Hát kirándultunk bokáig érő hóban fönn a hegyekben. A szünidő nagy részét inkább kirándulással töltöttük.*

– *Mit csinálsz iskola után?*

– *Hát tanulok vagy edzésre megyek, meg számítógépezek. Szeretek olvasni.*

– *Mit olvastál legutóbb?*

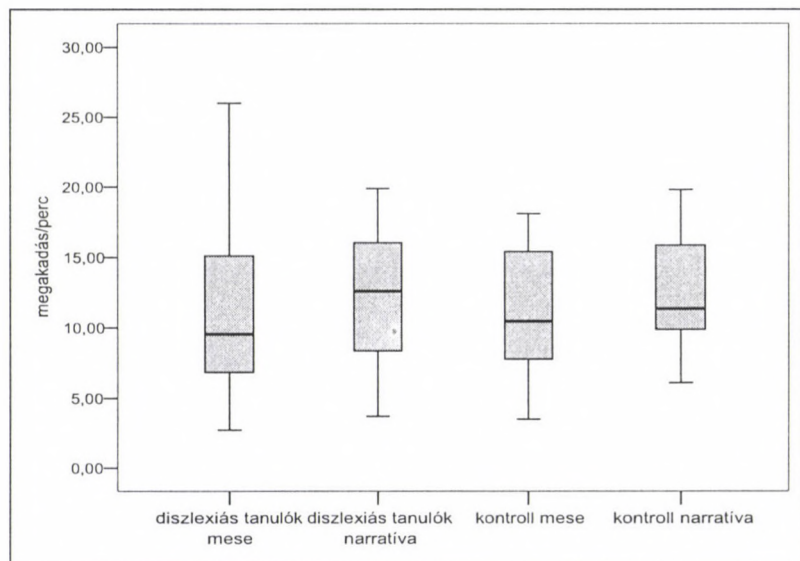
– *Mindent. Stephen Kinget.*

– *Mesélj róla, melyiket?*

– *Talizmán...sz...igen talizmán.*

A megakadásjelenségek vizsgálatával arra a kérdésre kerestük a választ, hogy a diszlexia miként befolyásolja a beszédtervezési folyamatokat. A tanulók között egyénileg nagy volt az eltérés abban, hogy milyen hosszan és szívesen beszéltek magukról; de csoportszinten a vizsgált diszlexiás és nem diszlexiás tanulók között nem volt szignifikáns különbség a közlések időtartamában. A diszlexiás tanulók átlagosan 4'31" (min.: 1'41", max.: 8'18"), a nem diszlexiások átlagosan 5'08" időtartamú (min.: 2'50", max.: 9'10") narratívákat alkottak (átl. elt.: 1,87, ill. 2,03). Az eredmények szerint a beszédidő hossza és a megakadások száma mindkét csoportban korrelációt mutat. A diszlexiás tanulók esetében a beszédidő hossza és a megakadások száma műfajtól függetlenül összefüggést mutat (mesében: $r = 0,620$, narratívában: $r = 0,741$; $p < 0,001$). Ez azt jelenti, hogy minél hosszabban beszél a beszélő egy témáról, annál valószínűbb, hogy közben többször keletkezik diszharmónia a beszédtervezés során. Ez az összefüggés a kontrollcsoportban is kimutatható a narratívákra ($r = 0,709$; $p < 0,001$).

A kutatás további kérdése, hogy a beszédműfajtól függően változik-e a beszédtervezés; erre a megakadásjelenségek jellemzői utalhatnak. Az eredmények azt mutatták, hogy a megakadások percenkénti gyakoriságában nem volt statisztikailag szignifikáns különbség a diszlexiás tanulók és a kontrollcsoport között, sőt az egyes szövegtípusokban adatolt megakadások gyakoriságban sem (3. ábra).



3. ábra

A megakadásjelenségek gyakorisága

Mindkét csoportban a narratívákban volt gyakoribb a megakadás, de a mesékre volt jellemzőbb az adatok valamivel nagyobb szóródása (1. táblázat). A diszlexiás diákok meséiben volt adatolható a két szélsőérték: volt olyan tanuló, aki egy perc alatt csak két megakadást produkált, de olyan is, akinél 26 fordult elő. A két csoportban meghatározott értékek a mesékben és a narratívákban is mintegy 85%-os átfedést mutatnak.

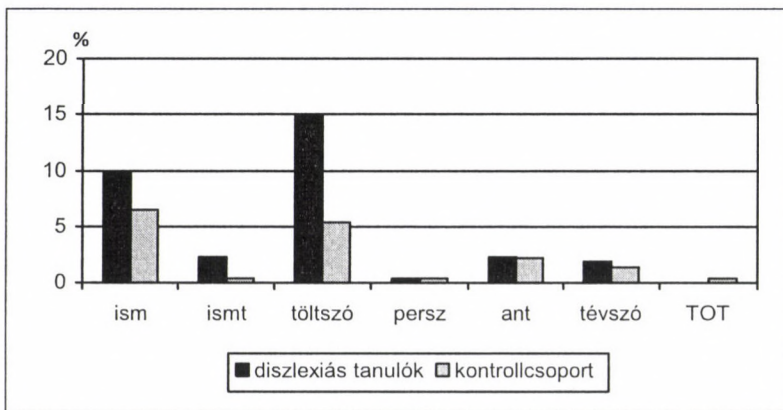
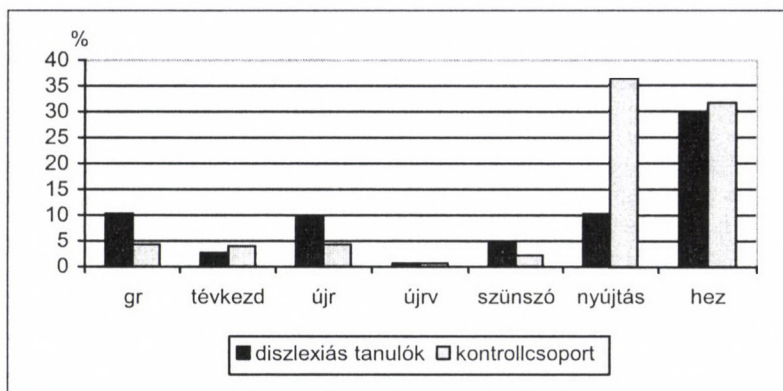
1. táblázat: A megakadásjelenségek gyakorisága percenként

Adatközlők	Szövegtípus	Átlag (db)	Minimum (db)	Maximum (db)
Diszlexiás tanulók	mesé	11,18	2,70	26,00
	narratíva	12,18	3,70	19,90
Kontrollcsoport	mesé	11,01	3,50	18,10
	narratíva	12,30	6,10	19,80

A beszédműfaj mellett a beszéd komplexitása is befolyásolja a megakadásjelenségek gyakoriságát. A statisztikai elemzés azt mutatta, hogy a KFM-érték és a megakadások száma összefüggést mutat: minél összetettebb a közlés, annál többször szakítja meg a folyamatosságát valamilyen bizonytalanság vagy hiba (Pearson-próba: $p < 0,001$, $r = 0,817$ a diszlexiás tanulók csoportjában, a kontrollcsoportban: $p = 0,018$, $r = 0,521$).

Elemeztük, hogy a tanulók spontán beszédében hogyan alakul az egyes megakadástípusok aránya, illetve van-e különbség a típusok tekintetében a diszlexiás diákok és a kontrollcsoport között. A mesékben összesen 15 megakadástípus fordult elő a diszlexiás diákoknál, 16 a kontrollcsoportban. A bizonytalanságok domináltak a beszédben (diszlexiás tanulók: 80,2%, nem diszlexiások: 87,1%). Ez arra utal, hogy az e korúaknak is a gondolatok nyelvi formájának tervezése jelenti a legnagyobb nehézséget (Gósy 2003b).

A diszlexiás diákok által létrehozott mesékben – csakúgy, mint a felnőtt beszélőknél általában – a leggyakoribb megakadásjelenség a hezitálás (4. ábra), amely multifunkcionális jelenség: időt biztosít a nyelvi tervezéshez, a lexmakereséshez, az önmonitorozáshoz; jelezhet a társalgás során beszéd-szándékot, továbbá fontos szerepe van a beszédfeldolgozásban (vö. Maclay–Osgood 1959, Fox Tree 2002, Horváth 2009), például: *a királyfi kita ő elhátározta; és ő egy várba mennek*. A hezitálás ebben a feladatban gyakran arra szolgált, hogy amikor a diák a következő képet megnézte, átgondolja, hogy mit lát rajta, és hogyan fűzze tovább a történetet: *ő még aznap éjjel beszél a medvével; és ő utána egy házat látunk öregemberrel*.



4. ábra

A megakadástípusok aránya a diszlexiás tanulók és a kontrollcsoport meséiben

Beszédtervezési nehézségek esetén a diszlexiás diákok gyakran ejtettek töltelékszót is, mint például: *ilyen, így, hát*. Ezek a szavak és kifejezések szervesen illeszkednek a közlésbe, nem adnak hozzá annak tartalmához; időtartamukat a beszélő arra használja fel, hogy megtervezze a következő gondolat nyelvi formáját, előhívja a szándékolt elemet a mentális lexikonból, vagy ellenőrizze saját közlését grammatikai, pragmatikai szempontból. Példák: *meg ő hát a medvével összebarátkozott; esküvő volt ilyen hercegeknél; hát szerintem itt a király beszél a fiúval, hát ellenkezik, vagy nem tudom; először így bemegy így a barlangba*. A diszlexiások beszédében a többi bizonytalansági megakadás (az ismétlés, a nyújtás, az újraindítás) aránya 10% körüli; a diszlexiás tanulók ezekhez hasonló gyakorisággal produkáltak valami-

lyen grammatikai hibát a mesék során. Ez az arány mindenképpen nagyinak mondható, mivel a nem diszlexiás beszélők grammatikai hibáinak aránya általában töredéke a bizonytalansági megakadásoknak (vö. Markó 2004, Szabó 2004, Horváth 2004). A jelen kutatásban a kontrollcsoport tanulói fele olyan ritkán produkáltak grammatikai hibát a mesék során, mint a diszlexiás diákok. Az eredmények szerint a diszlexiás tanulók nyelvi tervezési folyamatai gyakrabban hoznak létre a köznyelvi normától eltérő, grammatikailag hibás közléseket, ami arra utal, hogy diszlexia érinti ezt a beszédtervezési részfolyamatot is.

Az alábbiakban példák olvashatók a diszlexiás diákok által a mesealkotás során produkált, köznyelvi normának ellentmondó közlésekre:

búcsút vesz és fiú elmegy – a *fiú* főnév előtt nem szerepel a névelő

nem akarja hogy szerintem a királylánnyal összejöjjenek – a szerkezet szórendje nem megfelelő, ez eredményezi a hibás alakot. Elképzelhető, hogy a *szerintem* és a *hogy* lexémák felcserélődtek, de a metatézis eredménye a felszínen egy grammatikai, szintaktikai hiba.

el akarta kérni a lány kezét – a nem megfelelő igekötő használata (esetleg téves szótalálás eredményeként) hibás szerkezetet eredményez.

a paraszt végül sikerült felmenni a királyné szobájához – a közlésrészlet több morfológiai hibát is tartalmaz, hiányzik a *-nak* toldalékmorféma (*parasztnak*) és a személyrag az ige­névről (*felmennie*).

A nem diszlexiás tanulók meséiben is fordultak elő természetesen grammatikai hibát tartalmazó közlések – főként szám-személy egyeztetési tévedések – például:

boldogan élt mindenki, míg meg nem haltak

csatangolásai során eljutnak egy kristálypalotához

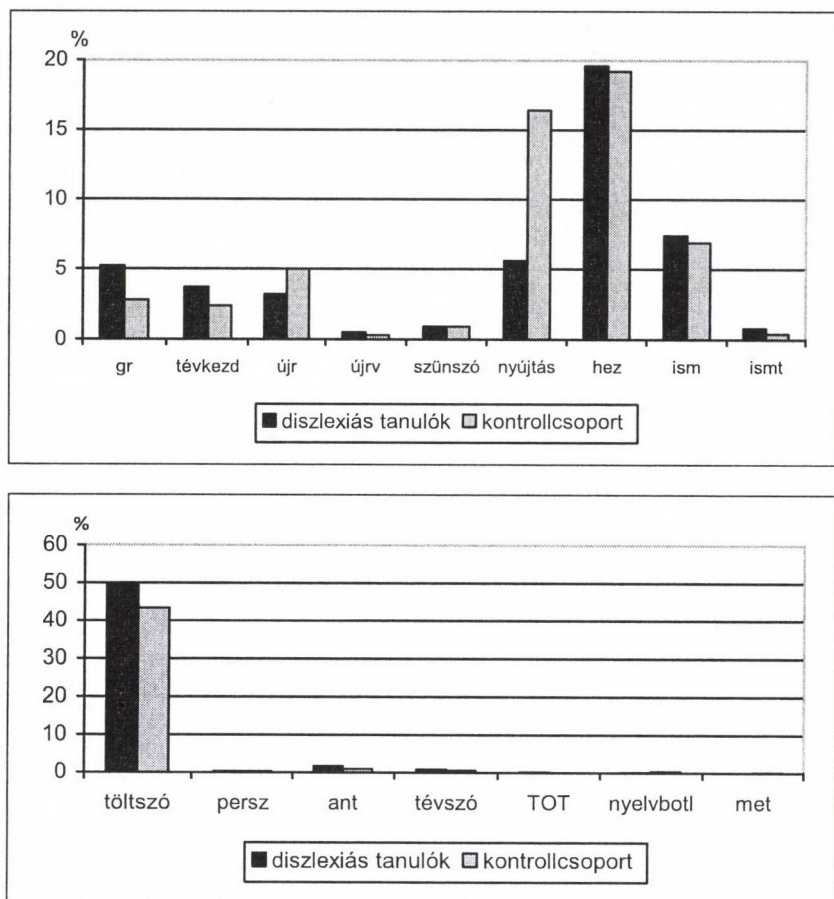
A nyújtás egy adott hang (általában névelő vagy szókezdő hang) időtartamának megnövelését jelenti, ezen időtartam alatt rejtett tervezés, illetve önmonitorozás zajlik. A nem diszlexiásoknál ez a leggyakoribb jelenség; a diszlexiás tanulók beszédében csak harmadannyi nyújtás fordult elő – ez a tény további beszédtervezési különbségekre utal. A nyújtást mint stratégiát a diszlexiás tanulók ritkán alkalmazták beszédtervezési nehézségeik áthidalására – valószínűsíthető, hogy egy aktivált lexéma bizonyos részének időtartam­növelése nem biztosít számukra elegendő időt a folyamatműködésekhez – avagy nem tudják jól értékesíteni ezt az időt a tervezésben –, ezért inkább a hezitálást és a töltelékszavakat preferálják.

A narratívákban 15 különböző jelenséget adatoltunk, bizonytalanságokat és téves kivitelezéseket egyaránt. A bizonytalanságok jelentek meg döntő többségben a narratívákban, csakúgy, mint korábban a mesékben: a diszlexiás

tanulók megakadásainak 87,3%-a, a nem diszlexiásokéinak 92,5%-a ebbe a kategóriába volt sorolható. A spontán monologikus közlésekben – témától függetlenül – tehát a bizonytalanságok dominálnak a beszédben, ahogyan ezt számos más kutatás is igazolta (Gósy 2003b, Markó 2004, Gyarmathy 2007).

A diákok spontán narratíváiban a legnagyobb arányban (diszlexiások: 49,8%, kontrollcsoport: 43,4%) töltelékszavak jelentek meg, ezek a mesékben csak a második, illetve harmadik leggyakoribb jelenségek voltak. A narratívákban a tanulók a töltelékszavakat preferálták a hezitálásokkal szemben, kétszer gyakrabban ejtettek valamilyen töltelékszót, mint kitöltött szünetet. Valószínűsíthető, hogy interjúhelyzetben, amikor magukról kellett beszélniük, jobban törekedtek a beszéd folyamatosságának fenntartására; míg amikor képek alapján kellett mesét mondaniuk, kevésbé tűnt fel nekik, kevésbé zavarta őket, hogy sokat hezitálnak. A másik lehetséges magyarázat, hogy többségük számára a képek alapján való mesealkotás nehezebb feladat volt (a képeket egyenként felfedve kellett összefüggő történetet mondaniuk), ezért a lehető legegyszerűbb stratégiát, a hezitálást preferálták – az interjúkban az artikulációsan összetettebb töltelékszavak is megfeleltek a diszharmónia áthidalására. Mindkét csoportban a második leggyakoribb jelenség a hezitálás (19,6%, ill. 19,2%), amely a mesékben még a legnagyobb arányban szerepelt. Ez az eltérés szintén a kétféle spontán beszéd témájának különbségéből adódik.

A narratívákban a két csoport között a nyújtások arányában találtunk jelentős különbséget (5. ábra). A nem diszlexiás tanulók – ugyanúgy, mint a mesékben – háromszor nagyobb arányban (16,5%) produkálták ezt a jelenséget, mint a diszlexiás tanulók (5,6%). A grammatikai hibák előfordulási arányában is jelentős a különbség: a diszlexiás tanulók ilyen jellegű tévesztései kétszer gyakoribbak, ami ismét megerősíti a meséknél tapasztalt tendenciát, miszerint a diszlexia befolyásolja a beszédtervezési folyamatokat, különösen a nyelvi tervezést, amely során a gondolatokhoz grammatikai formát rendelünk. Példák a grammatikai hibákra a diákok spontán narratíváiból: *olvasás, az egyre gyakrabban csináljuk; van egy ismerősöm, aki kommandós volt és azóta nagyon szeretnék lenni az* (diszlexiás tanulók); *ezt nem úgy érzem, mintha ezt muszáj megcsinálni; és akkor itt feküdtek mindenki a gyepen* (nem diszlexiás tanulók).



5. ábra

A megakadástípusok aránya a diszlexiás tanulók és a kontrollcsoport narratíváiban

Következtetések

A jelen kutatás során 20 diszlexiás és 20 nem diszlexiás elsős gimnazista spontán beszédét elemeztük. Az eredmények igazolták, hogy a diszlexia háttal van a spontán beszéd tervezésére. A megakadásjelenségek elemzésével a jelen korpuszon is igazoltuk azt a megállapítást, hogy a közlés hossza és összetettsége befolyásolja a jelenségek gyakoriságát. Minél hosszabb időtartamban beszéltek a diákok, és minél magasabb volt a komplexitásra vonatko-

zó mutató, annál többször szakította meg a közlést valamilyen megakadás-jelenség.

A kutatás eredményei igazolták a hipotéziseinket: eltéréseket találtunk a diszlexiás és a nem diszlexiás tanulók spontán beszédében. Ez arra utal, hogy a diszlexia nemcsak az olvasási képességet, hanem a beszédtervezési folyamatokat is befolyásolja. Egyes kutatók szerint a diszlexiás beszélők beszéd-produkciós nehézségei a fonológiai tervezési folyamatok érintettségével magyarázhatók (Catts 1986, 1989b). A diszlexiás diákok beszéde morfológiai és szintaktikai szempontból jóval kevésbé összetett, mint tipikus fejlődésű kortársaiké: egyharmaduk közlései nem érik el egy jól beszélő hatéves gyermek szintjét, a bővítmények használatában és a ragozásban egyaránt. A diszlexiásnak diagnosztizált tanulók beszéde a különböző szavak számában is elmarad kortársaikétól, ami a mentális lexikon hozzáféréseinek problémájára utal. A diszlexia érinti a beszédtervezési folyamatokat is, különösen a nyelvi tervezést, vagyis a gondolatok megfelelő grammatikai formájának létrehozását. Ennek következménye az, hogy a diszlexiás diákok beszédében jóval gyakrabban fordultak elő a köznyelvi normának ellentmondó, agrammatikus formák. A diszlexiás tanulók beszédtervezésére jellemző, hogy a nyújtások időtartamát a beszélők nem tudják jól felhasználni a tervezés során, ezért ritkán produkálják ezt a megakadást.

Minden tanulási nehézség esetén fontos, hogy megállapítsuk a zavar feltételezhető okát, mértékét, majd ennek alapján elkészítsük a fejlesztés tervét. A terápia azonban csak akkor lehet sikeres, ha megfelelő a diagnózis. A diszlexiásokkal folytatott kutatás eredményei felvetik a spontán beszéd elemzését mint újabb diagnosztikai lehetőséget a diszlexia és az olvasási nehézség elkülönítésére. A pedagógiában elterjedt gyakorlat, hogy azok a tanulók, akiket hivatalosan diszlexiásnak minősítenek, mentesülnek az írásbeli számonkérések, vizsgák alól. A kutatás pedagógiai tanulsága tehát az, hogy ha a szóbeli számonkéréskor a diszlexiás tanuló nehezen fejezi ki magát, az nem feltétlenül a készületlenségre utal, jelezheti a diszlexiás tanulóknál tapasztalt beszédprodukciós eltéréseket is.

Irodalom

- Blomert, Leo – Mitterer, Holger – Paffen, Christiaan 2004. In search of the auditory, phonetic, and/or phonological problems in dyslexia: context effects in speech perception. *Journal of Speech, Language, and Hearing Research* 47. 1030–1047.
- Csabay Katalin 1999. Az áldiszlexia mint korunk egyik járványveszélye. In dr. Salné Lengyel Mária (szerk.): *Fejlesztő-pedagógia – különszám. Diszlexia, diszgráfia, diszkalkulia, magatartászavar*. Mentor-Szanator Kft., Budapest, 42–44.
- Catts, Hugh W. 1986. Speech production/phonological deficits in reading-disordered children. *Journal of Learning Disabilities* 19. 504–508.
- Catts, Hugh W. 1989a. Defining dyslexia as a developmental language disorder. *Annals of Dyslexia* 39 (1). 50–65.

- Catts, Hugh W. 1989b. Speech production deficits in developmental dyslexia. *Journal of Speech and Hearing Disorders* 54. 422–428.
- Csépe Valéria 2005. *Kognitív fejlődés – neuropszichológia*. Gondolat Kiadó, Budapest.
- Csépe Valéria 2006. *Az olvasó agy*. Akadémiai Kiadó, Budapest.
- Csépe Valéria 2007. Azonos vagy különböző? Beszédeszélelési és olvasási zavarok: diszlexia és SLI. In Gósy Mária (szerk.): *Beszédeszélelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 164–181.
- Csiszár Orsolya 2007. Egy szóértési vizsgálat tanulságai. *Beszédkutatás* 2007. 134–141.
- Csiszár Orsolya 2009. Szövegértés és diszlexia. In Lengyel Zsolt – Navracsics Judit (szerk.): *Tanulmányok a mentális lexikonról. Nyelvelsajátítás – beszédprodukción – beszédpercepción*. Segédkönyvek a nyelvészet tanulmányozásához. Tinta Kiadó, Budapest, 301–310.
- Fox Tree, Jean E. 2002. Interpreting pauses and ums at turn exchanges. *Discourse Process* 34 (1). 37–55.
- Frith, Uta 1999. Paradoxes in the definition of dyslexia. *Dyslexia* 5 (4). 192–214.
- Fromkin, Victoria A. 1973. The non-anomalous nature of anomalous utterances. In Fromkin, Victoria A. (ed.): *Speech errors as linguistic evidence*. Mouton, The Hague, 215–242.
- Gerebenné Várbiro Katalin – Gósy Mária – Laczkó Mária 1992. *Spontán beszéd-megnyilvánulások szintaktikai elemzése DDS technika segítségével*. Kézirat. Budapest.
- Gósy Mária 2003a. Virtuális mondatok a spontán beszédben. *Beszédkutatás* 2003. 29–43.
- Gósy Mária 2003b. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127/3. 257–277.
- Gósy Mária 2004. A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye. *Beszédkutatás* 2004. 6–18.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Grácz Tekla Etelka 2007. Diszlexiás és tipikus fejlődésű gyermekek beszédfeldolgozásának vizsgálata. In Gósy Mária (szerk.): *Beszédeszélelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 202–213.
- Grácz Tekla Etelka – Gósy Mária – Imre Angéla 2007. Olvasási nehézség és diszlexia a beszédfeldolgozás tükrében. In Gósy Mária (szerk.): *Beszédeszélelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 214–228.
- Gyarmathy Dorottya 2007. Az alkohol hatása a beszédprodukción. *Beszédkutatás* 2007. 108–120.
- Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszédkutatás* 2004. 187–199.
- Horváth Viktória 2009. Funkció és kivitelezés a hezitációs jelenségekben. *Megjelenőben*.
- Kibby, Michael W. 1995. *Practical steps for informing literacy instruction: A diagnostic decision-making model*. International Reading Association, Newark, Delaware.

- Lieberman, Philip – Meskill, Robert H. – Chatillon, Mary – Shupack, Helaine 1985. Phonetic speech perception deficits in dyslexia. *Journal of Speech and Hearing Research* 28. 480–486.
- MacLay, Howard – Osgood, Charles E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15. 19–44.
- Mandel Glazer, Susan – Searloff, Lyndon W. – Gentile, Lance M. 1992. Reexamining reading diagnosis. In Mandel Glazer, Susan – Searloff, Lyndon W. – Gentile, Lance M. (eds.): *Reexamining reading diagnosis. New trends and procedures*. International Reading Association, Newark, Delaware.
- Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszéd kutatás* 2004. 209–222.
- Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszéd kutatás* 2005. 63–77.
- McLaughlin, Scott F. – Cullinan, Walter, L. 1989. Disfluencies, utterance length, and linguistic complexity in nonstuttering children. *Journal of Fluency Disorders* 14 (1). 17–36.
- Meixner Ildikó 1995. *Munkásságom tézis-szerű összefoglalása*.
<http://www.diszlexia.info/mexinertezis.htm>. Az utolsó letöltés ideje: 2009. október 8.
- Ransby, Marilyn J. – Swanson, H. Lee 2003 Reading comprehension skills of young adults with childhood diagnoses of dyslexia. *Journal of Learning Disabilities* 36 (6). 538–555.
- Rudel, Rita G. 1985. The definition of dyslexia: Language and motor deficits. In Duffy, Frank H. – Geschwind, Norman (eds.): *Neuroscientific approach to clinical evaluation*. Little Brown and Co., Boston, 33–53.
- Schwark, Rita (ed.) 1997. *Legasthenie*. Bundesverband Legasthenie e. V. Greifswald.
- Szabó Eszter 2004. „öhm, na hiszen, hogy is mondtam” – Megakadásjelenségek nyelvi játék közben. *Beszéd kutatás* 2004. 200–208.
- Vellutino, Frank R. – Fletcher, Jack M. – Snowling, Margaret J. – Scanlon, Donna M. 2004. Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry* 45 (1). 2–40.
- 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.

„NYELVBOTLÁS”-KORPUSZ

6. rész

Az első magyar valós idejű, jegyzeteléses technikával gyűjtött megakadás-jelenség-korpusz első részét 5139 tétellel 2004-ben publikáltuk (*Beszéd kutatás 2004.* 19–186); második részét 761 tétellel 2005-ben (*Beszéd kutatás 2005.* 145–173); harmadik részét 388 tétellel 2006-ban (*Beszéd kutatás 2006.* 231–247); negyedik részét 244 tétellel 2007-ben (*Beszéd kutatás 2007.* 187–198) és az ötödik részét 444 tétellel 2008-ban (*Beszéd kutatás 2008.* 221–240).

Eddig tehát összesen **6976** megakadásjelenséget elemeztünk; értelemszerűen ez a korpusz a hiba típusúak gyűjteménye. A kezdetektől számos, különféle szempontú tanulmány, disszertáció, szakdolgozat készült ebben a témakörben. Mindez a tapasztalat oda vezetett, hogy a „Nyelvbotlás”-korpusz osztályozási kategóriáit egyszerűsítsük. Ennek alapvető oka az, hogy a korpuszt használóknak nagyobb szabadságot biztosítsunk egy adott jelenség mélyreható elemzéséhez.

Az egyszerűsített osztályozás elve az, hogy minden esetben a felszínen tapasztalható jelenséget soroltuk kategóriákba, a finomelemzések az adott jelenség kutatójára várnak. A megakadásjelenségek tanulmányozása, illetőleg a kapott eredmények lehetővé tették továbbá az árnyaltabb definíció megfogalmazását. Ennek megfelelően a spontán beszéd artikulációs, percepciós folyamatosságát megakasztó, illetve a köznyelvi normának ellentmondó különféle jelenségeket nevezzük megakadásjelenségeknek. A fő kategóriák és meghatározásuk a következők:

1. Téves szó

A megfelelő szó helyett egy másik szó megjelenése a felszínen. Ez a kategória tartalmazza a freudi elszólásokat és a malapropizmusokat is, amelyek további elemzésekkel különíthetők el.

2. Grammatikai hiba

A köznyelvi normának ellentmondó morfológiai vagy szintaktikai struktúra.

3. Kontamináció

Két nyelvi jel (szó, szerkezet) összeolvadása, vegyülése.

4. Téves kezdés

Beszédhang vagy beszédhangsorozat ejtése, amely azonban nem szó.

5. Nyelvem hegyén van jelenség

A szándékolt szó jelentésének és morfológiai struktúrájának ismerete, de a fonológiai/fonetikai szerkezet kiejtésének (részleges) gátoltsága.

6. Perszeveráció

Egy kiejtett nyelvi elem szándéktalan, ismételt megjelenése a közlés egy későbbi időpontjában.

7. Anticipáció

Egy szándékozott nyelvi elem megjelenése a közlés egy korábbi időpontjában is.

8. Metatézis

Nyelvi elemek sorrendjének felcserélődése.

9. Egyszerű nyelvbotlás

A beszédtervezés artikulációs tervezési szintjén, ill. az artikulációban létrejött hiba, amely semmilyen egyéb okkal nem magyarázható. Altípusai: a) betoldás, b) helyettesítés, c) kiesés

10. Több típusba sorolható jelenségek

Azok az adatok kerülnek ebbe a kategóriákba, amelyek létrejöttében a felszíni elemzés alapján több rejtetten működő folyamat hibája is feltételezhető.

1. Téves szó

Megakadás

A jegyzetfoltban □ jaj, dehogy foltban, csak nézd, van egy folt a blúzodon – szóval a jegyzetboltban vettem.

A következő félórában Johann Sebastian Bach művei közül válogatunk. Elsőként hallgassák meg Antonio Vivaldi C-dúr Concertóját.

A középkori egyházi gyakorlatban fontos szerepe volt a letargiának.

A nős Odüsszeusz...

A szénsavmentes víz tele van mikroorganizmusokkal.

Az rögtön nekimegy □ nem neki megy, hozzá érkezik.

Be kell menniük a parókára.

Egy ilyen medence partján dobálnak csajozni.

Elismeréssel adózom az egész ellenő az egészszel szemben.

Én csak egy hangot egy gombot ütök be.

Ez egy hatalmas komplexus.

Ez máris más megvilágításba szerzi helyezi a mondatokat.

Ezek a szavak őő sorok jambikus lejtésűek.

Ezzel nem akartalak bántani □ bántani.

...hogyan Önök mivel szeretnének külön fogalmazni.

Járt ma itt a bádogos, mármint a köműves?

Kérem, készítsék el jövő hétre a hajónaplót! Jaj, a hospinaplót, elnézést.

Majd mutatok nektek egy pszichológiai értelmezést ebben az értelemben.

Szándékolt közlés

Bach–Vivaldi: C-dúr Concerto

liturgiának

a hős Odüsszeusz

mikroorganizmusokkal

parókiára

próbálnak

komplexum

foglalkozni

ebben a témában

Megakadás**Szándékolt közlés**

Mást csinál, mint amit tesz, akarom
mondani másat tesz, mint amit
mond.

...megcsapta a fejemet fületem...

Minden kellett hozzá, még a szüle-
tési anyanyelvi ☐ anya ☐ anya-
könyvi kivonat is.

Mindent megvesz magának, amit
megkínál megkíván.

Mindjárt megyek, addig tereged
már ki a rizses ruhákat.

Mire jössz vele?

Na, megyünk a kocsmába?

Nekem nem is volt soha doppingvá-
gyam ☐ mármint doppingvá-
dam.

Nem akarok agyhagymagyulladás-
kapni! Mit mondtam? Nem haj-
hagymát, mi?

Nem hallgatjuk meg a könyvet?

Pénteken nyílik meg a szakkiállítás
és vár vásár.

Politikai erőt lehet rá élni építeni.

Sokkal nagyobb pénz öö ☐ bevé-
telre tesz szert.

Teljesen regisztrálnak a sejtek.

Úgy olvassa a festményt, mint ahogy
a könyvet festik ☐ öö szóval ér-
tik!?

Veszek egy pillantást a képekre.

Zsombi, a drogkereskedő rendőrku-
tya ☐ vagy drogkereső rendőrku-
tya a második helyezett lett.

vizes

mész

a könyvtárba

kazettát

regenerálódnak

vetek

2. Grammatikai hiba**Megakadás****Szándékolt közlés**

A cikk, amelyet végül is hírt adott...

Annyit kérnék Önöknek

Az, amit marxizmus alatt értettek a
korai marxisták, azok nem Marx
nézeteit tükrözte.

amely

Önöktől

az nem

Megakadás

...és aztán összekeveredjük a két
gyilkosság miatt...
...hogy melyik azok a pályák...
...minden lehetőséget meghasznál
arra...
...mondanám, hogy vonzerő volt ne-
kik...
...nem feltétlenül érhetett célban...
Nem vagyunk beszélő viszonyban.
Röviden és érthetően felvázolom
önöket...

3. Kontamináció

Megakadás

33 éves voltam, amikor anyám meg-
született, amikor én megszület-
tem, szóval érted.
A barokk festészet éke a kép: láthat-
ják, a kislány körülbelül hatvanki-
lós pumpában ☐ ruhában pom-
pázott.
A technika ördöge bármikor fel-
mondhatja a szolgálatot.
...drogáll drogós állapotban...
Egy fél százalékos jegyet kérek!
El kell vallanom...
Elektronikus ☐ műtyűket ☐ mű-
szereket, ez az, lehetne árusítani.
Elnézést a késésért, de elaludtam a
buszt.
Már lábadok ki a gödörből, egyet
felfelé...
Meg kell mondjam, nekem tiszta a
lelkifurdalásom.
Mit vagy? Akarom mondani, mit
szeretnél és hogy vagy?
Trollamossal ☐ trillamossal ☐ na,
villamossal és trolival mentünk.
Tudjátok, mire fontolok ☐ gondolok
☐ mit fontolok, na!

Szándékolt közlés

összekeveredünk
melyek azok a pályák
kihasznál
vonzerejük volt
érhetett célba
beszélő viszonyban
önöknek

Szándékolt közlés

33 éves volt anyám, amikor megszü-
lettem

ruhában × pompázott

a technika ördöge beleszól × a műszer
felmondja a szolgálatot

félárú × ötven százalékos
be kell vallanom × el kell mondanom
kütyük × műszerek

lekéstem a buszt × elaludtam

lábadozok × lábalok ki

tiszta a lelkiismeretem × nincsen lelki-
furdalásom

mire gondolok × mit fontolok

Megakadás**Szándékolt közlés**

Tudod van az a kutyásos ☐ kutyusos mese.

kutyás × kutyusos

Vettek egy bírányt.

bárányt × birkát

4. Téves kezdés**Megakadás**

A délidá ☐ délibábok hőjét kéne megszerezni.

A focicsapat dupla ver ☐ veszteséget szenvedett.

A nagy ojal ☐ olajtársaságok és az általuk teremtet világ...

A néhány hetes cecem ☐ csecsemők szülei állnak a képen.

A szívdőh ☐ a szívdőglesztő sztár fénykorát a '80-as években élte.

A té ☐ voltak téves észlelések is.

...algoritmusok kialakulása ☐ kialakításával értelmezni kellett...

Azért észleltek több szünetet a szün a szövegben, mert...

...csoportosítha csoportosíthatjuk különféle módokon.

...de ilyenkor nov ☐ okt ☐ szeptemberben már jól fel kell öltözni.

Egy hívá ☐ egy hivatkozási alap...

En elengedtem, csak azért eve avatkoztam be.

En még ki akarom bot ☐ javítani a nyelvbtlásokat.

Es milyen bo ☐ bűdös a zoknim.

Ez nehéz, ne nagyon nehéz kérdés.

Felhívtad a pa ☐ a mamádát?

...hogya a legma legalacsonyabb iskolázottságúak...

Itt megfe ☐ megfigyeltek különböző érdekességeket...

...látszólag funkcionál ☐ funkciótlanok...

Más metódust próbálok meg a propa ☐ probléma elhárítására.

Megszólalt belőled a soven a soviniszta.

Mindig röhög, ré ☐ az interjúkban is mindig röhögött régen.

Nem jegye ☐ nem rendelsz jegyet?

Nem tanultak meg az ő gene ☐ korosztályuk angolul.

...némán, mint egy refl ☐ mon ☐ mu ☐ bemutatja...

Onnan megy masze ☐ masszírozni.

Például az ala ☐ a mondat első részében ismertnek véli az alanyt.

Statisztikával is izo ☐ izo ☐ igazolható a különbség.

Vannak nagyon szép fázisai ennek a kezdé öö fázisnak.

Viszont nem ettem meg a lektá ☐ a lekváros buktát.

5. „Nyelvem hegyén van” jelenség

Megakadás

...a tudatlan □ tudattalan □ tudatalatti...

Adnál egy □ ööö □ na, mi is az □ kanalat?

Annál az izénél, a □ na, izénél, munkásnál □ szerelőnél van a gépünk!

Értelmezze legyen szíves szociolo szocioli szociologikusan...

Ha kölcsönadod az izédet, amivel könyveket a könyvtárból, na □ tudod □ olvasójegyedet, akkor elintézem helyetted.

Már megint nézed a dobozt □ az izét □ azt □ mondjad már □ a tévét.

6. Perszeveráció

Megakadás

A kezemben volt a sminkecet, és kiesett

A polgáriasodásban nagy szerepe volt a berlini zsidóságban.

A politikai pártok patája után...

...a tanár és a tanuló szerepe, mert ez a könyv ezt is összeveszi...

Ady Endre drémonparádéja ő démonparádéja...

Bocs, de el kell lennem!

És akkor elbújtáj?

Ezt a beszélgetésben részt vevők hosszú sura □ sora igazolta.

Fel van bontva a tonthal! Szóval a tonhal.

Fizetni kell a hitelbírálátért.

Ha megadom a számod, felhívsz?

Két mondat van: az sz-típusú □ sz-típusú...

Majd a hallgató dekódal □ dekódol...

Nézd, egy sündisnó!

Országos artistaképző.

Sosem szórok szót a zöldségre.

Tessék beszálni!

Úgy tűnhet, hogy a fogalmakat felserélheti egymással.

Szándékolt közlés

kiesett

zsidóságnak

csatája

összeveti

mennem

elbújtál

hitelbírálátért

számom

sündisznó

artistaképző

sót

beszállni

felcseréli

Megakadás**Szándékolt közlés**

Úgy tűnik, mintha nemcsak a szerelmi bűbánat bűbánattal kapcsolatban lennének tapasztalatai.
Verbalításban jobbak a nőbbek nők.

7. Anticipáció**Megakadás****Szándékolt közlés**

A következő levél ööö szöveg az Ellopott levél.

A krapitalizmus mellett felhozott érvek...

kapitalizmus

A mágy másik felére megyek.

az ágy

Alig kepek fizetést.

kapok

Áruld már te, te mit csinálsz?

Áruld már el...

...és a Tőkéből vélt ☐ összerakni vélt ☐ gazdasági elképzeléseket...

Ha jól tudom, akkor két hét múrva lejárt volna a büntetése.

múlva

Hogy gondolom, illetve hogy gondold? Mert én úgy gondolom, hogy...

Holnap kétkor jönnek.

hétkor

...hozzáférést szegítő ☐ segítő szerkezet...

Jó nekem velem?

neked velem

Mi a probléma?

probléma

Nem, nem, nem elég konsztrasztos ☐ kontrasztos az elméletek közötti különbség.

...olyan előadóktól kérve segítságet, mint például...

segítséget

Temantikailag szemantikus ☐ vagyis tematikailag szemantikus kérdéskör következik.

8. Metatézis

Megakadás

A szakmája technéskertikus.
 Belelóg a szemem a hajamba.
 Egy válaszra ti ketten tudjátok a kérdést.
 És ezek után beletörtem a zárat a kulcsba.
 Ez károsíthatja a mesémet és a vájamat.
 Ide adod a kapalácsot?
 Kimentem a tetemőbe.
 Kucorral vagy édesítővel □ ööö cu-
 korral vagy édesítővel?
 Kucsd be kérlek!
 Ma lesz hullagesillás.
 Még hideg a dariátor □ a dariátor □
 a ra-di-á-tor, na.
 Még tart a lekrám?
 Megyek, bezárom a kaput, és kinyitom a kutyát.
 Nyáron szoktam sojtam jojózni.
 Reggelizett már a Frizsu?
 Tettél már kávét a cukrodba?

Szándékolt közlés

kertésztechnikus
 a hajam a szemembe
 kérdésre a választ
 a kulcsot a zárba
 a vesémet és a májamat
 kalapácsot
 temetőbe
 csukd be
 csillaghullás
 reklám
 bezárom a kutyát, és kinyitom a kaput
 sokat szoktam
 Fruzszi
 cukrot a kávédba

9. Egyszerű nyelvboltás

Megakadás

Ahhoz, hogy most tapasztalatot gyűj-
 tyenek gyűjtsenek...
 Arról van szó ööö arról van szó...
 Biztos, hogy van tapasztalatok...
 ...biztosítja az artikulációt lehetővé
 tesző levegőáramlást.
 Ezen a gépen írtam ezt a félrehallá-
 sok dokumentot.
 ...hogy az üzenet emberileg befo-
 gadhatóvá váljon...
 Kontolláld már magad egy kicsit!
 ...közel-kelti országok...
 Nem egyszerűen sruktuális vakság...

Szándékolt közlés

tapasztalatotok
 lehetővé tevő
 dokumentumot
 befogadhatóvá
 kontrolláld
 közel-keleti
 strukturális

Megakadás	Szándékolt közlés
Összevissza forgolódtunk, bementünk a behajtani tilosba fordító...	fordítva
Tegnap láttam a tévében a Kolcsonkért a Kölcsönkért a Kölcsönkért kastélyt.	
Tehát a befogó □ a befogadó miként szerepel a diskurzus világában?	
...terjedése erőteljes tencia.	tendencia

10. Több típusba sorolható jelenségek

Megakadás	Szándékolt közlés
A következő ülésen tudunk erről beszélni.	következő
...a másik pedig a megene □ megenele □ na, segítsenek, tegnap túl sok dolgozatot javítottam □ meg-e-le-ve-ni-tés.	
A szokástól elkérően azelőtt bokáig ért.	eltérően
A teniszbajnokság nyitó fordulójában...	fordulójában
...autópályamacica...	autópálya-matrica
Egye □ megye □ egy meggyes müzliszeletet kérek!	
Ennek van egy szociológiai következménye is.	
Ez meg annak tudható, hogy fiatal.	tudható be
Ez senkit se bátorítson el □ illetve bátortalanítson el.	
Nyitva vagyunk bármiféle más ötletre is.	nyitottak vagyunk
Olyan szép, amikor hullanak a fakelek □ falekevek □ falevelek.	
Ők vannak azok, akik □ ők azok, akik sokat együtt vannak.	
Tehát ott vannak a könyvespolcok □ folyton könyvespolcot akarok mondani □ könyvesboltok polcain.	
Tejet kérsz vagy tataót?	kakaót

Megakadás

Tele volt a macskának a kaja száraz-
zal.

Van egy olyan markoló, ami megeszi
□, dehogy, megemeli és felteszi a
betonkeverőt.

Szándékolt közlés

a macskának a tálja száraz kajával

A PEDAGÓGUSJELÖLTEK KAZINCZY-VERSENYE

A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 37. országos döntője

Kecskemét, 2009. április 17–18.

A nyelvújító Kazinczy Ferenc születésének 250. évfordulóján, 2009-ben a Kecskeméti Főiskola Tanítóképző Főiskolai Kar rendezte meg a Kazinczy Ferencről elnevezett 37. szép magyar beszéd verseny országos döntőjét. Az idei rendezvény újdonságnak számított, hiszen ebben az évben már nemcsak a hazai és határon túli pedagógusképző intézmények hallgatói küzdhettek a Kazinczy-emlékéremért, hanem a hittudományi és a jogi karok hallgatói is. A jeles eseményt negyvenhárom főiskolai és egyetemi hallgató és felkészítő tanáraik tisztelték meg jelenlétükkel. Közülük hárman a határon túlról érkeztek, harminchárman a pedagógusképző intézményeket képviselték, a versenykiírásban kibővített felsőoktatási intézmények közül pedig hét versenyző érkezett az alföldi városba.

A rendezvény hivatalos programja április 17-én a kar dísztermében Kuti István rektor-helyettes és Hegedűs Gábor dékán köszöntőjével vette kezdetét. A versenyt Gósy Mária, a versenybizottság elnöke nyitotta meg. Az ünnepélyes megnyitó után a versenyzők, a felkészítő kollégák és a rendezvény iránt érdeklődők tudományos előadásokat hallgathattak meg. Az előadások a hagyományoknak megfelelően kapcsolódnak a Kazinczy-versenyekhez, szakmai betekintést adnak a beszédkutatás egyes területeire. Elsőként Gósy Mária professzorasszony tartotta meg a *Szabálytalan „előzés” a beszédben* című előadását, majd Balázs Géza egyetemi tanár előadását hallgathattuk meg *Változások a média nyelvhasználatában* címmel. Ezután Horváth Viktória tudományos segédmunkatárs beszélt a *Hezitációs jelenségek a spontán beszédben* témáról, végül Bóna Judit egyetemi tanársegéd *A beszéd folyamat zavarairól* tartott előadást.

Délután a szabadon választott szövegek bemutatására került sor. A változatos szövegválasztás és a különböző felolvasási módok érdeklődést és az eseményhez méltó izgalmat váltottak ki a versenyzőkből, valamint az őket érdeklődéssel hallgató közönségből. A szakmai bírálóbizottság elnöke Gósy Mária (ELTE, Budapest) volt, a zsűri tagjai pedig Balázs Géza (ELTE, Budapest), Baukó János (Konstantin Egyetem, Nyitra), Bóna Judit (ELTE, Budapest), Havas Judit (Petőfi Irodalmi Múzeum, Budapest), Horváth Viktória (MTA Nyelvtudományi Intézet, Budapest), Molnár Ildikó (Pető Intézet, Budapest), Nábrádiné Jurasek Ibolya (Vitéz János Katolikus Főiskola, Esztergom), Raisz Rózsa (Eszterházy Károly Főiskola, Eger), és Steklács János (Kecskeméti Főiskola, Kecskemét) voltak.

A rendezvény első napját az esti állófogadás zárta, amely szintén a Kazinczy-verseny hagyományai közé tartozik. Ez az este lehetőséget adott az ismerkedésnek, a beszélgetéseknek is, és emlékeztetése tette a kétnapos rendezvényt. Vacsora közben hangulatos zenét és műsort biztosított a helyi néptáncsoport és zenekara, az est további hangulatát pedig a közös táncház fokozta.

Április 18-án, a verseny második napján délelőtt a kötelező szöveg, *Szabó Zoltán: Szerelmes földrajz* című műve egyik részletének bemutatása következett, erre a néhány hónappal korábban felújított Könyvtár olvasótermében került sor.

A verseny ünnepélyes eredményhirdetése délután a zsűrielnök szakmai értékelésével kezdődött. A zsűri észrevételeit összegezve a testület elnöke felhívta a figyelmet a helyes hangsúlyhasználatra, a keresztnévek és a családnevek főhangsúlyaira, a hangerő és hangszínezet használatára, a *s* és *sz* hangok helyes ejtésére. Az összefoglalóból az is kiderült, hogy az idei verseny színvonala az átlagosnál jóval magasabb volt. Ezután a jelenlevők Havas Judit tolmácsolásában meghallgathatták a szöveg egy lehetséges felolvasási módját. A kötelező szöveg nem volt könnyű, ezért a holtversenyes pontozás esetében a második forduló felolvasásainak értékelése határozta meg a jutalmazottak sorrendjét.

A zsűri döntése alapján 15 Kazinczy-érem kiosztására került sor. A 2009-es Kazinczy szép kiejtési versenyen Kazinczy-érmet kapott hallgatók névsora a következő: Uracs Miklós (Szegedi Tudományegyetem BTK), Bogár László (Kodolányi János Főiskola), Varga Nóra (SZTE JGYPK Hittudományi Intézet), Katona Ágnes (ELTE BTK), Nagy Edina (Zsigmond Király Főiskola), Pozsgai Zsófia (SZTE JGYPK Hittudományi Intézet), Pintér Noémi (Újvidéki Egyetem BTK), Rigó Balázs (ELTE BTK), Soós Enikő (Szegedi Tudományegyetem BTK), Bors Anikó (ELTE BTK), Túri Ágnes (Kecskeméti Főiskola), Scheffer Mónika (ELTE TÖFK), Kovács Vivien (PTE Illyés Gyula Főiskolai Kar), Lejtényi Emanuel (Esztergomi Hittudományi Főiskola), Orsós Brigitta (NyME Benedek Elek Pedagógiai Kar).

A zsűri elnök a Kazinczy Alapítvány nevében megköszönte a Kecskeméti Főiskolának a kétnapos rendezvény szervezését, lebonyolítását, majd a staféta átadására került sor. A verseny jelképe, a Kazinczy-plakett ezúttal Egerbe utazott, a 2010-es döntő leendő helyszínére.

Fehér Éva

Szabó Zoltán: Szerelmes földrajz részlet

A verseny kötelező szövege

Ha valaki céljául tűzné ki, hogy a jelesebb magyar költők hazai utazásainak vonalait térképen tünteti fel, aligha maradna betöltetlen folt. Költőink nemcsak politikai ösztönük helyességében, hanem utazásaik mennyiségében is lepipálják másfoglalkozású honfitársaikat.

Mennyit utaztak! Kazinczy Ferenc Tihanyrévnél fáklyafény mellett száll a hajóba, s ittas legények szövétnek mellett vezetik Szántódra, hol egy istenfi él, kit nem mulaszthat keblére ölelni. Jár Pannonhalmán, Egerben, Kassán, bejárja Nógrádot, Gömört és egész Erdélyt, erről könyvben be is számol. Kufsteinből Munkácsra menet, fogolyként utazza be Magyarországot nyugati szélétől a keletiig. Mikor Pozsonyba ér, ujját megcsókolja, majd a földhöz érinti azt, üdvözléssel a haza földjének.

Csokonai végigvendégeskedi a Dunántúlt, verset kiált a tihanyi ekhónak, iskolát jövendől Somogynak, a dámák farsangi diadaláról furcsa vitézi verseteket játszat Kaposvárt, s gondosan beszámol visszajöveteléről az Alföldön. Petőfi huszonhat évébe belefér a fél ország megismerése. Illetve: csaknem az egész ország megismerése. A versei alá írt helységnevekből egy igen figyelemreméltó helységnévtárt lehetne összeállítani. Alig van hely Magyarországon, melyet be ne járt volna, s a területet, melyet bejárt, egy legyezőformához hasonlítani. Forgója valahol a Kiskunságban van, egyik szára innen Pozsonyig, másik szára innen Segesvárig fut, pereme pedig a Kárpátok csipkés félköríve. Többnyire gyalog járta be e vidéket, két jó csikaján, vagyis az apostolok lován, valami fatális véletlen folytán rendszerint zimankós télben, havas esőben, gyakran betegen.

Ilyenféle gyaloglás Aranyak is kijut, mintha azt a lángelme kötelező érvényéül írta volna elő a kor. E keserves útjaikon ugyancsak realiztikus személyes élményeket kaptak a hazától, s ugyancsak éles emlékeket a tájról, mely az eső fátyola, a nap forrósága vagy a tél ködének szállongása alatt jelent meg a szemük előtt.

Osiris Kiadó, Budapest, 1999.

A 2009-es döntő kötelező szövegének értő-értető előadása

Mielőtt meghangosítanánk egy szöveget, először mindig meg kell értenünk azt. Csak ekkor lehetséges a szöveg helyes interpretációja, amely elősegíti, hogy a hallgató is biztosan megértse. A szöveg értelmezéséhez, befogadásához a tartalom mellett meg kell határozni a szöveg műfaját, stílusát, hangulát stb. is.

A szöveg interpretálásakor alapkövetelmény, hogy az olvasás ne mesterkéltséggel, inkább élőbeszédszerű legyen, a felolvasó ne változtassa meg természetes hangfekvését. Elengedhetetlen a jó beszédlevegzés, a tiszta artikuláció, a hibátlan olvasás, a helyes hangsúlyozás és intonáció, a megfelelő tempóváltások és a szünettartás, illetve a szövegegész gondolatainak érzékletes megjelenítése.

Szabó Zoltán *Szerelmes földrajz* című írásának idézett részletében a magyar költők hazai utazásairól szól. A szöveg bár négy bekezdésből áll, öt részre tagolódik, hiszen a harmadik bekezdésben két költőnk utazásairól is képet kapunk: Csokonai és Petőfi útjairól. A szöveg műfaja leírás, ezért az egyes bekezdésekben nem találunk tételekmondásokat; a leírás nem kívánja meg az összegzést, a bekezdések gondolatmenete ugyanis nem esik szét tételekmondatok nélkül sem. Tartalmi összegzésül a szöveg utolsó bekezdése szolgál. Az értelmezést nehezítik a régies kifejezések, ritkán használt szavak (például *szövétnék* 'fáklya'). A felolvasást, a helyes interpretációt pedig a hosszú, többszörösen összetett mondatok nehezítik meg.

Az első mondat három részre tagolódik, amelyet a központozás is jelez; a tagolódást főképp a szünettartással és az első szakasz végén ereszkedő, míg a második szakasz végén emelkedő intonációval valósíthatjuk meg. A második mondatban nehézséget okozhat a hangsúlyok helyes megvalósítása, a főhangsúlyokat a *helyességében* és a *menyiségében* szavak kapják.

A *Mennyit utaztak!* mondat meghangosításában figyeljünk arra, hogy intonációnk és hangszínezetünk ne csodálkozást fejezzon ki. A következő többszörösen összetett mondat tagmondatait szünettartással tagolhatjuk; a dallam az első három tagmondat végén lebegő; hangsúlyt kap a *fáklyafény*, a *szövétnék*, az *istenfi* és a *nem*. Ezután a Kazinczy által bejárt helyek felsorolása olvasható, amelynek elemeit nyomatékkal tagoljuk. A mondat második tagmondata előtt tartsunk szünetet. A bekezdés harmadik mondatát egy szólamként olvashatjuk fel, főhangsúlyt kap a *fogolyként*. A szakasz utolsó mondatában az első két tagmondat lebegő, a második kettő ereszkedő dallammal valósítható meg.

A következő mondat ismét többszörösen összetett, öt részre tagolható, amiben a vesszők segítenek. A felolvasásnál a tagolást kis szünetekkel, illetve az első és az utolsó tagmondat végén ereszkedő, a közbülső tagmondatok végén lebegő dallammal valósíthatjuk meg. Az erősebb nyomatékkal kiemелendő szavak a földrajzi nevek: *Dunántúlt*, *tihanyi*, *Somogynak*, *Kaposvárt* és *Alföldön*. Ezután Petőfi utazásairól olvashatunk. A két egyszerű mondatban a *fél*

és az *egész* kapják a főhangsúlyt. Megtévesztő lehet az *illetve* szó után álló kettőspont, azt sugallja, hogy tartsunk szünetet; ez azonban megtöri és megnehezíti a felolvasást, ezért olvashatjuk egy szólamban is az egész mondatot. Az ezt követő mondatban az *igen* partikula kapja a főhangsúlyt. A következő többszörösen összetett mondatokat az intonáció és a szünettartás mellett a tempóváltásokkal is tagolhatjuk. A hangsúlyos szavak az *alig*, *be*, *legyezőformához*, *Kiskunságban*, *Pozsonyig*, *Segesvárig*, *Kárpátok*. A bekezdés utolsó mondatában a gyalog értelmezőinek (*két jó csikaján, vagyis az apostolok lován*) a meghangosítása, illetve a mondat végén szereplő, jelzős szerkezetből álló felolvasás okozhat nehézséget. Az értelmezőket az intonációval, valamint tempóváltással tagolhatjuk; a *valami* előtt tartsunk szünetet, a felsorolás tagjainak felolvasásakor ügyeljünk arra, hogy a jelzőkre kerüljön a hangsúly.

Az utolsó, összegző bekezdést megvalósíthatjuk az előzőektől eltérő hangfekvésben is. A szakasz első mondatában hangsúlyt kap az *Aranyinak* és a *kötelező*. Az utolsó mondatban félreértésre adhat okot a kétszer is előforduló *ugyancsak* szó, amely itt nem 'is' jelentésben, hanem 'meglehetősen' jelentésben áll; ezért főhangsúlyt kap. A záró tagmondatban ismét a jelzős szerkezetek okozhatnak nehézséget, ezekben a jelzőkre helyezzük a nyomatéket.

Végezetül mindig tartsuk szem előtt, hogy egy szöveg interpretálásakor soha nem a felolvasó hangjának kell megszólalnia, hanem a szövegnek magának. Bár többféleképpen is felolvashatjuk a szöveget, mindig figyeljünk arra, hogy a meghangosítás adekvát legyen az értelmezéssel.

Bóna Judit

A MAGYAR ÉS A NEMZETKÖZI TUDOMÁNYOS ÉLET ESEMÉNYEI

Beszéd kutatás 2009 konferencia

Budapest, 2009. október 16–17.

2009 októberében került megrendezésre a Nyelvtudományi Intézetben a 8. Beszédkutatás konferencia, melynek témája a spontán beszéd volt. A kétnapos rendezvényen 35 előadást hallgathattak meg az érdeklődők, melyek más-más tudományterületek keretein belül vizsgálták a jelenséget.

Kenesei István intézetigazgató megnyitó beszéde után Sáry Gyula plenáris előadásában összefoglalta, milyen módszerekkel vizsgálták a beszéd és az agy kapcsolatát, mely agyterületek felelősek a beszédprodukciónak és a beszédértéssel összefüggő funkciókért. Az agyfélteke-dominancia vizsgálata, a Broca- és Wernicke-területekről kapott új ismeretek további kutatásokhoz szolgálhatnak kiindulópontul: többek között a kétnyelvűek beszédprodukciónak, a kétnyelvűség jelenségének jobb megismeréséhez. Ebben a témában több előadás is elhangzott: Judith Rosenhouse a magyar és a héber egymásra hatását mutatta be, Navracsics Judit a kódváltás és kódkeverés jelenségét tanulmányozta kétnyelvűeknél, Szende Virág pedig a magyart idegen nyelvként tanulókat beszédét vetette össze a nyelvtanulás korai és egy későbbi szakaszában az artikulációs és percepciók szint fejlődésének tükrében. Bánréti Zoltán és Mészáros Éva kutatásukban képekkel és hozzájuk kapcsolódó kérdésekkel vizsgálták, hogyan érinti az afáziás nyelvi sérülés a rekurzivitást attól függően, hogy a Broca- vagy a Wernicke-területen jelentkezik.

A beszéd minősége az életkor előrehaladtával változik: megkülönböztünk gyermek-, felnőtt- és öreghangot. Bóna Judit 8 idős és 8 fiatal női beszélő hanganyagát alapul véve a magyar nyelvre elsőként vizsgálta meg, hogyan függ össze a magánhangzók artikulációja a beszélő életkorával. Megállapította, hogy az életkor előrehaladtával szűkül az a frekvenciatartomány, melyben a magánhangzók realizálódnak, illetve igazolódott, hogy a magyar beszélők-nél is romlik a zöngé minősége idős korban, azonban a két életkori csoport adataiból megállapítható, hogy nagyok az egyéni különbségek. Balázs Boglárka előadásban a rekedtség hátterében megbúvó organikus és funkcionális zavarokra hívta fel a figyelmet.

A spontán beszéd jellemző sajátossága, hogy ép mentális működés és ép beszélőszervek mellett a fonéma gyakran nem az elvárt módon valósul meg az artikuláció során. Felvetődik a kérdés, hogyan képes a hallgató az eltérő ejtészvariációkat értelmezni. Gósy Mária az *akkor* szó ejtészvariációt vetette

össze és arra a következtetésre jutott, hogy egyazon beszélő két ejtése között is lehetnek nagy különbségek, ugyanakkor a különböző beszélők ejtése is lehet relatíve hasonló. Az elvárásokhoz képest a hangok temporális viszonyai-ban, a magánhangzók formánsszerkezetében, a mássalhangzók akusztikai szerkezetében, a zöngékezdési időben jóval nagyobb a variancia a spontán beszédben, ezért feltételezhető, hogy vannak olyan állandó jegyek, amelyek lehetővé teszik a lexikális hozzáférést. Gráczy Tekla Etelka, Markó Alexandra és Beke András szó belseji explozívák akusztikai vizsgálatával megállapította, hogy a képzési hely hátratulódásával nő a zöngékezdési idő, illetve az explozíva fonémák 25%-a nem zárhangként realizálódik. Beke András és Horváth Viktória a nazális koartikulációról tartott előadást. A magyarban a nazális *m* és *n* hangok eltérő mértékben változtatják meg, nazalizálják a szomszédos magánhangzót. Úgy tűnik, a nazalizáció a nyelv függőleges irányú mozgását érinti. Bárkányi Zsuzsanna, Kiss Zoltán és Mády Katalin a *v* kétarcú fonológiai viselkedéséhez hasonlóan azt tapasztalta, hogy szóvégi helyzetben a *z* is zöngétlenedik, a *sz–z* hangpár esetében azonban más összetevőknek köszönhetően nem szűnik meg a fonológiai kontraszt. Beke András és Szaszák György a magyar magánhangzók alaprealizációja és semleges ejtése közti különbségeket vizsgálta. A magyarban ezek a legbiztosabban a második formáns alapján különíthetők el. Vizsgálták továbbá, hogy a gépi osztályozás mely változók beiktatásával adja a legpontosabb eredményt. Bárkányi Zsuzsanna, Beke András, Bóhm Tamás, Csapó Tamás Gábor és Gráczy Tekla Etelka kutatásukban arra voltak kíváncsiak, hogy a szubglottális rezonanciák milyen mértékben alkalmasak a magánhangzók osztályozására. Vargha Fruzsina nyelvjárási korpuszon vizsgálta az *e* magánhangzó realizációját és olyan különbségeket vett észre, amelyek segítik az adott nyelvjárás jobb megismerését.

Minél többet tudunk meg arról, hogyan realizálódnak a spontán beszédben a mássalhangzók, hogyan változik a magánhangzók formánsszerkezete, annál több ismeretet lehet a gyakorlatban felhasználni. Rácz Zsuzsanna Zsófia egy épülő formánsadatbázis lehetőségeit mutatta be Olasz Gáborral és Bartalis Mátyással közös előadásában. Zainkó Csaba Csapó Tamás Gáborral és Németh Gézával közös kutatásukról számolt be, melynek célja, hogy életszerűbb szintetizált beszédet állítsanak elő. Tarján Balázs, Mihajlik Péter, Tüske Zoltán és Fegyő Tibor spontán és tervezett magyar beszéd gépi felismerésének különbségeit vizsgálták. Pápay Kinga az Egyesült Államokban egyre elterjedtebb VSA készülékekről tartott ismertetőt, melyek a hagyományos poligráfokkal szemben a beszédhang változásaiból mutatja ki a hazugság fiziológiai tüneteit.

Több előadás is foglalkozott azzal, hogyan jelöli a beszélő a spontán beszédben a szakaszhatárokat, illetve a megnyilatkozás végét, ezzel akár a beszélőváltás lehetőségét a diskurzusban. Mády Katalin előadásában ismertette, hogy a gyenge, az erős és a rendkívüli határokat mely jegyek alapján milyen

arányban képesek az emberek észlelni. A legbiztosabban a nyújtást ítélték határjelölőnek. Markó Alexandra eredeti spontán beszédfelvételek, akusztikailag szűrt és az eredeti felvételek lejegyzett változatán végzett percepciók tesztet. A részt vevő személyek csak az összes határ 14%-át jelölték mindhárom feladatban, a határészlelésben tehát a szintaxis és a szemantika meghatározóbb tényező, mint a prozódia, ugyanakkor az emberek szünetet és ereszkedő dallamot várnak ezeken a helyeken. Bata Sarolta spontán diskurzusokban vizsgálta, hogyan jelzik a felek a megnyilatkozás végét, kiemelten a hallgatás és az együttbeszélés gyakoriságát, időtartamát és a beszélőváltások időtartamát. Megfigyelte, hogy a hallgatások számában nincs nagyobb eltérés a diskurzus egyes szakaszaiban, azonban az utolsó másfél percben a felek hosszabb ideig hallgatnak. Az együttbeszélés a diskurzus végére egyre gyakoribb lett, bár a felek egyre többször vágtak egymás szavába, az esetek nagy részében ez együttműködő szándékkal történt. Jellemző volt, hogy az új beszélő valamivel előbb szólalt meg, mint ahogy az előző abbahagyta. A beszélőváltások gyorsaságában a magyar – Stivers eredményeit alapul véve – a japánhoz áll közel.

A spontán beszéd sajátossága, hogy a beszéd tervezése és kivitelezése egy időben zajlik. Ezen folyamatokban keletkezett zavarok elemzése segít mélyebb működésük megértésében. Gyarmathy Dorottya a perszeveráció jelenlétét vizsgálta aszerint, hogy mely nyelvi elemeket és mely szófajokat érint gyakrabban, valamint mekkora az az időbeli hatótávolság, amelyen belül a már kimondott elem újra aktiválódhat. Laczkó Mária középiskolások beszédében elemezte a különböző megakadástípusok előfordulását, Bakti Mária pedig szinkrontolmácsok spontán és tolmácsolt beszédét hasonlította össze, és azt tapasztalta, hogy a vizsgált személyek 75%-a több megakadást ejtett spontán beszédben, illetve más megakadástípusok domináltak a tolmácsolt és a spontán beszédben. A még folyamatban lévő beszédtervezésre utalnak a megakadásjelenségek mellett a töltelékszavak, töltelékelemek is. Dér Csilla Ilona ezen elemek előfordulási gyakoriságát és elhelyezkedését vizsgálta aszerint, hogy vajon csak a beszédtervezés megkönnyítését segítik elő, vagy más funkcióit is betöltenek.

Váradí Viola előadásában összefoglalta, milyen hasonlóságok és különbségek mutathatók ki a beszélő beszédében, ha hallott szöveget ismételt meg, vagy felolvas. Az olvasási hibák elemzése segíthet feltárni, hogy a hangos olvasás során a beszélőnek mely részfeladat okoz nehézséget. Imre Angéla és Bóna Judit 20–76 éves nők és férfiak hangos olvasását hasonlította össze. Az olvasás minőségét számos tényező befolyásolja – háttértudás, emlékezet stb. –, de szignifikáns különbség a nagymértékű szórás miatt csak a férfiak és a nők artikulációs tempójában mutatható ki. A bizonyos megakadások a beszélő bizonytalanságára utalnak, míg a helytelen tagolás és hangsúlyozás megértési problémát jelez. Fontos tanulság, hogy a gyakorlott olvasás nem feltétlenül jár jó szövegértési képességgel.

A pszicholingvisztika izgalmas kutatási területe az anyanyelv-elsajátítás. Koós Ildikó az anya és a csecsemő szemkontaktusait tanulmányozva érdekes információkkal szolgált a preverbális kommunikációról, melyek segítik a beszédfejlődés korai szakaszának mélyebb megismerését. A nemzetközi szakirodalom alapján úgy tűnik, az adott nyelvtől függ, hogyan érnek be a fonológiai tudatosság alsó szintjei. Míg az angolban a rímképzés, a magyar gyermekek esetében a szótagolás, a szótagokra bontott szavak felismerése indul meg először. Jordanidisz Ágnes 4–6 éves gyermekek ötletes szótagolási feladattal való vizsgálata során szignifikáns fejlődést mutatott ki a 4. és 5. életév között a szótagokra bontásban, míg hatéves kor után a hangokra való szegmentálásban észlelhető látványos fejlődés. Ha a gyermek valamilyen okból lemarad az anyanyelv-elsajátításban, beszédfejlődési zavarról beszélünk. Ennek általános jelei: szókincselmaradás, megértési zavarok, lassabb lexikai hozzáférés. Kas Bence és Lukács Ágnes nyelvfejlődési zavaros és ép fejlődésű óvodások vizsgálata során további nyelvspecifikus jegeket keresett, melyekkel könnyebben diagnosztizálható a zavar. Antalné Szabó Ágnes a tanórai kommunikációt tanulmányozta és a tanári értékelő megnyilatkozások kapcsán a pedagógusok gyakori hibáira hívta fel a figyelmet.

Gerebenné Várbiro Katalin előadásában a gyermeki spontán beszéd kutatásának módszertani kérdéseit vette sorra. Több előadó beszélt arról, hogy az adatközlők számának növelésével, a korpuszok bővítésével és a mintavételi technikák újragondolásával biztosabb eredmények érhetők el. Tatár Zoltán a fonetikai alapú beszélőazonosítás nehézségeit ismertette. Nikléczy Péterrel közös kutatásukban arra keresték a választ, hogy mely mintavételi eljárás lehet a leghasznosabb a kriminalisztikai fonetika számára.

A konferencia újdonsága a poszterszekció volt, melyben idén hét előadás szerepelt változatos témákban. Hegedűs Fanni Linda kétfázisú kísérletben vizsgálta a beszédhang és a személyiségjegyek kapcsolatát, Menyhárt Krisztina néprajzi kutatás keretében hasonlította össze a narratív műfajok temporális és szünettartási sajátosságait. Bíró Tamás a spontán beszédben előforduló performanciahibákat modellezte és mutatta be a holland zöngésségi hasonulás példáján, Oravecz Csaba és Sass Bálint a spontánbeszéd-adatbázisok építésének technológiai kérdéseit vették sorba. Neuberger Tilda azt vizsgálta, milyen mértékben javítják az óvodások és kisiskolások a megakadásaikat, Szentágotai Zita a sajátos nevelési igényű gyermekek iskolai kommunikációját és beszédfejlesztésük meglévő gyakorlatát tanulmányozta. Maurer Péter az egyén beszéde és a bűnözés kapcsolatát vizsgálta, kitérve azokra a környezeti tényezőkre, amelyek hozzájárulnak a bűncselekmény elkövetéséhez, illetve azokra, amelyek segítik a társadalomba való visszailleszkedést, a bűnisméltés lehetőségét csökkentik, de akár meg is akadályozzák, hogy az egyén bűnelkövetővé váljon.

A kétnapos rendezvény sok érdeklődőt vonzott, a színvonalas és sokféle témát érintő előadásoknak köszönhetően pedig mindenki megtalálta az érdek-

lődésének megfelelőt, új ismeretekkel is gazdagodott. Bár a spontán beszéd kutatásának csak az elején járunk, az élénk viták, hasznos észrevételek, meglepő eredmények a megkezdett munka folytatására ösztönöznek. Bebizonyosodott, hogy szükség van spontánbeszéd-korpuszokra, hiszen számos tudományterület kutatásaiban bizonyulhatnak hasznosnak.

Erdős Klaudia

A KÖTET SZERZŐI

Név	Intézmény	E-levél
Bakti Mária	SZTE JGYPK Modern Nyelvek és Kultúrák Tanszéke	bakti@jgytf.u-szeged.hu
Bata Sarolta	ELTE Fonetikai Tanszék	batasarolta@gmail.com
Bartalis Mátyás	BME Távközlési és Média-informatikai Tanszék	bartalis@tmit.bme.hu
Beke András	ELTE Fonetikai Tanszék	beke.andras@gmail.com
Bóna Judit	ELTE Fonetikai Tanszék	bona_judit@freemail.hu
Fehér Éva	Kecskeméti Főiskola	feher.eva@tfk.kefo.hu
Gósy Mária	MTA Nyelvtudományi Intézet	gosy@nytud.hu
Grácz Tekla Etelka	MTA Nyelvtudományi Intézet	graczi@nytud.hu
Gyarmathy Dorottya	MTA Nyelvtudományi Intézet	gyarmathyd@nytud.hu
Horváth Viktória	MTA Nyelvtudományi Intézet	horviki@nytud.hu
Imre Angéla	ELTE Fonetikai Tanszék	angelai@freemail.hu
Markó Alexandra	ELTE Fonetikai Tanszék	markoxa@ludens.elte.hu
Neuberger Tilda	ELTE Fonetikai Tanszék	tilda.neuberger@gmail.com
Olaszy Gábor	BME Távközlési és Média-informatikai Tanszék	olaszy@tmit.bme.hu
Rácz Zsuzsanna Zsófia	BME Távközlési és Média-informatikai Tanszék	zsuzska.racz@gmail.com
Szaszák György	BME Távközlési és Média-informatikai Tanszék	szaszak@tmit.bme.hu
Váradí Viola	ELTE Fonetikai Tanszék	varadiviola@gmail.com



2008 OKT 04

