

Dpl



BESZÉDKUTATÁS 2003

Elméleti és alkalmazott fonetikai
tanulmányok



MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest 2003

BESZÉDKUTATÁS 2003

**Elméleti és alkalmazott fonetikai
tanulmányok**

**Olaszy Gábor 60. születésnapjának
tiszteletére**

Szerkesztette:

Gósy Mária

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest 2003

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztő:
Menyhárt Krisztina

ISSN 1218-8727
© MTA Nyelvtudományi Intézet

Felelős kiadó: Kenesei István igazgató



Olaszy Gábor 60 éves

Szakmai életrajz

Olaszy Gábor a József Attila Gimnáziumban érettségizett Budapesten, 1961-ben. Ezt követően a Budapesti Műszaki Egyetem Villamosmérnöki Karán folytatta tanulmányait (1962–1967), ahol villamosmérnöki oklevelet szerzett. Az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriumában 1974 óta dolgozik. A BME Távközlési és Telematikai Tanszékén 1984-től vesz részt a kutatásban és az oktatásban, valamint a PhD-képzésben. 1989-ben megszerezte a nyelvtudomány kandidátusa fokozatot, majd 2003-ban az MTA doktora címet.

Kutatási területei: a beszéd akusztikai szerkezete, a szegmentális és a szuprasegmentális elemek kutatása, többnyelvű szöveg-beszéd átalakító beszéd szintetizáló rendszerek tervezése, a hullámforma-szintézis fonetikai tervezése.

Számos tudományos konferencián vett részt belföldön és külföldön, illetve tartott előadást.

Tagja az MTA Akusztikai Komplex Bizottságának (2003-tól titkára), valamint az ISCA International Speech Communication Associationnak.

Az International Journal of Speech Technology c. folyóirat szerkesztőbizottsági tagja.

Publikációi: megjelent 2 könyve, 8 könyvrésze, 56 tudományos cikke angolul és 39 magyarul. Nyolc magyar szabadalom társszerzője.

Olaszy Gábor válogatott publikációi

Olaszy Gábor (1999): Gépi beszéd-keltés információs rendszerekhez Magyarországon. Akusztikai Szemle III/1-3: 4-13.

Olaszy, Gábor – Németh, Géza (1999): IVR for Banking and Residential Telephone Subscribers Using Stored Messages Combined with a New Number-to-Speech Synthesis Method. In: Human Factors and Voice Interactive Systems. Ed.: Gardner-Bonneau, D. Kluwer Academic Publishers, 237-256.

- Olaszy Gábor (1999): Beszédatadbázisok készítése gépi beszédelőállításához. In: Beszédkutatás '99 Szerk.: Gósy Mária MTA Nyelvtudományi Intézet. Budapest, 68-89.
- Olaszy, Gábor – Németh, Géza – Olaszy, Péter – Kiss, Géza – Zainkó, Csaba – Gordos, Géza (2000): Profivox – a Hungarian TTS System for Telecommunications Applications. *International Journal of Speech Technology* 3-4: 201-215.
- Olaszy, Gábor (2000): The prosody structure of dialogue components in Hungarian. *International Journal of Speech Technology* 3-4: 165-176.
- Olaszy, Gábor – Koutny, Ilona – Olaszy, Péter (2000): Prosody prediction from text in Hungarian and its realisation in TTS conversion. *International Journal of Speech Technology* 3-4: 187-200.
- Olaszy Gábor (2000): Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. In: Beszédkutatás 2000. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 26-38.
- Olaszy Gábor et al. (2000): Profivox – a legkorszerűbb hazai beszéd-szintetizátor. In: Beszédkutatás 2000. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 167-179.
- Olaszy, Gábor – Koutny, Ilona (2001): Intonation of Hungarian questions and their prediction from text. In: *Proceedings of Prosody 2000*. Eds.: Puppel, S. – Demenko, G. Poznan, 179-196.
- Olaszy Gábor (2001): Prozodémák fonetikai reprezentációja. In: Beszédkutatás 2001. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 28-45.
- Olaszy, Gábor – Németh, Géza – Olaszy, Péter (2001): Automatic Prosody Generation – a Model for Hungarian. *Proc. of the Eurospeech 2001/1*: 525-528.
- Olaszy, Gábor (2002): The most important prosody patterns of Hungarian. *Acta Linguistica Hungarica* 49/3-4: 277-306.
- Olaszy, Gábor (2002): Model to predict Hungarian sound durations for continuous speech. *Acta Linguistica Hungarica* 49/3-4: 321-345.

- Nikléczy Péter – Olaszy Gábor (2002): Kempelen beszélőgépének rekonstruálása. In: Beszédkutatás 2002. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 5-17.
- Olaszy Gábor (2002): A magyar kérdés dallamformáinak és intenzitás szerkezetének fonetikai vizsgálata. In: Beszédkutatás 2002. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 83-99.
- Olaszy Gábor (2002): Az alapfrekvencia és a hangsúlyozás kapcsolata a magyarban. In: Kísérleti Fonetika, Laboratóriumi Fonológia 2002. Szerk.: Hunyadi László. Debreceni Egyetem Kossuth Kiadója. Debrecen, 73-84.

*Boldog születésnapot és további szakmai
sikereket kívánunk!*

A fonetikusok

Tartalom

Gocsál Ákos – Huszár Ágnes: Csábító hangok	9
Gósy Mária: Virtuális mondatok a spontán beszédben	19
Grassegger, Hans: A hangminőség szubjektív és objektív megítélése	44
Markó Alexandra: Az idegen nyelvi beszédtervezés stratégiái	57
Menyhárt Krisztina: Zöngésedési és zöngétlenedési folyamatok a /j/ fonéma realizációiban	75
Navracsics Judit: Szólehívás vizuális stimulus alapján	89
Nikléczy Péter: A zöngé periódusidejének funkciója a hangszínezetben	101
Olaszy Gábor: A magyar szóalakok hangidőtartam-térképei	113
Tóth László – Kocsor András: A Magyar Telefonbeszéd-adatbázis (MTBA) kézi feldolgozásának tapasztalatai	134
Varga László: Dallami meghatározottságok a magyar mondatban (a topik dallama)	147
Vicsi Klára – Szaszák György: A magyar nyelv kiejtésvariációi és felhasználásuk a beszédfelismerésben II.	163
A 15. Fonetikai Világkongresszus (Gósy Mária)	177
8. Beszédkommunikációs és Beszédtechnológiai Konferencia: Eurospeech '03 (Olaszy Gábor)	179
Megakadások a spontán beszédben konferencia (Menyhárt Krisztina)	180
A kötet szerzői	182

CSÁBÍTÓ HANGOK

Gocsál Ákos – Huszár Ágnes

Bevezetés

A beszéd akusztikumában nagy egyéni különbségek figyelhetők meg. Ezek fonetikai mérőeszközökkel jól kimutathatók, de a mindennapi megfigyelés számára is érzékelhetők. Ha megkérdezzük embereket, hogyan hat rájuk ismeretlenek hangja, sokszor szinesztéziás vagy metaforikus jelzők segítségével érzékeltetik benyomásaikat. Beszélnek például *éles, kemény, bársonyos, lágy, meleg, hideg, rideg* hangról. Ez azt mutatja, hogy a beszédhang akkor is emocionális reakciót vált ki a hallgatóból, ha nem látja a beszélőt. A hangjátékszínesz karakterformáláskor a saját beszédhangjában rejlő lehetőségeket aknázza ki, illetve alakítja úgy, hogy a megformálandó szereplőt idézze hallgatói elé. A színész által megjelenített karakterről a hang alapján a hallgatók egymással jól harmonizáló benyomásokat fogalmaznak meg (Fónagy 1995).

A mindennapi megfigyeléseken túlmenően azonban a beszéd paramétereinek és a beszélő személy bizonyos tulajdonságainak összefüggéseivel viszonylag kisszámú vizsgálat foglalkozott. Még kevesebb azon kutatások száma, amelyek arra vonatkoznak, hogy milyen beszédparaméterek alapján milyen személyiségjegyeket tulajdonítanak a hallgatók a nem látható, csak hallható beszélőnek.

A beszélő hangjának és bizonyos biológiai jellemzőinek összefüggéseivel foglalkozó kutatások közül megközelítésének és módszereinek komplexitásával kiemelkedik Gósy (2001) a testalkat és az életkor becslésére vonatkozó kutatása. Az bizonyítottan látszik, hogy a hangból viszonylag megbízhatóan lehet következtetni a beszélő életkorára (Gocsál 1998), viszont kevésbé megbízhatóak az illető testmagasságára és testtömegére vonatkozó becslések (Gósy 2001).

Más kutatások a beszéd bizonyos jellemzői – például gyorsasága – és a beszélő valós személyiségjegyei között próbáltak meg összefügg-

géseket megállapítani. A beszédtempónak és a személyiségjegyeknek az összefüggését bizonyította Feldstein és Sloan (1987) vizsgálata. Azok a kísérleti személyek, akik az Eysenck Personality Inventory alapján extrovertáltaknak minősültek, gyorsabban beszéltek, mint akik introvertáltaknak (idézi Langenmayr 1997, 269).

Más kutatások nem a beszélő valódi tulajdonságai és beszéde közti összefüggést vizsgálták, hanem azt, hogy a hallgatók milyen jegyeket tulajdonítanak a beszélőnek, függetlenül attól, hogy ez mennyire fedti a valóságot. A beszédtempó és a beszélő személynek tulajdonított jegyek összefüggéseire vonatkozó kutatások közül érdemes felidézni egy 1976-os, angol nyelvű beszélőkkel folytatott vizsgálatot, amelynek eredményei alapján a kísérleti személyek a gyorsabb beszédtempót és a tiszta hangképzést a fiatalabb életkorra jellemzőbbnek gondolták (Hartman-Danhauertől idézi Gósy 2001, 485). Magyar beszélők körében folytatott vizsgálatában Gósy összefüggést mutatott ki a beszédhang alaphangmagassága és bizonyos feltételezett biológiai jellemzők, így az életkor, illetve testalkat között: „Anyagunk alapján kimondható, hogy a nőknél a magasabb alaphangmagasság a fiatalabb életkorral, illetőleg a sovány vagy normál testalkattal mutat szoros összefüggést.” A férfiaknál a mélyebb alaphangmagasság alapján ugyanakkor fiatalabb életkorra következtetnek a megkérdezettek (Gósy i. m.)

Bonyolultabb összefüggésekre mutat rá Buller és Aune 1992-es vizsgálata, amelynek során a kísérleti személyeknek arra kellett választ adniuk, rokonszenves-e számukra a csak hangja alapján megismert beszélő. Ebben a vizsgálatban 263 kísérleti személlyel hallgattattak meg kilenc (mesterségesen) különböző sebességűvé alakított beszédmintát. Megvizsgálták a választ adó személyek beszédtempóját is. Az eredmény: a kísérletben megkérdezett személyek a saját beszédtempójukhoz hasonló tempóban beszélőket találtak rokonszenvesnek. Más kérdésre válaszolva ugyanezek a kísérleti személyek a gyorsabban beszélőket hitelesebbnek, kompetensebbnek vélték, mint a lassabban beszélőket (idézi Langenmayr 1997, 270).

Az eddigiekben ismertetettek szociálpszichológiai szempontból is értelmezhetők. Amikor egy másik emberrel kapcsolatba kerülünk, összbnyomást alakítunk ki róla. Az eleinte strukturálatlan benyomások ún. burkolt személyiségelméletté szerveződnek bennünk, ezek alapjául egy mentális sablon szolgál. A George Kelly által az 1950-es években kidolgozott elmélet (Forgács 1994) lényege az, hogy kétpólusú konstruktumokat – ezeket szemléletesebben mércéknek lehetne nevezni – „illesztünk” az észlelt személyre. Ilyen konstruktum például a „szeret – nem szeret”, „jópofa fickó – unalmas fráter”, „határozott – határozatlan” stb.

Kelly elméletének segítségével megvizsgálható az, hogy a beszéd fonetikai sajátosságai mennyiben felelősek a beszélőről kialakuló benyomásért. Ha beszédmintákat játszunk le hallgatóknak, majd a szemantikus differenciál módszerével értékeltetjük velük a hallott beszédmintát, akkor olyan észlelési mechanizmusokat indítunk meg bennük, amilyenek természetes körülmények között is működnek. A kísérlet során a kutató adja meg az értékelés szempontjait, míg spontán észlelés esetén ezek maguktól adódnak. További fontos különbség, hogy a kísérlet során felhasznált hangminták fonetikai tulajdonságait kvantitatív módon is tudjuk jellemezni, így egzakt összefüggéseket kereshetünk a beszéd fonetikai paraméterei és a kialakult benyomás egyes jellemzői között.

Gelfer (1993) vizsgálata során 17 tulajdonságpár alapján, a párok között elhelyezett 9 fokozatú skála segítségével értékeltette hallgatóival a lejátszott, női beszélőktől származó hangmintákat. Ennek során kiderült, hogy akár szakértők (beszédpatológusok), akár képzetlen személyek hallgatták a hangmintákat, az *alaphangmagasság* és a *beszédtempó észlelése* mindkét esetben fontos és egymástól elkülönülő jellemzője volt a hallgatók percepció mechanizmusainak. Gelfer kutatásából az is kiderült, hogy a beszédészlelés szempontjából gyakorlott személyek percepció folyamatai öt különböző dimenzió mentén jellemezhetők, míg a képzetlen személyek esetén csupán két dimenzió jelenik meg.

A jelen kutatás

Saját vizsgálatunk egy kísérletsorozat első része. Arra keressük a választ, hogy milyen személyiségvonásokat tulajdonítanak kísérleti személyek a hang alapján olyam beszélőknek, akiket nem látnak. Kísérletünkben női beszélők hangját ítélték meg férfi hallgatók.

A következő hipotéziseket fogalmaztuk meg:

1. A beszéd gyorsasága és a vélelmezett (szakmai) kompetencia között pozitív összefüggés áll fenn.
2. Az alaphang magassága és a személynek tulajdonított szexiség között pozitív összefüggés áll fenn.

Bár azt feltételeztük, hogy a hang többet is elárul a személyiségjegyekről és a beszélő személy pillanatnyi állapotáról, ezeknek a feltárására irányuló vizsgálatok komoly módszertani nehézségekkel járnak. Egyet kell értenünk Gósy Máriával abban, hogy „A beszéd számos további sajátosságról is tartalmaz információt, mint a szociális státus, a személyiség vagy az érzelmi állapot, ezeknek a megfeleltetését a beszéd egyes paramétereivel azonban egyelőre még nem sikerült elvégezni.” (Gósy 2001, 479).

A kutatás menete, a résztvevők és a kutatás módszerei

A kutatás első lépéseként 2-3 perces spontánbeszéd-mintákat rögzítettünk tizenegy női beszélőtől. Mindegyik beszélő középosztálybeli származású, magyar anyanyelvű, 20-26 év közötti főiskolai hallgató. A beszédminták rögzítésekor fontos szempont volt, hogy a beszélők egy számukra érzelmileg semleges témáról beszéljenek, ezért mindannyian az általuk jól ismert pécsi Széchenyi teret mutatták be.

A kutatás második lépésében a beszédmintákat lejátszottuk harmincegy 20-25 év közötti férfinak. Mindannyian főiskolai hallgatók, középosztálybeli származásúak, magyar anyanyelvűek. A hallgatóktól azt kértük, hogy egyes beszédminták meghallgatása után az 1. ábrán látható űrlapon a megfelelő skálaérték bekarikázásával jelezzék, milyennek gondolják a beszélőt hangja alapján. A hallgatók minden beszédmintához külön űrlapot kapott.

Az űrlapon található tulajdonságok – a későbbi feldolgozáshoz – két csoportra oszthatók. Hipotéziseinknek megfelelően megtalálhatók rajta a *vonzó nő*, illetve a *kompetens munkavállaló* tulajdonságai.

fiatal	1	2	3	4	5	nem fiatal
okos	1	2	3	4	5	nem okos
szexis	1	2	3	4	5	nem szexis
természetes	1	2	3	4	5	mesterkéltnak
nyugodt	1	2	3	4	5	nyugtalan
érthető	1	2	3	4	5	nehezen érthető

1. ábra

A hallgatók által kitöltött űrlap

A kutatás során megmértük az egyes beszédminták legfontosabb paramétereit is, azaz az alaphangmagasságot, a beszédtempót, az artikulációs tempót és az artikulációs hatásfokot (ez utóbbit vö. Gocsál 2000). Ezeket a Speech Analyzer v1.5 program segítségével határoztuk meg.

Eredmények

Összefüggések a vélt tulajdonságok és az egyes fonetikai paraméterek között

Az 1. táblázatban foglaltuk össze a hallgatóktól kapott válaszokat. A táblázat az egyes beszélőkre és tulajdonságokra vonatkoztatva a 31 hallgató által megadott skálaértékek átlagait mutatja be. Az űrlappal összevetve például megállapítható, hogy a hallgatók a legfiatalabbnak a 6. sz. beszélőt, a legidősebbnek a 2. számút tartották.

A 2. táblázat az egyes beszédminták akusztikai paramétereit tartalmazza. Az artikulációs hatásfokot Gocsál (2001) alapján számítottuk ki. Az artikulációs hatásfok egy arányszám, amely azt mutatja meg, hogy a beszédminta teljes időtartamát 1-nek véve, annak hányad részét fordította a beszélő artikulációra. Az alacsony hatásfokértékek

esetén a szünetek aránya magas, tehát ez az adat a szünettartások összidőtartam-arányának jellemzésére is alkalmas.

1. táblázat: A hallgatók által megadott skálaértékek átlagai

Beszélő	Fiatal	Okos	Szexis	Természetes	Nyugodt	Érthető
1.	2,25	2,93	3,22	2,64	2,54	2,32
2.	3,35	2,48	3,77	2,67	2,58	2,22
3.	2,03	2,58	2,9	2,32	2,8	2,71
4.	1,74	3,19	2,77	2,51	2,87	2,29
5.	3,32	2,87	3,61	2,22	2,06	2,48
6.	1,67	2,74	2,58	2,64	3,06	2,32
7.	2,25	3,64	3,25	3	2,8	3,22
8.	3,12	3,29	4	2,61	2,74	3,54
9.	2,67	2,41	3,83	3	2,51	1,96
10.	2,16	4,06	3,77	2,83	2,8	3,41
11.	2,03	2,38	2,35	2,41	2,25	2,03

2. táblázat: A kutatásban szereplő személyek spontán beszédének fonetikai paraméterei

Beszélő	Alaphang (Hz)	Artikulációs tempó (h/s)	Beszédtempó (h/s)	Artikulációs hatásfok
1.	200	13,29	7,8	0,59
2.	180	12,64	8,2	0,65
3.	188	11,62	8,28	0,71
4.	225	13	10,8	0,83
5.	170	14,23	11,12	0,78
6.	230	12,85	7,93	0,62
7.	220	12,33	8,38	0,68
8.	190	15,57	10,23	0,66
9.	195	13,01	9,31	0,72
10.	215	12,41	7,99	0,64
11.	200	11,57	9,2	0,80

Az akusztikai paraméterek és a hang alapján a beszélőre jellemzőnek vélt tulajdonságok között az összefüggéseket a korrelációs számítás módszerével határozzuk meg. A korrelációs számítást és a klaszteranalízist az SPSS 10.0 for Windows program segítségével végeztük.

Négy esetben kaptunk szignifikáns korrelációt. Az **alaphangmagasság** pozitív korrelációt mutatott a nyugodtsággal és az okossággal, negatív korrelációt pedig a fiatalsággal és a szexisséggel. A legerősebb korrelációt az alaphangmagasság és a fiatalság-ítélet között találtuk ($r = -0,489$), ennél gyengébb összefüggés mutatkozott az okossággal ($r = 0,216$) és a szexisséggel ($r = -0,233$). Még kisebb korrelációs együtthatót kaptunk a nyugodtság ($r = 0,189$) esetén. A korrelációs együttható előjeléből arra következtethetünk, hogy a magasabb alaphangú beszélőket a hallgatók nyugtalanabbnak, kevésbé okosnak, ugyanakkor fiatalabbnak és szexisebbnek vélték a mélyebb alaphangúaknál. Megfordítva ezt az állítást: a mélyebb alaphangú beszélők nyugodtságot, okosságot érzékeltettek magukról, miközben őket kevésbé tartották a hallgatók fiatalnak és szexisnek.

Az **artikulációs tempó** három tulajdonsággal mutatott szignifikáns korrelációt, mégpedig az érthetőséggel, a fiatalsággal és a szexisséggel. Mindhárom esetben $p < 0,01$ szintű szignifikanciát kaptunk, és a korreláció előjele pozitív volt. A legerősebb összefüggést az artikulációs tempó a fiatalsággal mutatta ($r = 0,345$). Gyengébb korrelációs együtthatót kaptunk a szexisség ($r = 0,282$) és az érthetőség ($r = 0,158$) esetén. Összességében elmondható, hogy a lassabban artikuláló beszélőket érthetőbbnek tartották a hallgatók, ugyanakkor meglepő és a korábbi vizsgálatoknak ellentmondó eredmény az, hogy a gyorsabb artikulációjú beszélőket idősebbnek gondolták. További érdekes adat az, hogy a hallgatók a lassabb artikulációjú beszélőket tartották inkább szexisnek, mint a gyorsan beszélőket.

Az **átlagos beszédtempó** két tulajdonsággal mutatott szignifikáns korrelációt. A beszélő vélt életkorával hasonló összefüggést találtunk, mint az artikulációs tempónál, azaz nagyobb beszédtempó értékek idősebb beszélő benyomását keltették ($r = 0,2$, $p < 0,01$). Ennél gyengébb, negatív korreláció mutatkozott a nyugodtsággal ($r = -0,111$,

$p < 0,05$). A nyugodtság azonban semmilyen összefüggést nem mutatott az artikulációs tempóval, ezért ebből arra következtethetünk, hogy az eltérés oka a kétféle tempóérték számítási módjában lévő különbség – a beszédtempó kiszámításakor figyelembe vesszük a szüneteket is, az artikulációs tempó esetén nem. A több szünettel beszélő személyeket tehát nyugodtabbnak vélték a hallgatók, függetlenül attól, hogy milyen artikulációs tempóval beszéltek.

Az **artikulációs hatások** minden tulajdonsággal negatív korrelációt mutatott, szignifikáns, ugyanakkor meglehetősen gyenge összefüggést azonban csak az érthetőség ($r = -0,151$, $p < 0,01$) és a szexiség ($r = -0,158$, $p < 0,01$) esetén találtunk. Ez azt jelenti, hogy a több szünettel beszélők érthetőbbek és szexisebbek voltak a hallgatók szerint.

A hallgatók benyomásainak vizsgálata

A kapott eredmények alapján felmerül az a kérdés is, hogy vajon a hallgatókban a kutatás során használt tulajdonságok az észlelési folyamatok során milyen rendszerbe szerveződnek (2. ábra). Ennek megállapításához a klaszteranalízis módszerét alkalmazzuk (vö. Falus–Ollé 2000, 269).

NYUGODTSÁG	-+-----+-----+			
TERMÉSZESSÉG	-+-----+-----+			
ÉRTHETŐSÉG	-----+-----+			+-----+
OKOSSÁG	-----+-----+			I I
SZEXISÉG	-----+-----+			+ I
FIATALSÁG	-----+-----+			

2. ábra

A klaszteranalízis eredményeként kapott dendrogram

Az ábráról leolvasható, hogy a legszorosabban a nyugodtság és a természetesség függ össze, azaz akit nyugodtnak tartottak a hallgatók, azt természetesnek is – ami arra utal, hogy tulajdonságok közt is szoros a kapcsolat. Hasonlóképpen erős összefüggés található az érthetőség és az okosság, illetve azok ellentétjei között. Megfigyelhető, hogy a négy tulajdonság az ábrán külön ágat alkot. A szexiség nem kapcsolódik ilyen szorosan egyik fenti tulajdonsághoz sem, nagyjából

azonos erősségű a korrelációja velük (a korrelációs együtthatója az érthetőséggel 0,198, az okossággal 0,178, a természetességgel 0,162, mindhárom esetben $p < 0,01$). A fiatalság kizárólag a szexisséggel mutat korrelációt ($r = 0,378$, $p < 0,01$), a többi tulajdonsággal nem.

Következtetések

Visszatérve kutatásunk hipotéziseihez, az alábbi megállapításokat tehetjük. Első hipotézisünk a vélelmezett kompetencia és a beszéd gyorsasága között nem igazolódott. Szignifikáns korrelációt találunk az artikulációs tempó és a beszélő személy szakmai kompetenciája (okossága, érthetősége) között. A lassabban artikuláló beszélőket (ők a „megfontoltak”) a hallgatók nagyobb valószínűséggel találták érthetőbbnek, okosabbnak, mint a gyors beszédűeket (a „sietősek”). A szünettartások nagyobb arányú előfordulása is elősegíti azt, hogy érthetőbbnek, okosabbnak tartsa a hallgató a beszélőt. A mélyebb alaphang az okosság benyomását keltette.

Második hipotézisünk – az alaphang magassága és a szexiség közötti pozitív összefüggés – igazolódott. A magasabb alaphangú beszélőt csakugyan szexisebbnek gondolták a kísérletben szereplő fiatal férfiak. Ez jól illeszkedik a fiatalosság és a szexuális vonzás bizonyított összefüggéséhez. Eredményeink szerint a kísérletünkben részt vevő férfiak a lassabban beszélő, több szünetet tartó nőket nagyobb valószínűséggel ítélték vonzóbbnak. Ezt úgy értelmezzük, hogy a megfontolt beszédstílusból sugárzó szakmai kompetencia és a szexiség nem zárják ki egymást, mint ahogy a közhelyes gondolkodás szökenő-viccekben megnyilvánuló paneljai sugalmazzák.

A statisztikai elemzés alátámasztotta azt a feltételezést, hogy a különböző beszélőktől származó hangminták különböző benyomásokat váltanak ki a hallgatókban, a verbális üzenettől függetlenül. Kimutattuk, hogy egyes akusztikai paraméterek hogyan befolyásolják a benyomások kialakulását. Különösen nagy jelentőségűnek tűnik a beszédszünetek szerepe. Az adataink azt mutatják, hogy a több szünettel, megfontoltabban beszélő nők kedvezőbb benyomást tudnak magukról kialakítani. Más oldalról megközelítve a kérdést, megállapít-

ható az eredmények alapján, hogy léteznek különböző sztereotípiáink, amelyeket azonnal aktiválunk, amint meghalljuk valakinek a hangját. Véleményt alkotunk a beszélő személyiségéről és különböző egyéb tulajdonságairól, anélkül, hogy ismernénk vagy akár csak látnánk őt. Kutatásunkban csak néhány egyszerűbben mérhető fonetikai paramétert vizsgáltunk meg, ám nagyon valószínű, hogy sok más beszédjellemző – kiváltképp a hangszínért felelős paraméterek – is fontos szerepet játszik ezekben a folyamatokban.

Megítélésünk szerint az ilyen jellegű kutatások lehetővé teszik a beszédkommunikáció és általában a személyközi kommunikáció tágabb, több szempontú vizsgálatát, s a hétköznapi kommunikációs helyzetek pontosabb megértését.

Irodalom

- Falus Iván – Ollé János (2000): Statisztikai módszerek pedagógusok számára. Okker. Budapest.
- Fónagy Iván (1995): A hangkarakterológia esélyei. ÁNYT. XVIII: 23-41.
- Forgács József (1994): A társas érintkezés pszichológiája. Gondolat. Budapest.
- Gelfer, M. P. (1993): A multidimensional scaling study of voice quality in females. *Phonetica* 50: 15-27.
- Gocsál Ákos (1998): Életkorbecslés a beszélő hangja alapján. In: *Beszéd kutatás '98*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 122-135.
- Gocsál Ákos (2000): A beszéd időviszonyai különböző életkorú személyeknél. In: *Beszéd kutatás 2000*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 39-50.
- Gocsál Ákos (2001): Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszéd kutatás 2001*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 61-72.
- Gósy Mária (2001): A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125: 478-488.
- Langenmayr, A. (1997): *Sprachpsychologie*. Verlag für Psychologie. Göttingen.

VIRTUÁLIS MONDATOK A SPONTÁN BESZÉDBEN

Gósy Mária

Bevezetés

A beszélők és hallgatók a kommunikáció során mintegy ösztönösen, ritkábban tudatosan betartják az ún. együttműködési elvet (Gósy 1999). Ennek az a lényege, hogy a beszélő igyekszik úgy beszélni, hogy a lehető legnagyobb mértékben biztosítsa a hallgató számára az elhangzottak feldolgozhatóságát; a hallgató pedig úgy igyekszik működtetni a beszédészlelési és beszédmegértési folyamatait, hogy az elhangzottakat mind pontosabban, a közlő eredeti szándéka szerint dekódolja. Az együttműködési elv azonban a mindennapi kommunikációban gyakran sérül, sokszor anélkül, hogy akár a beszélő, akár a hallgató tudatában lenne ennek. A beszédtapasztalat során kialakult korrekciós stratégiák – mindkét mechanizmus esetében – lehetővé teszi a kommunikációban résztvevők számára, hogy az együttműködési elv tökéletlen megvalósulása esetén is megtörténjen az információcsere. Különösen nagy a jelentősége mindennek a spontán beszédben, illetőleg a spontán narratívák feldolgozása során, mivel a spontán beszéd nem olyan tökéletesen szerkesztett, mint az előre megfogalmazott szövegek meghangosítása. Továbbá a tervezés és kivitelezés nagymértékű egyidejűsége, illetőleg átfedése következtében, a produktum nyelvileg nem feltétlenül mintaszerű.

Az elhangzó hosszabb közlések megértésének és értelmezésének egyik alapfeltétele a megfelelő tagolás. A gondolatfolyam nyelvi formába öntésekor – a beszédprodukció ún. mikrotervezési szakaszában (Levelt 1989) – a beszélő már igyekszik megtervezni a tagolásokat egyfelől a saját gondolatainak szemantikai, logikai váza mentén, másfelől a majdan elhangzó közlések fonetikai interpretációjának tervezésében. A spontán beszéd eddigi elemzési eredményei azt mutatták, hogy az artikulációs megvalósítás túlnyomórészt ösztönös, a beszélő nincs feltétlenül tudatában annak, hogy mely szegmentális vagy szupraszegmentális tényezőt alkalmazza tagoló funkcióban, illetőleg meg-

lehetősen nagyok az egyéni különbségek (vö. Kohler 1983). A hierarchia ún. alsóbb nyelvi szintjein a beszélőnek nincs nagy választási lehetősége a prozódiai megvalósítást illetően (vö. a szegmentális szerkezet fonológiai struktúrája). A magasabb szinteken azonban opcionális lehetőségei vannak, s ezeknek az aktuális együttes kiválasztása számos tényező függvénye. A korábbi elemzések kimutatták, hogy a mondat szintaktikai struktúrája hatással van ugyan a prozódiai megvalósításra, de nem határozza meg azt (Frazier et al. 2003).

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. Ennélfogva sokkal nagyobb mértékben képes a hangosítás jellemzőire figyelni, illetőleg azokat tudatosan megvalósítani. Megfelelő központoszással ellátott, felolvasott szövegekben a mondathatárok percepcióis jelölése 90%-osnál valamivel jobb egyöntetű eredményt hozott (Batliner et al. 1998).

Spontán beszédkor a beszélő a beszédprodukciós folyamatnak csak a legvégző szakaszában – az artikulációs tervezés szintjén – jut el a közlés számos artikulációs sajátosságának az összerendezéséhez, így például a tagoláshoz. Gyakorlatlan beszélők a verbális nyelv tagolásainak tudatos megvalósítására nemigen képesek, azt fiziológiai tényezők, például a légzés, sokkal inkább szervezik, mint a szöveg tartalmi és formai egységeinek tudatos jelzései. A hallgató ugyanakkor mindvégig arra kényszerül, hogy valamiképpen tagolja, bizonyos egységekre bontsa az elhangzottakat, s ezeknek az egységeknek mentén ismerje fel az összefüggéseket. A tagolási egység lehet mondatnyi hosszúságú közlés, avagy annál hosszabb, összefüggő szövegrész. Lehiste ez utóbbit ’bekezdésnek’ (vö. paragraph) nevezi (1979), annak ellenére, hogy tudatában van, a szakkifejezés elsősorban az írott szövegek bizonyos egységeinek (tipográfiai) elkülönítésére szolgál. Úgy véli, a spontán beszédben is léteznek ezek a nagyobb, összefüggő egységek, amelyeket magasabb szintű egységeknek nevez.

A mondatok, illetőleg a bekezdések mint egységek szintaktikailag és szemantikailag meghatározható összefüggésrendszert mutatnak, és a szuprszegmentumok jellegzetes érvényesülésén keresztül különül-

nek el (pl. Schafer et al. 2000). Ilyen szupraszegmentum a szünet (mint jelkimaradás), az alaphangmagasság változása, az intenzitás csökkenése, valamint a szünetet megelőző szó időtartamának növekedése. Az intenzitás csökkenését és a frekvencia változását mondathatárokon spontán diskurzusokban igazolták (Hird–Kirsner 2002). A mondatvég temporális változása, az utolsó szó (szavak) lassabb artikulációját más kísérletek is megerősítették (pl. Lindblom 1978). Lehiste és Wang eredményei szerint a bekezdéseket nagy arányban azonosan ismerték fel angol anyanyelvű hallgatók (1977).

A határjelzés kérdése a beszédben nemcsak a mondathatárokkal vagy az annál nagyobb egységekkel kapcsolatos, hanem például a szóhatárokkal is. A szóhatár jelzése univerzális, még inkább talán fiziológiai jellegű sajátosság, amely azonban aktuális megvalósulásában nyelvspecifikus sajátosságokat mutat és mind a beszélőknél, mind a hallgatóknál különbözőképpen érvényesül. A szóhatárjelzés megjelenhet a szegmentum, illetőleg a szótag időtartamának változásában, az alaphangmagasság módosulásában, sőt bizonyos koartikulációs hatásokban, marginálisan még jelkimaradásokban is. A szóhatár felismerése az anyanyelv-elsajátítás során alakul ki, elsősorban a disztribúciós szabályok, másrészt a fonotaktikai sajátosságok megtanulása révén (Brent–Cartwright 1996). Ez a kiindulás a szöveg nagyobb egységeinek szegmentálásához.

Német beszéd elemzése alapján arra a következtetésre jutottak, hogy a közlések szövegkohéziója különböző szintek egységeinek többé-kevésbé hierarchikus egymásra épülésén alapszik, s ezen a módon jönnek létre a nagyobb egységek (Kohler 1983, 129). A szövegkohézió nem a szintaktikai struktúrában vagy a szintaxis és a prozódia együttjárásában keresendő, hanem a szemantikai meghatározottságban, illetőleg a szemantika és a prozódia összefüggéseiben. Két mondat következhet egymás után anélkül, hogy bármilyen információs kapcsolat lenne közöttük, ugyanakkor a prozódiai megvalósításban nem különböznek a kohéziót mutató, összetartozó mondatkapcsolatoktól. A szemantikai kohézió önmagában nem eredményez szövegegységet, azt a beszélő szándéka szerinti szövegkohézió fogja létrehozni, amely a szemantikai és a prozódiai sajátosságok együttes megvalósulása.

Hipotézisünk szerint a spontán közlések egysége valamiképpen kötődve van a beszélőben, és a beszélő mint hallgató ezt az intuitív tudást felhasználja a beszédfeldolgozás során tagolásra. Hosszabb szöveg esetén ennek a legkisebb egységnek a mondatnak kell lennie, függetlenül attól, hogy a 'mondat' nyelvészeti definíciójával az átlagos beszélő/hallgató nyilvánvalóan nem rendelkezik. A 'mondat' szót tehát jobb híján használjuk a spontán beszéd alapegységeinek a meghatározására. Valójában virtuális mondatokról kell beszélnünk, a 'virtuális' jelzőt a 'benne rejlő' értelmében használva. Ezek a virtuális mondatok – feltételezésünk szerint – többé-kevésbé azonos paraméterekkel jellemezhetők, s ezeket a paramétereket a hallgatók képesek tagoló funkcióban használni spontán közlések hallgatásakor. A fonetikai és a pszicholingvisztikai szakirodalomban nagyobb szintaktikai/szemantikai egységek elkülönítésére a spontán beszédben a 'közlés' szót használják, amely állhat egyetlen szóból vagy – a tradicionális értelemben vett mondatot használva – akár 3-4 mondatból. A jelen kutatáshoz azonban a 'közlés' terminus használata túlzottan tág, nem definiálható egyértelműen; lényegében tehát ugyanolyan problémát jelent a használata, mint a mondaté. Másfelől pedig sokkal nehezebb egy adott kísérletben a résztvevőknek az instrukciót megfogalmazni. Azt kellene ugyanis mondanunk, hogy egy hosszabb közlést fog hallani, amelyben a kisebb (?) közléseket keressük... (vö. a 'megnyilatkozás' magyarázatát is: Wacha 1988). Nem véletlen, hogy a spontán beszéd elemzésével foglalkozók felvetik egy egészen másfajta grammatika, illetőleg grammatikai leírás lehetőségét a spontán beszéd elemzésére (Fábricz 1988). A továbbiakban tehát – jobb híján – maradunk a virtuális mondat kifejezés használatánál.

A kérdésünk az, hogy a beszélő vajon hogyan tagol az artikulációs kivitelezésben mondatszínten, továbbá a hallgatók feldolgozó mechanizmusa mennyire egyöntetűen működik az elhangzottak észlelésében? A hipotézisünk ugyanis az, hogy ha a mondattagolás relatíve nagymértékű egyöntetűséget mutat, akkor a virtuális mondatok kognitív szinten egyértelműen léteznek – függetlenül a terminus problémás voltától, avagy a grammatikai definíció nehézségeitől. Kimondható-e tehát, hogy a 'virtuális mondat' valóban létezik, akusztikai-fonetikai korrelátumai vannak, s a hallgatók ezek individuális stratégiai felhasznál-

nálását végzik a spontán beszéd szegmentálásakor? Kísérletünk célja annak a feltételezésnek az elemzése (igazolása vagy elvetése) volt, hogy léteznek virtuális mondatok, amelyeket a szünetek és az alaphangmagasság aktuális értékétől függően észlelnek a kísérleti személyek. Felmerülő kérdés az is, hogy van-e mindennek bármilyen kapcsolata a spontán szöveg megértésével.

Anyag, módszer, kísérleti személyek

A tesztanyag egy magnetofonszalagra rögzített spontán narratív szöveg volt. A beszélő 35 éves nő (beszédhibája nincs), a felvétel készítésének tudatában volt. A CSL 4300B típusú jelfeldolgozó segítségével végeztük a temporális sajátosságokra vonatkozó méréseket, továbbá az F0 változásainak objektív adatolását a szüneteket megelőző és követő szóban, valamint a kísérleti személyek által mondathatárnak jelölt valamennyi további helyen. Szemantikai és szintaktikai kritériumok alapján meghatároztuk a szöveg mondatait, a szerkezethatárokat, a grammatikai anomáliákat, a megakadásjelenségeket.

A szöveget lejegyeztük (helyesírásban), de központosítás nélkül, és ezt a két oldalas tesztlapot kapták meg a kísérleti személyek. A résztvevők feladata az volt, hogy kétszeri meghallgatás alapján jelöljék meg a tesztlapon egy függőleges vonallal azokat a helyeket, ahol véleményük szerint a beszélő szándéka szerint mondatvég volt. Semmilyen egyéb magyarázatot vagy definíciót a mondat fogalmára nem adtunk, s a kísérleti személyeknek lehetőségük volt akár az első elhangzás, akár a második elhangzás alkalmával jelölniük. Hangsúlyoztuk azonban, hogy spontán beszédet fognak hallani.

A kísérleti személyek – célzottan – két csoportba kerültek. Az egyik csoportot magyar szakos bölcsészhallgatók (átlagéletkoruk 21 év) alkották (ELTE, Budapest), valamennyien nők. A másik csoportot pedig mérnökhallgatók (átlagéletkoruk 21 év) a budapesti Műszaki és Gazdaságtudományi Egyetemről, valamennyien férfiak. Mindkét csoportban 28-an voltak. Feltételezésünk szerint – éppen tanulmányaik különbözősége miatt – a bölcsészhallgatóknak mások a kódolt ideáik a mondatról, mint a mérnökhallgatóknak. A feladat egyik csoportnak sem okozott nehézséget, s a hallgatók többsége mindkét csoportban már az első elhangzáskor jelölte a vélt mondathatárokat.

A mondatjelölést követően még egy feladatot kaptak a kísérleti személyek, amelyre azonban korábban nem készítettük fel őket. Őt, megértést ellenőrző kérdésre kellett válaszolniuk az elhangzottakkal kapcsolatosan. A kérdések a szöveg jelentősebb részleteire vonatkoztak, amelyeket az elhangzottak megértéséhez nélkülözhetetlenek ítéltünk. Négy kérdés megválaszolása egy-két szóval megoldható volt; egyetlen kérdés igényelt egy mondatnyi kifejtést (ez a kérdés: „*Hogyan zajlott az esküvő Egyiptomban?*”). A kérdéseket a kísérletvezető tette fel, és a résztvevőknek egy külön lapra kellett a válaszokat leírniuk.

A spontán szöveg fonetikai jellemzése

A rögzített hanganyag terjedelme 5 perc 10 másodperc, a kiejtett szavak száma 550. A beszélő artikulációs tempója: 107,84 szó/perc, beszédtempója 141,02 szó/perc. A szövegben előforduló 100 ms-nál hosszabb szünetek száma 135, időtartamuk szórása: 100-1350 ms, az átlag 532,44 ms. A csendes szüneteken túl hezitálások is találhatók a szövegben, ezek elsősorban az [ə] magánhangzó különböző időtartamú változatai, néhány esetben a [m] és [hm] is előfordul. A hezitálás és a néma szünet különféle kombinációja 15 alkalommal fordult elő.

1. táblázat: A spontán tesztszöveg szünetei (sz.=szerkezetháron, n.=nem szerkezetháron)

Szünet idő- tartama (ms)	Típusa (db)					
	néma		kitöltött		kombinált	
	sz.	n.	sz.	n.	sz.	n.
100-300	13	13	1	7		
310-500	12	8		7		1
510-800	24	5		9	3	1
810-1000	6	1			2	3
1010-1200	3	1		2	2	2
1210-1300	5	1		2	1	
Összesen	63	29	1	27	8	7

Az összes nem tartalmaz jel (különféle szünetek) időtartama: 71,88 s. Szerkezethatáron összesen 72 alkalommal, míg nem szerkezethatáron 63 esetben tartott szünetet a beszélő. Leggyakoribbak a néma szünetek voltak, majd a hezitálások, ezekhez képest ritkábban fordultak elő kombinált szünetek. Az 1. táblázat a szünetek időtartamának és számának összefüggéseit szemlélteti.

A beszélő dallamformái megfelelnek az elvártaknak, átlagos alaphangmagassága 198,5 Hz, az F0 terjedelme: 150-400 Hz. Az intenzitás szint stabilnak mondható, sehol nincsenek feltűnő erősödések vagy elhalkulások. A hangsúlyok megvalósítása nem mond ellent a nyelvi szabályoknak. A szöveg a néma és kitöltött szüneteken kívül tartalmaz megakadásjelenségeket, amelyek közül jellegzetesek a nyújtások, különösen a névelőkön, valamint a szóban tartott szünetek. A névelőnyújtásra 10, egyéb nyújtásra további négy adatunk van. Az [ɔ] és [ɛ] magánhangzókön tapasztalt nyújtások időtartama megközelíti a 300 ms-ot. A beszélőre nem volt jellemző a szünet előtti szó vagy szavak időtartamának megnyújtása; jellegzetesen az utolsó szótag magánhangzója szenvedett nyúlást, de ez nem érvényesült következetesen. Sok esetben éppen a szünet előtt tapasztaltuk a szavak vagy az utolsó szó tempójának növekedését. A szöveg a következő megakadásjelenségeket tartalmazza (példákkal). Ismétlés: *hogyan jön ide, hogyan jön ide*; egyszerű nyelvbőlcs (artikulációs kivitelezési probléma): *találkozóhoz találkozóhoz* vagy *máris és más nevük* (javítás nélkül); ismétlés morfológiai változtatással: *a neve* (452 ms néma szünet) *nevén keresztül* vagy *testrésznél testrésszel*; agrammatizmus: *sokkal magas szociális* (ti. sokkal magasabb szociális); szünet a szóban: *kultusz* (187 ms) *ból* vagy *nép* (354 ms) *zenében*; téves kezdés: *ter teherbeesés*; újrazdés: *pro prostituáltak*; változtatás: *rendszeres öö általános*.

Eredmények

A kapott adatokat a következő szempontok szerint elemeztük: a virtuális mondatok jelöléseinek sajátosságai (kvantitatív és kvalitatív eredmények); a jelölések és a szünetek összefüggésének elemzése; a jelölések és az alaphangmagasság összefüggésének elemzése; a szövegértés elemzése.

Virtuális mondatok jelölése

A két csoport összes virtuális mondatvég jelölésének száma 58. Ebből egyöntetű jelölés 49, ez azt jelenti, hogy az adott helyen mindkét csoportból jelöltek mondatvéget. Összesen 9 esetben tapasztaltunk különbséget a két kísérleti csoport között. A bölcsészek összesen 52 helyen, a mérnökhallgatók összesen 55 helyen észleltek virtuális mondathatárt a szövegben; a fiúk és a lányok között e tekintetben gyakorlatilag nincs különbség. Ezt a statisztikai elemzések is alátámasztották. A bölcsészhallgatók átlagosan 23,75 mondatot, a mérnökhallgatók átlagosan 26,11 mondatot jelöltek, a lányok csoportjában a szórás: 14-34, a fiúknál pedig 8-43. Nem választva szét a kísérleti személyek két csoportját, a virtuális mondatok átlaga 24,92, az átlagos eltérés pedig 6,14. Az adatokban feltűnő az, hogy a kísérleti személyek igen nagy arányban ítélték ugyanazon a helyen mondatvéget. Az egyöntetűen jelölt 49 lehetőségből 16 helyen teljesen azonos az ítéletek száma (ez a jelölések 32,7%-át jelenti)

Egy személy eltéréssel 28,57%-ban ismét egyöntetű a jelölés, két személy eltéréssel további 8,16%-ban. Ez összesen 69,38%, vagyis azt mondhatjuk, hogy az anyanyelvűek virtuális mondatészlelése csaknem 70%-ban azonosnak tekinthető (ezt azért is állíthatjuk, mert a két hallgató az összes résztvevőnek csupán 7,14%-a). Adataink azt mutatják, hogy 3-4 hallgató különbséget 8 helyen tapasztaltunk (ez 16,32%); s mindössze 14,3%-ban tekinthetők nagy különbségűnek a két csoport közötti eltérések (ez a szöveg hét helyén fordult elő). A számadatok arra engednek következtetni, hogy egyfelől a hallgatók ugyanazon paramétereket használhatták döntéseikben, illetőleg, hogy az azonos anyanyelvűek (s azonos életkorúak) nemtől függetlenül virtuálisan nagy hasonlósággal tagolják az elhangzó szöveget. A lányok és a fiúk jelölése között szignifikáns különbség nincs. Anyagunkban a lányok 3 helyen jelöltek mondatvéget, ahol a fiúk nem, a fiúk pedig 6 olyan helyen, ahol a lányok nem (ezek elemzésére visszatérünk).

A tesztanyagban 9 olyan helyet találunk, ahol a kísérleti személyek nem jelöltek virtuális mondatot, pedig szintaktikailag, illetőleg szemantikailag szerkezetathárta estek a beszélő által tartott szünetek. Ebből három eset valódi szerkezetathárta volt, két esetben az és kötőszóval elválasztva (a szünetidőtartamok: 360-650 ms). Hat esetben pedig alá-

rendelő összetett mondat határán fordultak elő a szünetek, vonatkozó névmási kötőszók, illetőleg a *hogy* szomszédságában. A relatíve hosszú szünetek ellenére – 187-860 ms, átlag 553 ms – a kísérleti személyek nem észleltek mondathatárokat.

A kísérleti tesztanyagot szemantikai és szintaktikai kritériumok alapján négy csoportba soroltuk azokon a helyeken, amelyeket a kísérleti személyek megjelöltek: a) valódi mondatvég, b) kötőszó következik, c) szerkezethatár, de nem tekinthető valódi mondatvégnak d) szerkezet belseje. (A továbbiakban a 'valódi mondat' kifejezést szemben a 'virtuális mondattal' akkor használjuk, ha egyébként félreérthető a szövegtörnyezetben, hogy a percepció ítéletről van szó, avagy nem.) Példákkal illusztráljuk az egyes kategóriákat, a központoszással segítve a szövegértést. Valódi mondatnak tekintettük a következőket: „*A másik jellegzetesség a mozgás, ami eltér az európai kultúrától.*” vagy „*Eredetileg termékenységi kultusból indult ki.*” Példák olyan közlésrészletekre, amikor kötőszó következik, és a kötőszó előtt mondatvéget jelöltek a kísérleti személyek (az egyértelműség kedvéért a kötőszót félkövérítettük): „*Az ókorban ezeknek szent helyük volt, tehát így sokkal magas szociális státuszban voltak.*” vagy „*Az nem elsősorban dallamon, hanem ritmuson alapul, és a ritmus is egészen más, mint mint az európai zenében.*” Szerkezetek, amelyek határán mondatvéget jelöltek a kísérleti személyek: „*Inkább nézzük a következő produkciót, /mondatvég jelölése/ nem felejtve a gyökereket, most elsősorban a szórakoztató célú hástáncot fogják látni.*” vagy „*Annyiban tér el, és ezt majd a következőkben tapasztalni fogják, /mondatvég jelölése/ a zene, a keleti zene az európai zenétől...*”. A szerkezeten belül ítélt mondatvégre példák: „*Egy másik aspektusával, de szintén a szüléssel kapcsolatos /mondatvég jelölése/ az úgynevezett berber törzsi szülési ceremónia.*” vagy „*A francia gyomortáncról kezdve a görög, úgynevezett ciftetelli, ami egy török ritmusnak a neve /mondatvég jelölése/ nevén keresztül...*”. A kötőszók, amelyek előtt mindkét csoportban mondatvéget jelöltek a hallgatók, mellérendelő kötőszók, leggyakrabban az *és* és a *tehát*. A *tehát* szó esetében feltételezhetjük azt is, hogy nem kötőszó funkcióban, hanem töltelékszóként használta a beszélő. A percepció eredmény azonban – az adatok szerint – független az esetleges funkciókülönbségtől.

A szünetek elemzése

A spontán beszéd szünetei kapcsolatban vannak a szintaktikai szerkezettel, ennek az összefüggésnek a pontos meghatározása azonban nem egyszerű (vö. Duez 1993). A pszicholingvisztikai kutatások kezdeti vizsgálataiban azt találták, hogy a mondatokat elkülönítő szünetek közel 80%-ának az időtartama 500 ms-nél hosszabb volt, míg a szerkezeteken belüli szünetek mintegy 66%-ánál rövidebb, mint 500 ms (Goldman-Eisler 1972). Francia és angol interjúk összehasonlítása azt mutatta, hogy a szünetek mintegy 70%-a nagyobb frázisegységek határára esett, s ezek hosszabbak voltak, mint a frázisokon belüliek (Grosjean-Deschamps 1975). A szünetészlelési eredmények azonban nem csupán a szintaktikai szerkezettel, hanem más prozódiai paraméterekkel is mutatnak összefüggést (Duez 1985; Gósy 2000).

A 2. táblázatban összegeztük az időtartamok és a virtuális mondatok jelölésének összefüggéseit (a néma szünetet két esetben hezitálás követte, ezt plusz jellel kapcsoltuk a néma szünet értékéhez), azokban az esetekben, amikor a szemantikai és szintaktikai elemzések alátámasztották, hogy „valódi” mondat hangzott el.

A szünetek szórása 205-1350 ms közötti, az átlag 793,458 ms. Az összes kísérleti személyt figyelembe véve a 24 lehetséges helyen átlagosan 79,97%-ban észleltek mondatot. Két esetben volt teljesen egyöntetű a vélemény, ekkor a szünetek időtartama 1212 ms (néma), illetőleg 1127 ms (néma szünet, amit egy 224 ms-os hezitálás követett) volt. További négy helyen 98%-on felüli egyöntetű jelölést kaptunk, a szünetidőtartamok itt is relatíve hosszúnak mondhatók: 680 ms-tól 1062 ms-ig, az átlag 946,09 ms. A 69% és 87% közötti egyöntetű ítéleteknél a szünetek átlaga 735,25 ms; az 50% és 59% közötti ítéletek esetén a szünetátlag 612 ms. Az 50%-osnál kisebb egyöntetű ítéleteket perцепciónál már egyáltalán nem tekinthetjük relevánsnak, ekkor a szünetek 291 ms-tól 423 ms-ig szórnak, az átlag: 323 ms. Az adatok alátámasztják a szünet időtartamának meghatározó szerepét a mondatvég észlelésében. Minél hosszabb szünetet tart a beszélő, annál inkább az a hallgató benyomása, hogy gondolati egységet zárt le, vagyis megjelenik a virtuális mondat vége.

2. táblázat: A szünetidőtartam és a mondatvég jelölésének összefüggése (az előfordulás sorrendjében)

Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
397	53,5	205	41,07
812	98,21	1040	94,64
996	87,5	1263	98,21
1038	91,07	710	98,21
812+250	96,42	1212	100
860	58,92	590	55,35
680	98,21	900	91,07
600	57,14	423	46,42
700	69,64	710	80,35
530	94,64	1150	91,07
290	39,28	1127+224	100
1224	92,85	300	85,71

A 3. táblázat a szünetek és a szerkezethatárok közötti összefüggést szemlélteti. A beszélő ezeken a helyeken már jóval rövidebb szüneteket tart, s két esetben megjelenik a hezitálás is (mindkettőnél a néma szünetet követően). A mondathatár-jelölések nagyobb szórást mutatnak, de a két csoport között itt sincs lényeges eltérés. A szünetidőtartamok szórása 150-830 ms, az átlag 465,87 ms.

3. táblázat: Szünetek és szerkezethatárok

Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
632+198	67,85	540+240	12,5
491	1,78	282	8,92
229	46,42	150	1,78
665	17,85	300	1,78

A kötőszó előtti szünetek is nagymértékben keltették a virtuális mondat végének benyomását a hallgatókban. Ezeknek a szüneteknek a szórása hasonló a szerkezethatáron tartottakéhoz, 150-800 ms, az át-

lagérték azonban hosszabb, ami azt jelenti, hogy a tartományon belül valamivel nagyobbak az időtartamok, mint a szerkezethatáron mértéké (vö. 4. táblázat). Az átlag: 509,87 ms.

4. táblázat: A mondatvég észlelése kötőszó előtt

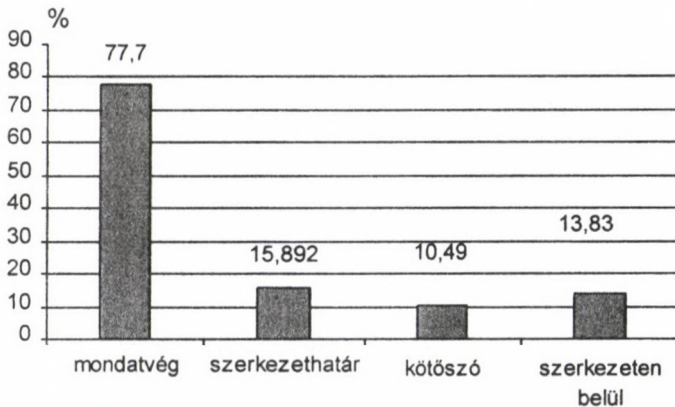
Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
515	14,28	nincs szünet	3,57
386	7,14	587	16,07
473	1,78	603	3,57
521+284	1,78	791	37,5
665	35,7	347	1,78
334	23,21	150	7,14
597	12,5	540	16,07
644	5,35	286	1,78
435	1,78		

Sem a mondathatároknál, sem szerkezethatároknál nem tapasztaltuk, hogy a hallgatók virtuális mondatot jelöltek volna és nem volt szünet. Az ilyen típusú szüneteket Duez 'szubjektív szüneteknek' nevezi (1993, 23). A hallgató valamilyen belső percepció tagolás következtében jelkimaradást érzékel annak valóságos akusztikai manifesztálódása nélkül. A kötőszók esetében azonban egy ilyen eset előfordult, ahol két hallgató is jelölt mondatvéget. A kötőszó itt a *tehát* volt. Összegeztük az egyértelműen szerkezeten belül megjelenő mondatjelöléseket (5. táblázat).

5. táblázat: Virtuális mondatok jelölése szerkezeten belül

Szünet (ms)	Mondatvég- észlelés (%)	Szünet (ms)	Mondatvég- észlelés (%)
452-498	8,92	nincs szünet	3,57
nincs szünet	3,57	584-450	64,28
679	8,92	633	3,57
257	3,57	705	14,28

A kísérleti személyek jelölték mondathatárt szintaktikai szerkezeten belül is, mindössze nyolc esetben, ez az összes határjelölés mintegy 14%-a, tehát elenyésző arány. Érdekes mégis, hogy mi eredményezheti néhány hallgatóban a közlésegység végének benyomását olyan helyzetben, amikor mind a szemantikai, mind a szintaktikai sajátosságok ellentmondanak ennek. A táblázatban látható szünetek átlaga 709,66 ms, vagyis relatíve hosszúnak mondhatók. A határértékek: 257-1034 ms. Két esetben a néma szünetet hezitálás is követi, ezáltal a szegmentális jelkimaradás időtartama jelentősen megnövekszik. A hallgatók több, mint 60%-a ott jelöl mondatvéget, ahol egy toldalékválasztási hiba miatt tart a beszélő hosszabb szünetet, és nem egyértelmű, hogy a folytatás önálló egység kezdete, avagy a megelőző szerkezet bizonytalan folytatása. Itt a beszélő alaphangmagassága leesik 170 Hz-re, s ez ismét a virtuális mondat befejezésének benyomását eredményezi (szemben a másik hosszú szünet esetén, ahol az alaphangmagasság kismértékű emelkedését tapasztaljuk). Két helyen jelölték a kísérleti személyek mondatvéget, ahol objektíve nem volt szünet, vagyis ismét előállt a 'szubjektív szünet' helyzete. Mindkét esetben a beszélő által „elkövetett” szintaktikai hiba okozhatta a zavart. Az 1. ábra a virtuális mondatvég jelölésének átlagait szemlélteti.



1. ábra

A szemantikai, szintaktikai kritériumok és a virtuális mondatok átlaga

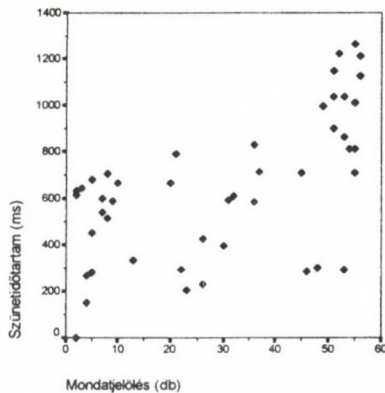
Ha a négy – szemantikai/szintaktikai kritériumok szerint elkülönített – helyzetet összehasonlítjuk, akkor közöttük jelentős különbséget tapasztalunk, ezt szemlélteti az 1. ábra. A statisztikai elemzéseket az ANOVA ún. Welch-próbájával végeztük, eszerint a szemantikai/szintaktikai csoportok szerint a különbség 1%-os szinten szignifikáns (72,867, $df_1=3$, $df_2=19,086$, $p < 0,0001$). A 4 kategória post hoc tesztekkel történt összevetése is megerősítette, hogy a „mondatvég” a döntő tényező. A hallgatók tehát igen nagy mértékben támaszkodnak a tartalomra és a formára együttesen a virtuális mondatathárók felismerésében.

Az adatok alapján elmondható, hogy a közlések tartalma meghatározó a virtuális mondat megítélésében, vagyis adott esetben mintegy felülbírálja a formát. A szünetek hossza ugyanakkor valamilyen mértékben befolyásolja a döntést. Kimondható, hogy minél hosszabb a szünet, annál jobban kelti a hallgatóban a mondatvég benyomását. Ez magyarázza azt a tényt, hogy a nyelvi és jelentésbeli egyértelműség ellenére a valóságos mondatvégek jelölésében is csak akkor kaptunk 90%-os vagy 100%-os egyöntetű ítéletet, ha a beszélő által tartott szünet tartama relatíve hosszú volt. Ez az érték átlagosan 900 ms. Korábbi szünetészlelési kísérletek – amelyekben funkciótól függetlenül kellett a beszélő által tartott szüneteket jelölni – azt igazolják, hogy már 500 ms tartamú néma szüneteket is képesek a hallgatók 80-100%-ban felismerni (Gósy 2000). Ez azt támasztja alá, hogy a szünet fontos, de nem egyedüli tényező a virtuális mondatok azonosításában. A szünet-időtartamok és a mondatthár-ítéletek közötti összefüggést korrelációszámítással ellenőriztük. A Spearman-féle korrelációs együttható pozitív és közepesen erős ($r=0,663$, $p=0,01$). Ez azt jelenti, hogy az esetek nagy hányadára igaz az, hogy a nagyobb szünetidőtartamhoz a virtuális mondatthárók nagyobb száma társul (vö. a 2. ábra pontdiagramját).

A szünetek időtartama különbözik a megállapított szemantikai, szintaktikai kategóriák szerint. Kimondható, hogy a szünetek időtartama azokon a virtuális mondathelyeken a leghosszabb, amelyek egyben szemantikai és szintaktikai kritériumok szerint is valóságos mondatvégek; az eredmény szignifikáns (Welch-próba: 5,305, $df_1=3$, $df_2=8,398$, $p=0,025$).

Az alaphangmagasság elemzése

A következő elemzett paraméter, az alaphangmagasság változása volt, illetőleg annak a mondatvég észlelésére gyakorolt hatása. Az alapkérdés itt az, hogy vajon a beszélő jelzi-e valamiképpen az alaphangjával az egység végét. Ha igen, a következő kérdés az, hogy vajon a hallgató képes-e ezt az információt feldolgozni. Holland anyanyelvűekkel végzett kísérletek azt igazolták, hogy mind a beszélő, mind a hallgató használja a befejezettség érzékeltetésére az alaphangváltozást. Ezek a változások a dallamcsúcsban, valamint az átlagos alaphangmagasságban és annak szórásában voltak megfigyelhetők (Swerts–Collier 1991). Német nyelvű kísérletek alapján az látható, hogy az alaphang frekvenciacsúcsának elhelyezkedése meghatározó a határjelzés szempontjából. Ha a hangsúlyos helyzetű szótagban nem a szótag végefelé jelent meg a frekvenciacsúcs (hanem az elején vagy a szótag közepén), akkor annak kisebb határjelző hatása volt az egyidejűleg megjelenő szünet ellenére, s így ekkor gyakoribbak voltak a megszakítások (Moosmüller 1997).



2. ábra

A szünetek tartamának és a virtuális mondatok jelölésének összefüggései

Saját anyagunkban a korábban felállított négy kategóriát tekintve azt látjuk, hogy a szemantikai és a szintaktikai kritériumok szerinti mondatvégek jelentős részénél a beszélő alaphangja csökken, vagyis a frekvencia egy magasabb értékről egy alacsonyabb értékre jut az utolsó szóban vagy az utolsó szótagban. Az anyagunkban található 25 mondatvég esetében ez 14 esetben valósul meg. Jellemzően esik az F0 190-200 Hz-ről 170-150 Hz-re. Más esetekben a beszélő lebegteti az alaphangját, vagyis sem csökkenés, sem emelkedés nem következik be, a frekvenciaérték vagy 160 Hz körüli, vagy jóval magasabb, 300 Hz körüli. Az utóbbi a tesztszövegben 4 esetben fordult elő, az F0 tartománya ekkor: 290-325 Hz. A fennmaradó 7 esetben a beszélő alaphangmagasságában bekövetkező változás emelkedés, amelynek a mértéke eltérő (de szökésnek, tehát igen rövid időtartam alatt nagy frekvenciaemelkedésnek nem tekinthető). A frekvenciaváltozás különbsége relatíve nagy tartományban szór: 80 Hz és 220 Hz között; változó a kiinduló frekvenciaérték, valamint az az időtartam is, ami alatt a változás bekövetkezik. A frekvenciataromány: 140-220 Hz, az időtartam: 150-300 ms. A 6. táblázat a frekvenciaváltozás függvényében szemlélteti a virtuális mondatokra kapott ítéleteket.

6. táblázat: Az F0 és a virtuális mondatok észlelése valódi mondatvég esetén

F0 (Hz)	Virtuális mondatvég észlelése (%)	
	átlag	szórás
ereszkedő	85,16	39,28-100
emelkedő	65,77	23,21-91,07
lebegő	72,22	46,42-94,64

A kísérleti személyek tehát felhasználják döntéseikben az alaphangmagasság változását vagy stagnálását. Legnagyobb mértékű a mondatvég jelölése az ereszkedő dallam esetében, majd a lebegő dallamnál, és ezekhez képest legkevésbé ítélnék mondatvéget, ha a dallammenet emelkedik. Az alaphangmagasság változását a másik három kategória esetében is elemeztük.

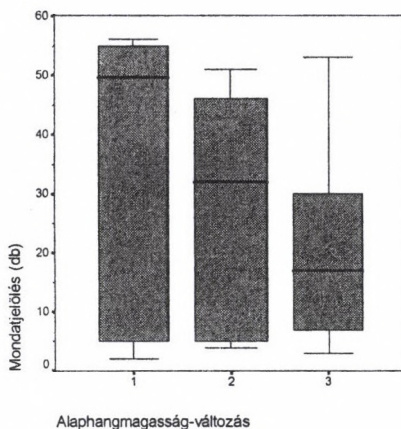
A szerkezetvégeket a beszélő döntően az alaphangmagasság növelésével vagy lebegtetésével jelzi, ez következik be az esetek 75%-ában, ebből 50%-ban történik a dallam alacsonyabb értékről magasabb értékre növelése. A különbségek itt 30-140 Hz közöttiek, s a változás kiindulása 170-310 Hz-es tartományban szór (a frekvenciaváltozás átlagosan 400 ms alatt következik be). Ha összevetjük ezeket az adatokat a valóságos mondatvégek akusztikai jellemzőivel, az látható, hogy szerkezethatáron a frekvenciaváltozás különbségei kisebbek, s a kiinduló érték tendenciaszerűen magasabb szerkezethatáron. A beszélő tehát kétféleképpen is jelzi, hogy másféle egység határán van (mint mondatok befejezésekor). A szerkezetvégek azonosítása virtuális mondatként meglehetősen alacsony, az ítéletek és az alaphangváltozás tekintetében nem mutatható ki semmiféle összefüggés.

Kötőszó előtt is az alaphangmagasság változásának mindhárom említett típusát megfigyelhetjük. Adataink szerint 50%-ban tapasztaljuk a frekvenciaemelkedést, közel 19%-ban a lebegést (magas frekvenciaértéken), a többi esetben pedig alacsony frekvenciaértéken enyhe csökkenést regisztráltunk. A dallammenet frekvenciaemelkedésének különbsége 90-160 Hz, tehát a tartomány szűkebb, mint a mondatoknál, de nagyobb, mint a szerkezethatárnál. A kiinduló frekvenciaérték szórása 150-260 Hz, vagyis nagyobb, mint a mondatok esetében, és szűkebb, mint a szerkezethatárok esetében tapasztaltak. A beszélő tehát sajáttságosan, de tendenciaszerűen jelzi, hogy másféle szerkezeti határon van, mint akár valódi mondatvég, avagy szerkezethatár esetében. Itt sem találtunk egyértelmű összefüggést a dallammenet és a virtuális mondatvég között. Kötőszó előtt például 37,5%-ban ítétek mondatvéget a hallgatók, amikor a szünetet (és a kötőszót) megelőző dallammenet 130 Hz-et emelkedett. Csaknem ugyanilyen arányban jelöltek mondatvéget (35,7%-ban) ott, ahol az F_0 értéke 100 Hz-nyit csökkent a szünetet, illetőleg a kötőszót megelőzően. Nyilvánvaló, hogy a dallamváltozás mint paraméter nem elsődrendű fontosságú a szövegegység elkülönítésében.

Szerkezetben belül három esetben fordult elő hangmagasság-emelkedés 60 Hz, 80 Hz, illetőleg 150 Hz különbséggel; a kiinduló értékek 150-250 Hz között mozogtak. Az összes többi esetben (ötben) az F_0 alacsony frekvencián volt vagy kismértékben csökkent átlagosan 170

150-250 Hz között mozogtak. Az összes többi esetben (ötben) az F0 alacsony frekvencián volt vagy kismértékben csökkent átlagosan 170 Hz-re. Az egyetlen nagyobb arányú, egységes mondatészlelés egy ilyen frekvenciaérték esetén következett be (64,28%-os egyöntetű ítélet).

A statisztikai adatok ezúttal nem mutatnak olyan összefüggést a hallgatók ítéleteivel, mint a szünetek és az előzetesen meghatározott kategóriák esetében. Az F0 értéke és a jelölt mondathatárok közötti korreláció negatív, gyenge és nem szignifikáns. Ez azt jelenti, hogy statisztikailag nem támasztható alá az, hogy az alacsonyabb frekvenciaérték mondathatár-ítéletet idézne elő. A középértékek ugyan eltérő tendenciát jeleznek, azonban az adatok erőteljes szórása miatt összefüggés nem mutatható ki. A tendencia arra utal, hogy a csökkenő F0 a mondatvég benyomását kelti (vö. 3. ábra).



3. ábra

Az alaphangmagasság (F0) és a virtuális mondatok észlelésének tendenciaszerű összefüggése (1=alacsony F0, 2=magasan lebegő F0, 3=emelkedő F0)

Az elemzett adatok arra utalnak, hogy a mondatvég összefüggésben van az alaphangmagasság csökkenésével, illetőleg az alacsony

F0-értékkel, míg a dallamemelkedések azt jelzik a hallgatónak, hogy az adott egységnek nincs vége, tehát a kísérleti személyek általában nem fognak mondatvéget jelölni. A beszélő azonban nem feltétlenül konzekvens a dallammegformálásban, a szemantikai, szintaktikai tervezés nem mindig jár együtt ugyanazzal a prozódiai megformáltsággal. Nem egy esetben azonban nemcsak a szünet előtti F0-változást, hanem a szünetet követő frekvenciaértéket is vizsgálunk kell, hiszen a hallgató a közlés folyamatosságában dönt, tehát minden lehetséges paraméter befolyásolhatja az ítéletét. Ha a szünetet követő szó első szótagjának frekvenciaértéke magasabb, mint a szünet előtti szó utolsó szótagjában mért érték, akkor előfordul, hogy mondatvéget jelöl a hallgató. Az ilyen jelölések anyagunkban azonban ritkák és csak néhány kísérleti személynél jelentkeztek.

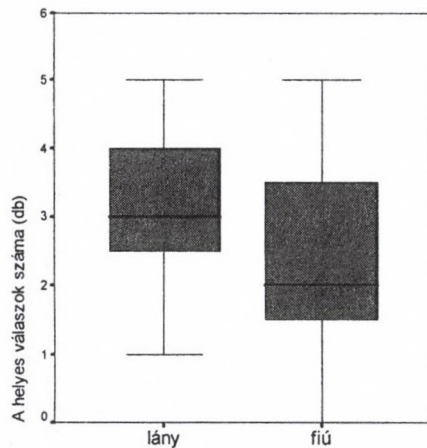
Az eddigi elemzésekből kihagytuk azokat a mondatjelöléseket, amelyek vagy csak az egyik, vagy csak a másik kísérleti csoportból kerültek ki. A következőkben ezt a 9 virtuálismondat-jelölést vizsgáljuk meg. A lányoknál 3 ilyen eset adódott, érdekességgéppen egyik helyen sem tartott szünetet a beszélő, sőt jelentős alaphangmagasság-változást sem mértünk. A három esetből kettő megfelel a valódi mondat szemantikai és szintaktikai kritériumainak, míg egy esetben szerkezethatárra esett a jelölés. A fiúk 6 olyan helyen jelöltek, ahol a lányok nem; közülük 4 kötőszó előtti helyzet, két esetben pedig szerkezeten belüli a mondatvég észlelése. A kötőszók egy kivétellel alárendelő tagmondatokat kapcsolnak össze (az egy kivétel az *és így tovább* közlésvég, amely egyben valódi mondatvég is; itt nyilvánvaló jelölési hibáról van szó). Mind a 6 helyen szünetet tartott a beszélő, ezek szórása 257-521 ms, az átlag pedig 420,66 ms.

Elemeztük a nemek szerinti esetleges különbséget. Hangsúlyozzuk, hogy tekintettel a kísérleti személyek kiválasztására, a két csoport nemcsak abban különbözik, hogy lányok és fiúk alkotják, hanem abban is, hogy – feltételezhetően – az ún. nyelvi tudatosság a magyar szakos bölcsészhallgatókban nagyobb, mint a mérnökhallgatókban. Ezért nagy különbséget feltételeztünk a két csoport között. A kapott adatok azonban ellentmondanak a feltételezésnek. A lányok virtuálismondat-jelöléseinek átlaga 23,75 (átlagos eltérés 10,16), a fiúké 24,81 (átlagos eltérés, 10,86). A két csoport között nincs szignifikáns kü-

lönbség. Hipotézisünk volt az is, hogy a lányok jelentősen több mondatthatárt fognak bejelölni, mint a fiúk; továbbá hogy a lányok és a fiúk között a szerkezeten belüli és a kötőszó előtti kategóriákban jelentős lesz a különbség. Adataink egyiket sem támasztották alá. A két csoport között statisztikailag semmilyen különbség nem volt igazolható.

Szövegértés

Az elhangzott szöveg értését a megértést ellenőrző kérdésekre adott válaszok elemzésével vizsgáltuk. A lányok általában több kérdésre válaszoltak helyesen (vö. 4. ábra), mint a fiúk; a virtuális mondatok jelölése és az értés között azonban statisztikai összefüggés nincsen. Az egytényezős ANOVA szignifikáns különbséget mutat a két csoport összevetésekor ($F(1, 54)=4,799$; $p=0,033$).



4. ábra

A spontán szöveg értése a két kísérleti csoportban (átlag és szórás)

Elemeztük az egyes kérdésre adott helyes válaszokat. A két csoport hasonlóan válaszolt helyesen az ötből három kérdésre; s ugyanaz a kérdés okozta valamennyi kísérleti személynek a legtöbb problémát.

Ez utóbbit nem egészen 20%-uk tudta csak jól megválaszolni. Nagy különbséget a bölcsészek és a mérnökhallgatók között két kérdésben tapasztaltunk. Az elsőre a lányok 75%-ban, a fiúk csak 32,24%-ban adtak korrekt választ; a negyedik kérdésben pedig 50%, illetőleg 85,7% a helyes válaszok aránya, ismételten a lányok javára. Mindkét csoportban akadt természetesen olyan személy, aki egyetlen kérdésre sem tudott válaszolni, de olyan is, aki az összes kérdésre hibátlanul felelt. Összevetettük a kérdésekre adandó válasz helyét a szövegben a kapott válaszokkal. Az első kérdés a spontán szöveg legelején, a második, a harmadik és a negyedik a szöveg első felében elhangzottakra kérdez rá, míg az utolsó kérdésre elvárt válasz a szöveg második felében jelenik meg. A 7. táblázat a kérdésre várt válasz szintaktikai/prozódiai/észlelési jellemzőit veti össze a helyes válaszok arányával (a „mondajelölés” a válasz elhangzását követő jelölést mutatja). Az első „válasz” a mondat közepén, az összes többi a mondat végén hangzott el.

7. táblázat: A megértést ellenőrző tesztben a kérdésekre várt válaszok szöveghelyeinek szünetjellemzői, a virtuális mondatok ítéletei és a helyes válaszok (K.=kérdés)

K.	A kérdésre várt válasz jellemzői		Helyes válaszok aránya (%) b.=bölcsészek, m.=mérnökhallgatók	
	szünet (ms) előtte/utána	mondatje- lölés (%)	b. – m.	összesen
1.	812/996	87,5	75 – 32,14	53,57
2.	601/703	66,07	64,28 – 60,7	62,5
3.	597/205	41,07	14,28 – 25	19,64
4.	1263/710	98,21	85,7 – 50	67,85
5.	590/791	37,5	67,85 – 71,4	69,64

Az adatok nem mutatnak egyértelmű összefüggést, vagyis a prozódiai tényezők nincsenek hatással a megértésre. Ugyanakkor az sem igazolható, hogy a virtuális mondatok jelölése miatt nem tudtak a kísérleti személyek a kérdésekre jól válaszolni, hiszen a mondatjelölés-

től függetlenek a megértést ellenőrző kérdésekre adott helyes válaszok.

Következtetések

A kutatások eddigi megállapításai szerint a spontán beszédben szoros (bár nem minden esetben tökéletesen megvalósított) a kapcsolat a prozódia és a szintaktikai szerkezet között (vö. pl. Cutler et al. 1997). Ezt a tényt német anyanyelvűekkel végzett agyi kiváltott potenciál vizsgálatok (ERP) is alátámasztották, amennyiben a prozódiai határok azonnal szintaktikai szerkezeteket aktiváltak (Steinhauser–Friederici 2001; Steinhauer 2003). Valahányszor a kísérleti személy intonációs határt észlelt, pozitív hullám jelent meg az ERP-ben (amit a szerző(k) CPS-nek neveztek, vö. closure positive shift). A kísérletben azonban nem spontán beszéd, hanem összetett mondatok képezték az ingeranyagot, amelyben a szemantikai, szintaktikai és prozódiai határok egyértelműen egybeestek. A kutatók alapkérdése az volt, hogy a prozódiai jellemzők vajon a szintaktikai struktúrával vagy a szemantikai sajátosságokkal mutatnak-e szorosabb kapcsolatot.

A szemantikai koherenciát igazolják angol nyelvű kísérletek a szintaktikai szerkezet és a prozódiai megvalósítás egységével szemben (Frazier et al. 2003). A szerzők kimondják, hogy adataik szerint nem létezik olyan prozódiai szerkezet, amely képes biztosítani a szintaktikai szerkezeti egységet a hallgató számára. A hallgatók a szintaktikai szerkezetek prozódiai megtörését nem is érzik természetellenesnek. A lényeg tehát az, hogy nem a szünet jelenléte vagy hiánya az, ami döntő, hanem az, hogy a létrejött prozódiai struktúra által kijelölt közlészakasz szemantikailag koherens-e vagy sem. A hallgatók 'nem természetesnek' ítélték olyan több mondatból álló szintaktikai egységeket, amelyek prozódiailag ugyan tökéletesek voltak, de szemantikailag kevésbé koherensek.

A spontán beszéd egységeinek jelölése, felismerése, illetőleg elkülönítése az automatikus beszédfelismerés egyik alapvető problémája (pl. Batliner et al. 1998; Gallwitz et al. 2002). Korpuszokon végzett vizsgálatok eredményei azt mutatták, hogy a prozódiai (és egyéb) kulcsok különböznek a szöveg típusától függően. A szünet és az alaphang sajátosságai meghatározóak például a hírek szegmentálásakor, ugyan-

akkor a párbeszédekben a szünetnek, a beszédhangok időtartamának és a szavak szegmentálásának van nagyobb jelentősége (Shriberg et al. 2000). A mesterséges beszédfelismerés eredményei szerint a következő paraméterek meghatározóak a határjelzésekben (digitális jelfeldolgozás, majd lineáris regressziószámítás, illetőleg a rejtett Markov-móddel alkalmazásával): F0 és változásai, szótag/szó intenzitása, temporális viszonyok (elsősorban a szótag/szó időtartama), a szót megelőző és/vagy követő szünet tartama, a beszédtempó normalizálása, a szóhangsúly jelenléte egy adott szón (Batliner et al. 1998). A humán beszédfeldolgozó mechanizmus részben más folyamatok működtetése, részben más feldolgozás révén dekódolja a spontán beszédben lévő egységeket, de természetesen ugyanazokat az akusztikai-fonetikai paramétereket használja fel.

A humán mechanizmus rendkívül rugalmas, nem konkrét értékek alapján, hanem viszonyok alapján dönt, amelyek egy bizonyos egységen belül, avagy egységek sorozatában realizálódnak. A bekezdés végét általában hosszabb szüneteknél jelzik a hallgatók, szemben a mondatvégekkel. A konkrét adatok mégis azt mutatják, hogy bekezdésvéget észleltek 1380 ms-os szünet esetén és mondatvéget 1390 ms-os szünetnél (Lehiste 1979, 196). A spontán beszéd bármely prozódiai sajátossága, amely a virtuális egység befejezését jelezheti, önmagában nem, csak bizonyos összefüggésekben bír határjelző funkcióval, s csak akkor, ha a szemantikai sajátosságok ezt a határt megerősítik.

Kísérletünkben a virtuális mondatok percepció határait a szünetek (néma vagy kitöltött) megjelenése, az alaphangmagasság csökkenése vagy lebegése és a szemantikai strukturáltság eredményezte. Megállapítható, hogy a 'virtuális mondat' valóban létezik (mind a beszélő produkciójában, mind a hallgató percepciójában), akusztikai-fonetikai korrelátumai vannak, s a hallgatók nagymértékben egyöntetűen ismerik fel ezeket a spontán beszédben. Az elhangzott közlések megértése azonban független a szegmentálási stratégiáktól és működésektől.

Irodalom

- Batliner, A. – Kompe, R. – Kiessling, A. – Mast, M. – H. Niemann – Nöth, E. (1998): M = Syntax + Prosody: A syntactic-prosodic labelling scheme for large spontaneous speech databases. *Speech Communication* 25: 193-222.
- Brent, M. R. – Cartwright, T. A. (1996): Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61: 93-125.
- Cutler, A. – Dahan, D. – van Donselaar, W. (1997): Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 40: 141-201.
- Duez, D. (1993): Acoustic correlates of subjective pauses. *Journal of Psycholinguistic Research* 22: 21-39.
- Fábricz Károly: A beszélt nyelvi szövegalkotás kérdéséhez. In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. MTA Nyelvtudományi Intézet. Budapest, 76-90.
- Frazier, L. – Clifton, Ch. Jr. – Carlson, K. (2003): Don't break, or do: Prosodic boundary preferences. *Lingua* 1: 1-25.
- Gallwitz, F. – Niemann, H. – Nöth, E. – Warnke, V. (2002): Integrated recognition of words and prosodic phrase boundaries. *Speech Communication* 36: 81-95.
- Goldman-Eisler, F. (1972): Pauses, clauses and sentences. *Language and Speech* 15: 103-147.
- Gósy Mária (1999) *Pszicholingvisztika*. Corvina Kiadó. Budapest.
- Gósy Mária (2000): A beszédészünetek kettős funkciója. In: *Beszéd kutatás 2000*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 1-15.
- Grosjean, F. – Deschamps, A. (1975): Analyse des variables temporelles du français spontané. *Phonetica* 28: 191-226.
- Hird, K. – Kirsner, K. (2002): The relationship between prosody and breathing in spontaneous discourse. *Brain and Language* 80: 536-555.
- Kohler, K. J. (1983): Prosodic boundary signals in German. *Phonetica* 40: 89-134.
- Lehiste, I. (1979): Perception of sentence and paragraph boundaries. In: *Frontiers of Speech Communication Research*. Eds.: Lindblom, B. – Ohman, S. Academic Press. London, New York, San Francisco, 191-201.
- Lehiste, I. – Wang, W. S.-Y. (1977): Perception of sentence boundaries with and without semantic information. In: *Phonologica 1976: Innsbrucker Beiträge zur Sprachwissenschaft*. Eds.: Dressler, W. U. – Pfeiffer, O. E.

- Institut für Sprachwissenschaft der Universität Innsbruck. Innsbruck, 227-283.
- Levelt, W. (1989): *Speaking*. MIT Press. Cambridge, Mass.
- Moosmüller, S. (1997): The relevance of fundamental frequency contour for interruptions: A case study of political discussions in Austria. In: *Communicating Gender in Context*. Eds.: Kotthoff, H. – Wodak, R. John Benjamins. Amsterdam, Philadelphia, 401-420.
- Schafer, A. J. – Speer, S. R. – Waren, P. – White, S. D. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29: 169-182.
- Shriberg, E. – Stolcke, A. – Hakkani-Tür, D. – Tür, G. (2000): Prosody-based automatic segmentation of speech into sentences and topics. *Speech Communication* 32: 127-154.
- Steinhauser, K. (2003): Electrophysiological correlates of prosody and punctuation. *Brain and Language* 86: 142-164.
- Steinhauser, K. – Friederici, A. D. (2001): Prosodic boundaries, comma rules, and brain responses: The Closure Positive Shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research* 30: 267-295.
- Swerts, M. – Collier, R. (1991): On the controlled elicitation of spontaneous speech. *Speech Communication* 12: 35-48.
- Wacha Imre (1988): Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz. In: *Beszélt nyelvi tanulmányok*. Szerk.: Kontra Miklós. MTA Nyelvtudományi Intézet. Budapest, 10-159.

A HANGMINŐSÉG SZUBJEKTÍV ÉS OBJEKTÍV MEGÍTÉLÉSE

Hans Grassegger

Bevezetés

Egy korábbi kutatásunkban harminckilenc személytől származó hangrészleteket becsültettünk meg harminc lehallgatóval. A kísérleti beszédanyag izolált magánhangzókat, valamint hosszabb közléseket (logatomokat, szavakat és mondatokat) tartalmazott. Az elsődleges cél az volt, hogy felderítse az izolált magánhangzók és az összefüggő beszéd esetleges eltérő hatását a zöngéképzésre, illetőleg a zöngé percepciók megítélésére (vö. Wrentschur 1998).

A jelen vizsgálathoz a fenti tanulmányból választottunk ki tizenöt személyt, hogy összehasonlítsuk a percepciók becslést az objektív elemzésekkel. Ez utóbbihoz a Kay Multidimensional Voice Programot (MDVP) használtuk. A kérdésünk az volt, hogy vajon van-e és milyen különbség a humán (percepciók) becslés és a mért akusztikai adatok között.

Anyag és módszer

A kísérleti személyek életkora 17 és 76 év között szórta; közülük hét volt férfi és nyolc nő (1. táblázat). A tizenöt személyt három csoportba osztottuk a klinikailag tesztelt, illetőleg diagnosztizált hangminőségük szerint. Az első csoportba a funkcionális diszfóniások (F), a másodikba az organikus diszfóniások (O), a harmadikba pedig a normál zöngéhangúak (N) kerültek.

Valamennyi résztvevő különböző hangképzési feladatot oldott meg, amelyeket magnetofonszalagra rögzítettünk: a) 1-től 10-ig számolás, b) három rövid mondat kiejtése, c) egy, illetve két szótagú szavak felolvasása, CV-szótagok ismétlése és izolált magánhangzók hosszan kitartott ejtése. A beszédjeleket Revox magnetofonnal (B-77-es típus) és egy a beszélő szájától 30 cm-es távolságra tartott AKG kondenzátoros mikrofon (W-23-as típus) segítségével rögzítettük. Az

eredeti hangfelvételeket DAT-fellevővel digitalizáltuk.

1. táblázat: A kísérleti személyek: azonosító, életkor, csoport, nem

N-csoport			F-csoport			O-csoport		
Beszélő	Nem	Kor	Beszélő	Nem	Kor	Beszélő	Nem	Kor
N1	férfi	24	F1	nő	24	O1	nő	53
N2	nő	17	F2	nő	43	O2	nő	76
N3	férfi	25	F3	nő	24	O3	férfi	71
N4	nő	24	F4	férfi	52	O4	férfi	57
N5	férfi	25	F5	nő	33	O5	férfi	49

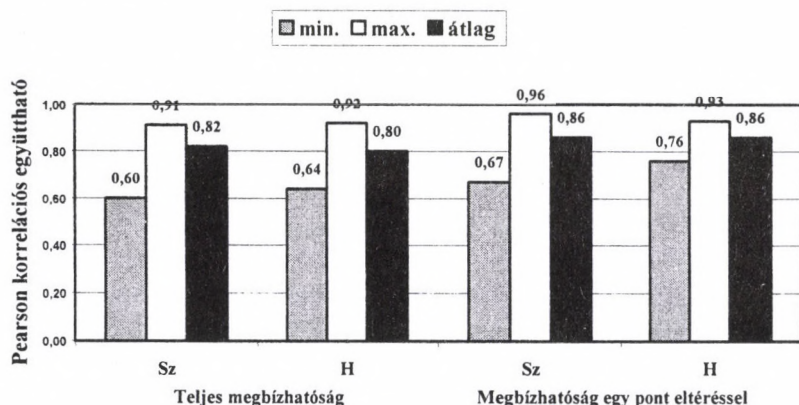
A kísérleti anyagot a lehallgató személyek véletlenszerű sorrendben hallgatták, zavarmentes környezetben. A lehallgató személyek részben beszédvizsgálatban járatos személyek, fül-orr-gégeszakorvosok és beszédterapeuták voltak (összesen 9-en), valamint 10 elméleti tudással ugyan rendelkező, azonban gyakorlati tapasztalattal nem bíró logopédushallgató. A kísérleti személyek feladata az volt, hogy megbecsüljék a beszélők hangminőségét a rekedtség, a hehezetesség és a bizonytalan hangképzés mentén egy vizuális-analóg skálán, amelyen a végpontok a következőképpen voltak megcímkézve: „egyáltalán nem rekedt/hehezetes/bizonytalan zöngéhang”, illetőleg „rendkívül rekedt/hehezetes/bizonytalan zöngéhang”. A statisztikai kezelhetőség céljából az adatokat egy hétpontos skálán digitalizálták, ahol az 1-es a normál hangminőséget, míg a 7-es az attól nagymértékben eltérő hangminőséget jelölte. Ahelyett, hogy bármilyen meghatározást adtak volna a minőségi paraméterekre vonatkozóan, arra kérték a kísérleti személyeket, hogy tapasztalataik és elméleti tudásuk alapján ítélik meg az elhangzottakat. Az ingereknek egy bizonyos száma – bár ezt a lehallgatók nem tudták – kétszer szerepelt, ez lehetőséget nyújtott a becslések egyéni megbízhatóságának kiszámítására.

Eredmények

Percepció elemzések

Az egyéni megbízhatóság megállapításához a Pearson korrelációs együtthatót használtuk. Az átlagos korreláció a szakértő megfigyelők esetében 0,82, míg a logopédushallgatóknál 0,80. Egy kiegészítő meg-

bízhatósági teszt révén, amely ± 1 pont eltérést engedélyez a becslési skálán a megismételt eseteknél (vö. Hammarberg 1986), a korrelációérték 0,86-ra emelkedett mind a szakértők, mind a diákok esetében. Így mindkét módszer elfogadható egyéni megbízhatóságot mutat a kísérleti személyek mindkét csoportjában (1. ábra).



1. ábra
A megismételt ítéletek korrelációja
(Sz=szakemberek, H=hallgatók)

A személyek közötti megbízhatóság megállapításához Pearson korrelációs együttható kiszámítása szükséges; először egy csoporton belül, azaz külön a szakértők és külön a diákok esetében, majd a kísérletben résztvevő egyének összességét véve alapul. Ezen számítások kielégítő eredményeket hoztak, az együtthatók átlaga 0,76 és 0,84 között mozog, bár nem egészen meglepő módon a szakértők közötti egyezés mértéke magasabbnak bizonyult, mint a nem szakértők közötti (2. táblázat).

Annak megállapítására, hogy a szakértők átlagos becslése jelentősen különbözött-e a logopédushallgatókétól, minden egyes minőségparaméterre t -próbát végeztünk. Mivel mindhárom paraméter esetében $p > 0,05$ lett a kapott érték, valamint a szakértők és a diákok együttes

egyének közti megbízhatósága is ezt támasztotta alá, a percepciók eredmények elemzése a kísérletben részt vevő összes személy ítéletein alapszik.

2. táblázat: A Pearson-féle r -együttható az egyes csoportokon belül, valamint az összes résztvevőnél

Érték	Szakemberek	Hallgatók	Teljes csoport
minimum	0,782	0,661	0,661
maximum	0,888	0,836	0,888
átlag	0,837	0,759	0,779

Ami az izolált ejtésben hosszan kitartott magánhangzók és az összefüggő beszéd közötti különbséget illeti, számítások egész sorát alkalmaztuk (vö. Wrentschur 1998, 180) arra, hogy meghatározzunk egy olyan beszédmintát, amely többé-kevésbé alkalmas diagnosztikai célokra. Ezen számítások egyike összehasonlította a szórásokat a magánhangzók és az összefüggő beszéd esetében, azaz ismételt szótagok, szavak és mondatok becslésében külön-külön minden egyes személy és paraméter esetében, azt feltételezve, hogy a szórások alacsonyabb értékei magasabb konzisztenciát jelölnek a beszédminta becslésében, így bizonyítva azt, hogy azok 'jobb minták' az alkalmazhatóság szempontjából.

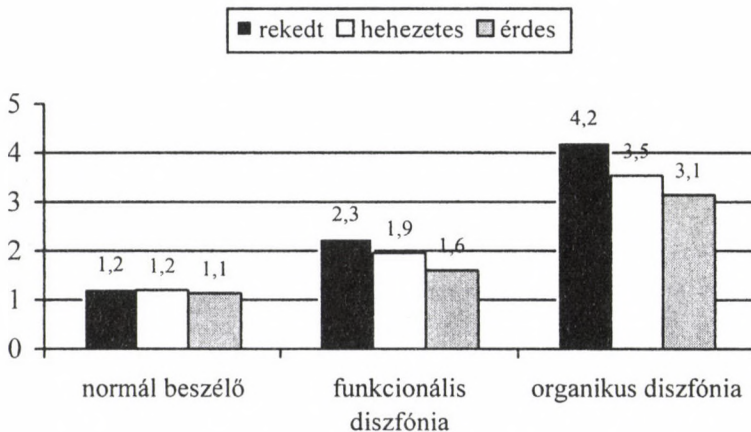
Mindezen túl a szórások és becslési átlagok – amely két érték alapesetben teljesen független egymástól – úgy tűnik, korrelálnak a beszédmintákban. A percepciók alapú becslések szórása a normál hangoknál alacsony volt (alacsony középértékkel), viszont a funkcionális és organikus diszfónia esetében magasabb lett (magas középértékkel, amely a hangrendellenesség növekvő értékét jelöli; vö. 3. ábra). Ennek alapján megállapítható, hogy a hangminőség perceptuális becslésének pontossága kevésbé függ a beszédminta típusától, sokkal inkább a beszélő hangrendellenességének súlyosságától (Kreiman et al. 1993). Minthogy semelyik másik fent említett számítás nem szolgált meggyőző bizonyítékkal arra nézve, hogy jelentős különbség létezne a hosszan kitartott magánhangzók, illetve az összefüggő beszéd perceptuális becslése között, a további elemzéseket az összes inger figyelem-

bevételével végeztük.

3. táblázat: A három tesztelt hangminőségítípus átlagos becslése

Csoport	Érték	Rekedt	Hehezetes	Érdes
normál beszélők	minimum	1,103	1,081	1,033
	maximum	1,359	1,332	1,230
	átlag	1,232	1,200	1,139
	szórás	0,088	0,087	0,074
funkcionális diszfónia	minimum	1,579	1,392	1,409
	maximum	3,038	2,376	2,100
	átlag	2,253	1,955	1,610
	szórás	0,490	0,379	0,259
organikus diszfónia	minimum	1,800	1,664	1,561
	maximum	6,352	5,360	4,798
	átlag	4,215	3,535	3,142
	szórás	1,650	1,307	1,146

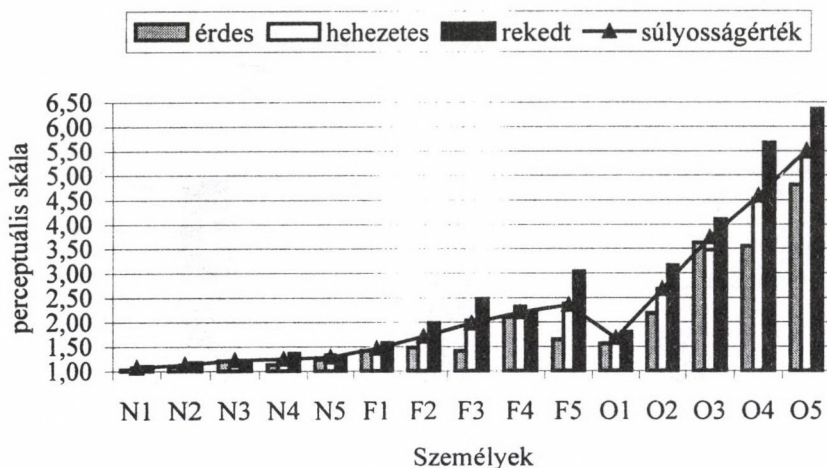
A 3. táblázat felsorolja minden egyes minőségi paraméter becslési átlagát, megadva mind a maximális, mind a minimális értéket, valamint az átlagot (vö. 2. ábra is), továbbá a szórásokat minden kísérleti csoport esetében. Ezek az eredmények azt sugallják, hogy a három tesztelt zöng hang megkülönböztethető az alkalmazott paraméterek segítségével. Az egyes lehallgató személyek esetében azonban a hangminőség megbecsülhetőségének vannak határai. Nem szabad figyelmen kívül hagyni, hogy a 3. táblázatban megadott értékek az összes kísérleti személy becsléseinek átlagán alapszanak, a minőségi paraméterek középértéke viszont az egy csoporton belüli személyek becslésein. Az egyes hangtípusok minimális és maximális paraméterértékei részlegesen átfedik egymást a normál és a funkcionálisan rendellenes hangok esetében, és ez még inkább igaz a funkcionális és az organikus diszfóniákra. Ez azt jelenti, hogy a minimálértékek az organikus diszfónia mindhárom perceptuális paraméterénél alacsonyabbak, mint a funkcionális diszfónia maximális értékei. Ez hasonló Wendler (1986) eredményeihez, jelezve azt, hogy a hangminőség megfelelő perceptuális osztályozása csak csoportosan lehetséges.



2. ábra

Az egyes hangminőség-paraméterek perceptuális átlagai
a kísérleti csoportok szerint

Az akusztikai elemzést megelőzően minden egyes beszélő estében középértéket számítottunk, átlagolva az összes perceptuális paraméter becslési értékét. A 3. ábra mutatja a beszélő alaphangjának általános kiértékelését – figyelembe véve olyan perceptuális faktorokat, mint rekedtség, hehezetesség és érdeség. Az általános súlyosság nemcsak megkülönbözteti a hangminőségeket, hanem az azonos minőségű hangokat fokozati osztályokba sorolja. Mindazonáltal, amint azt már kimutattuk a kísérleti csoportokon belüli (vö. 3. táblázat) minimum és maximum paraméterértékeknél, a súlyosságértékek is figyelemre méltó átfedést mutatnak a funkcionális és az organikus diszfónia esetében (vö. pl. az O1 kísérleti személy általános súlyosságértéke majdnem megegyezik az F-csoportéhoz tartozó legalsóbb értékekkel).



3. ábra

A hangminőség-paraméterek perceptuális becsléseinek átlagai és az általános súlyosságértékek kísérleti személyenként

Akusztikai elemzések

Akusztikai változók. Az elmúlt időszakban nagyszámú akusztikai mérést végeztek, hogy meghatározzák a hangminőséget jellemző paramétereket (pl. Eskenazi et al. 1990; Yiu et al. 2000). Annak ellenére, hogy az eredmények néha következtlenek a tekintetben, hogy mely akusztikai jellemző mely perceptuális tulajdonsághoz kapcsolható, néhány vizsgálatban erős korreláció mutatkozott egyrészt az észlelt rekedtség, hehezetesség és érdeesség akusztikai változói között. A leggyakoribb akusztikai változók egyike, amely kimutathatóan korrelál a rekedtséggel (Yumoto et al. 1984), de szintúgy a hehezetességgel (Eskenazi et al. 1990) és az érdeességgel (Deal–Emanuel 1978), az a frekvenciaperturbáció. A jelen vizsgálatban a hangminőségre kapott perceptuális értékeléseket objektív akusztikai paraméterekkel vetettük össze, mint az F0, a J%, a RAP- és a PPQ-érték. A beszélő átlagos alaphangja (F0) azon összes alaphangfrekvencia-értékek átlaga, amelyeket a periódusok összegzéséből adatoltak (szünetek nélkül). A J% a hangváltozékonyságot kifejező paraméter, amely relatív becslés eredménye

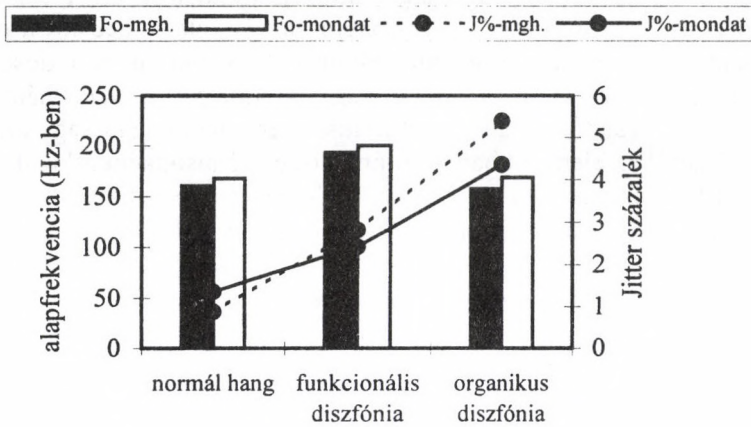
a periódusonkénti hangmagasság változékonyságáról az elemzett mintában. A RAP az átlagos perturbáció, amely relatív becslés a periódusonkénti hangmagasság változékonyságáról három periódusonként átlagolva. Az PPQ a hangmagasság-periódus perturbációs hányadosa, ami egy relatív becslés a periódusonkénti hangmagasság változékonyságáról az elemzett hangmintában öt periódusonként átlagolva.

A mért akusztikai paraméterekhez kétféle mintát használtunk. Egyfelől a kitartott magánhangzók közül az [a]-t, másfelől az „*Eine Maus saust aus dem Haus.*” mondatot (kivéve az O3 és az O4 kísérleti személyeket, akiknél a mondatokat két szótagos szavak sorozatával kellett helyettesíteni egyéb okok miatt). A 4. táblázatban összegeztük a kitartott magánhangzó és a tesztmondat esetében mért akusztikai változók átlagos értékeit és szórásait mindegyik csoportban.

4. táblázat: A mért akusztikai változók átlagos értékei és szórásai a kétféle megvalósításban

Típus	Előford.	Érték	F0	J%	RAP	PPQ
normál beszélő	mgh.	átlag	160,44	0,88	0,52	0,51
		szórás	48,22	0,31	0,21	0,19
	mondat	átlag	167,64	1,36	0,77	0,83
		szórás	50,47	0,56	0,33	0,33
funkcionális diszfónia	mgh.	átlag	193,43	2,81	1,68	1,65
		szórás	50,92	1,07	0,65	0,71
	mondat	átlag	200,09	2,42	1,42	1,46
		szórás	51,85	0,76	0,44	0,49
organikus diszfónia	mgh.	átlag	157,70	5,40	3,10	3,23
		szórás	36,65	3,93	2,21	2,43
	mondat	átlag	168,61	4,36	2,48	2,82
		szórás	41,95	3,14	1,75	2,18

Az egyes csoportok közötti különbségek az alaphangfrekvencia-átlagban (4. ábra) csupán a csoportok nemi összetételét tükrözik; az átlagértékek magasabbak a funkcionális diszfóniás csoportban, amelyben egy kivétellel mindegyik kísérleti személy nő.



4. ábra

Az alapfrekvencia-átlag és a hangváltozékonyság értékei a kísérleti csoportokban magánhangzó és mondat esetében

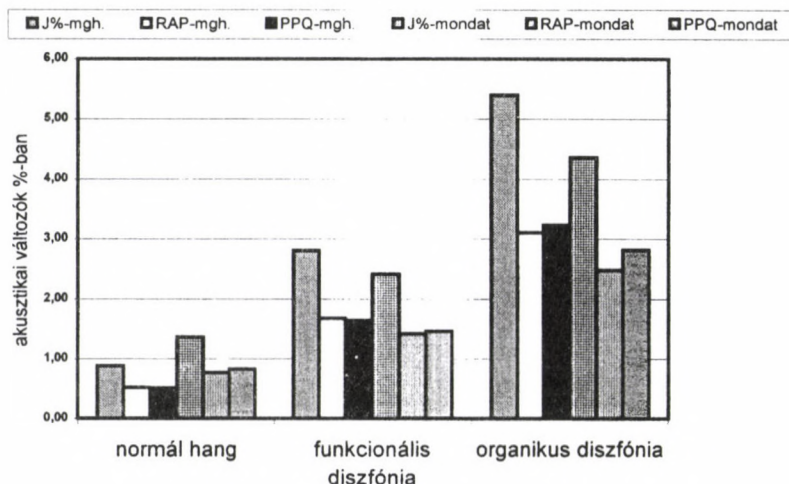
A realizációs feltételek közötti (azaz magánhangzó, illetve mondat) különbségekre vonatkozó akusztikai változók egyike sem (vö. a 4. ábrán az F0 és az 5. ábrán a J%, RAP és PPQ adatokat) érte el a kívánt szignifikanciaszintet. Ezért mindkét realizációs feltétel esetén a J%, a RAP és a PPQ adatait összevontuk minden kísérleti személy esetén, és egyszerű *t*-próbát végeztünk, hogy megbecsüljük a kísérleti csoportok közötti különbségeket. Az 5. táblázatban összesítettük a kísérleti csoportokból álló párok *p*-értékeit minden paraméterre vonatkozólag.

5. táblázat: A kísérleti csoportok páronkénti *p*-értékei

Csoportok	J%	RAP	PPQ
normál/funkcionális diszfónia	0,027	0,026	0,037
normál/organikus diszfónia	0,019	0,017	0,024
funkcionális/organikus diszfónia	0,265	0,275	0,265

Ezek az eredmények azt sugallják, hogy a zönghang változékonyságához kapcsolódó értékek általában különböznek a normális és a

diszfóniás beszélők esetében ($p < 0,05$ minden paraméternél), de a különböző típusú diszfóniák esetében az eredmény nem szignifikáns ($p > 0,26$).

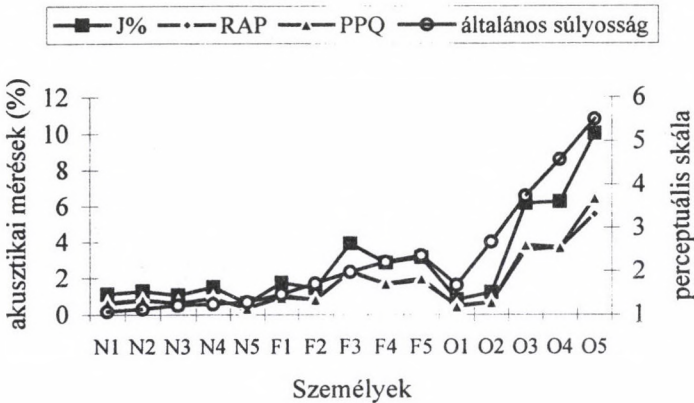


5. ábra

Akusztikai változók átlagai a kísérleti csoportoknál magánhangzó és mondat esetében

Akusztikai adatok, illetve általános súlyosság. Az egyedi adatok további vizsgálata átfedést tárt fel az akusztikai értékek és a hangminőség jellege között, mivel mindhárom diszfóniás személy (kettő közülük organikus) valamennyi paraméterértéke a nem diszfóniás hang értékeinek intervallumába esett (6. ábra). Ugyanezt tapasztaltuk az organikus diszfóniás csoportnál mutatkozó viszonylag magas szórásértékeknel mind a J%, a RAP és a PPQ változókban mind a magánhangzós anyag, mind a mondat esetében (vö. 4. táblázat). Mint azt a 6. ábrán láthatjuk, a kísérleti személyek diszfóniájának általános súlyossága nem korrelált a hang változékonyságát tükröző értékekkel. Ez különösen igaz az organikus diszfóniás személyek esetében. Megfigyelhetjük például a perceptuális értékelés eltéréseit a mért akusztikai érté-

kektől az O1/O2, illetve az O3/O4 beszélők esetében. Következésképpen a kétváltozós korrelációs elemzéssel csupán részleges korreláció volt kimutatható az általános súlyosság és a J% ($r=0,35$, $p<0,01$), a RAP ($r=0,25$, $p<0,05$), illetve a PPQ ($r=0,49$, $p<0,01$) változók között.



6. ábra

A hang változékonyságát jellemző értékek, illetőleg a diszfónia súlyosságának összefüggései

Következtetések

A zöngéhang normáltól eltérő vizsgált sajátosságai, a rekedtség, a hehezség és az érdeesség percepciók becslésében nagy egyéni és egyének közötti megbízhatóságot tapasztaltunk. Az összehasonlító vizsgálat eredményei az alábbiakban összegezhetőek. Sem a kitarzott magánhangzóminták, sem az összefüggőbeszéd-minták nem bizonyultak „jobb előjelzőnek” a hangminőség megítélésében. A perceptuális értékelések inkább a kísérleti személy hangrendellenességének súlyosságától függtek, mint attól, hogy melyik hangmintát hallották a kísérletben résztvevők. A perceptuális értékelés pontossága növekszik a megfigyelők számával, így kizárható az egyéni értékelési stratégiák és

kritériumok hatása; mindezek azonban tükröződnek az ugyanazon hangra kapott minimális és maximális értékekben.

A hang változékonyságát jellemző mért akusztikai értékek csak abban feleltek meg a perceptuális értékelés eredményeinek, hogy nem mutatott egyik sem szignifikáns különbséget a magánhangzó, illetve a mondat ejtése között. Az észlelt hangminőség-különbségeket a vizsgált akusztikai változók nem tükrözték következetesen, hiszen a hangváltozékonysághoz kapcsolódó értékek csak mérsékeltén vagy gyengén korreláltak az általános súlyossági értékekkel, és nem választották el a kétfajta diszfóniát. Ezzel szemben a normál hang mért akusztikai értékei szignifikánsan különböztek a diszfóniás hangokétól, ez meg egyezik Yiu et al. eredményeivel (2000), ahol a RAP és a J% kimutatták a diszfóniát összefüggő beszédben.

A viszonylag korlátozott méretű minta ellenére, amely a jelen vizsgálat nyilvánvaló gyengesége, mégis felmerül a kérdés: vajon a hangváltozékonyságot jellemző értékek a Kay MDVP programjában képesek-e megjeleníteni a különböző mértékűnek észlelhető diszfóniát. A jelen kutatás azt sugallja, hogy eltekintve a periodikusság mennyiségi meghatározásának problémájától a patológikus zöngéképzésben (amely természeténél fogva aperiodikus), akusztikai mérések aligha fogják helyettesíteni az emberi becslést, értékelést a hangminőség megítélésében.

Irodalom

- Deal, R. – Emanuel, F. (1978): Some waveform and spectral features of vowel roughness. *Journal of Speech and Hearing Research* 21: 250-264.
- Eskenazi, L. et al. (1990): Acoustic correlates of vocal quality. *Journal of Speech and Hearing Research* 33: 298-306.
- Hammarberg, B. (1986): Perceptual and acoustic analysis of dysphonia, Stockholm.
- Kreiman, J. et al. (1993): Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research* 36: 21-40.
- Wendler, J. et al. (1986): Classification of voice qualities. *Journal of Phonetics* 14: 483-488.
- Wrentschur, B. (1998): Die perzeptuelle Evaluation von Stimmstörungen, PhD-thesis. Univ. of Graz.

- Yiu, E. et al. (2000): Analysing vocal quality of connected speech using Kay's computerized speech lab: a preliminary finding. *Clinical Linguistics and Phonetics* 14/4: 295-305.
- Yumoto, E. et al. (1984): Harmonic-to-noise ratio and psychophysical measurement of the degree of hoarseness. *Journal of Speech and Hearing Research* 27: 2-6.

A tanulmány elsődleges fordítása angol eredetiből Romeisz Katalin és Kovács Pál egyetemi hallgatók munkája.

AZ IDEGEN NYELVI BESZÉDTERVEZÉS STRATÉGIÁI

Markó Alexandra

Bevezetés

Færch és Kasper (1983) definíciója szerint a **kommunikációs stratégiák** az egyén számára jelentkező problémák megoldására irányuló, részben tudatos tervek egy bizonyos kommunikációs cél elérése érdekében. A kommunikációs stratégiák két fő típusát különíti el a szakirodalom (pl. Færch–Kasper 1983; Tarone 1983) annak alapján, hogy milyen módon viselkedhetnek a nyelvtanulók, amikor kommunikációs problémákkal szembesülnek. Vagy a) az elkerülést választják a probléma kiküszöbölésére, azaz megváltoztatják a kommunikációs célt – ezt nevezzük **kerülő** vagy **redukációs stratégiának**; vagy b) a végrehajtást választják, úgy birkózva meg a feladattal, hogy alternatív tervet dolgoznak ki – ilyenkor **végrehajtó** vagy **parafrazeáló stratégiáról** van szó. A stratégiaválasztás természetesen függ a megoldandó probléma jellegétől, a nyelvtudás mértékétől, a nyelvhasználó egyéniségétől stb.

Færch és Kasper (1983) szerint a tervezési fázisban a nyelvhasználó kiválasztja azokat a szabályokat és elemeket, amelyeket a legmegfelelőbbnek tart annak a tervnek a megalapozásához, amelynek végrehajtása elvezet az eredeti célt megvalósító verbális viselkedéshez. (Az anyanyelvi kommunikációban a tervezési folyamatok általában tudattalanok és nagymértékben automatikusak, például ez is magyarázhatja az anyanyelvi transzfert a nem kielégítően automatizált idegen nyelvű beszédben.) A tervezési folyamat végeredménye a végrehajtási fázist ellenőrző terv. A végrehajtási fázis neurológiai és fiziológiai folyamatokból áll, a kommunikációs stratégiák szempontjából azonban ezek csak akkor válnak fontossá, ha az egyén a terv végrehajtásában problémát érzékel.

Raupach (1983) szerint a stratégiák vizsgálatának a középponti kérdése az, hogy a tanuló mit érez problémának a diskurzus tervezése közben. Ezért a kommunikációs stratégiák nyelvtanulói produkciókban történő elemzésének legalább két célja van: meghatározni a tanuló számára ne-

hézsséget jelentő műveleteket, valamint leírni azokat a stratégiákat, amelyeket ezek megoldására választ. A tanuló tervezési problémáinak meghatározásakor az elemző súlyos módszertani nehézségekbe ütközik. Mivel a tervezési problémákat nem lehet közvetlenül vizsgálni, mindenekelőtt azok a mutatók vizsgálhatók, amelyek a tanuló problémamegoldási kísérleteit tükrözik. Magától értetődik, hogy a döntési folyamatoknak csak egy kis hányada mutatkozik meg a szövegadatokban úgy, hogy adekvát interpretációt tegyen lehetővé. A legtöbb esetben azonban az elemző kénytelen a hezitációs jelenségekre és a bizonytalanság más, implicit jeleire támaszkodni.

„Mivel a valamilyen szinten történő tervezés [...] jelentős kognitív terhelést jelent azáltal, hogy például sok alternatíva áll rendelkezésre, vagy a személy számára nem ismerős a tervezési műveletek pillanatnyi sorrendje, a beszéd folyamatban szünetek fognak megjelenni” (Butterworth 1989, 253). A mondathatárokon megjelenő szünetek például időt adhatnak a beszélőnek a következő mondat megtervezésére, így a tervezés és a szünet mennyisége összefügg egymással. A szöveg tartalmának előzetes megtervezése következtében lecsökken a menet közbeni tervezési igény, s ezáltal a beszéd közbeni szünetek aránya.

Dechert (1983) szerint a köztes nyelvi kommunikációt korábban úgy tekintették, mint olyan izolált jelenséget, amely különösen a téves viselkedés fogalmához kapcsolódik. Természetesen használatosak például redukciós stratégiák a köztes nyelvi kommunikációban, de bizonyosan jelen vannak az anyanyelvi beszédprodukción is. Így pszicholingvisztikai nézőpontból csak egy nyelvi tervezési rendszer létezik, amelybe bepillantást nyerhetünk azáltal, hogy megvizsgáljuk a különböző (anya- és idegen) nyelvi beszédprodukción. A cél redukált (lehet) ahhoz viszonyítva, amelyet egy hasonló kommunikációs szituációban, de az anyanyelvén alkalmazna a nyelvhasználó.

A lehetséges startégiámarkereknek az anyanyelvben tapasztalható megoszlása is markánsan eltérhet az idegen nyelvi performanciában megjelenőtől. Ez azt jelezheti, hogy egy nyelv anyanyelvi beszélői másképpen tervezik a diskurzust, mint akik tanulják ugyanezt a nyelvet. Nemcsak a megoszlás, hanem néhány hezitációs jelenség és kommunikációs stratégia természete is különbözik az anyanyelvi és az idegen nyelvi performanciában (vö. Raupach 1983).

Az idegen nyelvi beszéd tervezésének jellemzőit mindezek miatt az anyanyelvi beszédtervezéssel való összehasonlításban kívántuk megvizsgálni, ehhez terveztünk kísérletet: (kvázi)azonos szöveget rögzítettünk anyanyelven és idegen nyelven. Az volt a hipotézisünk, hogy az anyanyelvi és az idegen nyelvi beszédtervezés folyamatai nagymértékben eltérnek egymástól. Ennek bizonyítékeként arra számítottunk, hogy az idegen nyelvű szöveg tartalmában (elsősorban a kerülő stratégiák használata miatt) redukált, a valódi beszédidő pedig rövidebb lesz, a tervezés időigényessége és bizonytalansága miatt több és hosszabb szünet tarkítja, és a tempója is lassabb lesz, mint az anyanyelvű szövegé.

Anyag és módszer

A felvételeket egy amerikai egyesült államokbeli születésű angol anyanyelvű és egy magyarországi magyar anyanyelvű beszélővel készítettük el. Az angol beszélő 28 éves nő, a kísérlet idején két éve tanult magyarul, célnyelvi környezetben; a magyar a harmadik idegen nyelve (előtte spanyolul és franciául tanult). A magyar kísérleti személy 20 éves nő, az angol első idegen nyelve volt, és kb. 8 évig tanulta. A középiskola befejezése óta franciául tanul, saját megítélése szerint a franciatudása jobb az angoltudásánál.

A beszédprodukciók összevethetősége érdekében olyan műfajú szöveg elmondására kértük a kísérleti személyeket – magyarul és angolul –, amely megfelel az alábbi követelményeknek:

- Mindkét kultúrában közismert, de nem szigorúan kötött formájú (nem „kívülről” mondható), így nagyjából tisztában lehetünk az eredetileg tervezett kommunikációs céloktól való eltérésekkel (szemben a saját gondolatok közlésével, amelyeknél az esetek többségében nem tárhatók fel az eredeti kommunikációs célok).

- Elmondatható más személyekkel is, így lehetővé válik a (különböző nyelvű) egyéni produkciók összevetése.

- Narratív jellegű, így a leggyakoribb monologikus kommunikációs forma vizsgálatát teszi lehetővé.

Mindezek alapján egy jól ismert, és már folklorizálódott Grimm-mesét választottunk ki, a *Piroska és a farkast*, amelynek szerkezete lehetővé teszi az összehasonlítást (mivel a történet bizonyos mozzanatai kötelezőnek tekinthetők). Mivel azonban az idegennyelv-oktatásban a mesemon-

dás nem játszik középponti szerepet, és maga az adott szöveg sem kizárólag a nyelvtanulók valószínűsíthetően elsajátított szókincséből építkezik, félő volt, hogy a beszédtervezés erőteljesen eltolódik a mentális lexikonban való keresés irányában, és így többé-kevésbé összefüggő szövegprodukciónak helyett elemezhetetlen, (néma és kitöltött) szünetekkel erősen tarkított, esetleg a felvételnél jelenlévők segítségét is igénybe vevő, így csak részben önálló produkciót kapunk. A beszéd aktuális tervezése közbeni segítségkérésért azért is kívántuk kirekeszteni a lehetőségek közül, mert – saját idegen nyelvi kommunikációs tapasztalatainkból tudjuk – ilyenkor gyakran előfordul, hogy a kommunikációs partner nem éppen azt a szót vagy szerkezetet adja meg segítségképpen, amelyre a beszélőnek szüksége van. Így vagy újabb segítségkérésre van szükség, vagy a beszélő alkalmazkodik a partnerhez, és megváltoztatja az eredeti közlési szándékát. A dialógusszituáció elkerülését azért is fontosnak tartottuk, hogy olyan tipikus parafrázisstratégiák ne jelenjenek meg a közlésben, amelyek egyrészt a beszédpartner reagálását válthatják ki (és ezáltal esetleg módosítják a tervezési folyamatokat), másrészt nem rögzíthetők a hangfelvételen (pl. gesztusok, arcjáték). Ennek kiküszöbölésére azt a megoldást választottuk, hogy megkértük a kísérleti személyeket, tervezék meg előre a történet idegen (magyar, illetve angol) nyelvű elmondását, és ha bármilyen hiányt érzlelnék a kompetenciájukban, a felvétel készítője megadja a szükséges információt. Jegyzetelhettek, és jegyzeteiket felhasználhatták a szöveg elmondásakor. (Annak érdekében, hogy a magyar beszélő számára valamelyest imitáljuk az angol beszélőnél fennállt idegen nyelvi környezetet, a mese elmondása előtt angolul kellett válaszolnia néhány kérdésre.) Az idegen nyelvű felvétel elkészültét követően rövid szünet után anyanyelven is elmondattuk a mesét (felkészülés nélkül).

(A szakirodalomban gyakorta alkalmazzák a képsorozatról beszéltetés módszerét – vö. pl. Dechert (1983) –, ezt azonban jelen helyzetben nem tartottuk alkalmazhatónak. Egyrészt azért, mert befolyásolja a redukció mértékét azzal, hogy ha valami megjelenik a képen, annak említését a beszélő valószínűleg nem tartja elkerülhetőnek. Másrészt – különösen az anyanyelvű produkcióban – az elbeszélés díszítését is visszafogja az, ha a beszélő azt gondolja, a képek leírását várják el tőle.)

Az angol beszélő produkcióit diktafonnal vettük fel, a magyar anyanyelvű beszélő produkcióját AVL1900 500 Ω -os irányított dinamikus

mikrofonnal közvetlenül számítógépre rögzítettük. A digitalizálás és az adatolás-címkézés Speech Analyzer v2.4b3.6 programmal készült.

Hogy objektívizálható legyen a különböző nyelvű beszédtervezés közötti eltérés, az elemzéseket konkrét, adatolható minőségi jellemzők alapján végeztük el a következő lépésekben:

1. a két-két szöveg főbb mennyiségi és tartalmi különbségeinek megállapítása;
2. a szövegek időszerkezetének feltárása;
3. a beszédszakaszok és a szünetek tipizálása;
4. a beszédtervezési folyamatokra utaló egyéb jelenségek (pl. megakadások) számbavétele.

Eredmények

A főbb **tartalmi** mennyiségi különbségek megállapításához az egyes beszélők anyanyelvi és idegen nyelvi szövegének történetmozzanatait feltettük meg egymásnak (beleértve a nem cselekményes részeket is, mint pl. Piroska nevének eredete, az erdei környezet leírása), de nem voltunk tekintettel az egyes mozzanatok nyelvi kidolgozottságának fokára. Beigazolódott az a hipotézisünk, hogy az idegen nyelvű szövegek (kb. egyharmaddal-egynegyeddel) kevesebb szerkezeti és jelentésmozzanatot tartalmaznak. Az angol beszélő angol szövegének 35 tartalmi elemével szemben a magyarban csak 24 van; a magyar beszélő magyar szövege 56, az angol szövege 41 mozzanatból áll. Az angol adatközlőnél csak egy olyan eltérés, amely a magyar szövegbeli többlet, a magyarnál 3 különbséget az angol szöveg „javára” tapasztaltunk. Ezek az adatok egyértelműen mutatják a redukciós stratégiák alkalmazását az idegen nyelvű produkciókban.

A szövegek **időszerkezetében** is jelentős eltéréseket találunk (1. táblázat). Az elemzésben a beszédszakaszokat mint két egymást követő szünet közötti beszédrészletet határoztuk meg, tekintet nélkül a szupraszegmentális konstruáltság más jegyeire (vö. Németh T. 1996; Varga 1999–2001 és Wacha 1988). Az összes szünetet figyelembe vettük, mivel véleményünk szerint a tervezési stratégiákról bármennyire rövid szünet is árulkodhat.

Az angol beszélő magyar szövege 20%-kal rövidebb az angolnál, ami ugyancsak redukcióra utal. Ezzel szemben a magyar beszélő szövegeinek

az időtartama közötti 30%-os eltérés épp fordított irányú: az idegen nyelvű szöveg hosszabb. Pusztán ez alapján a redukció jelenléte megkérdőjelezhető volna, vegyük észre azonban, hogy a beszédszakaszok összideje szinte azonos a magyar adatközlő két szövegében, vagyis a többletidőtartamot dominánsan szünetek töltik ki. Az ő idegen nyelvű produkciójában előforduló gyakori önkorrekciók és ismétlések okozzák az így is relatíve hosszú tiszta beszédidőt.

1. táblázat: A szövegek időparaméterei

Paraméterek	Angol beszélő		Magyar beszélő	
	Magyar	Angol	Angol	Magyar
Teljes időtartam	189,698 s	239,476 s	571,411 s	410,493 s
A beszédszakaszok összideje	132,904 s	179,827 s	283,726 s	272,500 s
A szünetek összideje	56,794 s	59,649 s	287,685 s	137,993 s
A beszédszakaszok és a szünetek idejének aránya	70%:30%	75%:25%	50%:50%	66%:34%

Mivel a szövegek eléggé hosszúak, és a spontán beszédben az artikuláció tisztasága sem mindig teszi lehetővé a másodpercenként elhangzó hangok számának biztos megállapítását, munkamegoldásként mintavételezés alapján számítottuk ki a hang/s-ban megadott **tempó**átlagokat. A szövegek minden tizedik beszédszakaszának tempóját (a beszédtempónál a követő szünetekkel együtt mérve) számítottuk ki, és ebből átlagoltuk a 2. táblázat eredményeit. A percenkénti szótagszámot természetesen a teljes szövegeken számoltuk. Ahogy vártuk, az idegen nyelvű szövegnek a tempóértékei minden esetben alacsonyabbak az anyanyelvűnél.

2. táblázat: A szövegek tempóadatai

Paraméterek	Angol beszélő		Magyar beszélő	
	Magyar	Angol	Angol	Magyar
Artikulációs tempó (hang/s)	9,24	12,02	9,43	12,05
Beszédtempó (hang/s)	6,71	8,42	4,36	8,19
Beszédtempó (szótag/perc)	162,66	210,50	102,63	203,07

A **beszédszakaszok** (és így a szünetek) számában az angol beszélőnél nincs lényeges eltérés, ezzel szemben a magyar adatközlőnél a sok ismétlés, újrakezdés következtében igen nagy a különbség (3. táblázat).

3. táblázat: A beszédszakaszok adatai

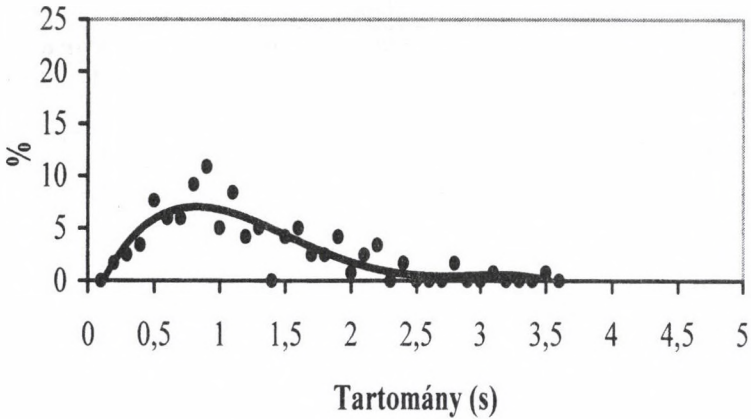
Paraméterek	Angol beszélő		Magyar beszélő	
	Magyar	Angol	Angol	Magyar
A beszédszakaszok száma	119	115	370	235
Minimális hossz	119 ms	140 ms	98 ms	46 ms
Maximális hossz	3499 ms	4986 ms	3706 ms	3179 ms
Átlag	1117 ms	1564 ms	765 ms	1160 ms
Szórás	641 ms	997 ms	459 ms	589 ms
A beszédszakaszok hosszából számítottelső kvartilis ¹	667 ms	728 ms	436 ms	718 ms
medián	921 ms	1331 ms	662 ms	1136 ms
harmadik kvartilis	1512 ms	2151 ms	984 ms	1517 ms

Az 1. és a 2. ábrán az angol beszélő magyar és angol nyelvű szövegét felépítő beszédszakaszok (100 ms-os léptékű) hossz szerinti százalékos eloszlása látható. A pontokra illesztett negyedfokú polinomiális trendvonal érzékelteti a fő tendenciát. Míg a magyar szöveg beszédszakaszainak egyötöde a 700–900 ms-os időtartamhatárok közé esik (1. ábra), addig az angol szövegben a 1300–1400 és a 600–700 ms-os tartományban fordul elő a legtöbb szakasz (2. ábra). Az angol szöveg beszédszakasz-időtartamai nagyobb változatosságot mutatnak.

Az angol beszélő magyar szövegében többször előfordul, hogy a szakasz pusztán egy toldalékmorfémából áll. Ez nyilvánvalóan abból fakad, hogy – a nyelv agglutináló volta miatt – a szöveg megtervezése több lépésben történik: a fő és az egyes toldalékok kiválasztása és az utóbbiak sorrendezése hosszabb és bonyolultabb folyamat, mint a beszélő (izoláló-flektáló) anyanyelvében (vagy a többi általa beszélt idegen nyelvben).

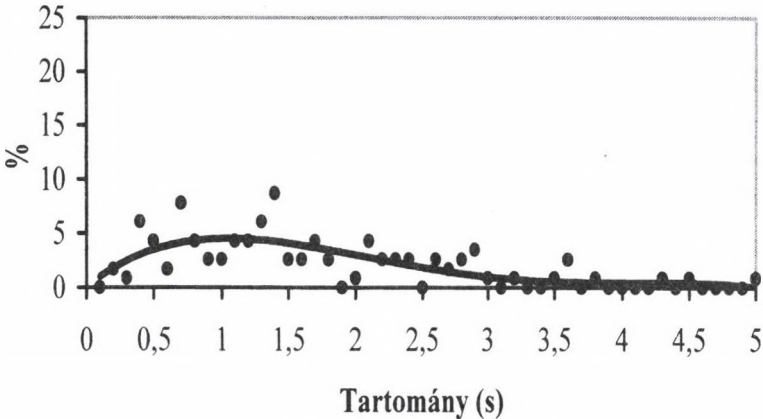
¹ A medián a rendezett minta közepén levő adat értéke (páros számú adat esetén a két középső átlaga). Sokkal kevésbé érzékeny a kilógó értékekre, mint az átlag, és ferde eloszlások esetén használhatóbb. Az első kvartilis a legkisebb és a medián (= második kvartilis) között középen elhelyezkedő adat számértéke, a harmadik kvartilis a medián és a legnagyobb érték között van középen.

Tipikus jelenség a magyar szövegben, hogy az igevonzat előtt jelenik meg a szünet, aminek valószínűleg a megfelelő szó vagy még inkább grammatikai eset megtalálásának nehézsége az oka.



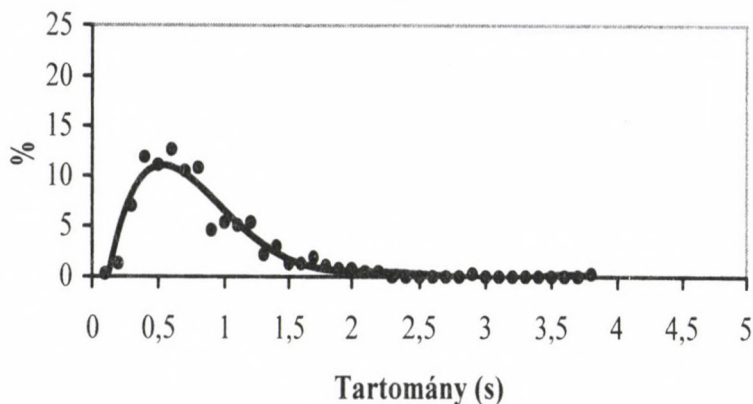
1. ábra

A beszédszakaszok százalékos megoszlása az időtartam függvényében az angol adatközlő magyar szövegében



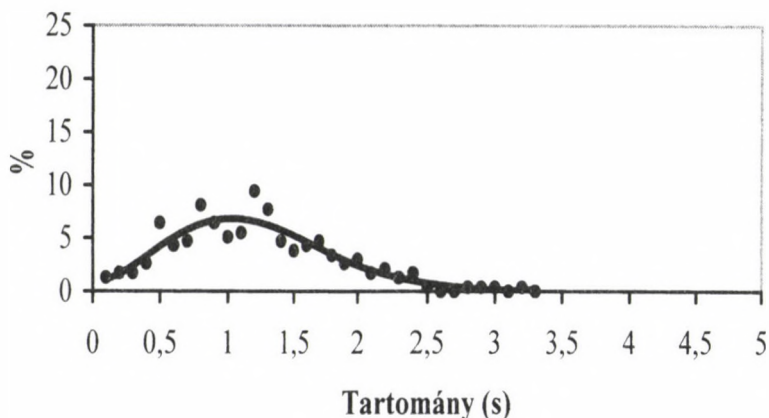
2. ábra

A beszédszakaszok százalékos megoszlása az időtartam függvényében az angol adatközlő angol szövegében



3. ábra

A beszédszakaszok százalékos megoszlása az időtartam függvényében a magyar adatközlő angol szövegében



4. ábra

A beszédszakaszok százalékos megoszlása az időtartam függvényében a magyar adatközlő magyar szövegében

A 3. és a 4. ábra a magyar beszélő angol és magyar nyelvű szövegét felépítő beszédszakaszok (100 ms-os léptékű) hossz szerinti százalékos eloszlását mutatja, a pontokra illesztett hatodfokú polinomiális trendvo-

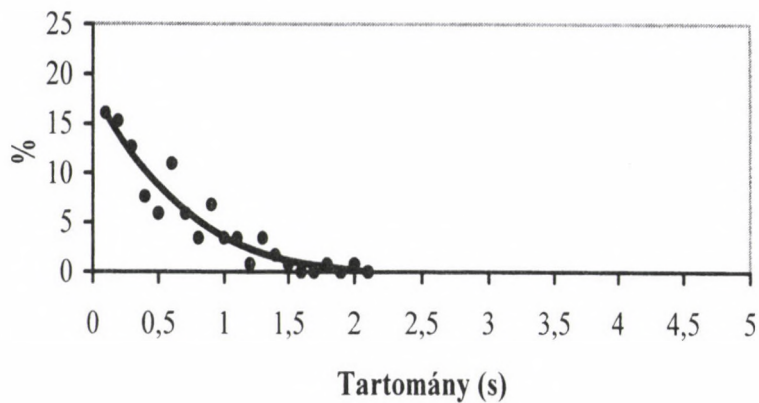
nallal. Az ő idegen nyelvű szövegében még dominánsabbak a rövid szakaszok: a legtöbb a 400–800 ms-os tartományba esik, és ezek többsége mindössze egy egy szótagos szót (kötőszót, névmást, névelőt stb.) tartalmaz. Ennek oka valószínűleg az angol nyelv izoláló jellegében és az idegen nyelven történő beszédtervezés nehezítettségében együttesen keresendő.

A **szünetek** hosszának objektív adataiban (4. táblázat) nem tapasztalunk akkora eltéréseket az angol beszélő két különböző nyelvű produkciója között, mint a beszédszakaszoknál. Láthatjuk azonban, hogy a kb. 20%-kal rövidebb tartamú angol szöveg összes szünetideje szinte megegyezik a magyar szövegével, vagyis a szünetek relatív többsége jellemzi a magyar beszédet – akárcsak a magyar beszélő esetében, akinél viszont jelentős eltérés tapasztalható a két produkció szünetezése között. Ez szintén az idegen nyelven zajló beszédtervezés nehezítettségét jelzi.

4. táblázat: A szünetek adatai

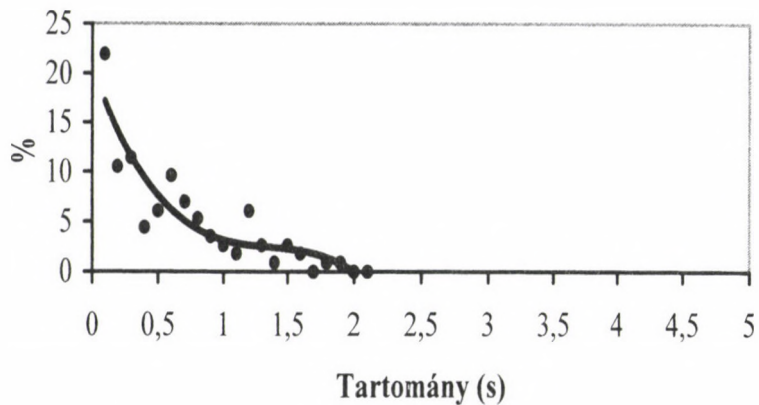
Paraméterek	Angol beszélő		Magyar beszélő	
	Magyar	Angol	Angol	Magyar
Minimális hossz	36 ms	28 ms	21 ms	31 ms
Maximális hossz	1919 ms	1871 ms	5183 ms	3195 ms
Átlag	481 ms	523 ms	778 ms	590 ms
Szórás	399 ms	453 ms	740 ms	531 ms
A szünetek hosszából számított első kvartilis	137 ms	120 ms	230 ms	149 ms
medián	390 ms	440 ms	561 ms	457 ms
harmadik kvartilis	703 ms	781 ms	1102 ms	855 ms

A két beszélő közötti talán leglényegesebb eltérés a szünetek számában, hosszában és típusaiban (néma, illetve kitöltött) mutatkozik meg. Az angol beszélő összes szünete rövidebb 2 másodpercnél, ezzel szemben a magyar kísérleti személynél még 3 s-nál hosszabb szüneteket is találunk. Ennek hátterében általános beszédtervezési különbségek húzódnak. (A szünetek megoszlását mutató 5–8. ábra pontjaira harmadfokú polinomiális trendvonalat illesztettünk.)



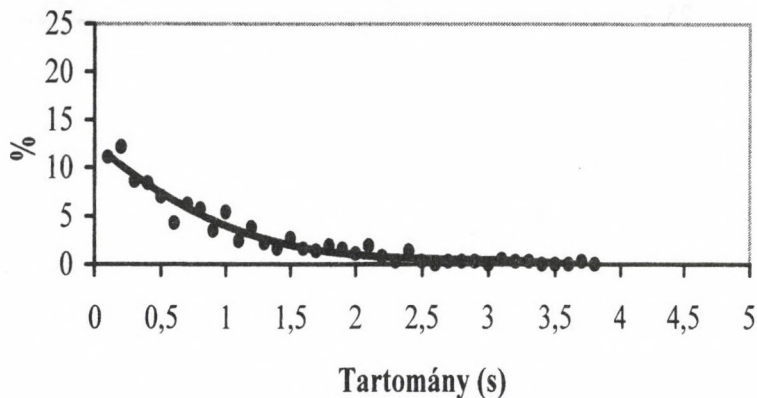
5. ábra

A szünetek százalékos megoszlása az időtartam függvényében
az angol adatközlő magyar szövegében



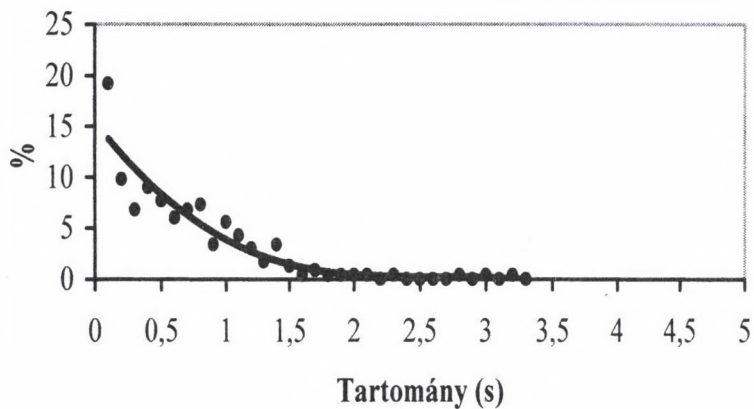
6. ábra

A szünetek százalékos megoszlása az időtartam függvényében
az angol adatközlő angol szövegében



7. ábra

A szünetek százalékos megoszlása az időtartam függvényében
a magyar adatközlő angol szövegében



8. ábra

A szünetek százalékos megoszlása az időtartam függvényében
a magyar adatközlő magyar szövegében

Az angol beszélő angol szövegében az összes szünet 75%-a (86 db) esik mondategységhatárra (Deme 1971), ez az arány a magyar szövegben csak 50% (59 db). A magyar anyanyelvű beszélő angol szövegében 30%-

nyi szünet (111 db) realizálódott mondategységhatáron, a magyar produkcióban ez az érték 59% (138 db). Mindez azt jelzi számunkra, hogy a szünetek tervezettebben valósulnak meg az anyanyelvi produkcióban, mint az idegen nyelvi beszédben.

Az angol adatközlőnél az angol szövegben a 100 ms alatti szünetek 44%-a (25-ből 11) esik mondategységhatárra, a magyar beszélő magyar nyelvű produkciójában ez 20% (45 szünetből 9). Ez az arány az angol beszélő magyar szövegében 26% (19 szünetből 5), a magyar személy angol szövegében pedig csak 5% (41 szünetből 2). Butterworth fentebb idézett megállapításából ennek az eltérésnek az alapján szintén az következik, hogy a mondategységek határán inkább van szükség hosszabb szünetekre az idegen nyelvi, mint az anyanyelvi tervezésben.

Mindkét anyanyelvű szöveg 100 ms alatti szüneteinek kb. egynegyede egyértelműen a követő szótag prominenciáját hivatott elősegíteni, az idegen nyelvű szövegekben nyilvánvalóan ilyen célú rövid szünetet nem találtunk. Ennek az lehet az oka, hogy anyanyelven jobban tudunk koncentrálni az árnyalatok érzékletes kifejezésére, a megfelelő hangsúlyviszonyok kialakítására, míg idegen nyelven a megfelelő szavak és szerkezetek előhívása jobban lefoglalja a figyelmet, és a szuprasegmentális eszközök adekvát használatával sem feltétlenül van tisztában a beszélő.

Kitöltött szünet, vagyis hezitálás az angol beszélő angol szövegében egyetlen egy fordul elő, egy téves kezdést követően. A magyar szövegben – számunkra meglepő módon – egyetlen hangos szünettel sem találkozunk. Ezt esetleg az előre tervezettség indokolhatja, de az angol minta alapján úgy tűnik, a beszélőnek az ismert témáról fogalmazott spontán megnyilatkozásaira nem jellemzők a hezitációk.

Az anyanyelvű szövegre a magyar beszélőnél sem jellemző a kitöltött szünetek felbukkanása, az angol beszédprodukcióban viszont kb. háromnegyed perc ezek összidőtartama; a 369 szünehelyből 105, vagyis közel 30% (legalább részben) kitöltött. A kitöltött szünetek ennél a beszélőnél tehát egyértelműen a tervezési nehézségekre utalnak, hosszúságukat tekintve is, mivel 229 ms és a maximális szünehossz között szóródnak.

Az idegen nyelvű beszédtervezés bizonytalanságát a fenti tartalmi-szerkezeti és temporális jellegzetességek mellett jól mutatják a tervezési zsákutca és a hangfelvételen megörökíthető nonverbális elemek.

A megakadások közül csak azokat vettük számításba, amelyeket javított a beszélő, mert ezek árulkodnak igazán a Færch és Kasper (1983) által a végrehajtási fázist ellenőrző tervnek nevezett működésről. Ezek az **önkorrekciók** természetesen nem mindig valódi javítások.

Az angol beszélő angol szövegében a produkció eleje és vége felé van két-két korrigált téves kezdés. Ezeknek az elhelyezkedése utalhat arra, hogy a beszélő még nem „melegedett be”, illetve a szöveg végéhez közelítve már lankad a figyelme. Az 5. beszédszakaszban a beszélő az eggyel későbbre tervezett szót kezd el kimondani, megáll, a következő szakaszban korrigál, majd megismétlődik a hibatípus (az anticipáció): *this little girl's w | name was Re | Little Red Riding-Hood*. A két szövegvégi tévesztés között is igen rövid idő telik el, ezeknél az történt, hogy a keresés az első alkalommal nem a várt eredményt hozta, ezt a beszélő már az artikuláció megindulása után konstataálta, és új keresést tartott szükségesnek, amely végül a megfelelő szót hívta elő. A 99. szakasz végén a 'megette' jelentésű angol ige jelent meg, de a beszélő a 'bekapta' szemantikai tartalmat érezte adekvátnak, ezért a teljes szerkezetet megismételve kicserélte a problematikus elemet: *and Little Red Riding-Hood screamed as the wolf [i:] | as the wolf gobbled her up*. Ugyanígy kerül a két szakasszal később megjelenő és köznapibb stilisztikai értéke miatt inadekvátnak ítélt *hunter* ('vadász') szó helyére a mesei stílusba jobban illeszkedő *woodsman*. (Itt jelenik meg két néma szünet között a szövegben az egyetlen hezitáció.)

A magyar személy anyanyelvű produkciója mindössze egy korrigált téves kezdést tartalmaz, amely egy tartalmi önellentmondás feloldását szolgálja: *[nem tudta, hogy kivel találkozik, ezért kedvesen köszönt neki:] jó reggelt, fark | jó reggelt kívánok*. Négy helyen találkozunk nyelvbottalással a hangfelvételen.

Az angol adatközlő magyar szövegében 9 önkorrekciót találunk, amelyek a szöveg különböző pontjain jelennek meg. Az angol példákkal ellentétben ezek között sem anticipáció, sem szemantikailag inadekvátnak minősített szótalálás nem szerepel, mindegyik a szóalak-megformálás valamely szintű problémájából adódik.

1. A morfológiai szinten jelentkezik a zavar az esetek két típusában:

a) A hasonló alakú szavak közül nem tud választani a beszélő a következő esetben: *[a farkas akarta a kosarat] Piroskatol ven | vinni*.

b) Többször okoz problémát a toldalékolás: *erdős* | *erdőben*; *halljal* | *halljalaká*^k; *ágybő* | *ágy* | *az ágyból*; *beszélge* | *c* | *nél*. Itt valószínűleg az játszódhatott le, hogy a (kijelentő módú) egyes szám második személyű igerag gyakorisága folytán a feltételes mód jelével kombinált változat elé került a keresésben. Az is elképzelhető azonban, hogy a beszélő magyar nyelvi kompetenciájának hiányossága következtében a **beszélgetsznél* formát állítja elő, azt tételezve helyesnek.

2. A fonetikai realizáció szintjén jelennek meg a következő korrigált anticipációk: *hon* | *hozned*, illetve kétszer *nam* | *nagymama*.

3. Bizonytalanságra vagy a szótalálás nehézségére vezethető vissza az ismétlés ebben az esetben: *meg* | *megölt*.

A magyar adatközlő idegen nyelvű produkciójában 16-szor fordul elő ismétlés, amely az esetek túlnyomó többségében egy szót érint. Ezek háttérben valószínűleg a beszélő bizonytalansága áll (vö. Gósy 2002).

A 20 önkorrekcióból 7 a névmáshasználatot érinti, ezek közül 5 esetben a grammatikai nem tévesztése miatt volt szükség korrekcióra (pl. *he* vs. *she*, *his* vs. *her*). A javítások másik egynegyedében az igei paradigmából történt választást bírálja felül a beszélő (pl. *is* vs. *was* vs. *were*, *go* vs. *went* vs. *gone*). Ezek egyértelműen a két nyelv grammatikai interferenciájából fakadnak. Újabb 4 esetben a szerkezet módosítását viszi véghez; közeli szemantikai-pragmatikai tartalmú szavak közül történt választást háromszor minősít hibásnak: *see* vs. *hear* (az előzmények „motiválják” a hibát), *some* vs. *little*, *here* vs. *out*; egy alkalommal pedig a főnév birtokos alakját cseréli fel alanyesetűre.

Az angol beszélő bizonytalanságára a magyar szövegprodukciónban más, nonverbális példákban is találunk bizonyítékot: három alkalommal sóhaj és/vagy nevetésszerű hang hallatszik a felvételen. Az első ilyen a *beszélge* | *c* | *nél* | *vele* szünetekkel széttagolt szerkezet, a második pedig a *halljal* | *halljalaká*^k toldaléktalálási problémája után fordul elő. A harmadikat közvetlenül megelőző szakaszban nincs hasonló korrekció, azonban két szakasszal korábban, de ugyanebben a mondategységben volt szükség kétszeres korrekcióra az *ágyból* szó helyes toldalékos alakjának létrehozásához. Ezeket egyértelműen a beszédformálásra való reflexióként értelmezhetjük. Ezzel szemben az angol felvételen összesen kétszer hallható vokális jelzések inkább a történetmondást színezik, hangulatfestő jellegűek. A magyar beszélővel készült felvételeken nem érhetők tetten ilyen jelzések.

Következtetések

Az adatokból az következik, hogy a kísérleti személyek idegen nyelvű beszédtervezésükben dominánsan redukciós stratégiákkal élnek – ezt mindenekelőtt a különböző nyelvű szövegek tartalmi struktúrájának összevetése támasztja alá. Ezeknek a stratégiáknak az alkalmazását tudatosnak tekinthetjük, hiszen az adatközlők lehetőséget kaptak az előzetes tervezésre, amelyhez segítséget is igénybe vehettek, mégis úgy tűnik az elemző számára, hogy az angol beszélő a történetet szinte a minimálisan kötelező tartalmi momentumokra redukálta, csak annyit mondott el, amennyit feltétlenül szükségesnek tartott, és a segítséget is csak ennek a szövegtartalomnak a megtervezéséhez vette igénybe. A beszélő globális célja a mese magyar nyelvű előadása, és ezt maradéktalanul teljesíti is, nem célja azonban olyan elemek bekomponálása a történetmondásba, amelyeket nem tart feltétlenül szükségesnek. A magyar beszélőnél a redukció mértéke nem ilyen nagy, az idegen nyelven sem csak a történetvázatot adja elő, olyannyira, hogy három esetben többletmozzanattal is találkozunk a magyar szöveghez képest. Megjegyzendő, hogy mindkét beszélő (a szűzsé elemei mellett) a szöveg szerves részének tekinti a meseformát, ez mind a bevezetésben, a lezárásban, a jellegzetes párbeszéd fordulatokban, illetve ezek élőbeszédszerűvé alakításában (köszönés, indulatszavak stb.) megnyilvánul. Úgy tűnik tehát, hogy ezek a formai elemek hozzátartoznak a megcélzott verbális viselkedéshez, így nem is törölődhetnek a redukció során.

A különféle temporális jellegzetességek egyértelműen bizonyítják, hogy az idegen nyelven zajló beszédtervezés nehezebb, ezért hosszadalmasabb folyamat. Az angol adatközlőnek magyar nyelven nemcsak a szótalálás problematikusabb (ahogy a magyar beszélő számára is az angol nyelvű szövegben), hanem a megfelelő toldalékok előhívása is, különösen a többalakúak közül a harmonizáló megtalálása, és a toldalékmorfémák adekvát sorrendjének összeállítása is folyamatosan nehéz feladat elé állítja a beszélőt – amint ezt az önkorrektciók típusai is mutatták. Bár elemzésünkben nem tértünk ki rá, az idegen nyelvű szövegekben rendkívül gyakoriak a szerkesztési hibák. Az előzetes tervezéskor a kísérleti személyek valószínűleg arra koncentráltak, hogy az általuk kifejezni kívánt tartalmat megfelelő szavakba öntsék, a struktúrák megtervezése ehhez képest háttérbe szorult. Ez szintén árulkodhat az idegen nyelvű beszédtervezés mechanizmusairól –

minél bonyolultabb műveleteket igényel egy-egy tervezési fázis, annál hosszabb időt vesz igénybe, és annál nehezebben is reprodukálható.

A beszédtervezés egyik legtranszparensabb mutatója, a szünet, nemcsak időtartama, hanem megjelenési helye és típusa miatt is meglehetősen árulkodó lehet. Eltérő szünetfunkciókat találtunk a különböző nyelvű szövegekben, különösen a rövid időtartamúaknál jelentősek a különbségek. Lényegesnek gondoljuk megvizsgálni a szünetek előfordulásának a mondat-egységhatárral való korrelációját a spontán beszédben, ahol ez szintén utalhat tervezési folyamatokra. Meglepőnek tartjuk, hogy a négy produkció közül háromban (szinte) nem is fordul elő hezitálás, a kísérleti személyek jóformán nyomdakészen fogalmaznak anyanyelvükön, alig akadnak meg, ami igen jó beszédteljesítményre utal, bár ebbe nyilvánvalóan belejátszhat a mese sémaszerű előhívhatósága a gyermekkori emlékekből.

A visszacsatolás működését az önkorrekciók és önreflexiók bizonyítják, amelyek – amellet, hogy sok minden megtudható belőlük a beszédtervezés nehézségeiről – az anyanyelven és idegen nyelven történő fogalmazás különbségére is utalnak. Bár kis mintán tudtunk csak összevetést végezni, a feltárt típusok jellegzetes módon különböznek nemcsak az anyanyelvi és idegen nyelvi tervezési folyamatokat illetően, hanem a különböző nyelvek tipológiai eltérései miatt is. Míg a biztos nyelvi kompetencia lehetővé tesz olyan tempójú beszédtervezést, amelynek eredményeként a beszélő már a következő szó artikulációját indítja el, ez egy kevésbé tudott idegen nyelven kevésbé valószínű. Ugyanígy másféleképpen jelentkeznek a mentális lexikonban való keresés nehézségei, hiszen míg az anyanyelvünkön nagyjából ugyanannak a fogalmi tartalomnak a kifejezésére több különböző stílusértékű szó közül választhatunk, s általában nem esik nehezünkre (például a műfaj alapján) kiválasztanunk a kontextusba megfelelőt, idegen nyelven egy bizonyos kompetenciaszint alatt nemcsak nem áll rendelkezésünkre a mentális lexikonban több lehetséges szó, amelyek közül választhatunk, de sokszor még a stílus- és fogalmi különbségekről sincs tudásunk. Az idegen nyelvű szövegekben előforduló korrekciókat áttekintve azt találjuk, hogy ezek a típusú tévesztések anyanyelvi szöveg alkotásakor nem jellemzőek (nemcsak a tipológiai eltérések miatt), illetve, még ha felszíni szerkezetükben hasonlóak meg is jelenhetnek, a mögöttes tényezők valószínűleg a legtöbb esetben egészen mások. Természetesen ez is – mint minden beszédtervezési folyamat – nagymértékben függ a nyelvtudás szintjétől.

Irodalom

- Butterworth, B. (1989): A beszédpszünetek adaléka. In: A beszédmegértés és a beszédprodukció pszichológiája. Szöveggyűjtemény. Szerk.: Pléh Csaba. Tankönyvkiadó. Budapest, 251-274.
- Dechert, H. W. (1983): How a story is done in a second language. In: Strategies in interlanguage communication. Eds.: Færch, C. – Kasper, G. Longman. London–New York, 175-195.
- Deme László (1971): Mondatszerkezeti sajátosságok gyakorisági vizsgálata (magyar szövegek alapján). Akadémiai Kiadó. Budapest.
- Færch, C. – Kasper, G. (1983): Plans and strategies in foreign language communication. In: Strategies in interlanguage communication. Eds.: Færch, C. – Kasper, G. Longman. London–New York, 20-60.
- Gósy Mária (2002): A megakadásjelenségek eredete a spontán beszéd tervezési folyamatában. Nyr. 126/2. 192-203.
- Németh T. Enikő (1996): A szóbeli diskurzusok megnyilatkozáspéldányokra tagolása. Akadémiai Kiadó. Budapest. (Nyelvtudományi Értekezések 142. sz.)
- Raupach, M. (1983): Analysis and evaluation of communication strategies. In: Strategies in interlanguage communication. Eds.: Færch, C. – Kasper, G. Longman. London–New York, 199-209.
- Tarone, E. (1983): Some thoughts on the notion of 'communication strategy'. In: Strategies in interlanguage communication. Eds.: Færch, C. – Kasper, G. Longman. London–New York, 61-74.
- Varga, L. (1999–2001): The unit of the Hungarian intonation. Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös nominatae. Sectio Linguistica tomus XXIV. Red.: Szathmári, I. ELTE Eötvös Kiadó. Budapest, 5-13.
- Wacha Imre (1988): Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In: Beszélt nyelvi tanulmányok. Szerk.: Kontra Miklós. MTA Nyelvtudományi Intézete. Budapest, 102-158.

ZÖNGÉSEDÉSI ÉS ZÖNGÉTLENEDÉSI FOLYAMATOK A /j/ FONÉMA HANGREALIZÁCIÓIBAN

Menyhárt Krisztina

Bevezetés

A /j/ fonéma fonológiai és a [j] mássalhangzó fonetikai szempontú osztályozása mind a mai napig számos vita forrása. A problémák jelentős része abból fakad, hogy az osztályba sorolási kísérleteknek gyakran keverednek fonológiai és fonetikai, artikulációs és akusztikai szempontok, illetve a hagyománykövető és az újító megközelítések.

A /j/ hagyományos magyar szakirodalmi besorolása alapján **zöngés, (medio)palatális réshang**, amelynek fonetikai jelölése [j] – vö. Kassai 1998; Laziczius 1943/1963; Szende 1976; Vértes O. 1982. Ez az értelmezés a 'réshang' fogalmának kiterjesztéséből fakad, amelyet így nemcsak a frikatíva, spiráns hangosztályra használnak, hanem belefoglalják a [j] és az [l] – oldalréshang – hangokat is (Siptár 1995; 2001). Azt, hogy a [j] hang akusztikailag különbözik a többi réshangtól, számos szerző elismeri, például: „A [j] spektrumának első fázisa az [i]-re emlékeztet, azzal a különbséggel, hogy a második és a harmadik formáns a [j] esetében egy kicsit magasabb.” (Magdics 1965, 63), vagy „A [j]-nek határozottabb formánsszerkezete van, mint a többi réshangnak. Az [i]-től szinte csak formánsainak szélesebb és ugyanakkor gyengébb volta különbözteti meg.” (Kassai 1998, 119). Más szerzők (például Dressler–Siptár 1989; Siptár 1995; 2001) szerint a /j/ alaprealizációja sem fonológiai, sem fonetikai tulajdonságai alapján nem tekinthető réshangnak. Ennek oka egyrészt, hogy a /j/ nem vesz részt az obstruensek zöngésségi hasonulásában, másrészt pedig képzésekor nem keletkezik a réshangokra jellemző zörej.

Felmerül a kérdés, hogyan is kellene akkor osztályozni a /j/ fonémát. A nemzetközi fonetikai szakirodalomban (vö. Pickett 1999; Stevens 1998) igen gyakran a 'glide', vagyis siklóhang, más néven félmagánhangzó elnevezést használják. Erre példákat magyar szerzők korai munkáiban is találunk, például Gombocz (1925/1940) a [j]-t dif-

tongus félhangzós elemeként jellemzi a [sa:j] vagy a [tɛj] szavakban, illetve Deme (1953, 41) szerint is a „[j] legtöbbször különböző tágaságú, általában illabiális, palatális félhangzók gyűjtőneve”. Ezt az álláspontot azonban számos fonológus és fonetikus nem osztja azon az alapon, hogy így diftongus keletkezne, holott a mai magyar irodalmi nyelv fonológiai rendszerében nincsenek diftongusok (vö. Kassai 1984; Siptár 1995). E tényt támasztják alá, hogy (i) a /j/ minden magánhangzó előtt és után is előfordulhat, így a diftongusok bevezetése megháromszorozná a magyar hangállományt, (ii) a /j/ kettőzhető, továbbá (iii) *-val*, *-vel* ragos alakjai [tɛj: ɛl].

A legegyszerűbben a /j/-t mint nem-nazális szonoráns mássalhangzót határozhatjuk meg (Siptár 1995). Ebbe az alosztályba tartozó mássalhangzókat a likvida elnevezéssel illetik, ide tartozik még az /l/ és az /r/ fonéma is. Gondot okozhat, hogy a /j/ semmilyen általános fonológiai vagy fonetikai indokok alapján nem sorolható egy osztályba az /l/ és az /r/ fonémákkal. Megoldásként bevezethető egy új alosztály – az approximáns, amelybe egyedül a /j/ fonéma tartozna (Dressler–Siptár 1989). Ez utóbbi nézetét későbbi munkáiban Siptár (vö. 1995; 2001) felülvizsgálja, és azt javasolja, hogy terjesszük ki a likvidák osztályát a /j/ fonémára is, így a szonoránsokat két csoportra lehet osztani: nazálisok és likvidák. Fonológiai szempontból ez a felosztás elfogadható, azonban mi a helyzet a [j]-vel fonetikai szemszögből, vagyis beszédhangként?

A képzés szerinti besorolási nehézségek kiküszöbölésére Laver (1994) a szájüregben keletkező szűkület fokozatai alapján osztályozza a mássalhangzókat. Megkülönböztet teljes szűkületet (zár), amikor a levegő kiáramlását megelőzően valamilyen akadályba ütközik, így jönnek létre a zárhangok. A második esetben kis rés keletkezik a szájüregben, ami méreténél fogva turbulens zörej megjelenését eredményezi (ez a réshangok képzési módja). Végül a harmadik fokozatnál a levegő minden turbulencia nélkül áramlik át a szájüregben keletkezett viszonylag nagyméretű résen keresztül. Így keletkeznek az approximánsok, vagyis a közelítőhangok. Összefoglalásként megállapíthatjuk, hogy fonológiai szempontból a magyar nyelv /j/ fonémáját a nem-nazális szonoránsok (zengőhangok) likvida vagy approximáns alosztá-

lyába sorolják (megközelítéstől függően), fonetikai szempontból azonban a [j] hang egyértelműen zöngés, palatális approximáns (vö. Crystal 1998; Gósy 2003; Siptár 2001).

A /j/ fonéma [j] hangrealizációja mellett a hangkörnyezet hatására azonban más hangmegvalósulások is létrejöhetnek. A [j] hang magánhangzó előtt vagy után jelentkezik, azonban a mássalhangzók között vagy mássalhangzó után abszolút szóvégen keletkezett változata tőle igen nagy mértékben eltérhet. Ezekben az esetekben valódi palatális réshang jön létre, amelynek lehet zöngétlen [ç] vagy zöngés [j̥] realizációja, attól függően, hogy zöngés vagy zöngétlen mássalhangzók között szerepel-e (vö. Siptár 2001). A szakirodalmi adatok alapján tehát megállapíthatjuk, hogy a magyar nyelvben a /j/ fonéma három hangban realizálódhat: zöngés, palatális approximánsként [j], illetve zöngés [j̥] és zöngétlen [ç] palatális réshangként.

A jelen tanulmány témája a /j/ fonéma hangrealizációinak elemzése akusztikai-fonetikai szempontból. Célunk viszonylag nagyobb korpuszon és több beszélő alapján vizsgálni a /j/ fonéma változatait a különböző hangkörnyezetekben, hogy objektív adatokkal támasszuk alá vagy cáfoljuk a korábbi – főleg elméleti jellegű, illetve csak korlátozott mértékben kísérleteken alapuló – kutatások eredményeit.

Anyag és módszer

A /j/ fonéma vizsgálatához egyszerű mondatokból álló korpuszt használtunk, amelyet hat személlyel – 3 nő és 3 férfi – olvastattunk fel.

1. táblázat: A résztvevők beszédsebessége (hang/másodperc)

Artikulációs tempó	Lassú	Közepes	Gyors
nők	10,7 h/s	11,2 h/s	12,4 h/s
férfiak	10,8 h/s	11 h/s	13 h/s

A résztvevők átlagéletkora 40 év volt, mindnyájan felsőfokú végzettséggel rendelkeztek. Artikulációs tempójuk alapján a kísérleti személyeket három csoportra osztottuk: lassú, közepes és gyors beszélők (1. táblázat).

A /j/ fonéma különböző megvalósulásait mondatba ágyazva vizsgáltuk a következő hangkörnyezetekben:

- zöngés mássalhangzó + /j/ + zöngés mássalhangzó, pl.: *Hívj de-rűs hangon.*;
- zöngés mássalhangzó + /j/ + magánhangzó, pl.: *Szívjon be sok friss levegőt.*;
- zöngés mássalhangzó + /j/ + zöngétlen mássalhangzó, pl.: *Hívj fel, ha tudsz valamit.*;
- zöngétlen mássalhangzó + /j/ + zöngétlen mássalhangzó, pl.: *Rakj fel mindent az asztalra.*;
- zöngétlen mássalhangzó + /j/ + magánhangzó, pl.: *Harapjon bele az almába!*;
- zöngétlen mássalhangzó + /j/ + zöngés mássalhangzó, pl.: *Lépj be a terembe!*;
- magánhangzó + /j/ + magánhangzó, pl.: *Olajos lett a ruhám.*;
- magánhangzó + /j/ + zöngés mássalhangzó, pl.: *Tegyél cukrot a tejbe!*;
- magánhangzó + /j/ + zöngétlen mássalhangzó, pl.: *Kinyílt az ajtó.*

Minden hangkörnyezettípusból nyolc különböző mondat szerepelt, így alakult ki a 72 mondatból álló vizsgálati anyag. Miután a kísérletben 6 személy vett részt, a /j/ fonémának összesen 432 realizációját vizsgáltuk. A mondatokat digitális hangfelvevő készülékkel rögzítettük, majd CSL 4300B digitális jelfeldolgozó segítségével elemeztük a /j/ fonéma különböző realizációinak akusztikai-fonetikai szerkezetét.

A zöngés [j], illetve [j̥] hangok esetében vizsgáltuk az első három formáns értékeit, míg a zöngétlen [ç] mássalhangzónál annak zörejelemeinek frekvenciaszerkezetét, illetve mindhárom realizációnál a hangok időtartamát. Összesen mintegy 2000 adatot dolgoztunk fel, emellett végeztünk statisztikai elemzéseket is (ANOVA).

Eredmények

Első lépében azt vizsgáltuk, hogy a /j/ fonéma lehetséges megvalósulásai – [j], [ɟ] és [ç] – hogyan valósulnak meg az eltérő hangkörnyezetek függvényében.

Adataink alapján megállapíthatjuk, hogy a /j/ fonéma magánhangzók mellett vagy között mindig zöngés approximánsként [j], zöngés mássalhangzók között csak a [ɟ] zöngés réshangként, míg zöngétlen mássalhangzók között kizárólag zöngétlen réshangként [ç] realizálódik. Ha a fonéma zöngés és zöngétlen (vagy fordítva) hangok között szerepel, mind zöngés, mind pedig zöngétlen változatban előfordul. Az összes elemzett adat között a zöngés [j] hang előfordulási aránya 80% volt, a zöngés réshangé [ɟ] 10,5%, míg a zöngétlen [ç] hangé 9,5%. Ez a tény a /j/ fonéma réshang-realizációinak korlátozott megvalósulását igazolja.

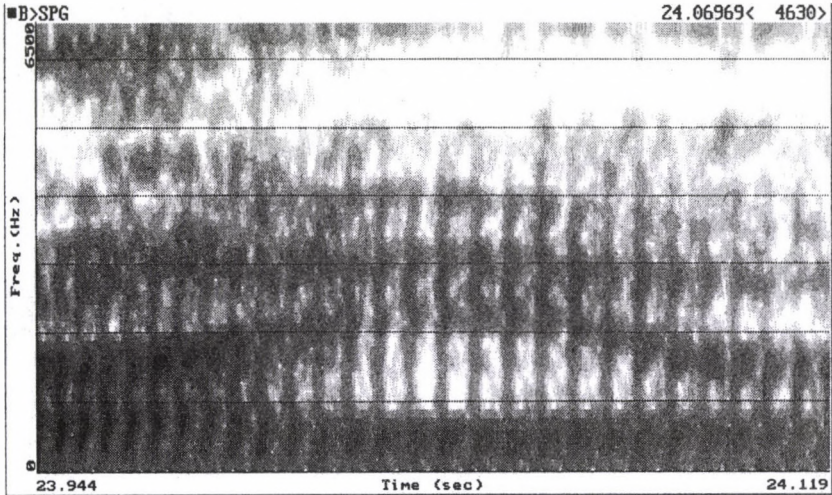
Adataink azt is megmutatták, hogy mássalhangzó-környezetben a /j/ fonéma igen gyakran – átlagosan 38,5%-ban – nem valósul meg. Az esetek egy részében csupán kiesik [le:pkɪ], [vi:vmɛg], míg más esetekben törlődése következtében a mássalhangzó-kapcsolatokra jellemző koartikulációs folyamatok jelentkeznek, például: [hɔrɔb:ɛɪ], [lɔgbɛɪ], [hi:f:ɛɪ].

Siptár (2001) a /j/ fonéma törlődésének okát a lezser, gyors beszédben látja, bár feltételezhető, hogy az artikulációs ökonómia is szerepet játszik. Miután a résztvevők artikulációs tempójuk alapján három csoportot alkottak, megvizsgáltuk a beszédtempó és a /j/ fonéma megvalósulási arányai közötti összefüggést. Megállapítottuk, hogy a /j/ legtöbbször – 59,3%, illetve 53%-ban – a lassú és a gyors beszélőknél törlődik (vagyis a „szélsőséges” esetekben), míg a közepes tempóban beszélőknél alig 6,3%-ban nem valósul meg.

A [j] zöngés, palatális approximáns akusztikai-fonetikai vizsgálata

Az első ábrán látható a zöngés [j] hangszínképe női ejtésben az *Olajos lett a ruhám.* példamondat egy részében. Jól megfigyelhető a [j] hang magánhangzószerű formánsstruktúrája, amire már Magdics

(1965) is felhívta a figyelmet. Emellett jellegzetes hangátmeneti szakaszok jelentkeznek az [ɔ] és az [o] hang felé, illetve látható, hogy a [j] formánsait kisebb intenzitás jellemzi, mint a környező magánhangzókét (vö. Vértes O. 1982; Gósy 2002).

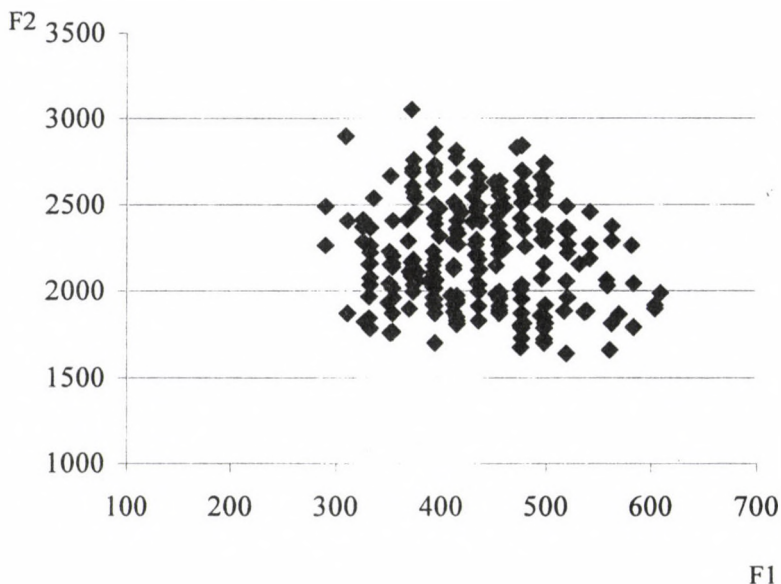


1. ábra

A [j] hang az [ɔjo] hangkapcsolatban

A 2. ábrán mutatjuk a [j] első és második formánsának viszonyító ábrázolását, magánhangzó melletti helyzetben az összes beszélőnél. A zöngés approximáns formánsadatai igen nagy területen helyezkednek el. Az F1 300 és 600 Hz között szór, átlagértéke 413 Hz, míg az F2 1700 és 2900 Hz között helyezkedik el, átlagosan 2227 Hz-en (az időtengely alapján mérve). Az F1 átlagos eltérése 24,4, az F2-jé pedig 142,2.

A nők és férfiak közötti eltéréseket is regisztráltuk: a női beszélők első formánsának átlagértéke 456 Hz, a férfiaknál ez 369 Hz, míg a második formáns 2372, illetve 2098 Hz-en helyezkedik el átlagosan. A nők adatai kevésbé szórnak, átlagos eltérésük 37,5 (F1) és 214,4 (F2), míg a férfiaké 57,2 és 340,5.



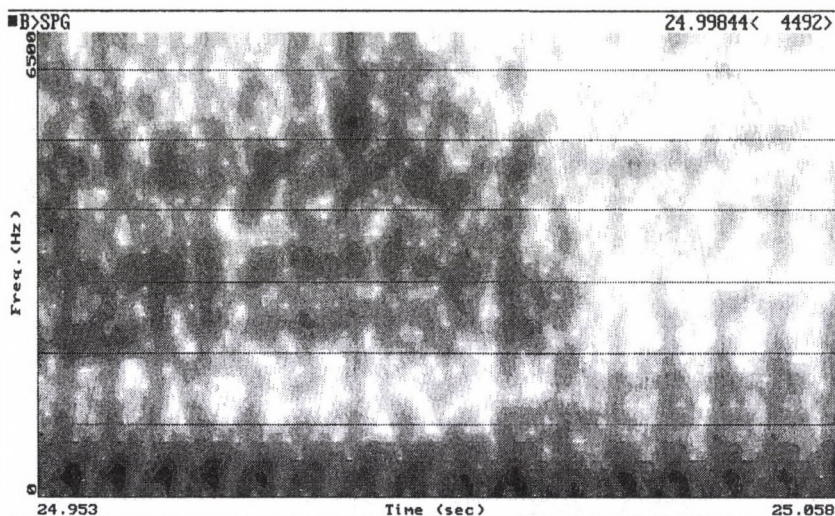
2. ábra

A [j] hang első és második formánsának eloszlása (Hz)

A [j] zöngés, palatális réshang akusztikai-fonetikai vizsgálata

A 3. ábrán látható a [j] hangszínképe a *Várj meg az állomáson*. példamondat egy részében. Jól megfigyelhető egyrészt a zöngés jelleg, másrészt pedig a kevésbé hangsúlyos formánsstruktúra, amelyhez a réshangokra jellemző zörejelemek társulnak. Az első formáns átlagosan 371,2 Hz-en, míg a második 2226,2 Hz-en helyezkedik el. Átlagos eltérésük az F1-nél 32,8, a F2-nél pedig 189,6 volt.

A különböző hangkörnyezetekben előforduló formánsátlagokat a nők, illetve a férfiak esetében a 3. táblázatban adtuk meg. Miután az előfordulások száma egyik beszélő esetén sem volt azonos, az átlagadatok között nagyobb eltérések fordultak elő. A férfiak formánsátlagai valamivel szélesebb tartományban szórtak, mint a női beszélők.



3. ábra

A [j] palatális zöngés réshang színképe női ejtésben
a [rjm] hangkapcsolatban

3. táblázat: A zöngés [j] formánsátlagai női és férfi ejtésben (Hz-ben)

Hangkörnyezet	Formáns	Nők	Férfiak
zöngés msh.+ [j] + zöngés msh.	F1	404	313
	F2	2552	2243
zöngés msh. + [j] + zöngétlen msh.	F1	342	333
	F2	2478	2021
zöngétlen msh.+ [j] + zöngés msh.	F1	451	341
	F2	2283	2021
átlagos eltérés	F1	50,9	52,4
	F2	149,8	198,2

Összevetve a [j] és a [j] hangok első és a második formánsának átlagértékeit, láthatjuk, hogy a különbség szignifikáns – F1: $p < 0,0001$; $F2 < 0,024$, illetve az átlagos eltérés a zöngés approximánsnál kisebb

határok között mozog mindkét formáns esetében (24,4/32,8; 142,2/189,6).

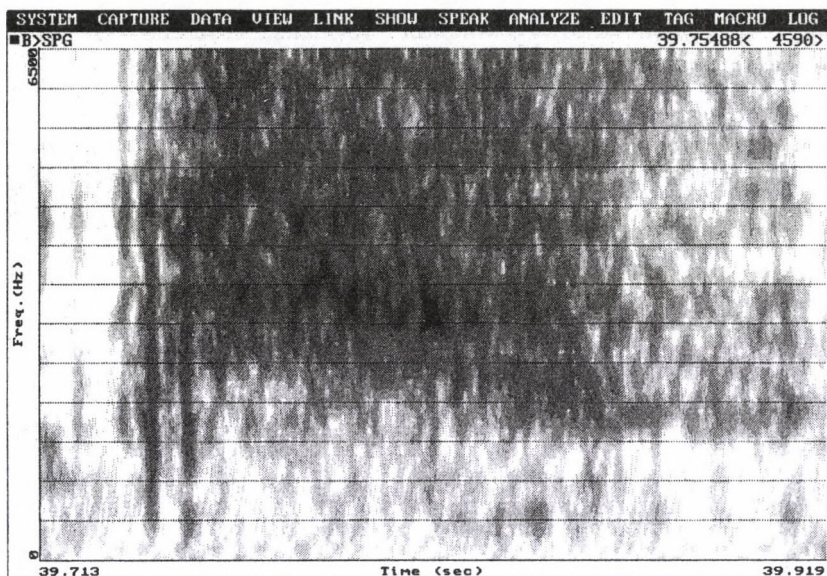
A [j] és a [j̥] hangok harmadik formánsának értékeit a különböző hangkörnyezetekben együttesen mutatjuk be a 4. táblázatban. A kapott adatok alapján megállapítottuk, hogy bár a lényeges eltérés nem mutatkozik aközött, hogy magánhangzó mellett vagy mássalhangzók között jelenik-e meg az elemzett hang ($p < 0,856$), a mássalhangzó-környezet adatai nagyobb arányban szórnak, ugyanúgy, mint az első két formáns esetében.

4. táblázat: Az F3 személyenkénti átlagértékei (Hz)

Személyek	Magánhangzó mellett [j]	Mássalhangzók között [j̥]
Nő 1.	3769,2	3416,6
Nő 2.	3555,4	3707,9
Nő 3.	3437,8	3481,8
átl. eltérés	331,5	
Férfi 1.	3427,8	3336,8
Férfi 2.	3471,8	3227
Férfi 3.	3327,6	3143,3
átl. eltérés	300,5	
csoportátlag	3498,3	3385,6
átl. eltérés	109,3	149,8

A [ç] zöngétlen réshang akusztikai-fonetikai vizsgálata

A 4. ábrán mutatjuk a [ç] zöngétlen, palatális réshang hangszínekét a *Rakj fel mindent az asztalra*. példamondat egy részében, női ejtésben. Az ábrán pontosan megfigyelhető a hang zöngétlen mivolta (a zönge hiánya), illetve a réshangokra jellemző zörejes jelleg. A vizsgált hang akusztikumát a zörejelemek frekvenciaszerkezetének elemzésével írtuk le (5. táblázat).



4. ábra

A [ç] zöngétlen réshang színképe a [kçf] hangkapcsolatban

5. táblázat: A zöngétlen [ç] akusztikai jellemzői (Hz-ben)

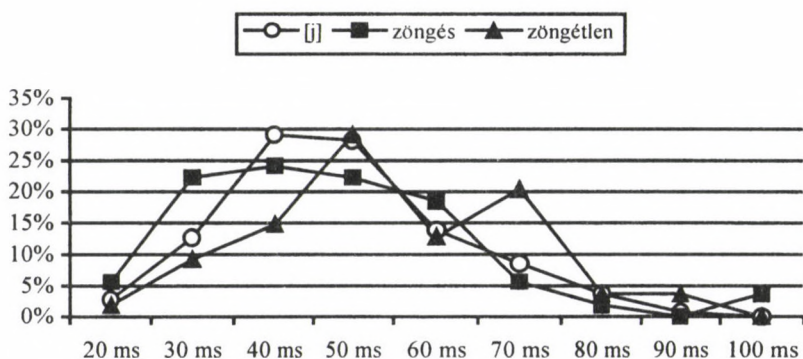
Hangkörnyezet	Ztl. msh.+[ç] + zöng. msh.	Ztl. msh. + [ç]+ztl. msh.	Zöng. msh. + [ç] + ztl. msh.
frekvenciaszerkezet	1266–5897	1349–6375	1077–6065
intenzív tartomány	3063,2	2980,9	3139,9

Az 5. táblázat adataiból látható, hogy a [ç] hang frekvenciatartománya igen széles – megközelítőleg 1000 és 6000 Hz között – mozog, beszélőtől és hangkörnyezettől függően. A nőknél ez a tartomány 1660 és 6375 Hz, míg a férfiaknál 1077 és 6209 Hz között helyezkedik el. A legintenzívebb része a nőknél átlagosan 3495 Hz-en, míg a férfiaknál 2948 Hz-en helyezkedett el, átlagos eltérése 217,9, illetve 388,8.

A vizsgált hangok időtartama

A három hang akusztikai-fonetikai vizsgálatakor foglalkoztunk az időtartam meghatározásával is. A korábbi kutatásokból ismert az a tény, hogy a [j] hang időtartam-megvalósulásai igen széles határok között mozoghatnak (vö. Magdics 1965; Kassai 1982; Olasz 1985; 2001). A jelen kutatásban kapott adatok is hasonló tendenciát mutatnak. Az 5. ábrán látható a /j/ fonéma hangmegvalósulásainak időtartam-eloszlása 10 ms-os időszavokban. A hangok időtartamát folyamatosan auditív és vizuális ellenőrzés mellett határoztuk meg.

A [j] hang esetében a legnagyobb arányban a 40 és 50 ms-os hangidőtartamok fordulnak elő, a [j]-nél ez a tartomány 30 és 60 ms közé esik, míg a [ç] hangnál két csúcspont is megfigyelhető: 50, illetve 70 ms-nál. A két zöngés hang időtartamátlagai között a különbség nem szignifikáns ($p < 0,579$), a zöngétlen réshang adatai azonban mindkettőtől szignifikánsan eltérnek – [ç]/[j]: $p < 0,0004$, [ç]/[j]: $p < 0,003$.



6. ábra

A három vizsgált hang [j, j̥, ç] időtartam-megoszlásai

Megállapítottuk a három elemzett hang személyenkénti átlagidőtartamait is (6. táblázat). Adataink nem mutatnak ugyan jelentős eltérést sem a személyek, sem pedig a különböző hangminőségek tekinteté-

ben, azonban jól látszik, hogy az átlagos eltérés értéke különbözik – legkisebb a zöngés approximánsnál és legnagyobb a zöngétlen rés-hangnál. Általánosságban elmondható, hogy itt is ugyanaz a tendencia érvényesül, mint a formánsadatoknál: a mássalhangzók közötti /j/ fonéma hangmegvalósulásainak időtartama szélesebb határok között mozog, mint a magánhangzók melletti, zöngés [j].

6. táblázat: A vizsgált hangok átlagos időtartamai (ms)

Személyek	[j] hang	[j̥] hang	[ç] hang
Nő 1.	50,3	42,6	48,9
Nő 2.	55,5	51,9	43,5
Nő 3.	52,9	45,9	66,4
Férfi 1.	46,1	50,95	50,4
Férfi 2.	55,3	50,2	65,1
Férfi 3.	55,4	39,9	71,9
csoportátlag	52,6	46,9	57,7
átl. eltérés	2,92	4,1	10,1

Összevetettük a kapott adatokat a korábbi kutatások eredményeivel. A legkorábbi vizsgálat (Magdics 1965) a [j] hangra 90 ms-os átlagot ad meg, Kassai (1982) pedig 73 ms-ot. Adatainkhoz mind időben, mind pedig időtartamértékben legközelebb Olaszy (2001) eredményei állnak, akinél a [j] 59 ms-os átlagban valósul meg. Ezek a tények jól mutatják azt a beszédsebesség-növekedést, amely az eltelt időszakban ment végbe.

Összefoglalás

Az elvégzett vizsgálat objektív adatokkal támasztotta alá azt a korábbi megállapítást (vö. Siptár 2001; Gósy 2002; 2003), amely szerint a magyar nyelvben a /j/ fonémának három különböző hangrealizációja jelenik meg. A vizsgált adatkörpuszban a leggyakrabban (az esetek 80%-ában) a [j] zöngés, palatális approximáns fordul elő, amelyet a magánhangzókéra emlékeztető formánsstruktúra jellemez. Közel azonos arányban (10-10%-ban) jelentkeznek a zöngés [j̥], illetve a zön-

gétlen [ç] palatális réshangok, amelyeknek megjelenése erősen függ a hangkörnyezettől. Az előbbire a nem túl hangsúlyos, de felismerhető formánsstruktúra mellett jelentkező zörej jellemző (zöngés-zörejes réshang), míg az utóbbinál egyértelműen a zöngétlen réshangoknál látható zörejesség mutatkozik.

A három hang időtartama igen tág határok között mozog, a két zöngés hang egymástól nem, azonban a zöngétlen palatális réshangtól szignifikánsan különbözik. A hangok időtartama jól tükrözi az utóbbi évtizedekben végbe ment beszédtempó-gyorsulást.

Miután a kísérletsorozat viszonylag korlátozott számú beszélőn és izolált ejtésű mondatkorpuszon alapult, illetve a három hangot csak szó belseji helyzetben vizsgáltuk, fontos lenne a jelen megállapításokat nagyobb spontánbeszéd-anyagon is ellenőrizni.

Irodalom

- Crystal, D. (1998): A nyelv enciklopédiája. Osiris. Budapest.
- Deme László (1953): A magyar nyelvjárások néhány kérdése. NytudÉrt. 3. Akadémiai Kiadó. Budapest.
- Dressler, W. U. – Siptár, P. (1989): Towards a natural phonology of Hungarian. *Acta Linguistica Hungarica* 39: 29-51.
- Gombocz Zoltán (1925/1940): Magyar fonétika. In: Összegyűjtött művei. II/1. Budapest, 3-61.
- Gósy Mária (2002): Fonetikai összefoglaló. Budapest. (Kézirat).
- Gósy Mária (2003): A magyar mássalhangzók osztályozásáról. In: Köszöntő könyv Kiss Jenő 60. születésnapjára. Szerk.: Hajdú Mihály – Keszler Borbála. ELTE Magyar Nyelvtudományi és Finnugor Intézet. Budapest, 180-186.
- Kassai Ilona (1982): A magyar beszédhangok időtartamviszonyai. In: Fejezetek a magyar leíró hangtanból. Szerk.: Bolla Kálmán. Akadémiai Kiadó. Budapest, 115-155.
- Kassai Ilona (1984): Kell-e a magyar köznyelvben diftongusnak lennie? *Nyelvtudományi Közlemények* 86: 152-154.
- Kassai Ilona (1998): Fonetika. Nemzeti Tankönyvkiadó. Budapest.
- Laver, J. (1994): The principles of Phonetics. CUP. Cambridge.
- Laziczus Gyula (1943/1963): Fonetika. Tankönyvkiadó. Budapest.
- Olaszy Gábor (1985): A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise. NytudÉrt. 121. Akadémiai Kiadó, Budapest.

- Olasz Gábor (2001): A beszéd akusztikai-fonetikai elemzése és modellezés különös tekintettel a korszerű beszédépítés követelményeire. Doktori értekezés. Budapest.
- Pickett, J. M. (1999): The Acoustics of Speech Communication. Allyn and Bacon. Boston.
- Stevens, K. N. (1998): Acoustic Phonetics. MIT Press, Cambridge, Mass.
- Siptár Péter (1995): A magyar mássalhangzók fonológiája. MTA Nyelvtudományi Intézet. Budapest.
- Siptár Péter (2001): Egy jottányi fonológia. In: Beszédkutatás 2001. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 1-16.
- Szende Tamás (1976): A beszéd folyamat alaptényezői. Akadémiai Kiadó. Budapest.
- Vértes O. András (1982): A magyar beszédhangok akusztikai elemzésének kérdései. In: Fejezetek a magyar leíró hangtanból. Szerk.: Bolla Kálmán. Akadémiai Kiadó. Budapest, 71-113.

SZÓLEHÍVÁS VIZUÁLIS STIMULUS ALAPJÁN – ANGOL–MAGYAR KÉTNYELVŰEK KÖRÉBEN

Navracsics Judit

Bevezetés

A nyelv állandó kapcsolata a külső világgal a jelentésben tükröződik. A jelentés az emberi agyban, a mentális lexikonban rögzül, és a világnak az általunk tapasztalt formáját írja le. A jelentés problémakörének fő kérdése az, miként tudjuk nyelvileg megvalósítani a nem nyelvi valóságot, tehát miként tudjuk tükröztetni a konkrét valóságot a feltételezettel (Bolinger 1975). A külső világot szegmentáljuk, a valóságot két részre osztjuk: a mindenki által és a csak magunk által láthatóra (objektív, szubjektív tapasztalatok). Minél tapasztaltabbak vagyunk, annál kevésbé lesz anyagivá (kézzel foghatóvá) a világ, annál jobban tudunk elvonatkoztatni.

A kultúra az egyén környezetének sajátosságát, szokásait jelenti, azokat a jellegzetességeket, amelyek az adott kultúrát elválasztják egy másik kultúrától (Hudson 1986). Vajon a nyelvek és a kultúrák mennyire különböznek egymástól?

Jól ismert a Sapir és Whorf által képviselt nyelvi viszonylagosság, relativitás, amely szerint a nyelv döntő mértékben meghatározza azt, hogy miként látjuk a valóságos világot: „A való helyzet az, hogy a világképet jelentős mértékben a nyelv határozza meg. Nincs két olyan nyelv, amely ugyanannak a társadalmi valóságnak a kifejezője volna. Azok a világok, amelyekben a különböző társadalmak élnek, különböző világok. Nem egyszerűen arról van tehát szó, hogy ugyanaz a világ fejeződik ki különböző módon két különböző nyelv esetében.” (Sapir 1971). A Sapir–Whorf elméletet azóta számosan bírálták, ellenkísérletekkel igazolták, hogy a tétel nem ilyen egyszerű. Maga Sapir és Whorf is többször változtatott nézetein. Valójában a világmindenséget a természeti környezetünk, a kommunikációs és kognitív szükségleteink szerint osztjuk fel, nem pedig a nyelvünk szerint (Hudson 1986).

Mivel minden nyelv egy sajátos, egyedi rendszer, más-másféle, nem pontosan azonos jelentéskategóriákat kínál a külvilág felbontására, tagolására. A külvilág azonban nemcsak jelentéskategóriákon keresztül közelíthető meg, hanem közvetlenül is. Ha megégetjük a kezünket, magunk is megtapasztaljuk, mit jelentenek a *meleg, forró* szavak. Míg tehát a jelentések világa népenként más és más, s elvben összehasonlíthatatlan és összemérhetetlen, addig a denotátumok, amelyek között az emberek élnek, voltaképpen ugyanazok az egész emberiség, illetve az emberiség igen nagy csoportjai számára (Antal 1978). Kiefer szerint a nyelvi jelentés szorosan összefügg a megismeréssel, vagyis azzal, hogy a bennünket körülvevő világot hogyan észleljük, hogy az észlelés eredményeit miként dolgozzuk fel, hogyan ábrázoljuk mentálisan (Kiefer 2000).

A fentiek alapján felvetődik a kérdés, hogy vajon a két- és többnyelvű emberek, akik két vagy több nyelvnek vannak birtokában és két vagy több kultúrát tudnak magukénak, miként osztják fel a külvilágot, a valóságdarabokat miként észlelik és nyelvileg hogyan fejezik ki. A kérdésre a választ két kísérlet – egy képmegnevezéses és egy szógyűjtéses teszt – eredményeinek elemzésével kíséreltük meg megadni. A vizsgálatokkal kettős célunk volt: (i) a magyar egynyelvű beszédmódban levő kétnyelvűek szólehívási mechanizmusainak (kódváltásainak) megfigyelése, (ii) annak vizsgálata, hogy a képeken ábrázolt denotátumokat miként nevezik meg a kétnyelvű, kétkultúrájú emberek, azaz közös-e a fogalmi szint.

Hipotéziseink a következők voltak:

1. Minél fiatalabb a kísérleti személy, annál kevesebb pontos (hiponim) megnevezést ad.
2. Biológiaiilag minél idősebb a kísérleti személy, minél nagyobb a kognitív érettsége, annál több szót gyűjt a képi stimulus alapján.
3. A kétnyelvű minél rövidebb ideje van Magyarországon, annál nagyobb lesz az általa alkalmazott kódváltások száma.

Kísérleti személyek, módszer

15 angol–magyar kétnyelvű személy adatait elemeztük. A kísérleti személyeket három kategóriába soroltuk, az életkoruknak megfelelően. Az első csoportot öt gyermek (6-14 évesek, átlagéletkoruk 10 év),

a második csoportot öt fiatal felnőtt (15-23 évesek, átlag életkoruk 19 év), míg a harmadik csoportot öt felnőtt (24-32 évesek, átlag életkoruk 29 év) alkotja.

A kísérleti személyek közül csak ketten nem élnek állandóan Magyarországon: Szandi, akinek édesanyja magyar (és a gyermekkel állandóan magyarul beszél), valamint Tibor, aki 18 éves koráig Magyarországon élt, azóta viszont (15 éve) az USA-ban, illetve az utóbbi három évben amerikai katonaként Németországban.

A résztvevők több vonatkozásban is eltérnek egymástól:

(i) 9 személy születésétől kezdve kétnyelvű, míg a többiek később váltak kétnyelvűvé (pl. 2, 10, 18, 24 éves korukban).

(ii) 10 személynek egynyelvűek a szülei, kétnyelvűségük tehát nem családi, hanem ideiglenes migráció vagy bevándorlás következménye (Skutnabb-Kangas 1984).

(iii) 5 személy csak a családban, 6 fő a családban és a közvetlen környezetben is, illetve 4 kísérleti személy a családon kívül, a mindennapi munkája során is használja az angolt.

Olivia, Christopher, Marie-France és Estella szüleikkel, Ati pedig kislányával otthon angolul beszél. Az iskolában, munkahelyen, a társadalmi élet különböző színterein azonban csak a magyar nyelvet használják.

Szandi és Tibor számára az angol a közvetlen környezet nyelve. Szandinak a magyar nyelvi inputot egyedül az édesanyja biztosítja, Tibor viszont kizárólag angol (illetve 3 éve német is) nyelvi környezetben él. Pál, Doris, Tim, angol szakos egyetemisták, tehát az otthonról (a korábbi közvetlen környezetből) hozott nyelvet az iskolában, tanulmányaik során is használják. Zsuzsi angol szakos tanárként hétköznapijainak nagy részét is angolul éli.

Tamás és Zsolt életében, amióta 3 éve Ausztráliából hazatelepült a család, nincs rendszerességgel jelen az angol, eltekintve az angol nyelvi óráktól, amelyeken az iskolában részt kell venniük. Kim és Jon szintén alkalmasszerűen tudja csak első nyelvét használni a magyar nyelvi környezetben (mindketten zenetanárok).

A kísérleti személyek két tesztet végeztek el. Az utasítások magyar nyelven hangzottak el, ezáltal a kétnyelvűeket magyar egynyelvű beszédmódba helyeztük. Nem korlátoztuk azonban a nyelvválasztásukat:

azt a nyelvet használhatták, amelyiken hamarabb megtörtént a képi stimulus által felidézett szó lehívása. Az első feladatban a képeken ábrázolt – alább felsorolt – dolgokat, jelentéseket kellett megnevezniük:

- 10 élőlény: madár, ló, béka, dinoszaurusz, bagoly, kakas, kutya, lepke, oroszlán, fa;

- 8 tárgy: busz, fotel, repülőgép, autó, templom, óra, karácsonyfadísz, nap;

- 6 foglalkozás: állatorvos, vadász, optikus, fogorvos, bűvész, zenész;

- 4 sportág: foci, súlyozás, síelés, rollerozás.

Az adott válaszokat két szempontból kategorizáltuk: a kódváltás és a megnevezés alapján (azaz a válaszuk mennyiben egyezik meg a kísérletvezető által etalonnak gondolttal).

A második feladat utasítása az volt, hogy gyűjtsenek még a képeken ábrázoltakhoz hasonlóakat, mindenegyus képnél 10 másodperc állt rendelkezésükre. Azt vizsgáltuk, hogy az egyes denotátumokhoz milyen típusú szavakat gyűjtöttek, a szólehívásokban mennyiben játszott szerepet a hiperonimia–hiponimia kapcsolat, milyen arányban szerepeltek mellérendelő viszonyban lévő megnevezések, szemben az alárendeltekkel és asszociatívokkal, valamint, hogy van-e összefüggés az életkor és a megnevezések jellege között. Ennél a feladatnál az egyes képekre adott összes válasz tükrében vizsgáltuk a kódváltás százalékos arányát. Feltérképeztük hogy melyik szónál történt a legaktívabb gyűjtés, milyen típusú asszociációkat lehet felfedezni. A képek megegyeztek az első feladatban bemutatottakkal.

Eredmények

A képmegnevezések elemzése

A képek megnevezésekor dominált a hiponim kifejezés, tehát a kísérleti személyek igyekeztek a lehető legpontosabban meghatározni a látottakat. Az élőlényeket, tárgyakat többnyire felismerték, és hasonló módon nevezték meg, azonban a sportágakat ábrázoló képek láttán többen figyelmen kívül hagyták az utasítást, miszerint a sportágat kellett volna megnevezni, helyette a sportolót nevezték meg.

A megnevezések típus és jel viszonylatának elemzése rámutatott, hogy azoknál a képeknél jelent meg kevés típus, amelyek a legkonkrét-

tabbak. Mivel a kísérleti személyek kétnyelvűek, a szemantikai reprezentációt és a lexikai lehívást külön választottuk, ezért az azonos szemantikai reprezentációt jelölő, de más nyelven lehívott szót 2 típusnak vettük.

A legnagyobb számú típus használata (10-10) jellemző a két sportág megnevezésénél (*rollerozás, súlyzózás*). Mind a két képi stimulus különböző szófajú elemek lehívását eredményezte: pl. igék (*rollerezik, korcsolyázik, játszik, futnak*), igéből képzett főnevek (*scootering, weight lifting, súlyemelő, weight lifter*), sőt, teljes szintagmák alkotását is (*going on the scooter, gyerek rollerezik*). A válaszadók megnevezései egyik esetben sem estek egybe a kísérletvezető által etalonnak tartott megnevezésekkel (*rollerozás és súlyzózás*). Problémásnak bizonyult a *karácsonyfadísz* és az *optikus* megnevezése is, amit szintén a magas típusszám jelez (9-9). Mindkét tárgynál percepciók különbségek tapasztalhatók: a karácsonyfadísz volt, aki *bombának* látta, az optikus pedig megosztotta a kísérleti személyeket: volt, aki *fodrásznak*, volt, aki *látszerésznek, szemésznek* vagy *kozmetikusnak* látta. Legnagyobb gyakorisággal mégis a *fodrász* szerepelt a megnevezések között. A *foci* és a *síelés* (8-8 típus) szintén sportágak, amelyek megnevezésekor a válaszadók megfélekeztek arról, hogy magát a sportágat kellett volna megnevezniük, ehelyett többnyire a sportolót nevezték meg (pl. *focista, footballer, soccer player, síelő, skier*), de itt is találhatunk szintagmatikus megnevezéseket (*kicking the ball, skating on the snow*).

A *madár* megnevezésekor (7 típus) különösen érdekes, hogy szinte kiegyensúlyozottan szerepelnek a hiperonim (*madár, bird*) és a hiponim (*golya, gém, heron, stork*) kifejezések. A képen egy gázlómadár látható, ezért a válaszadók kb. fele a hiperonim *madár, bird* szavakat hívta le, míg azok, akik a hiponim kifejezést választották, többnyire *golyának*, vagy *gémnek* látták, ami azt is jelenti, hogy a gázlómadarak legjellegzetesebb példányai a golyák és a gémelek.

Hat típus szerepel három foglalkozást jelölő képnél. Két kép megnevezésekor a leggyakoribb lehívás megegyezik az etalonnal (*zenész, állatorvos*), a harmadiknál az etalonként elvárttól (*bűvész*) eltér annyiban, hogy a kép hatására többen láttak varázslót, tehát a varázserő kifejezését a képben, így a legtöbb jel a *magician* típusnál jelent meg.

Mivel a zenész egy hegedűst ábrázol, előfordult, hogy szintén hiponim kifejezéssel éltek (*violinist, hegedűs*), sőt, egy asszociatív lehívás is történt (*cigány*). Az állatorvost ábrázoló kép egy esetben váltott ki hiperonim kifejezést (*physician*), két másik esetben pedig nem a foglalkozást nevezték meg (*pets, simogat*). Két tárgy, a *fotel* és a *templom* megnevezésekor szintén 6 típus fordul elő. A *fotel*nél egy (*szőnyeg*) nem tartozik az elvárt válaszok közé, a többi válasz mind ülőalkalmatosságot fejez ki. A *templom*, amely keretben van ábrázolva, több – a kép apróbb részleteire (*felhő, ablak*) vagy a megvalósításra (*kép, photo*), utaló – lehívást eredményezett. A *kutya* megnevezésekor csak egy esetben fordult elő, hogy a hiponim *farkaskutya* megnevezést használták, de akadt egy *goat* (kecske) is, ami téves lehívásnak számít. Az *autó* láttán egy kísérleti személy tartotta fontosnak az autó típusának megnevezését (*Rolls Royce*), egy *régi autó*, a többi pedig a magyar *autó*, illetve annak egy nyelven belüli, illetve nyelvek közötti szinonimái szerepeltek (*kocsi, car*).

Az 5 típussal szereplő képmegnevezések (kakas, nap, dinoszaurusz, fa, repülőgép) között már többnyire nyelvek közötti szinonimákat találunk, de még akadnak téves megnevezések: pl. *kakas* – *chicken*, *nap* – *sunflower*, *dinoszaurusz* – *kígyó*. 4-4 típussal találkozunk a *pillangó*, *óra*, *vadász* képeknél. Hiponim kifejezés a *falióra*, téves megnevezés a *catch a fish* és a *kalapál* (a vadásznál).

A 3, illetve 2 típussal szereplő *bagoly*, *fogorvos*, *ló*, *autóbusz*, *béka*, *oroszlán* képek egyértelmű nyelvek közötti szinonimákat hívtak le, a *ló* esetében egy gyermeknyelvi változat, a *horsie* is előfordul. Ebből arra is következtethetnénk, hogy ezeknek a szavaknak a lehívása a legkönnyebb, azonban, ha megnézzük az összes jel százalékos előfordulását, láthatjuk, hogy egyiket sem sikerült 100%-ban megnevezni (tehát nem mind a 15 kísérleti személy nevezte meg őket). 100%-os volt a megnevezés azonban a következő képeknél: *súlyozás*, *foci*, *síelés*, *madár*, *bűvész*, *kutya*, *autó*, *repülőgép*. Talán ezek a kategóriák azok, amelyek vagy a divat (*súlyozás*, *síelés*) vagy a gyermekkori emlékek (*foci*, *bűvész*) vagy a mindennapi élet által meghatározottan jelenlevők az életben és aktívak a mentális lexikonban (*madár*, *kutya*, *autó*, *repülőgép*).

A kódváltások elemzése

A kódváltás százalékos sorrendjét tekintve a következőket állapíthatjuk meg: a legtöbb kódváltás a *madár*, *bűvész*, *pillangó*, *kutya*, *súlyozás*, *fotel*, *kakas* és az *autó* képek hatására történt; csökkenően 53%-tól 40%-ig. Ez a sor a *fotel* és a *kakas* kivételével megegyezik a 100%-osan megnevezett képek sorával. Tehát ott, ahol mindenki megnevezte a képet, kódváltással is éltek, méghozzá nagy számban, amiből arra lehet következtetni, hogy a fogalmi rendszer közös (Snodgrass 1993). A legkevesebb arányú kódváltás a *béka*, *fogorvos*, *óra* képeknél fordult elő (ezek éppen a legkevesebb típust előidéző képek közé tartoznak). A magyar beszédmód tehát többnyire arra ösztönözte a kísérleti személyeket, hogy az ilyen egyértelmű képeknél csak nagyon ritkán éljenek a kódváltással. Ellentmondásos az eredmény, hiszen a *béka*, *fogorvos*, *óra* ugyanolyan szerves része a mindennapi életnek, mint azok a szavak, amelyek esetében gyakori kódváltás történt (*madár*, *bűvész* stb.). Helyesebb tehát, ha személyenként vizsgáljuk meg a kódváltási szokásokat.

Egyik hipotézisünk az volt, hogy minél rövidebb ideje van a kísérleti személy Magyarországon, annál több lesz a kódváltás a megnevezésekben. Öt kísérleti személy (Szandi, Olivia, Estella, Zsuzsi és Ati) esetében ez a hipotézis beigazolódott. Szandi ugyanis Angliában él és minden megnevezést angolul adott meg. Zsuzsi csak 4 éve van Magyarországon, válaszainak 68% angol nyelvű. Ati és Olivia 6 éve él itt, és mindössze 11%, illetve 4%-os az angol nyelvű szólehívásuk. Estella Magyarországon született és csak évente egyszer tölt bizonyos időt Angliában a nagyszülőknél, következésképpen nem is használt angol kifejezéseket.

Nem igazolt azonban a feltételezés a válaszadók 2/3-ánál. Az alábbi táblázatból kiderül, hogy – néhány kivételtől eltekintve – azok a kísérleti személyek éltek a kódváltás lehetőségével, akik az angolt nem csak a családon belüli kommunikációra használják, hanem a társadalom más színterein is. Ezzel magyarázható, hogy Szandi 100%-ban kódot vált, hiszen ő csak otthon, az édesanyjával beszél magyarul. Estella, aki csak az édesapjával beszél angolul, egyáltalán nem vált angolra, és Olivia is csak 4%-ban vált, hiszen a 8 éves kislány 6 éve magyar óvodába, illetve iskolába jár, az angolt csak otthon beszéli.

Ugyanezen okokból keveset vált Ati, Kim és Jon, hiszen a családi és baráti körön kívül nem használják az angol nyelvet. Érdekes Zsolt szólelvási stratégiája, aki, bár magyar szüleivel 3 éve költöztek haza, nem vált kódot (talán mert ő az idősebb), szemben öccsével, Tamással, aki 25%-ban vált. A legnagyobb arányban azok a kísérleti személyek váltanak, akik a családban is, valamint a családon kívül is rendszeresen használják mindkét nyelvüket: Pál, Doris, Tim angol szakos egyetemi hallgatók, Zsuzsi pedig angol szakos nyelvtanár. Két válaszadó nyelvválasztására nem tudunk magyarázatot adni: Christopher 90%-os és Tibor 0%-os kódváltásaira. Itt talán érzelmi okok lehetnek a háttérben, hiszen a kódváltás nem csupán nyelvi jelenség, sokszor fejezzük ki vele és általa a dolgokhoz való érzelmi hozzáállásunkat.

A szógyűjtés elemzése

Mennyiségi szempontból beigazolódott hipotézisünk, amely szerint a biológiai életkor emelkedésével nagyobb a gyűjtött szavak száma. Az első csoport összesen 113 szót tudott gyűjteni a megadott időn belül, a második 268-at, míg a harmadik 320-at.

Ha minőségileg is megvizsgáljuk a lehívott szavakat, akkor azonban a következőt tapasztaljuk: a fiatal felnőtt és a felnőtt csoport abban is különbözik a gyermekektől, hogy a konkrét megnevezéseken kívül asszociatív válaszokat is adnak. A felnőttek kevésbé tartották magukat a megadott utasításhoz: az általuk gyűjtött szavak nagy részben asszociatívok, gyakran az érzéseiket fogalmazták meg a képek látán.

A foglalkozásokat és sportágakat kifejező képek után általában fel kellett tenni a kérdést: milyen foglalkozásokat, illetve sportágakat ismer, ugyanis a szógyűjtés ezeknél a kategóriáknál nem volt aktív, többnyire semmit sem neveztek meg.

A gyermekek között csak a foglalkozást ábrázoló képek váltottak ki asszociatív válaszokat – tulajdonképpen a képen látható tevékenységet nevezték meg. Ezen kívül a *bagoly* láttán neveztek meg egy alkalommal tevékenységet (*reading*). A többi képre konkrét választokat adtak.

A két felnőtt csoport között nincs különbség: a fiatalabbak és az idősebbek közel egyformán gyűjtöttek konkrét és asszociatív válaszokat.

kat: a fiatalabbak válaszaiknak 60%-a, az idősebbek válaszaiknak pedig 64%-a konkrét.

A hiponímia–hiperonímia kapcsolat nem volt jellemző a gyűjtött szavak összességénél. Többnyire egyenrangú hiponím kifejezéseket neveztek meg a konkrét lehvások során. Ez alól csak az olyan tárgyak kivételek, mint pl. az autó láttán a *jármű*, a kutya láttán az *állat*, a pillangó láttán a *rovar* megnevezés. Ezek egy-egy esetben fordultak elő.

A legaktívabb gyűjtés az *autóbusz* (50), *madár* (45), *oroszlán* (37), *kutya* (37), *kakas* (36), *fa* (34), *fotel* (32) és *autó* (31) képeknél volt tapasztalható. Ezen képek által ábrázolt valóságdarabok a közvetlen környezetünkben vannak, mindennapi életünk intenzív részei. Ez alól egyetlen kivétel van a fent említett sorozatban, az *oroszlán*. Az *oroszlán* esetében 26:11 az arány a mellérendelő és az alárendelő, illetve asszociatív kifejezések között. Mellérendelő pl. a *párduc*, *tigris*, *elefánt*, *zsiráf*, *leopárd*, *jaguár* stb., tehát az állatkert lakói, a vadállat hiperoním kifejezéshez sorolható hiponím szavak. Alárendelő, asszociatív pl. a *félelmetes*, *szőrös*, *király*, *bunda*, *veszélyes*, *szép*, *strong* stb., azaz a denotátum külső és belső tulajdonságai alapján kialakult érzelmi hozzáállás kifejezése. Az *autóbusz* esetében 31:19 az arány a mellérendelő, illetve az asszociatív gyűjtések között. Mellérendelő pl. a *kocsi*, *long distance bus*, *short distance bus*, *coach*, *taxi*, *vonat*, *repülő trolibusz*, *villamos* stb., asszociatív pl. *bus driver*, *passanger*, *fare*, *street*, *ride*, *traffic jam*, *sok ember*, *kényelmetlen* stb. A *madár* esetében mind a 45 gyűjtés egymással mellérendelő viszonyban lévő megnevezés, pl. *golya*, *sirály*, *gém*, *kócsag*, *galamb*, *pigeon*, *robin*, *fecske*, *flamingó*, *daru* stb., nincs egyetlen asszociatív sem.

A távolabbi környezetünkhöz tartozó képek (*bagoly*, *ló*, *béka*, *pillangó*), valamint a ritkábban előforduló, de jó érzést kiváltó denotátumok (*templom*, *karácsonyfadísz*, *repülőgép*, *nap*) képeinek hatására 20 és 30 közötti a gyűjtött szavak száma. A *ló* (17:4) és a *béka* (15:6) esetében nagy az eltérés a mellérendelő és az asszociatív megnevezések száma között. A *bagoly* (9:10) és a *pillangó* (12:11) esetében az arány majdnem kiegyenlített, az asszociatív kifejezések többnyire az állatok tulajdonságaival kapcsolatosak (*bookish*, *okos*, *okostojás*, *hooting*, illetve *szép*, *színes*, *sárga*, *narancssárga*, *lila* stb.). A *templom* (6:20), *repülőgép* (10:20) és a *nap* (7:18) képek azonban lényegesen nagyobb

számú asszociatív lehvívást eredményeztek, amelyek a kísérleti személyek tájékozottságát, érzelmi kötődését mutatták ki az adott képek által reprezentált denotátumokkal szemben, pl. **templom**: *katolikus, pap, hívők, bor, kehely, mise, áldozás, ostya, békés, sötét, hatalmas, művészet, festészet stb.*, **repülőgép**: *landing, taking off, légi kisasszony, airport, crash, hosszú út, rossz ennivaló, megyek haza, jó mozi stb.*, **nap**: *napfény, nyár, meleg, utazás, nyaralni, Balaton, sunbathing stb.*

Legkevesebb gyűjtés a foglalkozást és a sportágakat kifejező képeknél történt. Ezért is kértük a kísérleti személyeket, hogy nevezzenek meg annyi foglalkozást, sportágat, amennyit csak tudnak. Így sor kerülhetett labdajátékok és foglalkozások lehvívására is, ami a képek által valami miatt nem volt motivált (*kézi-, kosár-, röplabda, tenisz, ping-pong, hockey, baseball, cricket*).

Kódváltás a szógyűjtésben

A legnagyobb mértékű kódváltást a *rollerozás* (80%), a *repülőgép* (57%), az *autóbusz* (52%) és a *súlyozás* (50%) képe okozta, míg egyáltalán nem történt kódváltás a *templom, óra, állatorvos, vadász, optikus, fogorvos* és *zenész* képek láttán.

A gyermekek csoportjában a szógyűjtés terén legaktívabb Olivia volt, azonban igen csekély mértékben, mindösszesen 3%-ban váltott át angolra. Marie-France ugyan fele annyi szót gyűjtött, de lényegesen magasabb százalékban váltott kódot: így a csoportban a kódváltás terén ő lett az első. Őt követi Christopher, aki az általa gyűjtött szavak 14%-át nevezte meg angolul (szemben az első feladat eredményeivel, amelynél 100%-os volt a kódváltása). Tamás nem váltott, Szandi pedig nem is gyűjtött szavakat.

A fiatal felnőttek csoportjában a legaktívabb szógyűjtő Doris volt, aki 60%-ban váltott kódot, őt követi Tim (65%-os kódváltással), majd Estella, aki egyáltalán nem váltott kódot, Pál, aki csak angolul gyűjtött és Zsolt, aki szintén nem váltott kódot.

A felnőttek csoportjában Kim a legaktívabb a szógyűjtés terén, de keveset váltott (17%), második helyen Jon, még kevesebb váltással (10%), majd Tibor (14%), végül Ati (32%) és Zsuzsi (89%).

Ennél a vizsgálatnál is kiderült, hogy a kódváltások számát, előfordulását elsősorban az határozza meg, hogy ki mennyit használja a másik nyelvét is a mindennapi életben, a családon kívül.

Következtetések

A két teszt eredményeinek kiértékelése során jól látható, hogy az életkor és a nyelvhasználat nagyban meghatározza, hogy milyen szavakat hívunk le adott képi stimulus alapján.

1. Az első hipotézisünk részben igazolódott: a gyermekek lényegesen kevesebb szót hívtak le, mint a felnőttek. Azonban megnevezéseik sokkal konkrétabbak, több a hiponim megnevezés, mint a felnőttek szólehívásai között.

2. A második hipotézisnek megfelelően az életkor növekedésével emelkedett a lehívott szavak száma, de egyben nőtt az asszociatív megnevezések aránya is: a kísérleti személyek tapasztalataikat, érzéseiket is megnevezték a képeken látható valóságdarabokkal kapcsolatban.

3. A harmadik hipotézis ellenőrzésekor kiderült, hogy a kódváltás gyakorisága nem attól függ, hogy mennyi ideje tartózkodik a személy Magyarországon, illetve mennyire kiegyensúlyozott kétnyelvű, hanem sokkal inkább a közvetlen nyelvi környezetétől, a két nyelv használatának gyakoriságától. Így történhetett meg, hogy a viszonylag rövid ideje Magyarországon élő és a magyart csak felnőtt korban elsajátító, de a mindennapi életben többnyire csak ezt a nyelvet használó személyek sokkal kevesebbet váltottak a magyarról angolra, mint azok, akik kiegyensúlyozott kétnyelvűek és az angol nyelvet is a hétköznapi nyelveként használják, mert angol szakos egyetemisták vagy tanárok.

A tesztek eredményei alapján megállapítható, hogy a képek által ábrázolt valóságot a kísérleti személyek egyformán észlelik, fogalmi rendszerük nyelvtől független, többnyire közös, csak az életkor és az élettapasztalat eredményez valamennyi különbséget a megnevezések között.

Irodalom

- Antal László (1978): A jelentés világa. Magvető. Budapest.
- Bolinger, D (1968): Aspects of Language. Harcourt, Brace & World, Inc. New York, Chicago, San Francisco, Atlanta.
- Hudson, R. (1986): Sociolinguistics. CUP. Cambridge.
- Kiefer Ferenc (2000): Jelentéselmélet. Corvina. Budapest.
- Sapir, E. (1971): Az ember és a nyelv. Gondolat. Budapest.
- Skutnabb-Kangas, T. (1984): Bilingualism or Not. Multilingual Matters. Clevedon.
- Snodgrass, J. (1993): Translating versus Picture Naming: Similarities and Differences. In: Schreuder, R. – Weltens, B. (szerk.): The Bilingual Lexicon. John Benjamins. Amsterdam, Philadelphia.

A ZÖNGE PERIÓDUSIDEJÉNEK FUNKCIÓJA A HANGSZÍNEZETBEN

Nikléczy Péter

Bevezetés

A beszédhangok képzésében jelentős szerepe van a hangszalagok által létrehozott, magas felhangtartalmú zöngerezgésnek, de a zöngeregzése önmagában nem tartalmazza az egyéni hangszínezet jellemzőit (Olaszy 1979). A hangszín kialakulása az illető személy artikulációs csatornájának, üregrendszerének alakjától és méretétől függ. Az általunk ismert személyeket az agyunkban elraktározott „hangképek” alapján ismerjük fel, amiben jelentős szerepe van az egyéni hangszínezetnek (Lavner et al. 1999). Kérdés az, hogy ezt az agyunkban tárolt egyéni hangszínt befolyásolja-e, ha a normál alaphangon beszélő személy hangját módosítjuk a zöngés hangok periódusidejének változtatásával. Vizsgálatunkkal arra próbáltunk választ kapni, hogy az alaphang mesterséges úton történő korrigálása az illető személy felismerését az észlelésben befolyásolja-e. Feltevésünk szerint, ha az adott hang kiejtésénél a személy artikulációs csatornája nem, csak az alaphangjának a frekvenciaértéke változik meg, akkor a megfelelő hangfekvésbe történő mesterséges transzformáció, lehetővé teheti a beszélő személy felismerését.

A kísérlethez használt eszközök

A kísérleti hanganyagot Sound Blaster hangeditáló programmal készítettük el, a műszeres kiértékelést KAY CSL 4300B és PRAAT analízáló programokkal végeztük el. Az eredeti hanganyagot stúdióban rögzítettük Sony MZ-R900-as típusú minidisc hangfelvevővel.

Anyag, a periódusok editálásának módszere, tesztelés

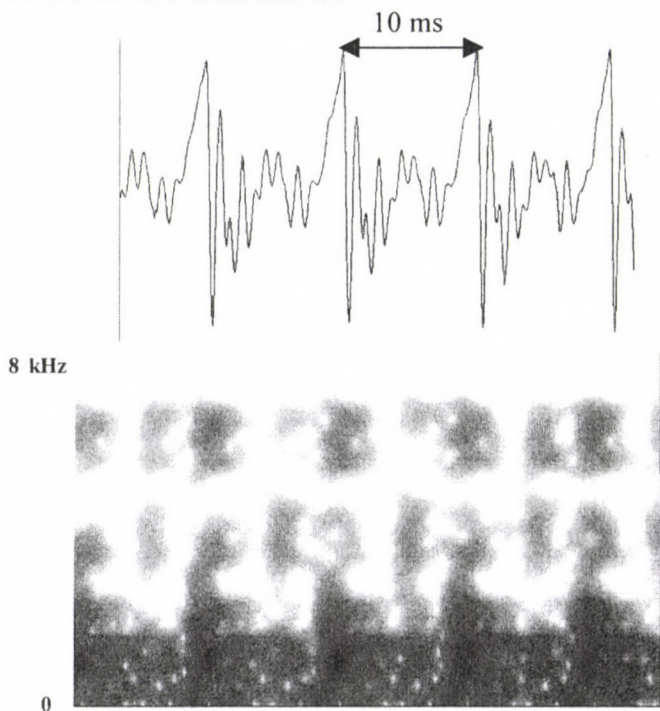
A kísérlethez használt hanganyag egy mondatrészlet volt, amelyet szándékosan úgy választottunk meg, hogy a percepció tesztelések során relatíve semleges szemantikai tartalmat képviseljen, továbbá hogy

rövidsége miatt a szupraszegmentális tényezők esetleges hatását a nullára redukáljuk. A hanganyag a következő volt: „*Lágyhang indításnál*”. A mondatrészletet egy 36 éves, magyar anyanyelvű, ép beszédű és hallású férfi mondta be. Kétféle bemondást rögzítettünk. Az első esetben normál ejtést kértünk, a második bemondáskor pedig mege-melt alaphangú artikulációt. A hanganyag időtartama mindkét esetben 1,7 s lett.

Az így kapott két hanganyag idődiagramjába mesterséges úton beavatkoztunk, úgy, hogy a periódusokban manuálisan változtatásokat hajtottunk végre. A hangkorrekciót a zöngés periódusok módosításával értük el oly módon, hogy (i) az egyes periódusok lecsengő szakaszát nulla intenzitású szünettel toldottuk meg, illetőleg (ii) a megfelelő periódusidő eléréséhez a kívánt hosszúságú részt kivágtuk. Ezzel azt értük el, hogy a hang karakterét meghatározó alacsonyabb frekvenciájú, nagy energiájú komponenseket nem változtattuk meg, ugyanakkor alacsonyabb, illetve magasabb hangfekvést tudtunk létrehozni.

A vizsgálat leírása előtt ismertetjük a spektrografikus elemzés azon részét, aminek alapján a hangkorrekciós eredményeket kiértékeljük. Mivel a beszédhangok is tartalmazzák a hangot kiejtő személy hangszínezetét, a spektrogramon elvileg azokat az egyéni jellemzőket is láthatjuk, amelyek az adott személyre utalnak. Megállapíthatjuk a beszéd alaphangértékét, vagyis a másodpercenként ismétlődő periódusok számát, a formánsok elhelyezkedését, sáv szélességét és intenzitását. Ha a hangszínképen ezeket az értékeket olvassuk le, azt láthatjuk, hogy ugyanazon személy különböző időben kiejtett azonos hangjánál az adatok változnak attól függően, hogy a személy az adott időpillanatban milyen alaphangértékkel ejtette ki a hangot. Nézzük meg elsőnek azt, hogy a spektrogram hogyan ábrázolja az egyes periódusokat. A periódusok megegyeznek egy függőleges vonallal, amelyek a formánshelyeknél feketedést mutatnak, illetve a vonal után (széles sávú elemzés esetén) kevésbé intenzív köztes résszel. A függőleges vonalak – melyek a hangszalagok nyitódásával megegyező időpillanatot realizálják (1. ábra) – a periódus kezdetén jelennek meg, majd a periódus köztes részén elhalványodnak. Az ábra felső részén a hang oszcillografikus rezgésképe, az alsó részén a spektrogramja látható. Ha ennek a hangnak csak három teljes periódusát vizsgáljuk, láthatjuk, hogy a

széles sávú spektrogram vonalas szerkezete csak a periódusok első intenzív részének frekvenciatartományát mutatja. Ennek alapján feltételezhető, hogy a periódus idődiagramjának lecsengő részében történő beavatkozás a hangszínekben nem okoz jelentős változást. Ha egy hangsor zöngés hangjait periódusonként rövidítjük meg a kis energiájú rezgések elhagyásával, magasabb vagy nulla intenzitású rész beiktatásával, akkor mélyebb hangfekvést tudunk létrehozni. Feltételezésünk szerint ez a beavatkozás a felhangok elhelyezkedésében és a percepció megítélésben változást hoz létre, a formánshelyek elhelyezkedésében azonban nem okoz eltérést.



1. ábra

Normál hangfekvésben ejtett [a:] hang
oszcillogramja és spektrogramja

A meghallgatásos tesztelés célja az volt, hogy megtudjuk, a zöngéváltozás hatással van-e a beszélő személyének felismerésére. A percepció kísérletben 4 fő vett részt, akik a hangsorokat bemondó személyt jól ismerték. A hangsorok meghallgatása során elsőként a megemelt hangú beszédet, majd a korrigált beszédet és végül a normál hangfekvésű beszédet hallgatták meg a tesztelő személyek. Minden elhangzó mondatrészlet után a résztvevőknek nyilatkozniuk kellett, hogy felismerik-e az elhangzott hangsor alapján a beszélő személyt.

A tesztanyag előállítás

A tesztanyagokat háromféleképpen állítottuk elő, és a kapott hanganyagokat műszeres analízissal ellenőriztük.

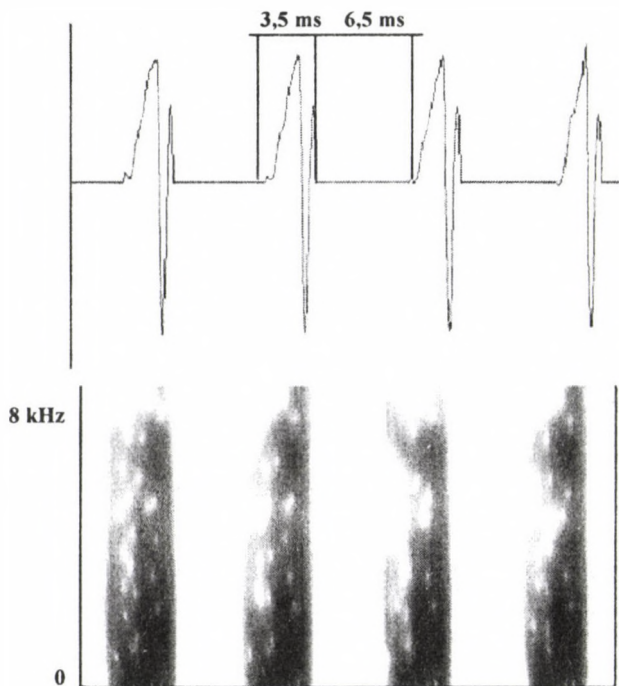
A hangedítálás formái a következők voltak:

1. Normál alaphangon ejtett férfi ($F_0=100$ Hz) hangsor periódusaiból csak az első intenzív rezgést hagytuk meg, a többi nulla intenzitásúra változtattuk.
2. A normál ejtésű férfihang hangfekvését megemeltük oly módon, hogy a nulla intenzitású rész időtartamát csökkentettük. A csökkentés eredményeképpen háromszoros alaphangértéken szólalt meg a hangsor ($F_0=300$ Hz).
3. A szándékosan torzított (háromszoros alaphangértékkal bemondott $F_0=300$ Hz) hangsor időtartamát úgy módosítottuk, hogy a periódusok végére megfelelő időtartamú nulla intenzitású részt vágjunk be. Ezzel az eredeti alaphangértéket állítottuk elő.

Az 1-es pont alapján elvégzett beavatkozásnál a meghagyott periódusrész időtartama 3,5 ms volt, a további rész 6,5 ms nulla intenzitású. A beavatkozás következtében a hangok időtartama és karaktere nem változott. A beszéd ugyan érdekesebbé vált, de az egyén hangja felismerhető maradt. A 2. ábrán látható a periódusok belsejében bekövetkezett változás, ami a spektrogram vonalas részében nem okozott jelentősebb eltérést. Ezzel magyarázható, hogy a személy megítélésében a periódus 2/3-ának elhagyása nem okozott gondot.

A tesztelők a hangsor alapján az illető személyt 100%-os eredménnyel felismerték. Természetesen azt sem lehet figyelmen kívül

hagyni, hogy a beavatkozás a zár- és zörejhangokat nem érintette. Megítélésünk szerint ez azonban a kísérlet eredményét nem befolyásolja, mivel a többi pontban elvégzett tesztelés esetében sem változtattunk ezeken a hangokon.



2. ábra

A normál hangfekvésben ejtett [a:] hang oszcillogramja és spektrogramja periódusmódosítás után

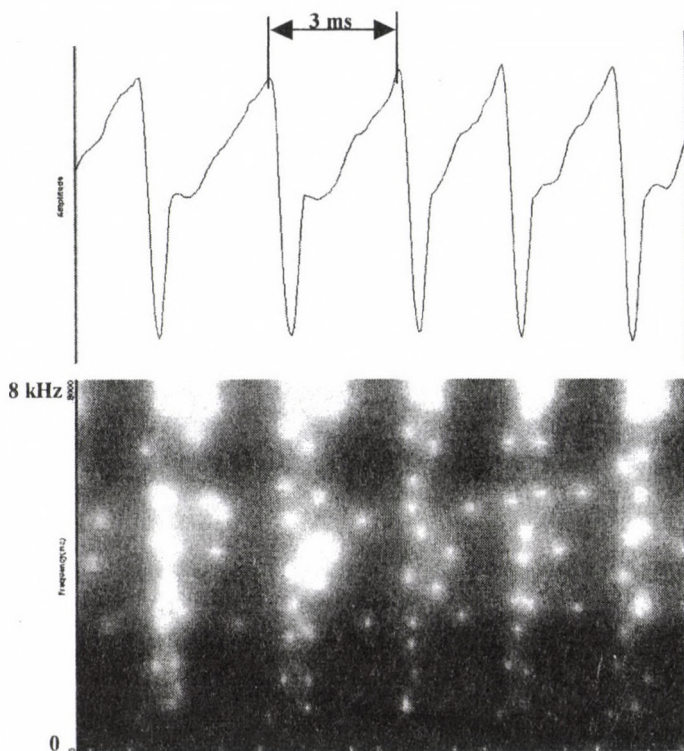
A 2. pontban leírt változtatást a periódusok nulla intenzitású részének elhagyásával hajtottuk végre úgy, hogy a megmaradt 3,5 ms időtartamú részek közvetlenül egymáshoz kapcsolódtak. Eredményképpen magas, közel 300 Hz-es ($F_0 = 286$ Hz) hangfekvésű hangsort kaptunk. A periódusidő 3,5 ms volt. Mivel a zöngés hangok a normál

tempójú kiejtéshez képest egyharmadra rövidültek, a hangsor időtartamát korrigálni kellett az eredeti hangsor időtartamára. Ezt úgy oldottuk meg, hogy a hang minden periódusát kétszer megismételtük. Az így elkészített hangsor alapján a tesztelő személyek a beszélő személyt nem tudták azonosítani. Mivel a hangsorról készített spektrogramikus vizsgálat a formánsszerkezetben és a formánsintenzitás-szerkezetben lényeges változást nem mutatott ki, megítélésünk szerint csak a beszéd alaphangjának lényeges emelkedése tette felismerhetetlenné a beszélő személyt. Kijelenthetjük tehát, hogy a beszélő személy megszokott hangjától eltérő szokatlan hangfekvés az egyén hangszínezetének változását okozza a hallgató számára. A 3. ábrán a módosított, [a:] hang spektrogramját láthatjuk. A formánsszerkezetben döntő változás nem következett be, tehát a személy felismerését csak a csökkentett periódusidő befolyásolta.

A 2. számú kísérletből kiindulva, feltételezésünk szerint, ha a hangsor pozitív irányú korrekcióját hajtjuk végre, tehát a szándékosan magasan kiejtett hangsort redukáljuk le normál alaphangra, akkor az így kapott hangsorral már a személy egyéni hangszíne is reprodukálásra kerülhet. Ebből kiindulva végeztük el a 3. pontban jelzett kísérletet.

A 3-as számú hangkorrekciós vizsgálatnál az előbbiekkal azonos hangsort ejtett ki a beszélő személy lényegesen megemelt alaphangon. A normál hangfekvésénél kb. 3-szor magasabb alapfrekvencia-értékű beszédet tesztelve, a hallgatók számára a beszélő személy nem volt azonosítható. Az előbbi kísérlethez hasonlóan tehát az alaphang-változtatás a személy felismerhetőségét megghiúsította. A megemelt hangú beszédnél a mért periódusidő 3 ms volt, tehát a normál hangfekvés létrehozásához a periódusidőt 7 ms tartamú nulla intenzitású résszel kellett meghosszabbítani.

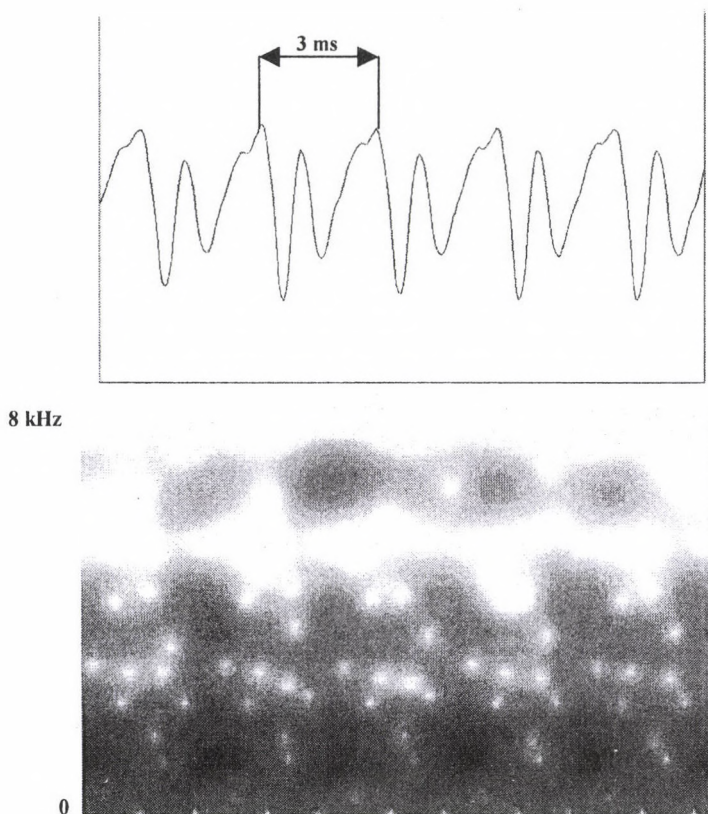
A beavatkozás következtében létrejött beszéd így alaphangértékben megegyezett a személy eredeti F₀-jával, de időtartama a periódusok számának háromszoros értéke miatt hosszabb lett. A hangsor időtartamának korrekcióját így itt is el kellett végezni, de természetesen negatív irányban. Például a bemutatott hang normál kiejtésben 17 periódust tartalmazott, míg a megemelt változat 54-et. Ez esetben tehát



3. ábra

Az időben korrigált hang oscillogramja és spektrogramja
8 kHz-es tartományban
(háromszoros alaphangérték)

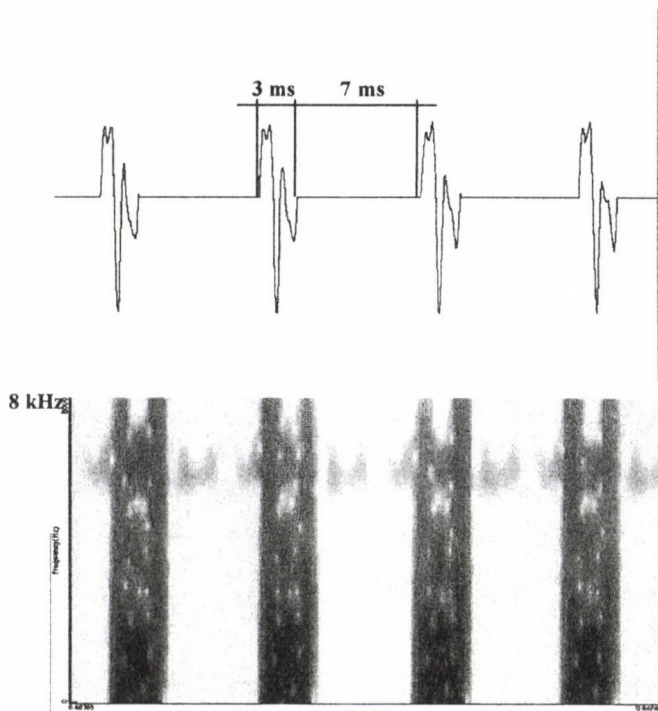
minden második és harmadik periódust kivágtuk. A zöngés hangok időtartama így megegyezett az eredeti, normál alaphangon kiejtett hangsorral. Az elvégzett teszt eredménye alapján a tesztelők 70%-os valószínűséggel azonosították a személyt, illetve úgy ítélték meg, hogy a hallott hanganyag alapján nem zárják ki az azonosságot.



4. ábra

A magas alaphangon ejtett beszéd
($F_0=300$ Hz) oszcillogramja és spektrogramja

A tesztelési eredményeknek a 6., 7. és 8. ábrán látható spektrografikus képek felelnek meg. Bemutatjuk a természetes hangfekvéssel bementett hangsort (6. ábra), a magas alaphangon ejtett hangsort (7. ábra) és a megemelt hangon ejtett hangsorból redukált természetes hangfekvésű (8. ábra) hangsort 0-8 kHz-es tartományban.

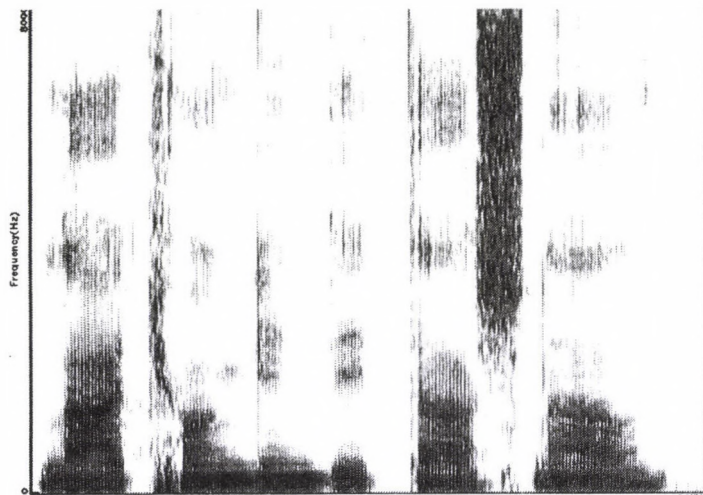


5. ábra

A magas alaphangú beszéd oszcillogramja és spektrogramja
transzformálás után

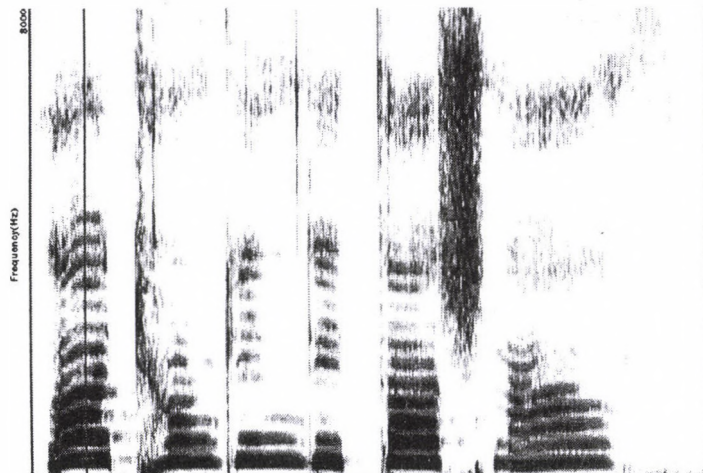
Jól látható az ábrákon, hogy a természetes alaphangú és a korrigált alaphangú hangsor spektrogramja nagyon hasonló, főbb akusztikai kulcsait tekintve megegyeznek, viszont a megemelt hangú hangsor hangszínképe eltérő.

A hangsor [a:] hangjának három különböző alaphangértékű variánsának átlag-energiaspektruma látható a 9. ábrán. Jól kivehető, hogy a magas hangfekvéssel ejtett [a:] hang energiaspektruma (nyíllal jelölt) eltér a másik két görbétől, ami a hangon belüli energiaviszonyok különbségére utal.



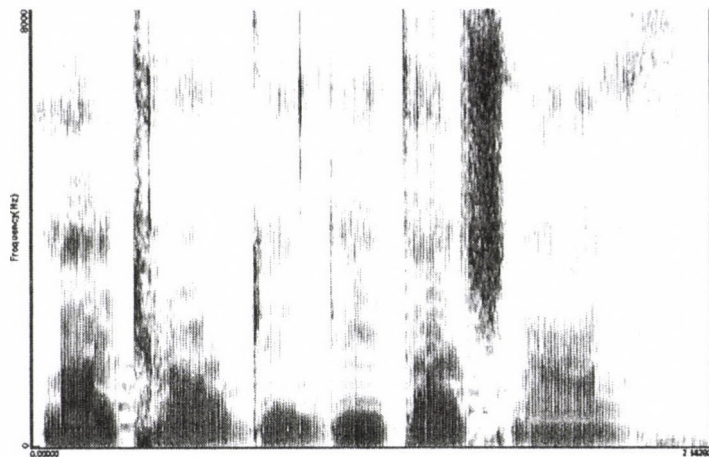
6. ábra

A normál alaphangon kiejtett hangsor spektrogramja 8 kHz-es tartományban



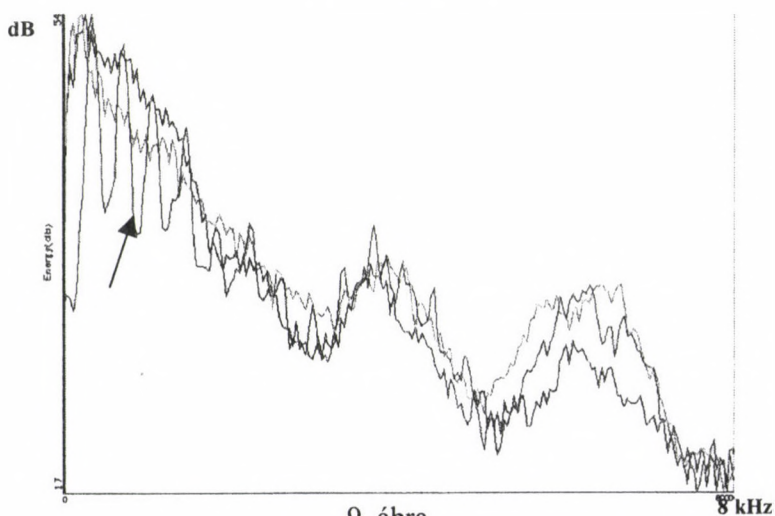
7. ábra

A magas hangfekvésben ejtett hangsor spektrogramja



8. ábra

A magas alaphanggal ejtett normál hangfekvésre korrigált hangsor
spektrogramja



9. ábra

Természetes hangfekvésben ejtett, magas hangfekvésű és korrigált
[a:] hang átlag-energiaspektruma

Következtetések

A kísérlet eredményei alapján kimondhatjuk, hogy míg a szándékosan torzított (megemelt alaphangú) beszéd a beszélő személy felismerését megnehezítette, illetve sok esetben lehetetlenné tette, addig a korrekciós beavatkozás a személy kilétének megítélését 70%-os felismerési eredményre módosította. Az egyéni hangszínezet megítélésében tehát nagy jelentősége van a személyre jellemző hangfekvésnek, és kisebb a jelentősége az artikulációs csatorna hangszínezet befolyásoló tényezőjének.

A kísérletnek az ad gyakorlati jelentőséget, hogy a személy felismerését sok esetben megnehezíti vagy lehetetlenné teszi a szándékosan mélyített vagy magasított beszédhang (Zetterholm 2003). Ha ez az eltérés az egyénre jellemző hangfekvéshez képest nagy (esetleg 2-3-szoros), a személy kilétének meghatározását sem percepciósi meghallgatással, sem spektrografikus elemzéssel nem lehet elvégezni. Ezekben az esetekben jelen kísérlet eredményei szerint (főleg a kriminalisztikai gyakorlatban) hatásosan alkalmazható a fent ismertetett eljárás. Meg kell azonban jegyeznünk, hogy a kísérletet stúdióban készített hangfelvételekkel végeztük el, így az szűkített spektrumú hangfelvételek esetében (pl. telefonvonalról rögzített hangfelvétel) csak korlátozottan alkalmazható.

Irodalom

- Lavner, Y. – Rosenhouse, J. – Gath, I. (1999): The prototype model in speaker verification. In: *Proceedings of Eurospeech '99*. Budapest, 771-774.
- Olaszy Gábor (1979): A zöngé szerepe az egyéni hangszínezet kialakításában. *Magyar Fonetikai Füzetek* 4: 137-145.
- Zetterholm, E. (2003): *Voice imitation*. Lund University. Lund.

A MAGYAR SZÓALAKOK HANGIDŐTARTAM-TÉRKÉPEI

Olaszy Gábor

Bevezetés

A beszélő által kiejtett hangsorok időszerkezetének kialakítását az anyanyelvi szabályok alapján automatikusan végezzük az artikulációs mozgásokkal egyidejűleg. Ismeretes, hogy a beszédhangok időtartamának kialakítását sok tényező befolyásolja mind szegmentális, mind pedig szupraszegmentális szinten. Korábban már végeztünk olyan kísérleteket, hogy meghatározzuk a beszédhangok hangkapcsolati szintű specifikus időtartamait (Olaszy 2000), és ezzel közelebb kerüljünk a beszéd időszerkezetének modell szintű leírásához. A specifikus időtartamok önmagukban jól jellemzik a vizsgált nyelv hangidőtartam-szerkezetét, azonban folyamatos beszédre való alkalmazásukkor a beszéd egysíkú, monoton hangzást ad, mivel nem valósulnak meg a magasabb szinten működő hangidőtartam-módosító mechanizmusok, vagyis a szó szerkezetéből adódó, és a mondat szintjén működő úgynevezett véglegesítő tényezők. Ebben a tanulmányban olyan kísérletekről számolunk be, amelyekben továbbléptünk, és meghatároztuk a magyar szóalakok szintjén működő hangidőtartam-módosító tényezőket. Külön szóalaknak nevezünk minden olyan szót, amelynek írott formája egy vagy több beszédhangnak megfelelő betűben különbözik. Tehát különböző szóalaknak tekintjük például a *fecskét* és a *fecskék* szót. Bevezettük a „szó hangidőtartam-térképe” fogalmát (HT), amelyikkel egységesített formában tudjuk a szó szintjén működő módosító tényezőket leírni, osztályozni. A HT és a specifikus időtartamok segítségével kiszámíthatók az adott szó hangjainak időtartamai. Feltételezésünk, hogy a magyar szóalakok véges számú HT segítségével jellemezhetők, így összeállítható egy olyan HT-tár, amelyikkel bármely szó hangidőtartamai mérés nélkül meghatározhatók.

A vizsgálat módszere

A kísérletek során iránymutatásként felhasználtuk a korábbi szakirodalmi adatokat, valamint hagyományos időtartam-méréseket is végeztünk. Arra, hogy a szó szintjén már működnek olyan befolyásoló tényezők, amelyek az időszerkezetre vonatkoznak, már Collier (1992, 207) is utal: „feltételezhetjük, hogy a szó szintén releváns elemként van jelen a beszéd temporális szervezésében”. Munkánk során arra törekedtünk, hogy a kísérletek megismételhetők legyenek. Ezért főleg az analízis szintézissel eljárásra alapoztuk a vizsgálódásainkat. Ennek lényege, hogy szintetizált mintákon, percepciós ítéletekre alapozva végeztünk hangidőtartam-módosításokat, majd ezeket osztályoztuk, csoportosítottuk és szabályrendszerben összesítettük. A vizsgálat fontos szempontja volt, hogy a szavak időszerkezetét folyamatos beszédbe ágyazva vizsgáltuk, így a kapott eredmények a folyamatos beszédre (nem önállóan kiejtett szavakra) vonatkoznak. A szintézis alapját ugyanaz a kísérleti összeállítás képezte, amellyel a specifikus időtartamokat meghatároztuk (Olaszy 2000). E szerint a szintetikus beszéd férfihangon szólalt meg 12-13 hang/s-os artikulációs sebességgel. A mondatokat enyhén eső intonációval szólaltattuk meg. A mondat szavait hangsúlyozás nélkül valószínűztük meg. Az analízis szintézissel módszerrel először a szerző végzett méréseket olyan egyértelműnek látszó jelenségek megerősítésére, illetve elvetésére, amelyeket a korábbi kutatásokból lehetett származtatni. A mérések alapján megfogalmazott szabályokat beépítettük a szintetizáló rendszerbe és ezek után következtek a több személlyel végzett meghallgatásos tesztek. A mondatokat többször is meghallgathatta a négy tesztelő személy (részletesen lásd Olaszy 2000). Feladatuk az volt, hogy jelöljék be a szövegben, ha szerintük valamelyik hang túl hosszú, illetve túl rövid. A szövegtörzs első része számukra 45 kijelentő mondatból állt (mondatonként 5-10 szó), második része egy A4-es oldalnyi szabadon választott szöveget tartalmazott. A hangra adott „túl hosszú”, illetve „túl rövid” megjelölést akkor fogadtuk el, ha azokat legalább hárman ugyanazon hangsorrészre egyformán bejelölték. A tesztelésből kapott jóváhagyó, illetve nem helyeslő válaszok alapján módosítottuk, bővítettük a szabályokat, egészen addig, amíg elérkeztünk egy olyan állapothoz, hogy a szövegekben már csak olyan jelölések voltak, amelye-

ket egy, illetve két személy jelölt be ugyanarra a pontra. Ezekben az esetekben a szerző döntötte el, hogy kell-e módosítani a hangidőtartamon, avagy sem. A percepciók tesztjeink befejezése és a korrekciók elvégzése után szabályokat fogalmaztunk meg, majd a szabályok alkalmazásával szintetizáltunk 15 kijelentő mondatot. A kísérletnek ezen a pontján már számszerűsítve is összevetettük a szintetizált mondatok és a természetes ejtésű ugyanolyan mondatok időszerkezetét. Ennek során vizsgáltuk mind a tendenciákat, mind pedig az egyedi időtartamokat. Az eltérésekből következtettünk a szabályok működésének helyességére, illetve gyenge pontjaira. A munka önmagába visszatérő, több lépcsős elemzési, értékelési és javítási folyamat volt. A végleges szabályokat a szerző határozta meg. A szabályok szavakra történő alkalmazásából születtek meg a HT-k, vagyis a szóra jellemző hangidőtérképek.

Az előkészítő kísérleteket 1999-2000-ben végeztük, a HT-k meghatározására és működésük ellenőrzésére 2002-2003-ban került sor.

Szabálymeghatározás a korábbi vizsgálati eredmények alapján

A korábbi kutatásokból ismert, hogy a hangidőtartamokra ható legfontosabb módosító tényezők a következők: a szó hossza, a szó helyzete a hangsorban, és a hangsúly. A szó hossza tekintetében az a tendencia, hogy minél hosszabb a szó, annál rövidebbek benne a beszédhangok (Gombocz 1909; Tarnóczy 1974). Kevesebb szótagszám nagyobb időtartamátlagot eredményez (Kassai 1979). Ezt a kiegyenlítő törvényével magyarázzák, amely szerint a produkció során az a törekvés, hogy a rövidebb és hosszabb hangsorokat nagyjából azonos idő alatt ejtsük ki. Fónagy (1959) kimutatta, hogy ez a csökkenés versmondás esetén 6 hangnál hosszabb hangsorokban már nem folytatódik. Ezt igazolták későbbi általánosított kísérletei is (Fónagy–Magdics 1960), amelyekben nem versmondást, hanem normál beszédet vizsgáltak. A beszédhang hangsorbeli helye szempontjából Magdics (1965) szerint a mondat elején rövidebbek a magánhangzók, mint mondat közepén, illetőleg a végén; az utóbbi helyzetben a leghosszabbak. Ugyanez a tendencia vonatkozik a mássalhangzókra is. Magdics tehát úgy találta, hogy a mondat kiejtése során a beszédtempó fokozatosan lassul. Kassai (1979) a következő eredményre jutott: a hangok a hang-

sor belsejében a legrövidebbek, hosszabbak hangsor elején és leg-hosszabbak a hangsor végén, tehát a tendencia: csökkenő a mondat belseje felé, majd növekvő a mondat vége felé. A hangsúly befolyását Kassai a következőképpen összegezte: a rövid magánhangzók megnyúlnak hangsúlyos helyzetben, a hosszúak változatlanok maradnak. Magdics is ilyen tendenciát állapított meg, az arány nála 1:0,8. Fónagy (1958, 16) ugyanakkor azt írta: „Semmivel sem nehezebb azonban olyan eseteket vagy esetek egész csoportját idézni, amikor a hangsúly nem esik egybe a nagyobb hangossággal, a magasabb hangfekvéssel, a nagyobb időtartammal.” A hangsúly hangidőtartamra vonatkoztatott hatása tehát nem tisztázott egyértelműen. A teszteredményeink szerint a magánhangzónak csak egy része nyúlik ebben a helyzetben a nagyobb része nem, sőt rövidül a specifikus időtartamhoz képest. Ez utóbbit erősítik meg Kovács (2002) legújabb mérései is.

Elsőként tehát megismételtük Gombocz Zoltán múlt század eleji kísérletét, és ennek alapján alkottunk szabályokat a hosszú magánhangzók viselkedésére különböző hosszúságú szavakban. A második lépésben az [l] és [r] hangok nyújtó hatását egyértelműen ki tudtuk mutatni az analízis szintézissel módszer segítségével, sőt az is kiderült, hogy ez a nyújtó hatás hangfüggetlen. A továbbiakban figyelembe vettük Olaszy (1993) korábbi mérési eredményeit is, az ezekből alkotott szabályokat is beépítettük a rendszerbe. Ezek után méréseket végeztünk a szerző által felolvasott mondatokban az első hangsúlyos szó első magánhangzójának viselkedésével kapcsolatban. A mérések alapján Magdicsnak azt a megállapítását fogadtuk el mérvadónak, hogy a mondat elején a hangok rövidebbek, mint a közepén, illetve a végén. Ezt az elvet alkalmaztuk a több szótagú szavakra is. Mindezek után végeztük a meghallgatásos kísérleteket több személy részvételével.

Hangidőtartam-módosító szabályok a szó szintjén

Az alábbi szabáyleírásokban a következő hangokra, hangcsoportokra alkottunk hangidőtartam-módosító szabályokat:

– az [a:, ɔ, o, u, y, i, e:, ø, ɛ] magánhangzók mindegyikére szó eleji, szóvégi és szó belseji helyzetre a mássalhangzó-környezet, valamint a szó hosszúságának a függvényében;

- a rövid mássalhangzókra VCV, VCCV, VCCCV, VCCCCV helyzetre;
- a VV, VVV kapcsolódásokra;
- a hiányzó hosszú hangokat a rövid hang időtartamának kétszeres értékre való nyújtásával jellemeztük.

A fentiek alapján tehát szinte bármilyen szöveges állomány bármely pontja vizsgálhatóvá vált. Az időtartam-módosítást egy M szorzófaktorral fejezzük ki. A hang szó szintű végleges időtartamát tehát a következő képlet segítségével számítjuk ki a specifikus időtartamból:

$$t(\text{szóban}) = t(\text{spec.}) \times M(n)$$

Az n változóval azt fejezzük ki, hogy a szorzófaktor értéke különböző lehet (ahogy a szabályokból látható). Az eredmény ms-ban adja a hangidőtartamot. A szó minden hangjára meghatározzuk a szorzófaktorot. Az M -ek számsorozata alkotja a szó hangidőtérképét (HT). Ugyanazon értékű M hatása nem feltétlenül ugyanazt a fizikai módosítást jelenti, hiszen a specifikus hangidőtartamok is megfelelő eloszlás szerint alakulnak az őket megelőző, és követő hang függvényében.

A szó eleji rövid magánhangzók időtartam-módosulásai

A szó első rövid magánhangzója időtartamának kialakítására meghatározott M szorzófaktorokat az 1. táblázatban adjuk meg. Egyedi szabályokban írtuk le a rövid magánhangzók időtartamát az [l, r] hangok előtt. A percepció teszt eredményei kimutatták, hogy VC helyzetben az [l, r] mássalhangzók időtartam-módosító hatása magánhangzó-specifikus. Külön szabályokat fogalmaztunk meg arra az esetre is, amikor szó előtt hangsúlytalan elem van (pl. névelő). Nézzük meg, hogy a kapott adatok milyen tendenciákra világítanak rá. A magánhangzók időtartam-módosulása $M=0,8-1,4$ -es tartományban valósul meg (1. ábra). A módosulás iránya és nagysága hangfüggő. A legkevesbé módosul az [o, i, ɔ, u], és szinte kizárólag rövidülnek. Az [l, r] nyújtó hatása a legnagyobb a [ɛ]-nél, kisebb az [ø]-nél, a többieknél nincsen ilyen hatás. Azokban az esetekben, amikor a magánhangzó után [l:] hang van (például *felláztak*) a nyújtás elérheti az 1,4-es faktort is. A mássalhangzó-környezet tehát az [ɛ] hangra gyakorolja a legnagyobb

hatást. Összegezve a rövidüléseket és a nyúlásokat azt mondhatjuk, hogy ezek az [ε]-ben valósulnak meg a legszélesebb skálán.

1. táblázat: A rövid magánhangzók időtartam-módosulása a szó első szótagjában

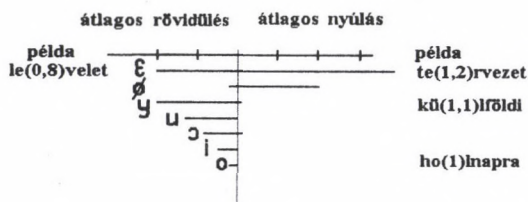
	Hangkapcsolat	Magánhangzók						
		[i]	[u]	[y]	[o]	[ɔ]	[ε]	[ø]
1	# C V C1	1	0,8	0,8	0,8	0,8	0,8	0,8
2	# C V C1 C	1	1	1	1	1	0,8	1
3	# C V C2	1	0,8	0,8	0,9	1	0,8	1
4	#□ C V C1	0,8	0,8	0,8	1	0,9	0,9	0,8
5	#□ C V C1 C	1	1	0,8	1	1	0,9	1
6	#□ C V C2	0,8	0,8	0,8	1	1	0,9	1
7	# V C	0,8	0,8	0,8	1	0,9	1	1
8	# V C1 C	0,8	1	1	1	1	1	1
9	#□ V C	1	1	0,8	1	0,9	1	1
10	#□ V C1 C	1	1	1	1	1	1	1,1
11	# V C2 C	1	1	0,8	1	1	1,3	1
12	# C V C2 C	1	1	1	1	1,1	1,3	1,1
13	#□ V C2 C	1	1	1	1	1	1,3	1
14	#□ C V C2 C	1	1	1,1	1	1	1,3	1
15	# V C3	1	1	0,8	1	1	1,4	1,2
16	#□ V C3	1	1	0,8	1	1	1,4	1,2
17	# C V C3	1	1	0,8	1	1	1,4	1,2
18	#□ C V C3	1	1	0,8	1	1	1,4	1,2

V = az adott magánhangzó; C = bármely mássalhangzó

C1 = bármely mássalhangzó, kivéve [r, l]; C2 = [r, l]; C3 = [l:]

□ = hangsúlytalan elem (főleg névelő) a szó előtt;

= szókezdő helyzet; szám = M szorzófaktor (például 0,8)



1. ábra

A szó eleji magánhangzók hangidőtartam-módosulásai

A szó belseji rövid magánhangzók időtartam-módosulásai

A szó belsejében a rövid magánhangzókat rövidíteni kellett, hogy a szó kiejtésében ne legyen érezhető bizonyos lomhaság, visszahúzás, erőltetett artikuláció. Ez korábbi kísérletek eredményeivel is egyezik (Olaszy 1993). A szabályok hangfüggősége itt nem mutat olyan széles skálát, mint a szó eleji magánhangzók esetében, amikor megindítjuk a szó artikulációját. A hangfüggőség csak annyit jelent, hogy az [i] hangot külön kell kezelni, mivel az önmagában is mindig rövid (a legrövidebb átlagértékkel rendelkezik a rövid magánhangzók között). A szabályokat a 2. táblázat mutatja.

2. táblázat: A rövid magánhangzók tartammódosulásai szó belsejében

	Hangkapcsolat	V1	V2	Megjegyzés
1	...CVC1...	0,9	1	
2	... CVC1C1...	1	1	
3	...CVC2C1	1	1	
4	...C3VC2C1	1,2	1,1	
5	...CVV...	1	1	
6	egyforma V-k a szóban a második szótagtól	0,9...0,8...0,8	..0,9...0,9...0,9.	ha nem CC, vagy V követi és nincs előtte hosszú magánhangzó

V1= bármely rövid magánhangzó; V2= [i];

C = bármely mássalhangzó

C1 = bármely mássalhangzó, csak nem C2 és nem C3;

C2 = [r, l]; C3 = [n, ʒ, c]

szám = M szorzófaktor (például 0,8)

A 2. táblázat szerint a CVC1C1 helyzetben lévő rövid magánhangzók időtartama nem rövidül, a specifikus értéken marad. Itt külön alkategóriát kell megkülönböztetnünk, amikor a magánhangzót megelőző mássalhangzó palatális, a magánhangzót követő CC kapcsolat első eleme pedig az [l, r] hang. Ebben az esetben az [l, r] nyújtó hatása szó belseji helyzetben is érvényesül. Például a *megnyergelte* szó magánhangzóinak időtartam-módosulási képe az 1. és 2. táblázat szabályai

alapján a következő lesz (a magánhangzók időtartam-korrigáló szorzó-faktorait zárójelben adtuk meg).

m e(0,8) g ny e(1,2) r g e(1) l t e(1)

Itt a második [ε] hang 1,2-es értékét a 2. táblázat 4. szabálya alapján írtuk elő. A CVV kapcsolatban a magánhangzó specifikus időtartama nem változik. Amennyiben a szóban ugyanazon rövid magánhangzó többször fordul elő (például: *alagutaknak, belefeledkezett*), akkor a szó második ilyen magánhangzója 0,9-es szorzófaktorral rövidül, a továbbiak 0,8-as szorzófaktorral, ha nem CC előtt állnak, illetve nem VV kapcsolatot valósítanak meg. A szó utolsó rövid magánhangzója a specifikus értékén marad, nem változik.

a(1) l a(0,9) g u t a(1) k n a(1) k

b e(0,8) l e(0,9) f e(0,8) l e(1) d k e(0,8) z e(1) tt

A hosszú magánhangzók időtartam-módosulásai

A hosszú magánhangzók fontos szerepet töltenek be a magyar beszéd időtartam-szerkezetének kialakításában. A rájuk kapott bonyolult időtartam-eloszlási térképek is ezt sugallják. A megállapított rövidítések, illetve nyújtások itt fizikailag nagyobb értéket jelentenek, mint a rövid magánhangzóknál, hiszen a hang eleve hosszabb. Például egy 160 ms-os [a:] hangnál 10%-os rövidítés férfihang esetén 2 periódus is lehet, női hangnál pedig ennek a kétszerese. A percepciók teszt eredményei azt mutatták, hogy a hosszú magánhangzók időtartam-kialakításának szabályai sokkal bonyolultabbak, mint gondoltuk, és másfajta szervezettséget mutatnak, mint amilyeneket a rövid magánhangzóknál láttunk. Az egyedüli közös vonás a hosszú és rövid magánhangzók között, hogy a hosszú magánhangzók is (főleg az [a:] és az [e:]) jobban nyúlnak [l] és [r] hangok előtt, mint más mássalhangzók előtt. A hosszú magánhangzó időtartam-módosulását a szó szótagszáma befolyásolja, amit Gombocz (1909) is megállapított. Eredményeink szerint további két esetet is figyelembe kellett venni, nevezetesen, hogy hány és milyen hosszúságú magánhangzó van a szóban és végül azt, hogy melyik szótagban helyezkednek el. E három szempont szerint összegeztük a végeredményeinket. Az eredmények alapján külön

kategóriába kellett sorolnunk az [a:] hangot és egy másikba az [e:] -t, továbbá egy harmadikba a többi hosszú magánhangzót. Ezen felül külön kategóriába kellett sorolni azokat az eseteket, amikor csak egy, illetve két hosszú magánhangzó van a szóban és egy másik továbbiba, amikor ennél több. Ez utóbbit – eredményeink szerint – az előbbi két esetre kialakított szabályok alkalmazásával le lehetett vezetni.

Az [a:] hang időtartam-szerkezete

A tudományos folytonosság kedvéért elvégeztük Gombocz 20. század eleji kísérletét, amit már Tarnóczy (1974) is megismételt. Az [a:] hang időtartam-csökkenését a három kísérlet szerint a 3. táblázat mutatja. A három mérési eredmény tendenciájában hasonló.

3. táblázat: Az [a:] hang időtartamának csökkenése a szótagszám függvényében

Mintaszó	Gombocz	Tarnóczy	Olaszy (szintetizált minta, 2000-ben)
tát	272 ms	210 ms	217 ms
tátog	242 ms	180 ms	182 ms
tátogat	209 ms	140 ms	160 ms
tátogatók	190 ms	120 ms	145 ms
tátogatóknak	182 ms	110 ms	128 ms

Az [a:] hang időtartam-módosító faktoraira vonatkozó jelenlegi eredményeket a 4., 5. és 6. táblázatban foglaltuk össze. A 4. táblázatban azokat az időtartam-módosító szorzófaktorokat rendszereztük, amelyek az [a:] hang időtartam-módosulásait fejezik ki, ha a szóban csak egyszer szerepel, a többi magánhangzó pedig rövid. A főbb tendenciák a következők.

- Az [a:] hang időtartama az első szótagban a szóhosszúság növekedésével arányosan csökken. Négy szótagnál hosszabb szavakban már nem csökken tovább.
- A hang időtartama attól is függ, hogy hányadik szótagban található. Az előbbi csökkenő tendencia – a 4. szótagig – akkor is megtalálható,

ha az [a:] hang nem az első szótagban van, csak ekkor a rövidülés nem olyan markáns, mint az első szótag esetében.

c) Hosszú szavakban, ha az [a:] a 4. szótag utáni részben szerepel, akkor időtartama stabilan rövidebb, mint más esetekben.

d) Ha az [a:] az utolsó szótagban szerepel, akkor hossza fokozatosan csökken a szó hosszúságának a függvényében.

e) Az [l, r] mássalhangzók nyújtó hatással vannak az [a:] hang időtartamára minden helyzetben. Ez a nyújtó hatás csökken, ha a szó hosszabb.

4. táblázat: Az [a:] hang időtartam-módosulása, ha egy [a:] hang van a szóban és nincs benne több hosszú magánhangzó

Az [a:] helyzete a szóban	Környezet	A szó szótagszáma					
		1	2	3	4	5	6
1. szótag	VC1	–	1	0,9	0,85	0,8	0,8
	VC2	–	1,3	1,2	1,2	1,1	1
2. szótag	VC1	–	–	0,9	0,85	0,85	0,8
	VC2	–	–	1	1	1	1
3. szótag	VC1	–	–	–	0,9	0,8	0,8
	VC2	–	–	–	1	1	1
többi szótag	VC1	–	–	–	–	0,8	0,8
	VC2	–	–	–	–	1	1
utolsó szótag a szóban	VC1	1,2	1	0,95	0,95	0,95	0,9
	VC2	1,3	1,3	1,3	1,2	1,1	1,1

V = az [a:] hang

C1 = bármely mássalhangzó, csak nem [r, l] ; C2 = [r, l]

szám = M szorzófaktor (például 0,8)

Tehát az [a:] hang időtartama mind az első szótagban, mind a további háromban a szó hosszúságának függvényében folyamatosan csökken. Fontos megállapítás, hogy az utolsó szótagban is markáns időtartam-csökkenés van jelen a szóhosszúság függvényében. Ez az

artikulációs ökonómiával magyarázható. Ha az [a:] a szó elején van és hosszú a szó, akkor ezért rövidítjük meg a hangot, ha a szó utolsó szótagjában van, akkor pedig egyrészt azért, mert az utolsó szótagra már nincs energia, másrészt, mert ez a szótag már nem hordoz lényeges információt, egy befejező szakaszt képvisel. Ezt két példán illusztráljuk. A *láthatatlan* szóban a magánhangzók időtartam-korrigáló szorzófaktorait zárójelben adtuk meg. Az [a:] szorzófaktorát (0,85) a 4. táblázat 4 szótagos oszlopából vettük. A szó magánhangzóinak M faktora azt mutatják, hogy az elején a tempó gyorsabb, majd lassul:

l á(0,85) t h a(0,9) t a(1) t l a(1) n

A *felebarát* szóban az utolsó szótagbeli [a:] szorzófaktorát a 4. táblázat utolsó szótag sorából vettük:

f e(1) l e(0,9) b a(0,9) r á(0,95) t

Az [a:] hang időtartama másképpen alakul, ha több hosszú magánhangzó van a szóban, és ezek közül az [a:] az első (5. táblázat). Ha az [a:] az első szótagban van, akkor időtartama hasonló képet mutat, mint a 4. táblázatban, ha a második szótagban van, akkor időtartama lényegesen lerövidül (5. táblázat). Az [l, r] mássalhangzók nemcsak az első, de a második [a:] hangban is érvényesítik nyújtó hatásukat. Ha a második [a:] az utolsó szótagban szerepel, akkor a rövidülése kissé erősebb, mint, amikor csak egy [a:] van a szóban. Ez valószínűleg annak tudható be, hogy az első [a:] -ra fordított artikulációs energia felemészti a másodikkra fordítandó energia egy részét, így kevesebb marad a másodikkra. Abban az esetben, ha több [a:] hang van a szóban akkor az 5. táblázat szabályait kell használni az első két [a:] időtartamának megállapítására, a további [a:] hangokra pedig azokat a szabályokat kell értelemszerűen alkalmazni, amelyek a második [a:] -ra voltak előírva. Vegyük például a *báránykájánál* szó magánhangzóinak időtartam-képét. Az 5. táblázatból az 5 szótagú szóra vonatkozó oszlopot kell kiválasztani. Az első [a:] -ra, mivel az [r] hang előtt van 1,1-es nyújtó szorzófaktorot kell alkalmazni. A másodikkra, mivel az a második [a:] hang és a második szótagban található a 0,85-ös időtartam-csökkentő szorzófaktorot kell alkalmazni.

5. táblázat: Az [a:] hang időtartam-módosulása, ha több hosszú magánhangzó van a szóban és az [a:] az első

Az [a:] helyzete a szóban	Környezet	A szó szótagszáma					
		1	2	3	4	5	6
1. szótag	VC1	–	1	1	0,9	0,8	0,8
	VC2	–	1,1	1,1	1,1	1,1	1
2. szótag	VC1	–	–	0,95	0,9	0,9	0,85
	VC2	–	–	1	1	1	1
többi szótag	VC1	–	–	–	0,85	0,85	0,85
	VC2	–	–	–	1	1	1

V = az [a:] hang bármelyike

C1 = bármely mássalhangzó, csak nem [r, l]; C2 = [r, l]

sám = M szorzófaktor (például 0,85)

A harmadikra, negyedikre, mivel azok nem az első két szótagban szerepelnek, az 5. táblázat harmadik sorát kell alkalmazni, itt a szorzófaktor 0,85 lesz. Az ötödik [a:] az utolsó szótagban szerepel és [l] hang előtti, tehát 1-es szorzófaktor fog kapni. A szó hosszú magánhangzóinak M faktoraival a kép a következő lesz:

b á(1,1) r á(0,9) ny k á(0,85) j á(0,85) n á(1) l

Az [a:] hang időtartama, másképpen alakul, ha több hosszú magánhangzó van a szóban és ezek közül az [a:] nem az első. Ekkor általánosságban 0,85-ös szorzófaktorot kap minden esetben, ha pedig [l, r] előtt van, akkor 0,95-öst.

Az [e:] hang időtartam-szerkezete

Az [e:] hangra vonatkozó hasonló mérési eredmények fő tendenciájukban hasonlóak, mint az [a:]-nál. Az [e:] hang időtartama is csökken a szóhosszúság növekedésével, azonban nem olyan markánsan, mint az [a:]-é. Ha az [e:] az utolsó szótagban szerepel, akkor hossza szintén kissé csökken a szó hosszúságának a függvényében. Hosszú szavakban az [e:] hang a 4. szótagtól számított szótagokban stabilan rövid, hasonlóan az [a:]-hoz. Az [l, r] mássalhangzók nyújtó hatással vannak

az [e:] időtartamára is. Ez minden helyzetben érvényes, foka azonban enyhébb, mint az [a:] hangnál.

A rövid magánhangzók hosszú párjainak időtartam-szerkezete

A rövid magánhangzók hosszú párjainak időtartam-módosulásait külön kategóriaként kezeltük. Az eredmények a következők. Az [i:, u:, y:, o:, ø:] hangok hasonlóan viselkednek, mint az [a:] és [e:] hangok, csak a nyúlások, illetve rövidülések nem olyan markánsak. Például a *kihívásáról* szó magánhangzóira vonatkozó M faktorok képe a következő lesz.

k i(1) h i(0,9) v á(0,85) s á(0,85) r ó(0,85) l

A mássalhangzók időtartam-módosulásai

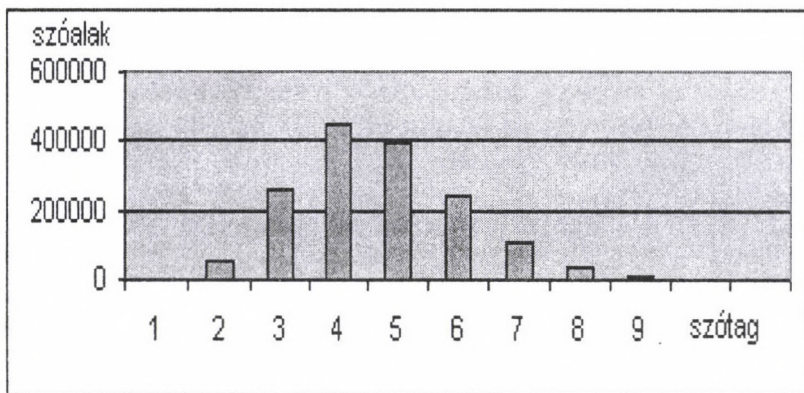
#CV, VCV, illetve VC# helyzetben a mássalhangzók nem módosulnak, a specifikus időtartamaikkal valósulnak meg. A VCCV helyzetű hangkapcsolatokban a mássalhangzók rövidítő szorzófaktorot kapnak, a redukciós hatás eredményeképpen. Általánosságban 0,9-es faktortal szorozzuk meg a CC kapcsolat mindkét tagját. Kivételek: a rövidülés nem vonatkozik az [r] hangra, az [nk], [ng], [ln] és [lm] hangkapcsolatokra, ezekben az esetekben a specifikus időtartamot kell megtartani.

A hármas mássalhangzó-torlódások hangjai erőteljesebben rövidítendőek, mint a CC-k. Itt az általános szabályt a következőképpen határoztuk meg: 0,6-0,6-0,8-as szorzófaktor-sorozatot alkalmazunk a C1-C2-C3 hangokra a helyes időtartamok beállítására. Összesen 312 különböző CCC kapcsolatot vizsgáltunk meg, és az egyedi vizsgálatok eredményeképpen számos kivételt kellett tennünk az alapszabály alól. Néhány általános szabályt adunk itt meg a kivételekre. Amennyiben C1=[l, r, n, j], akkor azt nem rövidítjük, ha C1=[r] és C2=réshang, akkor egyiket sem rövidítjük. Ha két zárhang találkozik bármelyik kombinációban, akkor az elsőt nem rövidítjük.

A négyes mássalhangzó-torlódások ritkák a magyarban. Összesen 86-féle CCCC kapcsolódást vizsgáltunk meg, ezeknél az általános rövidítési szabály a következő: minden hangot 0,6-os szorzófaktorral rövidítünk. A kivételek hasonlóak, mint a CCC kapcsolatoknál.

A magyar szóalakok statisztikai eloszlásai

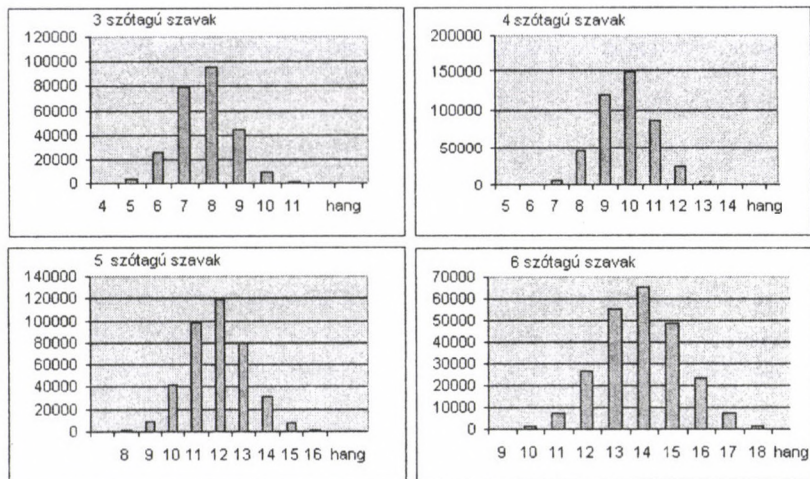
A célkitűzésünk a továbbiakban az volt, hogy a fenti szabályokkal megadjuk minden magyar szóalak hangjainak az M szorzófaktor-sorozatát, vagyis a szóra jellemző HT-t. A kiinduló adatállomány egy 50 millió szavas szövegtörzs volt (Németh–Zainkó 2002), ebből válogattuk ki a különböző szóalakokat, melynek eredményeképpen 1,5 millió elemű állományt kaptunk. Ebben az állományban tehát minden szó legalább egy hangnak megfelelő betűben különbözött egymástól. A célkitűzés eléréséhez a szavakat szótagszámuk, majd hangszámuk szerinti adathalmazokra kellett bontani. Első lépésben a szótagszámok szerint vizsgáltuk meg a korpusz szóanyagát. A kapott eloszlás képe azt mutatja, hogy a magyarban a 4 és 5 szótagú szavakból van a legtöbb, majd a 3 és 6 szótagúak következnek (2. ábra).



2. ábra

Magyar szóalakok szótagszám szerinti eloszlása az 1,5 millió szóalakot tartalmazó vizsgált korpuszban

Ahhoz, hogy a HT-eket meghatározhassuk minden szóalakra, tovább kellett bontani a szavakat a hangok szintjéig. Ennek eredményéből mutat be részleteket a 3. ábra.



3. ábra

Szóalakok hangsúly szerinti eloszlása

A szóalakok és hangidőtérképek

A több lépcsős felbontás után rendelkezésünkre állt az a szóállomány, amelyen a HT-k meghatározását elvégezhetjük. Erre speciális szoftvereket fejlesztettünk ki. Ennek lényege az volt, hogy a fenti időtartam-módosító szabályokat alkalmazva minden szóalakra meghatároztuk azokat az M szorzófaktorokat, amelyekkel kialakultak a szó hangjainak végleges időtartamai. Így a korpuszunkban szereplő 1,5 millió szó mindegyikére megkaptuk a HT-nek az adott szóra jellemző számsorozatot, tehát 1,5 millió HT-t. Ezek között természetesen vannak egyforma HT-k is, ezek csoportjait és a hozzájuk tartozó szavakat szeretnénk a vizsgálat célja szerint csoportokba szedni. Példaként bemutatunk néhány HT-t és a hozzájuk tartozó szavakból néhányat az 5. táblázatban.

A kutatásnak ezen a pontján jutottunk el ahhoz az állapothoz, hogy megvizsgálhattuk a hipotézist, mely szerint a szóalakok számánál lényegesen kisebb számú HT-vel leírhatók a magyar szóalakok hangjainak időtartamai.

5. táblázat: Különböző HT-k és a hozzájuk tartozó mintaszavak

HT és mintaszavak 10 hangú 3 szótagos szóra	HT és mintaszavak 8 hangú 4 szótagos szóra	HT és mintaszavak 13 hangú 5 szótagos szóra
1-1-0,9-0,9-1-0,6-0,6-0,8-1-1	1-1-0,8-1-0,9-1-1-1	1-0,8-1-1-0,9-0,9-0,95-1-1-0,9-0,9-0,9-1
célkönyvtár fagylaltban gázmaszkban háztömbhöz háztömbnek kontextus köznyelvben központban száznolcszor száznolcvan	adakozik alagutak amelyeken anyajuhot ecsetemet egyetemes idehozod igyekezel odajöhet ugyanazok	hazugságokból nehezményeznék
HT és mintaszavak 9 hangú 3 szótagos szóra	HT és mintaszavak 10 hangú 4 szótagos szóra	HT és mintaszavak 11 hangú 6 szótagos szóra
1-1-1-1-1-0,9-0,9-1,1-0,9-0,9	1-0,8-0,9-0,9-0,9-1-0,8-2-1-1	1-1-1-0,9-1-0,9-1-0,9-1-1-1
garantált gyerektánc harapdárd kicincált kiszolgálász kivizsgált leromlást maradványt sikoltást	becstelennek deszkalappal helynevekkel kedvesebben meglepetten megszerezzük rejtegettük szemgödörrel	aláírásokat automatákon kirohanásai kiszámítania várományosai

Szortírozó algoritmust készítettünk a HT-k összehasonlítására, amellyel az egyforma szótagszámú és hangszámú szóalakok HT-it gyűjtöttük csoportokba. Ez más szóval azt jelentette, hogy megkerestük mindazon szavakat, amelyek ugyanazzal a HT számsorozattal rendelkeztek. Ennek eredményéből mutat adatokat a 6. táblázat. E szerint tehát 1 332 936 szóalak időszerkezetének leírása 314 951-féle Ht-vel megoldható. Ez a szóállomány összegére vetítve 23,6%-ot jelent. Az adatok tehát igazolják a hipotézis gondolatát. Mint látható, az egyszer előforduló HT-k száma meglehetősen magas. Amennyiben megvizsgáljuk, hogy az egyszer előforduló HT-k miben különböznek a hoz-

zájuk legközelebb állótól, és azt is megnézzük, hogy besorolhatók-e abba, ha a különbözőséget adó hang M faktorát megfelelően módosítjuk, akkor ez a szám valószínűleg tovább csökkenthető. Ezt a módosítás utáni percepció vizsgálatával el lehet dönteni.

6. táblázat: A leggyakoribb szóalakok HT-inek száma

Szó	3 szótagú	4 szótagú	5 szótagú	6 szótagú
szószám a korpuszban	260 204	443 524	391 535	237 673
HT-k száma	12 148	58 190	122 064	122 549
egyszer előforduló HT-k száma	4749	28 045	65 990	65 942

A HT-k működésének percepció tesztelése

A szóidőtérképek minőségét ellenőrző teszt összeállításakor a problémát – a HT-k nagy száma miatt – a tesztelésre kijelölt szavak számának meghatározása és a stimulusok kiválasztása jelentette. Úgy döntöttünk, hogy vizsgálatunkat a legtöbb szóalakot tartalmazó 4 szótagú korpuszon végezzük el. Véletlenszerűen választottuk ki a percepció teszt stimulusait az összes 4 szótagú szó (443 452 darab) halma-zából, és így az 58 190-féle időtérképből is. Ez a módszer biztosította, hogy a percepció teszt anyagában az időtérképek változatosságát megfelelően reprezentáljuk. A véletlen mintavételezésnek köszönhetően a stimulusok eloszlása hasonló volt az eredeti korpuszéhoz. A tesztben 40 személy vett részt három csoportban (20, 10, 10 fő). Az adatközlők egyetemi hallgatók voltak (átlagéletkoruk 22,3 év). Minden kísérleti személy 200 különböző szót hallott és értékelt. A tesztelés számítógépen történt, az adatközlők fejhallgatón hallották a szavakat és a következő instrukciót kapták:

„Döntse el, hogy elég természetesnek találja-e a szó ritmusát! Minősítse a hangsort az alábbi 3 osztályzat egyikével:

3-as osztályzat = a szó ritmusa elég természetes;

2-es osztályzat = a szó ritmusa furcsa, de nem nagyon zavaró;

1-es osztályzat = a szó ritmusa rossz, magyartalan.

A szavak között előfordulhatnak nagyon ritka vagy esetleg nem is létező alakok. Ha ilyennel találkozunk, akkor ezt a 'Megjegyzések' címszó alatt szereplő a 'A szó nem érthető vagy nem magyar' felírra kattintva jelezze. Az osztályzat mindig csak a hangok időtartamára, a szó ritmusára vonatkozzék!"

A kísérleti személyek egy-egy szót kétszer hallgathattak meg, és módjuk volt jelezni, ha magát a szót találták furcsának. Ezen opció beépítését azért találtuk szükségesnek, hogy a hangsor esetleges idegensége a lehető legkevésbé befolyásolja az időviszonyok megítélését.

Mivel nehéz volt előre megjósolni, hogy mennyire fognak szórni az ítéletek, a lehallgatandó szómennyiséget dinamikusan határoztuk meg. Az alsó határt 4000 szóban állapítottuk meg. Ezt hallgatta meg a 20 személy. A következő lépésben 10 fő hallgatta meg a 200 szót, majd az utolsó 10 fő szintén. Ezzel a módszerrel nyomon kívántuk követni, hogy mennyire módosul az ítéletek átlaga, ha növeljük a meghallgatandó szavak számát. Mivel az átlagok a 4000, 6000 és 8000 válaszból számítva nem módosultak számottevően (7. táblázat), a lehallgatandó szavak számát nem növeltük tovább. Ennek megfelelően a továbbiakban szereplő számadatok és eredmények 8000 válaszra vonatkoznak.

7. táblázat: A lehallgatási teszt eredményei

Az ítéletek száma	Az ítéletek átlaga a kijelölt 3-as skálán
4000	2,594
4000+2000	2,5895
4000+2000+2000	2,608125

A tesztelés szempontjából fontosnak tartottuk, hogy a meghallgatandó szavak hossza (a hangok száma a szóban) ugyanolyan eloszlást mutasson, mint amilyen az egész 4 szótagos korpuszé (3. ábra). A 8. táblázat a hangszámokénti megoszlást mutatja. E szerint a tesztanyagban is a 10 hangból álló 4 szótagú szavakból volt a legtöbb. A percepció teszt anyag 4928 különböző időtérképet ölelt fel.

8. táblázat: 4 szótagú szavak hangszámának eloszlása a tesztanyagban

Hangok száma a szóban	A szavak száma a teljes korpuszban (%)	A szavak száma a percepció teszt anyagában (%)
6	0,14996	0,075
7	1,852963	1,8375
8	10,24981	10,3125
9	27,45957	27,3375
10	33,49652	32,7875
11	19,46502	20,3
12	5,96299	6,05
13	1,165402	1,075
14	0,173638	0,2
15	0,018717	0,025

Eredmények

A teszt eredményei azt mutatják, hogy a kísérleti személyek nagy arányban (68,4 %) elfogadták, megfelelően természetesnek ítélték a szavak ritmusát (9. táblázat), azaz a hangidőtérképekkel előállított szavak hangzása megfelel a magyar köznyelvi időtartam-struktúráknak.

9. táblázat: A percepció teszt értékelése

Megítélés	Százalékos arány	Kumulatív százalékos arány
1-es osztályzat	7,9	7,9
2-es osztályzat	23,7	31,6
3-as osztályzat	68,4	100

A megjegyzéseket vizsgálva azt találtuk, hogy az idegennek tűnő szóalakok között voltak valóban furcsa szóösszetételek, például *asszonynéni*, *barhentszoknya*, *békepapi*, *bozótfutók*; ritka szavak például *ámbitusán*, *famulusnak*; idegen szavak és nevek, például *csehovian*, *beatlesért*, *bestselleri* stb. Összesen 604 ilyen jelölés volt a teszt végén, ami a teljes válaszok 7,5%-a.

Összegzés

A fent ismertetett kutatással egyrészt igazoltuk a hipotézist, másrészt elértek azt, hogy minden magyar szóalakra meg tudjuk adni a szót alkotó hangok várható időtartamát a HT és a specifikus időtartam segítségével, adott artikulációs sebességre. Ezzel közelebb kerültünk a magyar beszéd időszerkezeti képének teljes körű modellezéséhez. A fenti módszer előnyei: személyfüggetlen, bármikor megismételhető, hangzó formában is ellenőrizhető, minden hangra ad adatot minden hangkörnyezetben, különböző célzatú statisztikai vizsgálatokra is alkalmas. A hangidőtartam-módosító szabályok összességének vizsgálatából az látszik, hogy a szó szintjén a legtöbb esetben rövidíteni kell a specifikus időtartamon, a nyújtásra csak az [l, r] hangok előtt kerül sor. További részeredmény, hogy a szó hangidőtartam-szerkezetének kialakításában olyan tényezők is lényeges szerepet játszanak, hogy milyen hangok vannak a szóban, és azok hogyan csoportosulnak a szón belül. A kapott adatok számos más kutatás számára jelenthetnek jó kiindulási alapot (például: a magyar mint idegen nyelv tanítása, beszéd-sérültek oktatása, időmodell-tervezés, fonetikai kutatások). A továbbiakban tervezzük az egyszer előforduló HT-k szerkezeti vizsgálatát, valamint a percepció tesztek kiterjesztését más szótagszámú szócsoporthoz is.

Köszönetnyilvánítás

Köszönetemet fejezem ki mindazoknak, akik e kutatást munkájukkal segítették. Elsősorban a BME Távközlési és Médiainformatikai Tanszékén dolgozó Zainkó Csabának és Kiss Géának, akik a munka statisztikai és beszédtechnológiai programozási munkáit végezték, továbbá Kovács Magdolnának a Debreceni Egyetem Általános és Alkalmazott Nyelvészeti Tanszék munkatársának, aki a HT-k percepció teszzeit szervezte, végezte, adatolta. Külön köszönet jár Abari Kálmánnak, aki általános programozói munkájával segítette a kutatást és létrehozta az internetes lekérdező rendszert. Ezt a kutatást az OTKA 037288 támogatta.

Irodalom

- Collier, R. (1992): A comment on the prediction of prosody. In: *Talking Machines: Theories, Models and Designs*. Eds.: Bailly, G. – Benoit, C. – Sawallis, T. R. Elsevier Science Publishers, 205-208.

- Gombocz Zoltán (1909): A magyar beszédhangok időtartamáról. Nyelvtudomány II. Budapest, 93-100.
- Kassai Ilona (1979): Időtartam és kvantitás a magyar nyelvben. NytudÉrt. 102. Akadémiai Kiadó, Budapest.
- Fónagy Iván (1958): A hangsúlyról. NytudÉrt. 18. Akadémiai Kiadó, Budapest.
- Fónagy Iván (1959): A költői nyelv hangtanából. Budapest.
- Fónagy Iván – Magdics Klára (1960): Beszédsebesség, szólam ritmusérzék. MNy 56: 450-458.
- Kovács Magdolna (2002): Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában. Debreceni Egyetem, Kossuth Egyetemi Kiadó. Debrecen.
- Magdics Klára (1966): A magyar beszédhangok időtartama. NyK. 68: 125-139.
- Németh G. – Zainkó Cs. (2002): Multilingual statistical text analysis, zipf's law and Hungarian speech generation. Acta Linguistica Hungarica. 49/3-4: 385-405.
- Olaszy Gábor (1993): Hangidőtartamok számítógépes elemzése a beszéd ritmikai szerkezetének vizsgálatához. In: Beszédkutatás'93. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest, 116-127.
- Olaszy Gábor (2000): Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. In: Beszédkutatás 2000. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 26-38.
- Tarnóczy Tamás (1974): A magánhangzók vizsgálatának akusztikai problémái. ÁNyT. X: 181-196.

A MAGYAR TELEFONBESZÉD-ADATBÁZIS (MTBA) KÉZI FELDOLGOZÁSÁNAK TAPASZTALATAI

Tóth László – Kocsor András

Bevezetés

A magyar telefonbeszéd-adatbázis (MTBA) olyan nagyméretű, telefonos beszédkorpusz, amely a magyar nyelvű fonetikai, beszédtechnológiai kutatások és fejlesztések támogatására készült. Az adatbázis 500 adatközlő hangfelvételeit tartalmazza, amelyek jelentős részét, beszélőnként 12 mondatot és 4 szót fonetikai szinten annotáltunk és szegmentáltunk. A feldolgozás során számtalan érdekes fonetikai és fonológiai jelenséggel találkoztunk, ami annak köszönhető, hogy a mondatok összeállításánál a hangkapcsolatokban való gazdagságra törekedtünk. Jelen tanulmányban a fonetikai szintű szegmentálás nehézségeiről, tapasztalatairól és érdekességeiről számolunk be.

A feldolgozott hanganyag tartalma és felépítése

Az MTBA magyar telefonbeszéd-adatbázis az Oktatási Minisztérium IKTA-3 pályázatának keretében készült a BME Távközlési és Telematikai Tanszékének és a Szegedi Tudományegyetem Számítástudományi Tanszékének együttműködésében 2000 szeptembere és 2002 augusztusa között. Az adatbázis 200 vezetékes és 300 mobilhívás felvételét tartalmazza. A telefonálók által bemondandó szövegeket úgy állítottuk össze, hogy az alkalmas legyen számítógépes beszédfelismerő-rendszerek betanítására. Ezért a szöveganyag tartalmazza például a legfontosabb magyar településneveket, a legjelentősebb intézmények neveit, dátumokat, családneveket, számneveket stb. Ebből fonetikai szintű feldolgozást adatközlőnként 12 mondaton és 4 szón végeztünk, így a továbbiakban csupán az adatbázisnak erről a feldolgozott részéről esik szó. Az adatbázis egyéb részeinek leírásáról vö. Vicsi (2002).

A fonetikailag feldolgozott mondatok és szavak azzal a céllal kerültek az adatbázisba, hogy azokon a beszédfelismerő-rendszerek fo-

netikai modelljeit lehessen betanítani. Ezért a mondatokat úgy kellett összeválogatni, hogy azokban a magyar beszédhangok, valamint bi- és trifónok (kettes és hármas hangkapcsolatok) kellően nagy számban forduljanak elő. A szöveganyag összeállítását végző budapesti partner e célból egy automatikus fonetikus átíróprogramot készített, amelynek segítségével ellenőrizte, hogy az adatbázisba bevett mondatok hangstatisztikája megfelel-e az említett elvárásnak. Összességében 1992 darab, újságcikkekből származó mondatot válogattunk össze az adatbázisba, és ezek mindegyikét összesen háromszor olvasták fel. A beszélőnként további 4 szóra azért volt szükség, hogy ily módon növelni tudjuk az olyan ritka hangkapcsolatok gyakoriságát, amelyek a mondatokban nem fordultak elő kellő számban.

A kézi feldolgozás feladata

Az adatközlők által felolvasott 12 mondat és 4 szó kézi feldolgozását egy speciálisan e célra készült program segítségével végeztük, amely képes a hangfájlok hullámformájának és spektrumának megjelenítésére, továbbá tetszőleges hangrészlet lejátszására is. E három információ alapján kellett a munkát végzőknek a hangmintát fonetikai szinten szegmentálni és címkézni. Címkézésre a nemzetközileg elfogadott SAMPA kódolást (vö. Sampa 2002) használtuk, ami az IPA számítógépes karakterekhez igazított változata. A magyar nyelvhez összeállított SAMPA kódtábla összesen 58 különböző fonetikai címkét tartalmaz. Ennél finomabb részletességű jelölés nem volt célunk, ugyanis a feldolgozás így is több mint egy évet vett igénybe.

A fonetikus átírat elkészítésének felgyorsítására a program mindig felkínál egy javaslatot. Ezt a korábban már említett átíróprogram készíti el a mondat ortografikus lejegyzéséből kiindulva. Az algoritmus képes kezelni az alapvető hasonulási szabályokat, viszont nyilvánvaló módon nem tudja megjósolni, ha a beszélő mást mondott, vagy más-hogyan ejtett ki valamit. Ezért a gép által felajánlott átírást csak javaslatnak tekintettük, és szükség esetén módosítottuk.

A szegmentálási feladat ismertetésének befejezéseként szeretnénk hangsúlyozni, hogy az adatbázis legfontosabb tervezett alkalmazása a gépi beszédfelismerő-rendszerek tanítása, ezért a szegmentálás során végig az automatikus felismerők igényeit tartottuk szem előtt. A gépi

tanulás hatékonyságának növelése szempontjából a legfontosabb, hogy az azonosan címkézett szegmentumok minél jobban hasonlítsanak egymásra, a különböző címkéjük pedig minél jobban eltérjenek. Ez a törekvésünk az oka annak, ha bizonyos esetekben valamelyest eltértünk a fonetika által javasolt címkézési vagy szegmentálási szabályoktól. Úgy véljük azonban, hogy ezek az eltérések nem számottevőek.

Általános tapasztalatok

A telefonhívások megszervezésénél gondosan ügyeltünk arra, hogy az ország minden területéről érkezzenek hívások, méghozzá életkor és nem szerint a magyar lakosság eloszlásának megfelelően. A felvételekhez az adatközlők megkapták a felolvasandó mondatok/szavak listáját és egy rövid tájékoztatást az adatbázisról. A felvételek feldolgozása során azt tapasztaltuk, hogy több beszélő is szebben artikulált, mint az természetes lett volna. Ezért az adatbázison végzett nyelvészeti vagy beszédtechnológiai vizsgálatok során tanácsos tekintetbe venni, hogy olvasott és nem spontán beszédről van szó¹. Ennek legfeltűnőbb következménye az volt, hogy az automatikus átíró által feltételezett hasonulásokat gyakran kellett „visszaalakítanunk” az átíratban. Talán a rögzítésre kerülés tudatával magyarázható az is, hogy rendkívül kevés a tájnyelvi jellemzőket tartalmazó felvétel: csak néhány esetben találoztunk például *ö-ző* vagy zárt *ë-t* használó beszélővel.

További általános tapasztalatunk volt, hogy nagyon kevés (a 8000-ből csupán 65) felvételnél találtunk olyan nagy mértékű háttérzajt, ami a szegmentálási munkát nehezítette. Valószínűleg ez sem felel meg a természetes telefonálási körülményeknek: a hívók többsége feltehetően keresett egy „csendes, nyugodt” helyet a hívás lebonyolítására. A feldolgozás során a háttérzajnál jóval több gondot okozott a sok esetben fellépő torzulás, illetve a telefonok recsegése. Az utóbbi kezelésére egy speciális szimbólumot vezettünk be, így a recsegés miatt használhatatlan beszédrészeket a fonetikai kódolás alapján kiszűrhetők,

¹ Az artikulációs gondosság két végletes példája lehet a 409. (túlargikuláló) és a 328. (alulargikuláló) adatközlő.

eltávolíthatók. Ugyanezt a jelet használtuk a különféle „beszélőkeltette zajok” – köhögés, krákogás, levegő kifújása – címkezésére is.

Mivel a felolvasandó szöveg összeállításánál a fonetikai tartalom volt a fő szempont, a mondatok többsége eredeti szövegkörnyezetéből kiragadva teljesen értelmetlenné vált, amihez még az is hozzáadódott, hogy többnyire viszonylag hosszú mondatokról volt szó (30-50 beszédhang). Ezért a beszélők gyakran tévesztettek a mondatok felolvasásakor, újra kezdtek egy-egy szót. A félbeszakadt szavakra szintén speciális jelölést használtunk, így a feldolgozás során ezeket is figyelmen kívül lehetett hagyni.

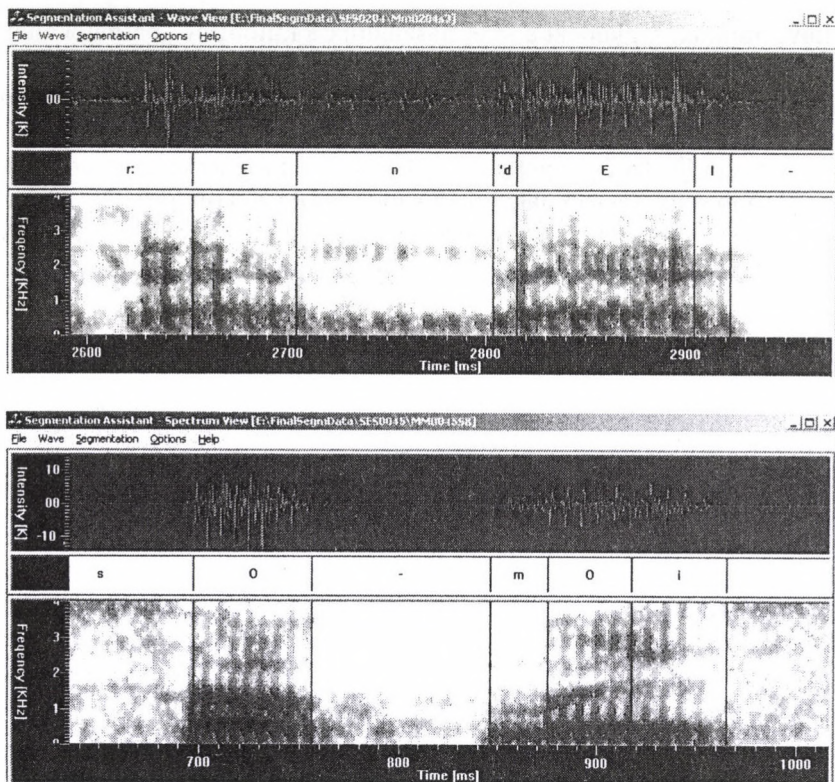
A szegmentálást nehezítő jelenségek

A szegmentálás során az egyes fonetikai szegmentumok kezdő- és végpontját egyetlen időpillanat hozzárendelésével kell megadni. Azonban az egyes fonetikai jegyek nem teljesen szinkronban kapcsolódnak ki és be, ami a hangok szintjén koartikuláció formájában jelenik meg. Ilyenkor a beszédhangok határát – jobb híján – általában az átmeneti rész közepére húztuk be. Ezzel a módszerrel a koartikulációt tudtuk ugyan kezelni, de néhány esetben feloldhatatlan nehézségekkel szembesültünk. Ezeket az eseteket, és az általunk alkalmazott áthidaló megoldásokat ismertetjük az alábbiakban.

A zárhangok és affrikáták zár/zörej felbontása

A zárhangokat és az affrikátákat a fonetika három képzési szakaszra bontja: a záralkotás, a zár és a felpattanás szakaszára (Roach 1991). Ezek a részek akusztikailag egymástól élesen elválnak és viszonylag önálló életet élnek, például a zár eltűnhet, ha a felpattanó zárhang nazális hang után áll – a nazálisból ugyanis közvetlenül fel lehet pattintani a zörejrészt (vö. 1/a ábra). Ugyanígy nincs értelme zárrészt bejeleníteni egy szövegkezdő zöngétlen zárhang esetén (a zár és a szó előtti csend nem különböztethető meg).

Más esetekben viszont csak a zár marad meg – a zörejrészt nem ejtjük ki. Tipikus példa a zárhang–nazális kapcsolat (1/b ábra). Ezért az eredeti SAMPA jelöléseket módosítottuk, ahogy az az ábrák jelölismódján is megfigyelhető.



1. ábra

a) Nazális utáni zárhang zár nélkül; b) nazális előtti zárhang zörej nélkül: az [m] előtti [k] hang zöreje hiányzik

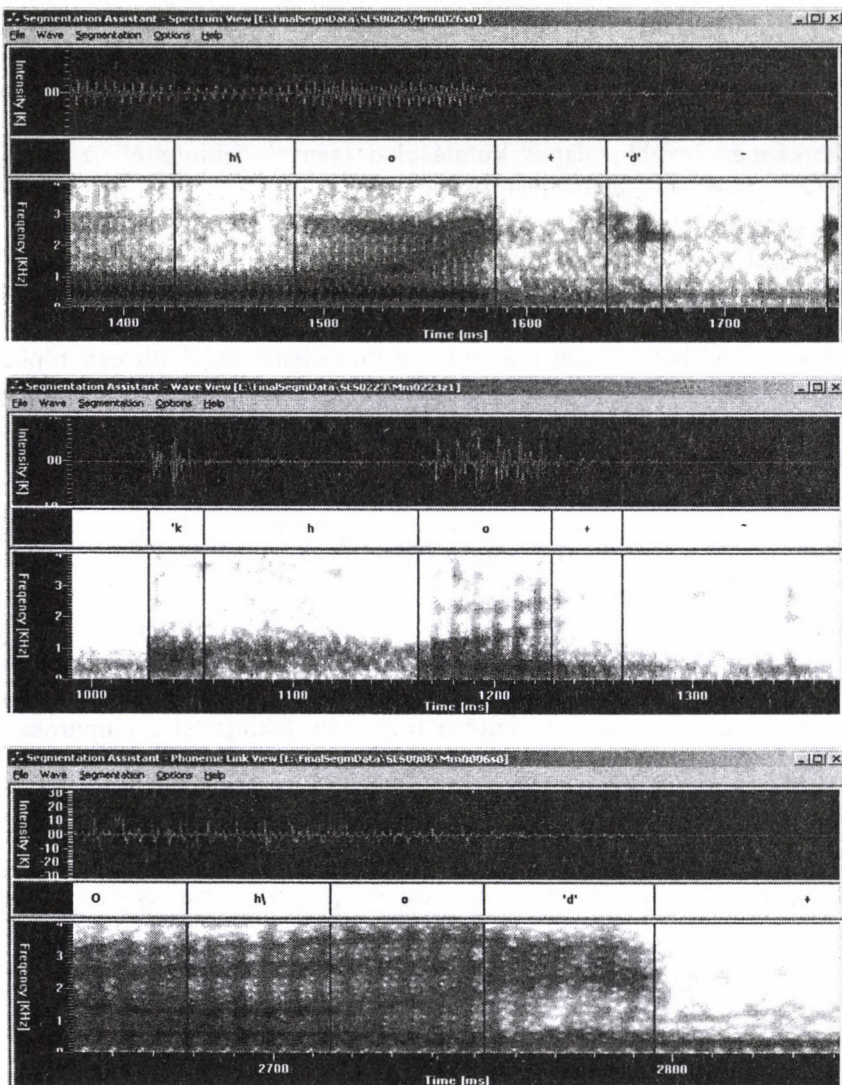
A zárhangokat és az affrikátákat felbontottuk két részre: a +, illetve – jeleket bevezettük a zöngés/zöngétlen zár jeleként, a hangokra vonatkozó jeleket pedig csak a zörejrészek címkézésére használtuk (egy kezdő aposztróffal kiegészítve, az eredeti jelöléstől való eltérést hangsúlyozandó). A záralkotás szakaszát nem illettük külön jellel, mivel általában nem önálló szegmentumként jelentkezik, hanem a megelőző hang képzését (formánsszerkezetét) módosítja. Ez azonban bizonyos hangkiesések esetén olyan szegmentálási eredményhez vezetett, ami

fonetikai szempontból semmiképpen nem kielégítő. Úgy gondoljuk, hogy a jelölésrendszert a hangkapcsolatokra és speciálisan a felpattanók kapcsolataira nézve további hasonló adatbázisok készítése esetén felül kell majd vizsgálni. A hangátmenetek akusztikai szerkezete mindenképpen további alapos kutatásokat igényel, tekintettel az átmenetek beszédpercepciók jelentőségére.

A hangkiesések problémája

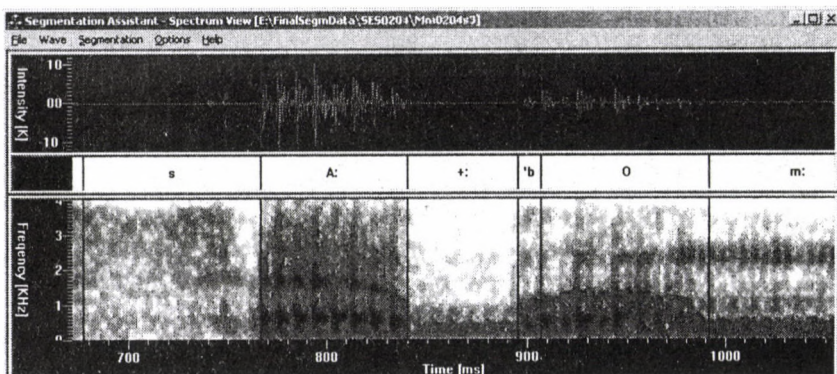
Az „átmenet-megfelezős” stratégia áthidalja ugyan a koartikuláció okozta határbehúzási nehézséget, de továbbra is feltételezi, hogy a két szomszédos beszédhang mindegyike megjelenik legalább egy röpké időintervallum erejéig. Az igazi nehézség akkor jelentkezik, amikor egy hang teljesen eltűnik, pusztán néhány jegye marad meg, amelyek a szomszédos hangok képzését módosítják. Ilyenkor a fonetikai szegmentumokra bontás egyszerűen lehetetlen, de legalábbis nem ad meggyőző eredményt. A vizsgált anyagban ilyesmivel ritkán találkoztunk, tekintve, hogy olvasott és nem spontán beszédről volt szó. Azonban néhány esetben kényszermegoldásokat kellett találnunk. Az egyik klasszikus példa a *hogy* szó. Pontos artikuláció esetén az [o], a zár [+] és a zörej [‘d’] elválik, és jól bejelölhető a spektrumon (2/a ábra). Azonban szép kiejtés esetén is megfigyelhető a második formáns gyors emelkedése az [o] kiejtése folyamán, tekintettel a hamarosan következő palatális zárra.

Ez a formánsemelkedés olyan markáns, (és a *hogy* szó általában annyira megjósolható), hogy folyamatos beszédben a zörej realizálására nincs szükség kommunikációs szempontból (2/b ábra). Egy másik ejtéstípusnál még a zár sem tökéletes, és a [d’]-re pusztán egy felszaladó formáns utal az [o]-ban. Ilyenkor a második formáns többnyire eléri a [d’]-nek megfelelő locusokat (2/c ábra), ezért ilyenkor ahhoz az (az ábrán is látható) jelölésmódhoz folyamodtunk, hogy kitöröltük a zárat, de a [‘d’] jelű szegmentumot meghagytuk. Sokszor azonban az egész szó olyan rövid, hogy azt éreztük volna a leghelyesebbnek, ha mind a [+], mind a [‘d’], azaz a teljes [d’] kimarad. Habár lehallgatáskor ezekben az esetekben is egyértelműen *hogy*-ot hallunk, valójában nincs olyan időintervallum, amelyet [d’]-nek lehetne jelölni.



2. ábra

A *hogy* háromféle ejtésben/jelölésben a) a [d'] zár- és zörejrésze is megjelenik, b) a zörej elmaradása, c) a zár elmaradása



3. ábra.

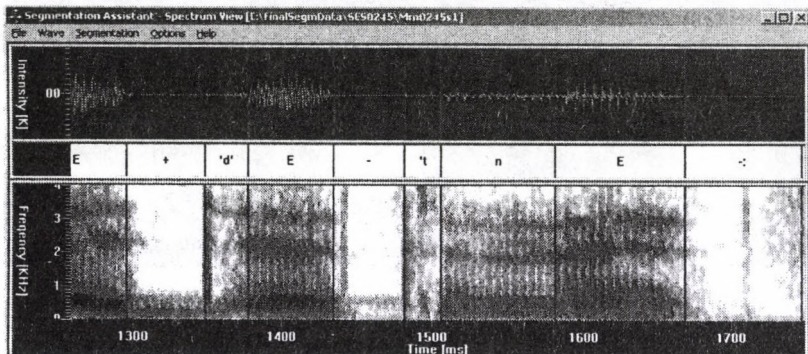
A [g] és [b] találkozása (az *országba* szóban) – az [A:] formánsai a [g]-hez igazodnak, de a [g] zöreje hiányzik, így hosszú zárat kellett jelölnünk

A másik szituáció, ami hasonló nehézségeket okozott, a zárhang–zárhang kapcsolatok esete volt. Ezek kiejtésekor a legritkább esetben található meg mindkét hang zörejrésze. Ehelyett általában csak a második zöreje marad meg, és a két zár egyetlen hosszú zárrá egyesül (vö. 3. ábra). Ennek ellenére lehallgatáskor tisztán érzékeljük mindkét zárhangot – ugyanis a megelőző magánhangzó formánsai az első felpattanó zörejének locusait veszik célba. Azonban hiába tudjuk mindezt, ha tartani akarjuk magunkat a szegmentum szintű jelöléshez, akkor az első zörejt egyszerűen nincs hova elhelyezni, így marad az ábrán is látható jelölésmód: hosszú második zárhang feltüntetése két rövid helyett. Ez az a szituáció, amikor a zárképzés önálló szegmentumként való kezelése segített volna, ugyanis akkor legalább ez a szegmentum megmaradhatott volna az első zárhangból.

A spontán beszédre jellemző jelenségek

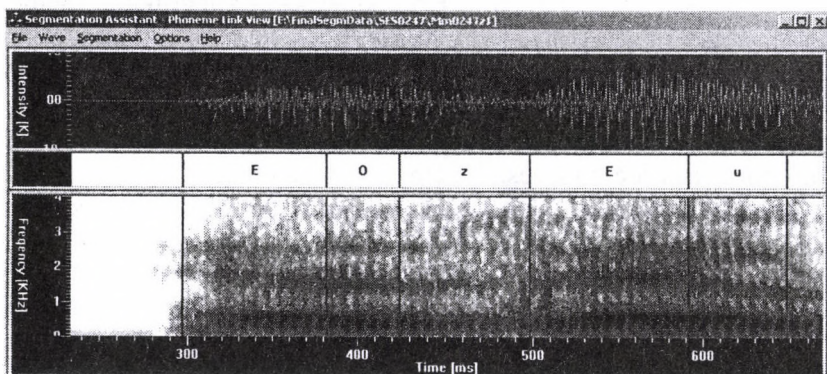
Azt gondoljuk, hogy spontán beszéd és a megértést nagyban segítő szöveggörnyezet esetén a nagyfokú, akár több beszédhangra is kiterjedő hangkiesés, elmosódás sokkal gyakoribb. Ez a jelenség azokat a

mondatépítő szavakat, kötőszavakat érinti leginkább, amelyek a leginkább megjósolhatóak a szöveggörnyezet alapján (vö. a *hogya* példáját).



4. ábra.

Az *egyetlenegy* szóban a *len* összeolvadt egy leginkább [n]-ként címkézhető hanggá

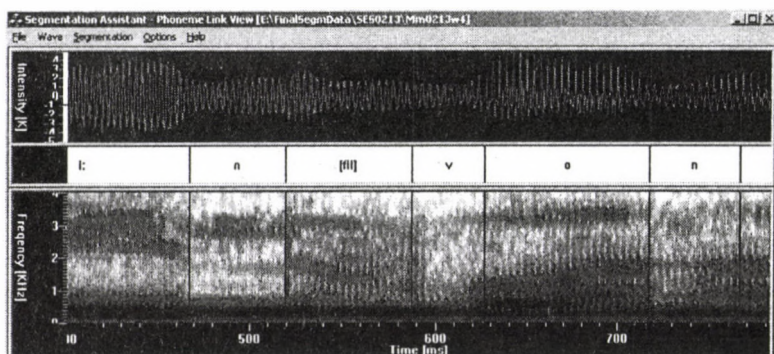


5. ábra.

A mondatkezdő *ez az* összemosódása az „Ez az európai...” kezdetű példamondatban

E szavak lehallgatása esetén az a jellemző, hogy a mondatban tökéletesen érthetőek, kiragadva őket viszont nem, vagy alig azonosítható-

ak, és sem a spektrumon, sem a hullámformán nem lehet megtalálni a belőlük „elnyelt” hangokat. Két példát láthatunk a 4. és az 5. ábrán. Előbbi esetében az *egyetlenegy* szó *len* szótagja összeolvadt egyetlen szegmentummá, amelyet *n*-ként címkéztünk. Még erőteljesebb összeemosódás történik az 5. ábrán, ahol a mondatkezdő *ez az* valami olyan egységgé olvad össze, ahol az eredeti hangok egyikét sem lehet pontosan bejelölni. Mivel ilyenkor is kellett valamilyen szegmentumokat bejelölnünk, nyilvánvalóan kényszermegoldásokhoz kellett folyamodnunk.



6. ábra.

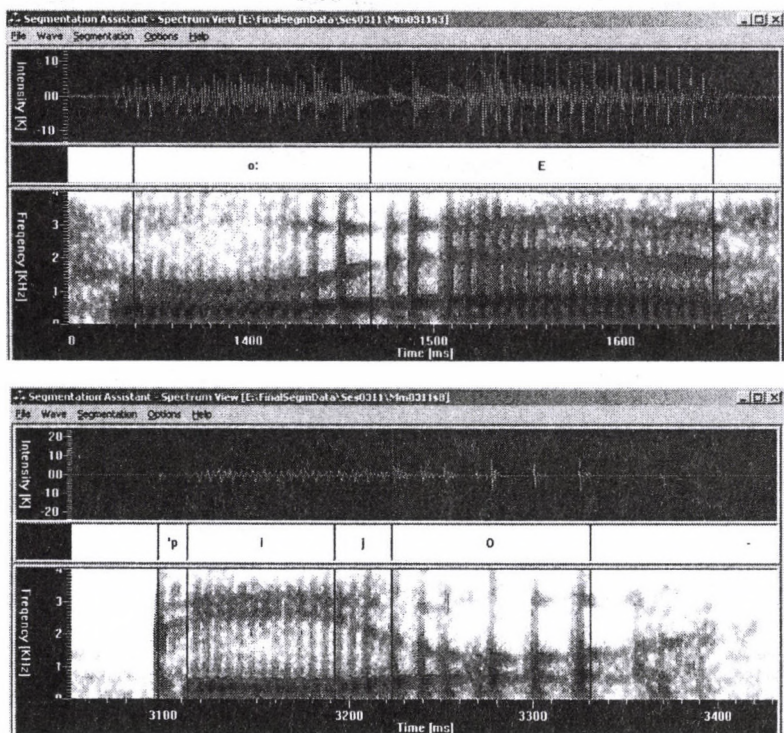
Az [n] és [v] találkozása a *színvonalas* szóban. Az [n] zárfelpattanását külön szegmentumként jelöltük.

Mint azt már említettük, külön jel szolgált az úgynevezett „beszélőkelte zajok” címkézésére. A fonetikai kódkészlet emellett tartalmaz egy speciális jelet a kitöltött szünet (hezitálás) megjelölésére is. Kitöltött szünetek csak ritkán fordultak elő az anyagban, azonban ezzel a jelöléssel – [fil] – címkéztünk fel olyan szegmentumokat, amelyek a beszéd részét képezték, de nem voltak pontosan meghatározhatók. Erre példát mutat be a 6. ábra, ahol egy [nv] hangkapcsolat szerepel a szóban. Ennek elvileg [Fv]-vé kellett volna alakulnia, ahol az [F] az [n] foghang változata. Az adatközlő azonban túl szépen akart beszélni, mindkét hangot kiejtette, ezért viszont kénytelen volt a két

hang közé beszúrni egy szegmentumot. Ez valójában az [n] zárfelpattanása, de semmiképpen sem akartuk az [n] részeként jelölni, mivel nagyon ritka jelenség, és így a statisztikai úton működő beszédfelismerőket megzavarhatná. Ezért [fil] címkével láttuk el.

További érdekességek

A pontos hanghatárok megtalálását rendkívül meg tudja nehezíteni az ún. laringalizált vagy „csikorgó” beszéd – „creaky voice”, „vocal fry” (Durand–Siptár 1997). Ilyenkor a hangszalagok szabályos periodikus nyitódása–záródása helyett egyes periódusok kimaradnak, és a spektrogram „szakadozottá” válik (7/a ábra).



7. ábra

a) Csikorgó beszéd; b) a szóhatáron fellépő hasonló jelenség

Ez az emberi beszédpercepciót egyáltalán nem nehezíti, azonban a számítógépes beszédfelismerő algoritmusok működését erősen zavarhatja. Úgy tapasztaltuk, hogy a fenti jelenség természetes módon jelentkezik a szóhatároknál: a szóhatáron a beszélő lecsökkenti a gégefőben a nyomást, de nem teljesen, így tökéletes csend helyett a hangszalagok továbbra is kinyílnak, de az elégtelen nyomás miatt nem szabályos periódusokban, hanem csak néha-néha. Szünet helyett a két szó között jellegzetes szaggatott jelszakasz keletkezik (7/b ábra). Ezt a jelenséget igen gyakran tapasztaltuk és nem csak mondathatároknál. Érdekes kérdésként vetődik fel, hogy vajon beszédérzékelésünk ki tudja-e használni a szóhatárok megtalálásánál.

További érdekességgént találtunk egy olyan hangváltozatot, amelyet a SAMPA kódtábla nem tartalmaz, pedig erősen eltér a hang szokásos kiejtésétől, és viszonylag sokszor fordult elő. Ez a változat a zöngétlen [l] volt, amely jellemzően a *-tlan/-tlen* végződés esetében jelentkezett. Megjegyezzük, hogy a SAMPA táblázatban külön feltüntetett zöngétlen [j] viszont jóval kevesebbszer fordult elő, mint ahogy az automatikus átíró jósolta: az általa javasolt esetek többségében át kellett javítanunk a címkét zöngés [j]-re.

Végül megemlítjük, hogy az adatbázisban mind az automata által javasolt, mind a valódi kiejtésnek megfelelő fonetikus átírat megtalálható, így a kettő közötti eltérés statisztikai módszerekkel könnyen számszerűsíthető. Ezért reményeink szerint az adatbázist nem csak a beszédtechnológiával foglalkozó mérnökök, hanem a nyelvészek is nagy haszonnal tudják majd felhasználni.

Irodalom

- Durand Jacques – Siptár Péter (1997): Bevezetés a fonológiába. Osiris kiadó. Budapest.
- Gósy Mária (1995): Szükséges és szükségtelen hangátmenetek. In: Beszédkutatás '95. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 20-31.
- Roach, P. (1991): English Phonetics and Phonology: a Practical Course. Cambridge University Press. Cambridge.
- Sampa (2002): <http://www.phon.ucl.ac.uk/home/sampa/home.htm>.

Vicsi Klára – Tóth László – Kocsor András – Gordos Géza – Csirik János
(2002): MTBA – Magyar nyelvű telefonbeszéd-adatbázis. Híradástechnika 8: 35-39.

A szerzők köszönetet mondanak Sejtes Györgyinek nyelvészeti tanácsaiért.

DALLAMI MEGHATÁROZOTTSÁGOK A MAGYAR MONDATBAN (A TOPIK DALLAMA)

Varga László

Bevezetés

Meggyőződésem, hogy a *megtervezett* magyar mondatban jobbról balra ható hanglejtésbeli összefüggések állnak fenn, vagyis a mondat utolsó résztömbjének dallama (hanglejtéskontúrja) korlátozza az azt megelőző résztömbök választható dallamát. Dolgozatomban ezt az álláspontomat szeretném részletesen körüljárni, és empirikus érvekkel alátámasztani.

Emlékeztetőül felidézem: a *résztömbök* (pontosabban *tonoszintaktikai résztömbök*) olyan főhangsúlyos kifejezések, amelyek a mondat szintaktikai összetevőiből szerveződnek, de mivel a nem-főhangsúlyos összetevőket magukba vonják, nem mindig azonos kiterjedésűek a szintaktikai összetevőkkel. Szintaktikai besorolásuk alapján a résztömbök többsége váltótömb, kisebb része ikertömb, némelyikük pedig váltótömb is, ikertömb is lehet (vö. Varga 1993, 93-109; 1994, 525-543; 2002, 97-121). Dolgozatom a váltótömbök egy speciális fajtájával, a topikokkal foglalkozik részletesebben. É. Kiss (1992; 1998) szerint a topik (T) pozíció a magyar mondat azon strukturális pozíciója, amely megelőzi a predikátumot (kommentet), és nulla, egy, vagy több topik-szerepű összetevő (T-összetevő) tölti ki. A T-összetevő mellékmondat is lehet.

A topik prozódiaját különféle szempontok alapján vizsgálták és számos megállapítást tettek például a topik utáni szünet lehetőségéről, a topik és a komment hangsúlyviszonyairól (vö. pl. Elekfi 1968; É. Kiss 1987-88; Varga 1987-88; Fónagy 1998), de kevesebb szó esett a topik lehetséges dallamairól, még kevesebb e dallamok meghatározottságáról. A jelen dolgozat erre az utóbbi kérdéskörre összpontosít.

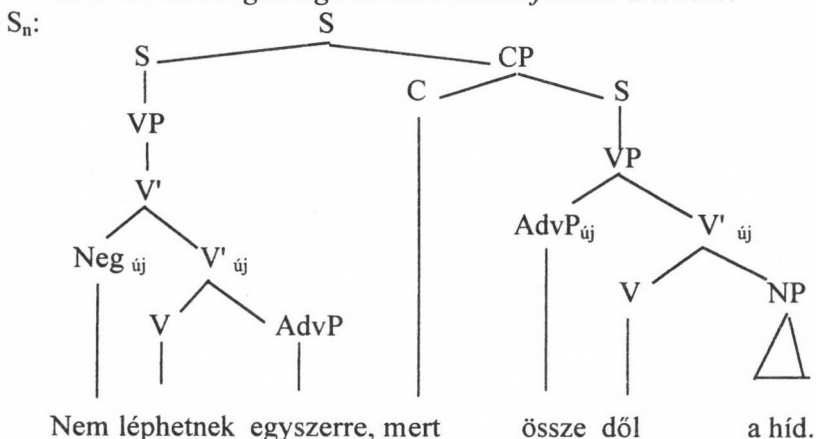
Mivel itt és most váltótömböket kívánok vizsgálni, ezért csak olyan mondatokat veszek tekintetbe, amelyekben ikertömb nincs. Hasonló

okokból eltekintek az olyan mondatoktól is, amelyekben szervetlen tömb (például vokatívusz, idéző mondat, stb.) szerepel (ezekről bővebben vö. Varga 1993, 90-92; 1994, 522-524; 2002, 93-97). A tárgyalt példák tehát minden esetben csak váltótömbökből állnak.

A szintaxistól a proszódiai formáig

Feltételezem, hogy amikor egy mondatot kimondunk, a fókuszstruktúrával gazdagított *szintaktikai felszíni szerkezet* (S_n) szolgál bemenetül a proszódiai megformáláshoz, vö. (1). A fókuszstruktúrával való gazdagítottság annyit jelent, hogy a felszíni szerkezetben az információs szempontból új, ismert, illetve szembeállított elemek valahogyan meg vannak jelölve (vö. Selkirk 1984, 206).

(1) Fókuszstruktúrával gazdagított szintaktikai felszíni szerkezet



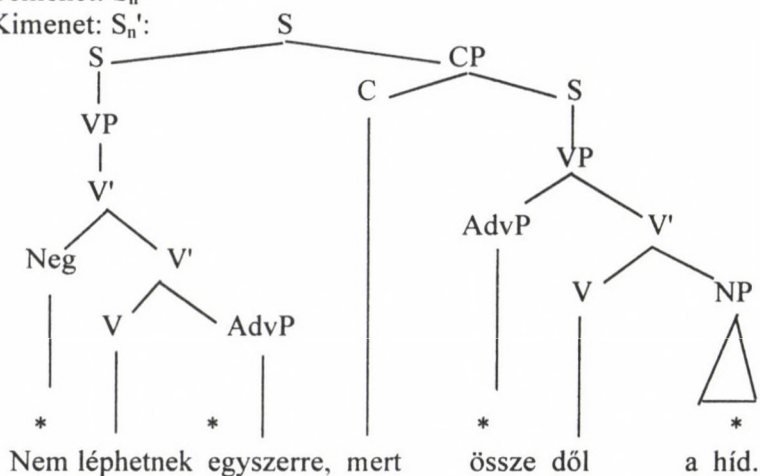
A proszódiai megformálás két szakaszból áll. Az első a *hangsúlyrögzítés szakasza*. Ebbe érkezik a fókuszstruktúrával ellátott szintaktikai felszíni szerkezet, és itt történik meg e szerkezet hangsúlyainak véglegesítése. A hangsúlyrögzítés szakaszában bizonyos szó eleji szóttagok megtartják a fonológia lexikális rétegéből magukkal hozott potenciális főhangsúlyukat, vagyis akcentusjelüket (*). Az akcentusjel egyfelől azt jelzi, hogy az őt hordozó szóttag hangsúlyos, másfelől azt, hogy majd a dallami tagolás szakaszában ez a szóttag egy bizonyos

dallam (hanglejtéskontúr) kiindulópontja lesz. A hangsúlyrögzítés szakaszában persze az is előfordul, hogy bizonyos szókezdő szótagok elveszítik a fonológia lexikális rétegéből magukkal hozott főhangsúlyukat, vagyis deakcentuálódnak, mellékhangsúlyossá redukálódnak. A hangsúlyredukció fordítottja is megtörténhet: előfordulhat, hogy egy funkcionális szó, amely hangsúlytalanul érkezik a fonológia lexikális rétegéből, a hangsúlyrögzítés szakaszában főhangsúlyossá válik, mert a mondat fókuszpozíciójában foglal helyet vagy szembeállított. A hangsúlyrögzítés eredményeként ún. *hangsúlyozott szintaktikai felszíni szerkezet* (S_n') jön létre, vö. (2b).

(2) *Hangsúlyrögzítés szakasza:*

a. Bemenet: S_n

b. Kimenet: S_n' :

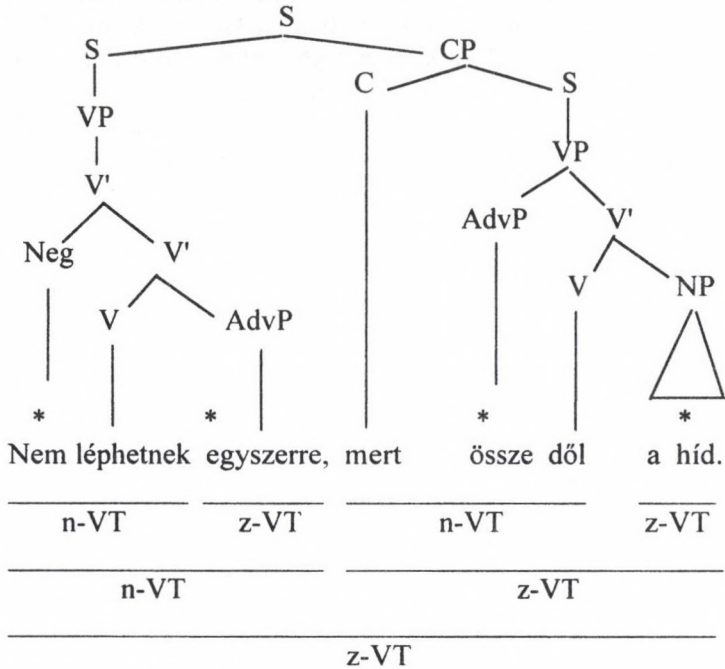


Ez a hangsúlyozott (akcentuált) szintaktikai felszíni szerkezet lép be a prozódiai megformálás második szakaszába, a *dallami tagolás szakaszába*. Itt először a résztömbök kijelölése történik. Például a (2b) alatt bemutatott szerkezet kívülről nézve egyetlen váltótömb, ám befelé két váltótömbre bomlik, majd ezek két-két kisebb váltótömbre bomlanak tovább, vö. (3b).

(3) *Dallami tagolás szakasza:*

a. Bemenet: S_n'

b. Résztömbök azonosítása:



(ahol n-VT = nyitott váltótömb, z-VT = záró váltótömb)

(Ha egy szintaktikai összetevő főhangsúly nélkül érkezik a dallami tagolás szakaszába, akkor nem minősül külön váltótömbnek. Ezért a (3b) példában az akcentus nélküli *léphetnek* ige a tagadószóval kezdődő váltótömb részévé válik, míg az akcentus nélküli *dől* ige az *össze* szóval alkotó váltótömböt; az akcentus nélküli *mert* és *a* szavak pedig az *összedől*, illetve a *híd* váltótömbjébe integrálódnak.)

A váltótömbök (VT-k) olyan akcentuált résztömbök, amelyek között a beszélő komplementáris viszonyt hoz létre, vagyis azt jelzi, hogy az összetartozó (egy mátrixtömbbe tartozó) VT-k közül a nem utolsó helyen állók kiegészítést, folytatást kapnak (nyitott VT-k), és csak az utolsó helyen álló VT (a záró VT) teszi az összetartozó VT-k

sorozatát teljessé, befejezetté. Minden tömb (mondat vagy váltótömb), amely váltótömbökre bomlik, az őt alkotó váltótömbök mátrixtömbje. A mátrixtömb jobbszélső dallama a mátrixtömb uralkodó dallama, és az őt tartalmazó váltótömb pedig a záró váltótömb. Ezt előzik meg a testvérei: nulla, egy, vagy több nyitott váltótömb a mátrixtömbön belül.

A dallami tagolás során a következő lépés a mondat, mint mátrixtömb uralkodó dallamának a meghatározása, vagyis egy dallam hozzárendelése a mondat jobbszélső főhangsúlyos szótagjához, vö. (3c). Ezt a jobbszélső dallamot a mondat kommunikatív típusa (közvetlen illokúciója) és speciális attitűdje dönti el, és ilyen értelemben a beszélő szabad döntését tükrözi. Tulajdonképpen **ez az egyetlen dallam, amelyet a mondatban a beszélő szabadon választ**. A további dallamkijelölések már nem szabadok, mert a jobbszélső dallam korlátozza a testvére(i) dallamát. Minden váltótömbsorozatban a jobbszélső váltótömb jobbszélső dallamától függ, hogy az egyes testvérváltótömbök jobbszélső főhangsúlyos szótagja milyen dallamot kap. Ezért ezek a dallamok függő dallamok. A függő dallamok kiosztását a *melódiaszabályok* vezérlik: ezek a szabályok határozzák meg, hogy milyen dallamok adhatók vagy adandók a testvértömbök főhangsúlyos szótagjainak. A melódiaszabályok kisebb része kötelező, nagyobb része választható. Amikor egy választható melódiaszabály működése elmarad, nem kerül függő dallam kiosztásra, és ilyenkor a főhangsúlyos (akcentusjellel ellátott) szótag egyelőre dallam nélkül marad. A függő dallamok kiosztásának folyamata a legfelső szinten kezdődik és minden egyes alacsonyabb szinten megismétlődik, amíg el nem érjük a legalacsonyabb váltótömböket, vö. (3d).

Miután a dallamokat kiosztottuk, az addig esetleg dallam nélkül maradt, akcentusjellel ellátott szótagok automatikusan féleső dallamot fognak kapni, mert a féleső dallam a főhangsúlyok default dallama, vö. (3e). (A féleső dallam olyan eső dallam, amelynek a vége nem ér le a beszélő hangterjedelmének alsó határára.)

A dallamok után a szünetek kiosztása következik, de ez a jelen vizsgálódás szempontjából érdektelen. A szünet jele a □, vö. (3f). A dallami tagolás végeredményeként a *hangsúlyozott-intonált szintaktikai felszíni szerkezet* (S_n) jön létre, amely azonos a mondat szintű mö-

göttes fonológiai reprezentációval (P_1 -gyel), azaz $S_n'' = P_1$ (vö. Selkirk 1984, 34). A dallami tagolás most ismertetett lépéseit a (3c-g) foglalja össze. Helyszűke miatt itt már a szintaktikai szerkezetet mutató ágrajzot és a VT-struktúrát nem mellékelem újra. Ezek természetesen továbbra is jelen vannak minden lépésben, és csak a kimenetben törölődnek.

(3) *Dallami tagolás szakasza (folytatás):*

c. A mondat uralkodó dallamának meghatározása:

* * *

| Nem léphetnek egyszerre, mert összedől a `híd. |

d. A legmagasabb szintű nyitott VT jobbszélső dallamának meghatározása:

* *

| Nem léphetnek egyszerre, | mert összedől a `híd. |



e. Az alacsonyabb szintű nyitott VT-k jobbszélső dallamainak meghatározása:

| 'Nem léphetnek egyszerre, | mert 'összedől a `híd. |

f. Szünetkiosztás:

| 'Nem léphetnek egyszerre, | □ mert 'összedől a `híd. | □

g. Kimenet: $S_n'' = P_1$:

| 'Nem léphetnek egyszerre, | □ mert 'összedől a `híd. | □

Valahányszor a kiosztott dallam más mint féleső, a dallam intonációs egységhatárban (|) fog végződni. A nem-féleső dallamok ugyanis úgy vannak meghatározva, hogy azok minden esetben egy-egy into-

nációs egység végső, terminális dallamai. Ezzel szemben a féleső dal-
lam csak akkor minősül egy intonációs egység végső dallamának, ha
szünetben végződik. Ha nem végződik szünetben, az intonációs egy-
ség skálájába tartozik (vö. Varga 2001, 308; 2002, 54-61).

A topik mint váltótömb, és a feltételezett melódiaszabályok

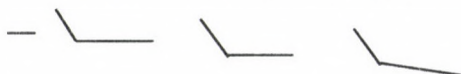
Most térjünk rá szűkebb témánkra, a topik dallami meghatározott-
ságára, illetve arra a kérdésre, hogy a melódiaszabályok hogyan állít-
ják elő a topikok dallamát. A példákban a T-összetevőt dőlt betűs sze-
dés jelzi.

Mivel a T-összetevők váltótömbök, a váltótömbökre vonatkozó ál-
talanos és speciális melódiaszabályok érvényesek rájuk. Ezek ismerte-
téséhez a következőket kell tudnunk. Ha a topik olyan komment test-
vércsomópontja, amelynek uralkodó dallama elől eső (azaz eső, féleső
vagy eső-emelkedő) dallam, és ez a komment nem pusztán egy tag-
mondat, hanem a mondat egész végén van, tehát abszolút mondatvégi
helyzetű, akkor azt mondjuk, hogy **a komment uralkodó dallama α -
típusú**. Ilyenkor bármilyen, *több főhangsúlyt tartalmazó* nyitott váltó-
tömb, tehát a többakcentusú T-összetevő is, opcionálisan lebegő dalla-
mot (azaz emelkedő, magas szinttartó, illetve ereszkedő dallamot)
kaphat, ahogyan az (4a) példa mutatja. Az (4b) a szabályalkalmazás
elmaradását szemlélteti, így a *szer-* szótagon az automatikus féleső
dallam is megjelenhet. Példáinkban az α -típusú dallamot az eső, a le-
begő dallamot pedig az emelkedő képviseli.

(4)a. | A 'kirándulás 'szervezője | ` megszökött. |



b. | A 'kirándulás 'szervezője ` megszökött. |



Ha a topikban *egyetlen főhangsúlyos szótag* van, akkor α -típusú
uralkodó dallam előtt két esetben kaphat lebegő dallamot: egyrészt ak-

kor, ha mellékmondat, vö. (5a); másrészt akkor, ha ismert, vagy valamivel szembeállított bővítmény, vö. (6a). Mivel a szabályok opcionálisak, elmaradásuk esetén a topik az automatikus féleső dallamot fogja kapni. Ezt a lehetőséget mutatják az (5b) és (6b) alatti változatok.

(5)a. | *Amikor* 'megjöttem, | `jó volt a lift. |



b. | *Amikor* 'megjöttem, | `jó volt a lift. |



(6)a. (János bácsi szereti,) | 'Angéla | `utálja. |



b. (János bácsi szereti,) | 'Angéla | `utálja. |



Úgy tűnik, hogy más lényeges melódiaszabály, amely a topik szempontjából releváns lenne, nincsen. Ezért a topik *minden egyéb esetben* az automatikus féleső dallamot kapja, hiszen nem működött rajta melódiaszabály. Ezek az esetek többfélék. Egyrészt közéjük tartoznak az *olyan T-összetevők, amelyek egy főhangsúlyt tartalmaznak és új, nem-szembeállított információt hordoznak, α-típusú uralkodó dallam előtt*, például (7).

(7) | Egy 'ismerősöm | `kivándorolt. |



Másrészt idetartozik *minden olyan T-összetevő, amely nem α-típusú uralkodó dallamú komment előtt áll*. Ilyen a topik, ha a komment uralkodó dallama *lebegő* (azaz emelkedő, magas szinttartó vagy ereszkedő), mint ahogy a (8) mutatja. Másodszor ilyen a topik akkor,

amikor a komment uralkodó dallama *lebegő-eső* (azaz emelkedő-eső, magas szinttartó-eső, ereszkedő-eső), mint ahogy a (9) alatt látjuk. Harmadrészt ilyen a topik akkor is, amikor a komment uralkodó dallama *elől eső* (tehát eső, féleső vagy eső-emelkedő) ugyan, de ez az elől eső dallam nem a mondategész végén, hanem belsejében áll, mint ahogy a (10) alatt látható. Minden ilyen esetben a topik az automatikus féleső dallamot kapja. A féleső dallam végső megjelenése azonban már nem külön szabály, hanem automatikus konvenció eredménye.

(8) | És ha a 'Balaton ' kiszárad?|



(9) | A 'Balatonon ^megnőtt a forgalom?|



(10) | Ha a 'Balaton ' kiszárad is, én `szeretni fogom.|



Empirikus alátámasztás

A topikok dallamára felállított szabályaink jól megfelelnek a valóságban talált topikdallamoknak. Az empirikus vizsgálathoz kétféle módszert alkalmaztam. Először megvizsgáltam a T-összetevők dallamát egy kb. 50 perces összefüggő narratíváról készült hangfelvételen. A hangfelvétel Dickens „Karácsonyi ének prózában” c. művét tartalmazza, Benedek Marcell fordításában, Töröcsik Mari tolmácsolásában. Ebben az anyagban összesen 237 T-összetevő szerepel. Ezeknek közel 87%-a olyan kommenthez tartozik, amelynek uralkodó dallama α -típusú eső. A többi T-összetevőnek kb. 10%-a emelkedő dallamú kommenthez tartozik, 3%-a pedig olyan féleső dallamú kommenthez, amely féleső dallam a mondategésznek nem utolsó dallama. Az alábbi táblázatokban a beszürkített sorok mutatják azoknak az eseteknek az adatait, amelyeknek százalékos részesedése az összes (tehát

237) topik viszonylatában az 5%-ot nem érte el. Ezeket az alacsony értékeket nyugodtan ignorálhatjuk.

1. táblázat: Topikok α -típusú, eső dallamú komment előtt

Topik dallama	Előfordulás száma	Százalékarány összes topikhoz képest (%)	Százalékarány az itt vizsgált topikokhoz képest
Féleső	97	40,93	47,1
Lebegő	104	43,88	50,4
Eső	4	1,69	1,9
Eső-emelkedő	1	0,42	0,5
Összesen	206	86,92	99,9

Az 1. táblázat olyan T-összetevők adatait tartalmazza, amelyek a mondategész végén eső dallamot hordozó komment testvérei. Ilyen volt 206 topik, vagyis az összes topik 86,92%-a. Az itt vizsgált topikoknak 47,1%-a a defaultnak minősülő féleső dallamot kapta, ugyanakkor 50,4%-uk lebegő dallamú lett. (Ezek százalékaránya az összes topikhoz képest 18,99% emelkedő, 13,08% magas szinttartó, 11,81% ereszkedő.) Az ide tartozó topikok között akadt még eső és eső-emelkedő is, de ezek nem fordultak elő értékelhető mennyiségben. Az eredmények összhangban állnak azzal, amit vártunk.

2. táblázat: Topikok információértéke α -típusú, eső dallamú komment előtt

Topik dallama	Új információ	Ismert információ	Szembeállított információ
Féleső	51 = 52,5 %	38 = 39,1 %	8 = 8,2 %
Lebegő	38 = 36,5 %	56 = 53,8 %	10 = 9,6 %

Az α -típusú, eső dallamú komment előtti T-összetevőket abból a szempontból is megvizsgáltam, hogy az összefüggő narratív szövegen belül új, ismert, vagy szembeállított információt hordoznak-e. A 2. táblázat ezek dallami kifejeződését mutatja. A táblázatból látható, hogy a féleső és a lebegő dallamok minden funkcióban előfordulnak,

de a félesők többsége új, a lebegők többsége pedig ismert információt hordozó T-összetevő dallama.

A következő, 3. táblázat olyan T-összetevők adatait tartalmazza, amelyek emelkedő dallamú komment testvérei. Ezek száma 24, az összes topik 10,12%-a. Közülük csak a féleső dallamú topik fordult elő értékelhető mennyiségben, ezek százalékaránya az összes topik viszonylatában 6,33 %, az itt vizsgált topikokhoz viszonyítva 62,5%. Mivel az emelkedő dallamú komment előtti topikok dallamára nézve nem létezik melódiaszabály, érthető, hogy az esetek túlnyomó többségében az ilyen topikok a féleső (default) dallamot kapták, vagyis azt, amit vártunk.

3. táblázat: Topikok nem- α -típusú (emelkedő dallamú) komment előtt

Topik dallama	Előfordulás száma	Százalékarány összes topik-hoz képest	Százalékarány az itt vizsgált topikokhoz képest
Féleső	15	6,33%	62,5%
Lebegő	7	2,95%	29,1%
Eső	1	0,42%	4,2%
Emelkedő-eső	1	0,42%	4,2%
Összesen	24	10,12%	100%

Viszont problémának tűnhet az, hogy néhány esetben más dallamot, például emelkedő dallamot is kapott az ilyen topik. Ez valószínűleg a művészi megformálással magyarázható: mintha ez a konfiguráció, éppen mivel előre megtervezett mondatban valószínűtlen, a mesélés spontaneitását, a mesélő izgalmát volna hivatott kifejezni. Ezért kaphat emelkedő dallamot a topik (*a jelenés*) a következő példában:

(11) *A 'jelenés* |□ 'láncát csörgetve, 'hátrálva ' távozott, |□ 'minden lépésére 'nyílt egy 'kicsikét az 'ablak, |...

Végezetül a 4. táblázat olyan T-összetevők adatait mutatja, amelyek féleső dallamú komment előtt állnak, de ez a komment nem a mondat egész végén helyezkedik el, mert utána új tagmondat követke-

zik. A korpuszban ezek száma elenyésző, mindössze 7, vagyis az összes topik 2,94%-a. Egyik dallammal sem fordulnak elő értékelhető mennyiségben.

4. táblázat: Topikok nem- α -típusú (féleső dallamú) komment előtt

Topik dallama	Előfordulás száma	Százalékarány összes topik-hoz képest	Százalékarány az itt vizsgált topikokhoz képest
Féleső	3	1,26 %	42,8 %
Lebegő	4	1,68 %	57,1 %
Összesen	7	2,94 %	99,9 %

Láthattuk, hogy az összefüggő hangfelvétellel az adatok esetlegesége miatt nem mindig nyertünk használható eredményeket. Ezért egy szisztematikusabb adatnyerő eljárást dolgoztam ki. Egy tíz mondatból álló sorozatot szerkesztettem, amelyet a (12) alatt mutatok be.

(12)a. (Képzeld!) *Izabella* itt marad.

b. Angelika elmegy, *Izabella* nem.

c. (Te mondd!) *Izabella* itt marad?

d. Ha *Izabella* itt marad, végünk van.

e. *Izabella* itt marad, Angelika nem.

f. (Képzeld!) *Izabella barátai* itt maradnak.

g. Angélák elmennek, *Izabella barátai* nem.

h. (Te mondd!) *Izabella barátai* itt maradnak?

i. Ha *Izabella barátai* itt maradnak, végünk van.

j. *Izabella barátai* itt maradnak, Angélák nem.

Az első öt mondatban a vizsgált topik az *Izabella* szó. A (12a)-ban és (12b)-ben ez a topik α -típusú, eső dallamú komment előtt áll; a (12a)-ban új információt hordoz, a (12b)-ben szembeállított. A (12c)-ben emelkedő-eső dallamú komment előtt áll, a (12d)-ben és (12e)-ben viszont lebegő dallamú komment előtt. A második öt mondatban a vizsgált topik az *Izabella barátai* frázis, melyet ugyanolyan kontextusokban helyeztem el, mint az első öt mondatban az *Izabella* szót.

Azért voltam kíváncsi ennek a hosszabb topiknak a realizációira, mert feltételeztem, hogy ezt a felolvasók két főhangsúllyal mondják majd ki, és érdekelt, hogy ez befolyásolja-e a választható dallamokat. A (12d) és a (12i) alárendelt mellékmondatai (*Ha Izabella itt marad*, illetve *Ha Izabella barátai itt maradnak*) egészükben is topikok, és ezért ezeket külön is vizsgálni kell.

A mondatokat 10 (fiatal és középkorú) magyar anyanyelvű beszélővel olvastattam fel. Az eredményeket az (5)-(7) táblázatok mutatják. Az emelkedő, ereszkedő és magas szinttartó dallamokat, mint eddig is, „lebegő” dallamokként összevontam. A megcsillagozott szótagok főhangsúlyosak.

5. táblázat: Topikok α -típusú, eső dallamú komment előtt

* <i>Izabella</i> (új): (12a)	10	8 féleső = 80 % 2 lebegő = 20 %	20	14 féleső = 70 % 6 lebegő = 30 %
* <i>Izabella</i> (szembeáll.): (12b)	10	6 féleső = 60 % 4 lebegő = 40 %		
* * <i>Izabella barátai</i> (új): (12f)	10	4 féleső+féleső = 40 % 6 féleső+lebegő = 60 %	16	12 féleső + lebegő = 75 % 4 féleső + féleső = 25 %
* * <i>Izabella barátai</i> (szembeáll.): (12g)	6	6 féleső+lebegő = 100 %		
* * <i>Ha Izabella itt marad</i> (új/ism.) (12d)	10	10 féleső+lebe- gő = 100%	20	20 (féleső+) féleső + lebegő = 100%
* * * <i>Ha Izabella barátai</i> <i>itt maradnak</i> (új/ism.) (12i)	10	10 féleső + féleső+lebegő = 100%		

Az 5. táblázat 4. sorában az előfordulások száma azért 6, mert a 10 felolvasó közül négyen az *Izabella barátai* szintagmát egyetlen fő-

hangsúllyal ejtették ki, mintha a *barátai* szó ismert információ lett volna.

A 6. táblázat 2. sorában az előfordulások száma azért 9, mert a 10 felolvasó közül egy az *Izabella barátai* szintagmát egy főhangsúllyal ejtette ki.

6. táblázat: Topikok lebegő-eső dallamú komment előtt

* <i>Izabella</i> (új) (12c)	10	3 féleső = 30 % 7 lebegő-eső = 70 %
* * <i>Izabella barátai</i> (új): (12h)	9	5 féleső+féleső = 55,5 % 4 lebegő-eső + le- begő-eső = 44,4 %

A 7. táblázat 4. sorában az előfordulások száma azért 9, mert a 10 felolvasó közül egy az *Izabella barátai* szintagma után itt nem lebegő, hanem eső-emelkedő dallamú kommentet használt.

7. táblázat: Topikok lebegő dallamú komment előtt

* <i>Izabella</i> (új/ismert): (12d)	10	10 féleső = 100 %	20	20 féleső = 100 %
* <i>Izabella</i> (szembeáll.): (12e)	10	10 féleső = 100 %		
* * <i>Izabella barátai</i> (új): (12i)	10	10 féleső+féleső = 100 %	19	19 féleső + féleső = 100 %
* * <i>Izabella barátai</i> (szembeáll.): (12j)	9	9 féleső+féleső = 100 %		

Összegzés

A jelen dolgozat abból a feltevésből indult ki, hogy a magyar mondatban jobbról balra ható, szabályokba foglalható hanglejtésbeli össze-

függések állnak fenn, majd ezt a feltevést *a topikok dallamviszonyaira vonatkozóan* empirikus vizsgálatokkal támasztotta alá.

A topikok feltételezett dallamszabályait először egy ötvenperces narratíváról készült hangfelvétel adataival szembesítettem. Ebben az anyagban a topikok közel 87%-a olyan kommenthez tartozott, amelynek uralkodó dallama α -típusú (azaz eső, féleső, vagy eső-emelkedő) volt. Ilyen komment előtt a topikok közel fele féleső dallamú, fele pedig lebegő dallamú volt (vö. 1. táblázat), és a féleső dallamú topikok többsége új, a lebegő dallamúak többsége pedig ismert információt hordozott (vö. 2. táblázat). A topikok mintegy 10%-a emelkedő dallamú komment előtt szerepelt, és ilyen komment előtt döntő többségük féleső dallamú volt (vö. 3. táblázat). A nyert adatok megfeleltek a szabályaink által megjósoltaknak, sajnos azonban a narratíva eset-leghességei miatt több kérdésben nem született értékelhető eredmény (vö. 4. táblázat). Ezért egy célzottabb adatnyerő eljáráshoz folyamodtam: olyan tízmondatos korpuszt olvastattam fel tíz anyanyelvi beszélővel, amely már több vizsgálandó szempontból értékelhető eredménnyel járt. A (topikokra felállított) melódiaszabályok jól vizsgáztak a gyakorlatban: általában a várt eredményeket kaptuk. Amint az 5. táblázat mutatja, α -típusú, eső dallamú komment előtt az egy főhangsúlyt tartalmazó főnévi topik 70%-ban féleső és 30%-ban lebegő dallamú. A két főhangsúlyt tartalmazó főnévi frázis pedig 75%-ban féleső + lebegő és 25 %-ban féleső + féleső dallamú. A mellékmondati topik 100%-ban (féleső +) féleső + lebegő dallamú. A 7. táblázat tanúsága szerint a lebegő dallamú kommentek előtti topikok 100%-ban féleső, illetve két főhangsúly esetén 100%-ban féleső + féleső dallamúak. Egyedül a 6. táblázat okoz némi meglepetést: lebegő-eső dallamú komment előtt azt tapasztaljuk, hogy a valóságban nem csupán a várt féleső, illetve féleső + féleső dallamú topikok lehetségesek, hanem a nem várt lebegő-eső dallamú topikok is. Ez a megoldás a hitetlenkedő kérdések jellemzője: úgy hangzik, mintha a topikokat külön is kérdezné a beszélő. Lehet, hogy ez az attitűdváltozat egy új melódiaszabályt kíván.

Irodalom

- Elekfi László (1968): Mondatformák és mondathangsúlyok fonetikai vetülete. *Nyelvtudományi Közlemények* 70: 329-350.
- É. Kiss Katalin (1987-88): Még egyszer a magyar mondat intonációjáról és hangsúlyozásáról. *Nyelvtudományi Közlemények* 89: 1-52.
- É. Kiss Katalin (1992): Az egyszerű mondat szerkezete. In: *Strukturális magyar nyelvtan*, 1. kötet, *Mondattan*. Szerk.: Kiefer Ferenc. Akadémiai Kiadó. Budapest, 79-177.
- É. Kiss Katalin (1998): *Mondattan*. In: É. Kiss Katalin – Kiefer Ferenc – Siptár Péter: *Új magyar nyelvtan*. Osiris Kiadó. Budapest, 15-184.
- Fónagy, I. (1998): Intonation in Hungarian. In: *Intonation Systems: A Survey of Twenty Languages*. Eds.: Hirst, D. – Di Cristo, A. Cambridge University Press. Cambridge, 328-344.
- Selkirk, E. O. (1984): *Phonology and Syntax: The Relation between Sound and Structure*. The MIT Press. Cambridge, Massachusetts.
- Varga László (1987-88): Hozzászólás egy hangsúlytanulmányhoz. *Nyelvtudományi Közlemények* 89: 53-66.
- Varga László (1993): A magyar beszéddallamok fonológiai, szemantikai szintaktikai vonatkozásai. *Nyelvtudományi Értekezések* 135.
- Varga László (1994): A hanglejtés. In: *Strukturális magyar nyelvtan 2. Fonológia*. Szerk.: Kiefer Ferenc. Akadémiai Kiadó. Budapest, 486-549.
- Varga László (2001): A magyar intonációs frázis és annak kettébontása. In: *Újabb tanulmányok a strukturális magyar nyelvtan és a nyelvtörténet köréből*. Kiefer Ferenc tiszteletére barátai és tanítványai. Szerk.: Bakró-Nagy Marianne – Bánréti Zoltán – É. Kiss Katalin. Osiris Kiadó. Budapest, 304-323.
- Varga, L. (2002): *Intonation and Stress: Evidence from Hungarian*. Palgrave, Macmillan. Houndmills, Basingstoke.

A MAGYAR NYELV KIEJTÉSVARIÁCIÓI ÉS FELHASZNÁLÁSUK A BESZÉDFELISMERÉSBEN II. *

Vicsi Klára – Szaszák György

A szövegkörnyezet figyelembevétele az ejtésvariáció-vizsgálatokban

Az előző tanulmányban bemutattuk a magyar beszédhangok kiejtésvariációira kapott általános eredményeket, amelyeket az MTBA telefonbeszéd-adatbázis vizsgálata alapján kaptunk. Az összefüggések pontos leírásához azonban szükséges a beszédhangok ejtésvariációinak elemzése a hangkörnyezet figyelembe vételével is (Kiefer 1994), ugyanis könnyen lehetséges, hogy egy adott beszédhang bizonyos ejtésvariációban csak adott hangkörnyezetbe kerülve jelenik meg nagyobb eséllyel. Ha célunk szabálybázis készítése az ejtésvariáció modellezéshez, akkor minél nagyobb a figyelembe vett hangkörnyezet, annál specifikusabb szabályokat kapunk, amelyek azután a hierarchiában is magasabb szinten helyezkednek el (Strik–Cucchiaroni 1998).

Kettős hangkapcsolatok vizsgálatánál feltételként kikötjük, hogy csak bizonyos megadott beszédhang előtt vagy után álló hangok kiejtését vizsgáljuk úgy, hogy közben nem foglalkozunk a feltételként megadott hang kiejtésének vizsgálatával. Ezzel a feltétellel azután elkészítjük a kiejtési mátrixot (vö. Vicsi–Szaszák 2002), amelyet azonban most a kezelhetőség miatt kissé leegyszerűsítünk. A vizsgálat lezárultával töröljük a mátrix minden olyan sorát, amelyben egy általunk megadott küszöbértéknél kevesebb a sorba bejegyzett előfordulások összege, tehát gyakorlatilag ez a hang a feltételként megadott hanggal olyan kevésszer alkot hangkapcsolatot, hogy annak vizsgálata felesleges, de mindenképpen csak pontatlanul lenne végezhető. Továbbá töröljük a mátrix minden olyan oszlopát, amelyben kizárólag

* A Magyar nyelv kiejtésvariációi és felhasználásuk a gépi beszédfelismerésben c. tanulmányt terjedelme miatt két részben közöljük (vö. Beszédkutatás 2002) Az első rész a beszéd kiejtésvariáció fogalmát, vizsgálatának elvi és gyakorlati kérdéseit tárgyalja, valamint bemutatja a magyar beszédhangok ejtésvariációit.

nulla értékeket találunk. Ily módon a kapott feltételes kiejtési mátrix kezelhető méretű lesz, és csak a releváns információkat tartalmazza. Az eredményekből itt most azokat az eseteket ismertetjük, ahol az eddigiekben bemutatott kiejtési mátrixhoz képest jelentősebb eltérés tapasztalható az egyes hangok kiejtésében akkor, ha adott másik hang előzi meg vagy követi azokat.

Az adott beszédhangot megelőző hangok kiejtésének vizsgálata

Az eredményeket az 1. táblázatban mutatjuk be. A táblázat második oszlopában áll a vizsgált kimondandó hang, a harmadik oszlopban pedig a tapasztalt ejtésvariációk, amennyiben azok gyakoriságukban eltérnek a kiejtési mátrixban kapottaktól. Az első oszlop a feltételt tartalmazza, az itt szereplő hangnak követnie kell a vizsgált hangot. (A táblázatokban a beszédhangok jelölésére SAMPA-szimbólumokat használunk (Roach et al. 1996), a szövegben szintúgy, de a hang jelét ekkor szögletes zárójelbe tettük. A magyar beszédhangok SAMPA jelölésrendszerét lásd a függelékben.)

Az eredményekből kiemeltünk néhány fontos esetet:

- Nazális hangok előtt az amúgy sem túlságosan gyakran kanonikus kiejtésű zárhangok kiejtése igen bizonytalan, ha nem kanonikus alakú a kiejtés, akkor zömmel a felpattanási zörej marad el, de a zárszakasz megtalálható. A felpattanó zárhangok közül a legstabilabb kiejtésű [t] hang is csak 70-80% gyakorisággal kanonikus kiejtésű, ha nazális beszédhang előtt áll. Különösen kritikus a [b-m], [d-m], [d'-m] és a [d-n] hangkapcsolatokban a [b], [d], illetve a [d'] hangok kiejtése.
- Az [l] hang előtt a [d] hang, [h] hang előtt a [t'] hang kiejtése mindig kanonikus alakú.
- A zárhangok kiejtése az őket követő mássalhangzó függvényében igen változó lehet, a kanonikus alakútól eltérő kiejtés általában ott gyakoribb, ahol a két hang egymás utáni képzése nehezebb, például a [d-r] hangkapcsolatban.
- Az [r] hang előtt az egyébként többnyire kanonikus kiejtésű [n] hang az esetek több mint 20%-ában kiesik. Szintén gyakori (13,95%) az [J] hang [n]-nel való helyettesítése [t'] hang kiejtése előtt.

- A [b] és [d] hangok előtt a [Z] hang kiejtése többnyire kanonikus alakú, szemben az általános esettel, ahol a [Z] hang kanonikus kiejtésének előfordulási aránya 87% körüli.

Látható, hogy sok esetben azoknak a hangkapcsolatoknak a kiejtése problémás, amelyekben a szomszédos hangok képzési helye azonos vagy igen közeli.

1. táblázat: Egy adott hang előtt álló beszédhangok ejtésvariációi

Következő hang	Vizsgált hang	A vizsgált hang jellemző ejtésvariációi és azok gyakorisága, ha a vizsgált hangot az első oszlopban megadott hang követi
+ ^h b	r	r 100%
	Z	Z 98,32%; S 1,68%
- ^h ts	l	l 100%
	n	n 100%
- ^h tS	r	r 100%
	m	m 100%
	S	S 94,00%; S: 4,00%; * 2,00%
+ ^h d	Z	Z 94,55%; S 3,64%
f	s	s 100%
+ ^h g	r	r 100%
	Z	Z 94,55%; S 3,64%
+ ^h d'	l, r	l, r 100%
	J	J 81,16%; J: 1,45%; n 17,39%
+ ^h d'	Z	Z 95,83%; S 4,17%
h	n	n 95,68%; egyéb 4,32%
	- ^h t'	- ^h t' 99,10%
	- ^h ts	- ^h ts 14,29%; ** 42,86%; - ^h :k 4,76%; egyéb 38,60% (oka a ch hangként vagy k hangként való kiejtése)
j	r	r 94,39%; * 2,80%; j: 1,87% (ekkor a jr helyett hosszú j)
	S	S 96,30%; S: 3,70%
	- ^h k	- ^h k 93,75%; - ^h p 3,12%; - ^h t 3,12%

Következő hang	Vizsgált hang	A vizsgált hang jellemző ejtésvariációi és azok gyakorisága, ha a vizsgált hangot az első oszlopban megadott hang követi
- ^h k	l	l 95,89%; * 3,86%
	r	r 100%
	j	j 100%
	S	S 100%
l	- ^h k	- ^h k 98,00%
	r	r 93,48%; * 6,52%
	S	S 97,85%; Z 1,08%
m	- ^h ts	- ^h ts 91,67%; * ^h ts 8,33%
	+ ^h b	+ ^h b 50,00%; +* 42,50%; * ^h b 2,50%; +* 2,50%
	+ ^h d	+ ^h d 66,04%; * ^h d 1,92%; ** 7,55%; +* 23,08%
	+ ^h g	+ ^h g 57,89%; * ^h g 9,65%; ** 0,88%; +* 30,70%
	+ ^h d'	+ ^h d' 50,26%; * ^h d' 19,37%; +* 29,84%
	- ^h k	- ^h k 73,02%; * ^h k 1,59%; ** 3,70%; - ^h 19,58%
n	- ^h t	- ^h t 81,63%; ** 2,38%; - ^h 19,58%
	s	s 95,45%; s: 4,55%
	+ ^h d	+ ^h d 68,09; +* 28,09%
	+ ^h g	+ ^h g 75,53%; * ^h g 6,91%; ** 1,06%; +* 16,49%
	+ ^h d'	+ ^h d' 68,09%; +* 29,09%
	- ^h k	- ^h k 81,05%; ** 4,58; - ^h 9,15%

n	- ^h t	- ^h t 70,39%; ** 1,21%; - * 27,19%
	r	r 94,29%; * 5,14%
	J	J 90,00%; n: 10,00%
J	+ ^h g	+ ^h g 57,14%; * ^h g 14,29%; ** 5,71%; +* 17,14%; + ^h d' 2,86%
	- ^h k	- ^h k 86,96%; ** 8,70%; más 4,35%
	- ^h t	- ^h t 86,36%; -* 13,64%
	r	r 96,97%; * 3,03%
	- ^h p	r 100%
- ^h p	r	r 100%
	m	m 94,51%; n 5,49%
	S	S 100%
r	+ ^h b	+ ^h b 91,89%; * ^h b 2,70%; +* 5,41%
	+ ^h d	+ ^h d 82,86%; ' ^h d 12,86%; +* 1,43%
	+ ^h g	+ ^h g 97,50%; * ^h g 1,67%
	- ^h k	- ^h k 98,42%; * ^h k 1,05%
	- ^h p	- ^h p 80,92%; * ^h p 12,98%; ** 1,53%
	n	n 77,92%; * 20,78%

r	- ^h t	- ^h t 96,03%; * ^h t 1,32%; ** 1,32%; -* 1,32%
S	- ^h k	- ^h k 84,91%; * ^h k 5,66%; -* 7,55%
	- ^h p	- ^h p 80,43%; * ^h p 6,52%; ** 4,35%; -* 4,35%
	j	j 100%
	l	l 97,60%; * 2,40%
	r	r 96,67%; * 2,67%
s	J	J 95,83%; * 4,17%
	j	j 88,57%; * 11,43%
	m, n	m, n 100%
- ^h t	j	j 100%
	m	m 100%
- ^h t'	J	J 86,05%; n 13,95%
v	+ ^h d	+ ^h d 96,61%; +* 1,95%
	+ ^h g	+ ^h g 83,59%; +* 12,50%
	- ^h k	- ^h k 88,64%; -* 5,68%
	- ^h t	- ^h t 96,48%; * ^h t 1,95%
	l	l 97,34%; * 2,33%
z	+ ^h b	+ ^h b 83,33%; * ^h b 4,17%; +* 12,50%
	+ ^h g	+ ^h g 83,33%; * ^h g 8,33%; ** 2,78%; +* 5,56%

Az 1. táblázat a mássalhangzó–mássalhangzó hangkapcsolatokban mutatta be a jelentősebb eltéréseket a hangkiejtési mátrixhoz képest. A mássalhangzó–magánhangzó hangkapcsolatokra általában jellemző, hogy bennük a mássalhangzók kiejtése kevés kivételtől eltekintve általában a beszédhang kiejtési mátrixban már megadott ejtésvariációk szerint alakul, az egyes ejtésvariációk gyakoriságában sincs jelentősebb eltérés, illetve ritkán van jelentősége a mássalhangzó kiejtésének szempontjából annak, hogy milyen magánhangzó követi. Az eddig elmondottak alól néhány kivétel azonban akad, ezek:

- A zárhangoknak magánhangzók előtt gyakorlatilag csak két ejtésvariációjuk van, a kanonikus kiejtésű alak, amikor a zárszakasz is és a felpattanási zöreje is megtalálható, illetve amikor a zárszakasz kiesik.
- Az [2] hang előtt a [b] hang zárszakasza 50%-os gyakorisággal kiesik. Hasonlóan [o] hang előtt a [d] hang zárszakasza esik ki 25,81%-os gyakorisággal, [u] hang előtt szintén a [d] zárszakasza marad el 24,18% gyakorisággal.

- Az [A:] hang előtt a [Z] hang az esetek 35,48%-ában hosszú [S:], 9,68% gyakorisággal rövid [S] ejtésvariációban hangzik el, és csak 51,61% a [Z] hang kanonikus kiejtésének a gyakorisága.
- Az [e:] hang előtt a hosszú [l:] hang 8,82% gyakorisággal megrövidül, egyébként kanonikus kiejtésű.

Egy adott beszédhangot követő hangok kiejtésének vizsgálata

A 2. táblázat foglalja össze röviden azokat az eseteket, ha feltételünk az, hogy a vizsgálandó hangot valamely adott hang előzze meg.

2. táblázat: Megadott hangot követő beszédhangok ejtésvariációi

Előző hang	Vizsgált hang	A vizsgált hang jellemző ejtésvariációi és azok gyakorisága, ha a vizsgált hangot az első oszlopban megadott hang előzi meg
+ ^h b	l	l 100%
	m	m 100%
	r	r 100%
	z	z 100%
- ^h ts	h	h 52,38%; * 47,62% (utóbbi, ha a ch kapcsolatot h-nak ejti)
	- ^h k	- ^h k 100%
	m	m 100%
	v	v 100%
- ^h tS	- ^h k	- ^h k 100%
+ ^h d	l	l 96,30%; * 3,70%
	r	r 98,63%; * 1,37%
	n	n 97,87%; * 2,13%
	m	m 100%
	v	v 100%
f	- ^h t	- ^h t 92,59%; * ^h t 7,41%
+ ^h g	j	j 100%
	r	r 100%
	m	m 100%
	n	n 100%
	J	J 97,14%; * 2,86%
+ ^h d'	l, r	l, r 100%
	m, n	m, n 100%
J	l, r	l, r 100%
	n	n 100%

Előző hang	Vizsgált hang	A vizsgált hang jellemző ejtésvariációi és azok gyakorisága, ha a vizsgált hangot az első oszlopban megadott hang előzi meg
J	+ ^h b	+ ^h b 96,43%; ** 3,57%
	+ ^h d	+ ^h d 90,32%; * ^h d 1,08%; ** 7,53%
	- ^h k	- ^h k 96,88%; más 3,12%
	- ^h t	- ^h t 100%
	S	S 100%
- ^h k	j	j 93,75%; * 3,12%; más 3,12%
	m	m 100%
	J	J 100%
	h	h 97,32%; h\ 1,34%
l	+ ^h b	+ ^h b 100%
	+ ^h d	+ ^h d 96,45%
	+ ^h g	+ ^h g 98,81%
	+ ^h d'	+ ^h d' 100%
	- ^h k	- ^h k 98,98%
	- ^h t	- ^h t 98,43%
	- ^h tS	- ^h tS 88,54%; * ^h tS 2,08%; - ^h t 3,12%; egyéb 6,25%
	m	m 100%
	J	J 100%
	f	f 100%
	S	S 100%
m	+ ^h b	+ ^h b 49,01%; * ^h b 50,20%
	- ^h k	- ^h k 97,73%
m	- ^h p	- ^h p 93,47%; ** 2,20%

	- ^h t	- ^h t 98,31%
	- ^h tS	- ^h tS 97,78%
	l, r, j	l, r, j 100%
	n	n 100%
	s, S	S, S 100%
n	+ ^h d	+ ^h d 40,75%; * ^h d 56,95%
	- ^h t	- ^h t 94,18%; * ^h t 2,62%; - * 1,27%
	- ^h ts	- ^h ts 98,31%
	- ^h tS	- ^h tS 97,78%
	r	r 100%
	z	z 100%
N	- ^h k	- ^h k 90,20%; * ^h k 3,81%; ** 2,01%; - ^h * 1,38%
J	+ ^h d ^h	+ ^h d ^h 69,70%; * ^h d ^h 27,27%; ** 1,52%
	- ^h k	- ^h k 94,87%; - ^h t 2,56%
	- ^h t	- ^h t 93,22%; - ^h * 3,39%
	- ^h t ^h	- ^h t ^h 90,48%; - ^h t 4,76%
	l, r	l, r 100%
	m	m 100%
	n	n 90,00%; * 10,00%
	S, v, z	S, v, z 100%
r	+ ^h b	+ ^h b 100%
	+ ^h d	+ ^h d 98,64%
	+ ^h g	+ ^h g 100%
	+ ^h d ^h	+ ^h d ^h 100%
	- ^h k	- ^h k 96,91%
	- ^h p	- ^h p 92,31%
	- ^h t	- ^h t 96,80%; * ^h d 1,92%
	- ^h t ^h	- ^h t ^h 97,92%
	- ^h ts	- ^h ts 96,63%; ** 1,12%
	- ^h tS	- ^h tS 95,00%; * ^h tS 5,00%
r	l	l 100%

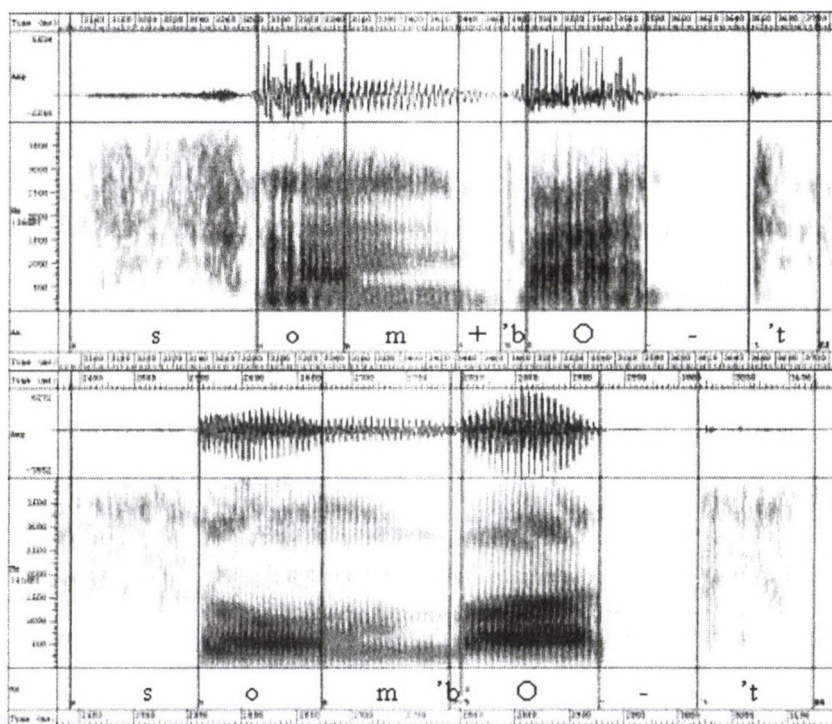
	m, n	m, n 100%
	f, S	f, S 100%
S	- ^h k	- ^h k 99,38%
	- ^h p	- ^h p 98,44%
	- ^h tS	- ^h tS 95,25%
	j	j 100%
	m	m 100%
s	- ^h k	- ^h k 99,00%
	- ^h p	- ^h p 100%
	- ^h t	- ^h t 96,65%
	r	r 100%
	m, n	m, n 100%
	f, v	f, v 100%
	h	h 94,03%; h\ 2,99%
- ^h t	l	l 100%
	J	J 100%
	f	f 100%
- ^h t ^h	f	f 100%
	h	h 98,20%
z	+ ^h b	+ ^h b 98,36%; * ^h b 1,64%
	+ ^h d	+ ^h d 89,66%; * ^h d 6,90%; + ^h * 1,38%
	+ ^h g	+ ^h g 100%
	l	l 100%
	r	r 96,15%; * 3,85%
	m	m 100%
	v	v 100%
Z	+ ^h b	+ ^h b 98,33%; * ^h b 1,11%
	+ ^h d	+ ^h d 88,00%; * ^h d 4,00%; + ^h * 4,00%
	+ ^h g	+ ^h g 97,14%
	+ ^h d ^h	+ ^h d ^h 100%

A 2. táblázat eredményeiből is szeretnénk néhány lényegesebbet külön kiemelni:

• Bizonyos hangok után az igen bizonytalan kiejtésű zárhangok és affrikáták viselkedése is stabil, ilyen például a [ts-k], [tS-k], [S-k], [s-k] hangkapcsolatokban a [k] hang kanonikus kiejtése vagy [z] után a [g], [r] után a [b], [j] után a [t], [s] után a [p] hang stb. biztosan kanonikus kiejtése. E hangpárok képzési helyei egymástól elég távoliak, így vél-

hetően nem okoz gondot a második hang kiejtése közvetlenül az előző hang után.

• Nazális hangok után a zöngés zárhangok kiejtésében az esetek igen jelentős részében elmarad a zárszakasz, és csak a felpattanási zörej hallatszik. Példát e variációra az 1. ábrán mutatjuk be. A kritikus hangkapcsolatok ebből a szempontból az [m-b], [n-d] és az [J-d'], ezen hangkapcsolatok mindkét tagjának képzési helye mindig azonos.



1. ábra

A *szombat* szó [b] hangja kanonikus kiejtésben (felső spektrogram), és a zárszakasz elmaradásával (alsó spektrogram)

• Az [l-tS] hangkapcsolatokban viszonylag sok esetben (3,12%) a [tS]

hangot külön [t] és [S] hangokként ejtik ki a beszélők. Gyakran az írásképpen is így jelenik meg ez a hangkapcsolat, például: *Oltsa el a lámpát!*. A felszólító mód miatt ilyenkor gyakran az utolsó szótagot is erősebben hangsúlyozzuk, ez szintén hozzájárulhat a fenti jelenséghez, ebben az esetben pedig az ejtésvariáció prozódiai jegyekkel való összefüggésére találtunk példát, ami egy külön vizsgálat tárgya lehet.

- A likvidák bizonyos mássalhangzók után ejtve igen stabilan kanonikus kiejtésűek, például [z] hang után ilyen az [l], vagy az [m] hang után az [l], [r] és a [j] hangok kiejtése. Máshol pedig kieshetnek, igaz nem túl gyakran, például [z] után az [r] hang 3,85% gyakorisággal kiesik.

- A nazális hangok, különösen az [m] kiejtése számos mássalhangzó után szintén minden esetben kanonikus, ám [J] után az [n] hang 10% valószínűséggel kiesik, [d] után szintén az egyébként viszonylag stabil kiejtésű [n] hang eshet ki. A fenti példákban a kapcsolatba került hangok képzési helye ismét igen közel van egymáshoz.

A magánhangzó–mássalhangzó hangkapcsolatokban is ritka a jelentősebb eltérés a beszédhang kiejtési mátrixához képest a mássalhangzó ejtésvariációinak gyakoriságában. Van azonban néhány kivétel:

- A rövid zárhangok bizonyos magánhangzók után hangkapcsolatokban viszonylag stabil kanonikus kiejtésűek, más esetben ejtésvariációik és azok gyakoriság szerinti eloszlása nem mutat jellemző jegyeket és közel áll a hangkiejtési-mátrix megfelelő értékeihez. A [t'] hang ejtésvariációi az a hang előtt állva: [-'t'] 61,90%; [-:'t] 19,05%; [-*] 2,38%; [*'t'] 4,76%. A [h] hangé az a [O] hang előtt [h\]: 83,05%; [h] 10,57%.

- Általában elmondható, hogy hosszú magánhangzót követően stabilabb a mássalhangzók kiejtése, mint rövid magánhangzók után.

- Az [o:] hang után a hosszú [m:] hang 4%-os gyakorisággal megrövidül.

- Az [i] hang után a [ts] csupán 58,62% gyakorisággal kanonikus kiejtésű, 20,69% gyakorisággal kiesik, szintén 20,69% gyakorisággal egyéb más módon módosul.

Zárhang-zárhang hangkapcsolatok ejtésvariáció-vizsgálata

Már a kiejtési mátrixból, de különösen a hangkapcsolatokat vizsgáló mátrixokból látszott, hogy a felpattanó zárhangok egymással alkotott hangkapcsolatainak kiejtése gyakran eltér a kanonikus alaktól. Mivel a mai beszédfelismerőkben éppen a zárhangok felismerése a legkevesbé pontos, ezért ezeket a hangkapcsolatokat külön is megvizsgáltuk, egyszerre odafigyelve mindkét hang viselkedésére a hangkapcsolatban. Emiatt az utóbbi két, hangkapcsolatok kiejtését vizsgáló 1. és 2. táblázatban nem is tértünk ki e hangkapcsolatokra.

Az elemzést csak azokra a hangkapcsolatokra végeztük el, amelyek legalább húszszor fordulnak elő az MTBA szövegében. Az eredmények a 3. táblázatban láthatóak.

3. táblázat: A felpattanó zárhangok egymással és egyes affrikátákkal alkotott hangkapcsolatainak ejtésvariációi

Eredeti HK	A hangkapcsolat (HK) kiejtésének alakja Az egyes esetek gyakorisága százalékban megadva							
	+d+b	+**b	+:**b	*d+b	~**b	+d~b	+d*b	egyéb
+d+b	56.50	5.69	20.73	9.35	0.00	0.41	0.81	6.50
	+g+b	+**b	+:**b	*g+b	~**b	+g~b	+g*b	egyéb
+g+b	47.02	2.82	26.96	21.00	0.00	0.00	0.31	1.88
	+d'+b	+**b	+:**b	*d'+b	~**b	+d'~b	+d'*b	egyéb
+d'+b	63.27	0.00	6.12	30.61	0.00	0.00	0.00	0.00
	+g'+d	+**d	+:**d	*g'+d	~**d	+g~d	+g*d	egyéb
+g'+d	58.93	5.36	16.07	16.07	0.00	0.00	1.79	1.79
	+d'+g	+**g	+:**g	*d'+g	~**g	+d'~g	+d'*g	egyéb
+d'+g	80.00	0.00	4.00	8.00	4.00	0.00	4.00	0.00
	-k'-p	-**p	-:**p	*k'-p	~**p	-k'~p	-k'*p	egyéb
-k'-p	50.65	11.69	24.68	0.00	1.30	3.90	1.30	6.49
	-k'-t	-**t	-:**t	*k'-t	~**t	-k'~t	-k'*t	egyéb
-k'-t	77.62	1.67	12.38	4.29	0.24	0.95	0.24	2.62
	-k'-ts	-**ts	-:**ts	*k'-ts	~**ts	-k'~ts	-k'*ts	egyéb
-k'-ts	76.47	5.88	11.76	0.00	0.00	2.94	0.00	2.94

	-l'-tS	-**tS	-:***tS	*l'-tS	~***tS	-l'~tS	-l'*tS	egyéb
-l'-tS	62.50	6.25	14.06	4.69	0.00	6.25	3.12	3.12

	-l'-t	-**t	-:***t	*l'-t	~***t	-l'~t	-l'*t	egyéb
-l'-t	73.49	0.00	18.07	20.01	0.00	20.01	3.61	2.41

	-l'-k	-**k	-:***k	*l'-k	~***k	-l'~k	-l'*k	egyéb
-l'-k	81.67	5.00	8.33	3.33	0.00	0.00	0.00	1.67

	-l'-tS	-**tS	-:***tS	*l'-tS	~***tS	-l'~tS	-l'*tS	egyéb
-l'-tS	89.19	2.70	5.41	0.00	0.00	0.00	2.70	0.00

	-l'-k	-**k	-:***k	*l'-k	~***k	-l'~k	-l'*k	egyéb
-l'-k	89.30	1.53	1.97	3.93	0.00	1.53	0.22	1.53

	-l'-p	-**p	-:***p	*l'-p	~***p	-l'~p	-l'*p	egyéb
-l'-p	68.97	17.05	6.90	0.00	0.00	13.79	0.00	17.05

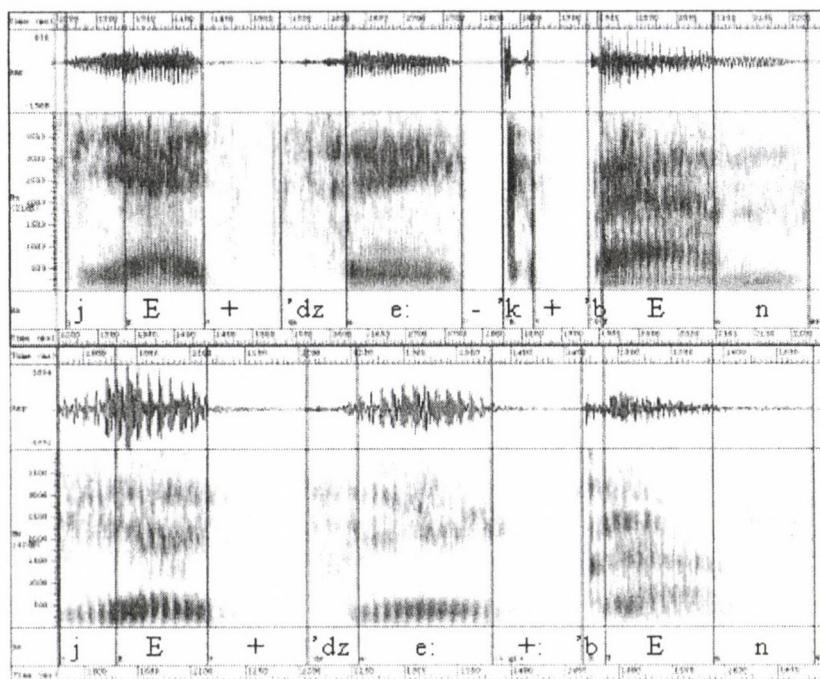
	-l'-s	-**s	-:***s	*l'-s	~***s	-l'~s	-l'*s	egyéb
-l'-s	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	-l'-p	-**p	-:***p	*l'-p	~***p	-l'~p	-l'*p	'd'*-p	egyéb
-l'-p	56.52	4.35	15.22	15.22	0.00	0.00	0.00	4.35	4.35

	-l'-k	-**k	-:***k	*l'-k	~***k	-l'~k	-l'*k	'd'*-k	egyéb
-l'-k	75.00	0.00	8.00	8.00	0.00	0.00	0.00	6.00	3.00

	-l'-t	-**t	-:***t	*l'-t	~***t	-l'~t	-l'*t	'd'*-t	egyéb
-l'-t	15.00	3.75	40.00	12.50	0.00	0.00	0.00	22.50	25.06

A táblázat eredményeiből látható, hogy a [t'-t] hangkapcsolat kivételével mindenhol leggyakoribb a hangkapcsolat mindkét tagjának kanonikus kiejtése, de igen jelentős az eltérő ejtésvariációk száma és gyakorisága is. Minden esetben tipikus ejtésvariáció a hangkapcsolat első tagjának kiesése és ezzel párhuzamosan a zárszakasz megnyúlása, amelyet a hangkapcsolat második tagjának zárfelpattanása követ. Erre példát a 2. ábrán mutatunk be. Ezen kívül a hangkapcsolat típusától függően további 2-3 egyéb ejtésvariáció is megfigyelhető. Külön hangsúlyozást érdemel a [ts-k] hangkapcsolat állandó kanonikus kiejtése, amelyre értelemszerűen a fenti megállapítások nem vonatkoznak.



2. ábra

A *jegyzékben* szó [k-b] hangkapcsolata két különböző ejtésvariációban. Az első esetben a [k] nem zöngésedik. Az alsó spektrogramban a zöngésedett [k] zár-felpattanása elmarad.

A magyar kiejtésvariációk vizsgálati eredményeinek további felhasználási lehetőségei

Az ejtésvariációk vizsgálatával kapott eredmények a gyakorlatban két főbb vonalon használhatóak fel. Az első lehetőség a felvételek szegmentálása és címkézése során használt fonetikai átírási szabályrendszer, illetve esetlegesen a szimbólumkészlet kontrollálása. Ekkor az ejtésvariáció-vizsgálatok eredményeit visszacsatoljuk a fonetikus átírás végrehajtásához, hogy az a későbbiekben a szükségesnek ítélt változtatások végrehajtása után minél pontosabban készülhessen el az egyes felvételekre. Az eredmények felhasználásának második, alapve-

tő lehetősége a szerzett ismeretek beszédfelismerés során történő hasznosítása, ekkor növelhető a felismerés hatékonysága, ha az egyes lexikon elemek többféle ejtésvariációját is figyelembe vesszük.

A fonetikus átírás szabályrendszerének bővítése

Az MTBA adatbázison végzett ejtésvariáció-vizsgálatok alapján megállapítható, hogy a bemondott szöveg kanonikus fonetikai átírásához használt szabályrendszer alapvetően jó, azonban néhány ponton mégis érdemes lenne kiegészíteni. A kiegészítés azokon a pontokon mérlegelendő, ahol bizonyos hangok kanonikusnak tekintett ejtésvariációja kevésbé gyakori, mint egy a kanonikustól eltérő más ejtésvariációjuk. A jövőben célszerű a leggyakrabban előforduló esetet tekinteni kanonikus ejtésvariációnak, és a fonetikus átírás szabályrendszerét ennek megfelelően módosítani. Az eddig elmondottak alapján legalább az alábbi két esetben szükséges a kanonikus fonetikai átírás kiegészítése:

1. A nazális hang után következő, ahhoz közeli képzési helyű zárhang zárszakasza elmarad ([n-d] kapcsolatnál a [d] hang zárszakasza gyakrabban marad el (57%), mint ahányszor a beszélők kiejtenék 40%). Ezért az [m-b], [n-d] és az [J-d'] hangkapcsolatoknál érdemes lenne a nazális hang, majd az azt közvetlenül követő zárfelpattanás együttesét külön hangként felvenni.
2. Kérdéses, vajon érdemes-e a [h] hang három eltérő változatát megkülönböztetni (jelenleg [h], [h\] és [x] változatokat különítenek el). Ezek megkülönböztetése felesleges, tekintve, hogy a kiejtés során e változatok instabilan viselkednek, egymásba rendszeresen átcsúsznak.

Az eredmények felhasználása a beszédfelismerésben

A jövőben az ejtésvariáció-vizsgálatok eredményeit folyamatos gépi beszédfelismerő felismerési hatékonyságának növeléséhez is fel lehet majd használni (Riley et al. 1998). A magyar nyelvű folyamatos gépi beszédfelismerőbe is beépíthetők az ejtésvariáció-vizsgálat elvégzésével kapott eredmények, illetve a belőlük készített ejtésvariációs modellek, természetesen a magyar nyelv sajátosságait szem előtt tartva.

Egy lehetséges eljárás ezek alapján a következő lehet: az ejtésvariáció-vizsgálatok eredményeiből szabálybázist lehet konstruálni, amely a magyar nyelv főbb kiejtési szabályait tartalmazza. A szabálybázis felhasználásával azután az egyes lexikonelemekre előállítható kanonikus kiejtésük, illetve a lexikonelemek egyes ejtésvariációi valószínűségeikkel együtt, például ejtésvariációs gráfok formájában. Ezt követően a lexikon kibővíthető az egyes lexikonelemeknél tapasztalt ejtésvariációkkal. A lexikon bővítésénél célszerű mértéktartónak maradni, hogy az elemszám drasztikus növekedése miatt ne állhasson elő az egyes lexikonelemek gyakoribb összetévesztése. Így célravezető lehet, ha csak a leggyakrabban előforduló lexikonelemekre veszünk fel ejtésvariációkat vagy ezek számát is próbáljuk korlátozni úgy (Schiel et al., 1998), hogy minden ejtésvariáció esetén megköveteljük annak egy minimális, például 10%-os előfordulási gyakoriságát.

Irodalom

- Kiefer Ferenc (szerk.) (1994): Strukturális magyar nyelvtan. 2 kötet. Fonológia. Akadémiai Kiadó. Budapest.
- Riley, M. – Byrne, W. – Finke, M. – Khudanpur, S. – Ljolje, A. (1998): Stochastic pronunciation modelling from hand-labelled phonetic corpora. In: Modeling Pronunciation Variation for ASR. Rolduc, 109-116.
- Roach, P. S. – Arnfield, W. – Barry, J. – Baltova, M. – Boldea, A. – Fourcin, W. – Gonet, R. – Gubrynowicz, E. – Hallum, L. – Lamel, K. – Marasek, A. – Marchal, E. – Meister, E. – Vicsi, K. (1996): BABEL: An Eastern European Multi-language database. International Conference on Speech and Language.
- Schiel, F. – Kipp, A. – Tillmann, H. G. (1998): Statistical modeling of Pronunciation variation, It's not the model, it's the data. In: Modeling Pronunciation Variation for ASR, Rolduc, 131-136.
- Strik, H. – Cucchiari, C. (1998): Modeling Pronunciation Variation for ASR: Overview and comparison of methods, In: Modeling Pronunciation Variation for ASR, Rolduc, 137-144.
- Vicsi Klára – Szaszák György (2002): A magyar nyelv kiejtésvariációi és felhasználásuk a beszédfelismerésben I. In: Beszédkutatás 2002. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest, 216-233.

Függelék

4. táblázat: A magyar nyelv beszédhangjainak és néhány tipikus realizációjuknak SAMPA jelölésrendszere

Betűk	Példák	IPA	SAMPA
a	ha	ɒ	ɒ
á	hát	a:	ɑ:
e	vet	ɛ	ɛ
é	vét	e:	e:
i	hű	i	i
í	szít	i:	i:
o	sok	o	o
ó	sók	o:	o:
ö	köt	ø	2
ő	sőt	ø:	2:
u	fut	u	u
ú	kút	u:	u:
ü	süt	y	y
ű	fűt	y:	y:

Betűk	Példák	IPA	SAMPA
p	pad	p	p
b	bab	b	b
t	tél	t	t
d	dél	d	d
k	kép	k	k
g	gép	g	g
c	cél	tʃ	tʃ
dz	bodza	dʒ	dʒ
cs	cső	tʃ	tʃ
dzs	dzsém	dʒ	dʒ
ty	tyúk	c	t'
gy	gyár	j	d'
f	fél	f	f
v	vér	v	v
sz	szép	s	s
z	zaj	z	z
s	só	ʃ	ʃ
z	zár	ʒ	ʒ
h	hét	h	h
r	rész	r	r
l	lép	l	l
j	jön, lyuk	j	j
m	méz	m	m
n	néz	n	n
ny	nyom	ɲ	J

A /h/, /j/, /n/ és /m/ fonémák megkülönböztetett ejtés-változatai (allofónjai)

j	kapj	ç	x
---	------	---	---

m	kémény, hamvas, hosszú, hosszú	ɱ	F
n	ing, lök	ɳ	N

h	doh	x	x
h	ühet	ɣ	x
h	lehet	f	h\

A hosszan ejtett hang jele minden esetben a megfelelő SAMPA szimbólum után tett kettőspont.

Köszönetnyilvánítás

A magyar ejtésvariációk kutatását a Magyar Telefonbeszéd Adatbázis felhasználása tette lehetővé, amely adatbázist az IKTA pályázat keretében hoztuk létre.

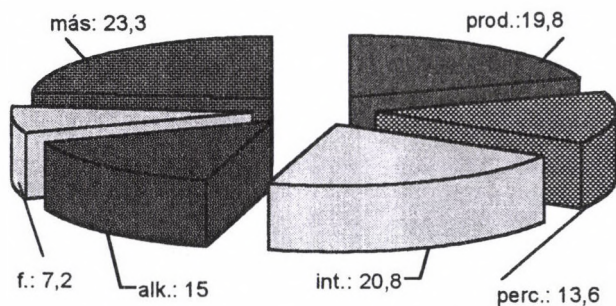
A 15. Fonetikai Világkongresszus International Congress of Phonetic Sciences

A Fonetikai Világkongresszus a beszéddel foglalkozó kutatók, oktatók, diákok és (könyv- és eszköz)kiállítók négy évenként megrendezésre kerülő nagy seregszemléje. A világkongresszusok történetének kezdete 1932-re nyúlik vissza, amikor először rendeztek ilyen eseményt Amszterdamban, s a résztvevők elhatározták, hogy bizonyos idő elteltével rendszeresen találkozni fognak, és ismertetik egymással legfrissebb eredményeiket. Azóta ez az esemény a fonetikusok legnagyobb és szakmailag legelismertebb találkozójává vált. A résztvevők száma általában nagy, és az előadások tudományos színvonala is magas. A világkongresszus 20. századi helyszínei a következők voltak: London (1935), Genf (1938), Helsinki (1961), Münster (1964), Prága (1967), Montréal (1971), Leeds (1975), Koppenhága (1979), Utrecht (1983), Tallinn (1987), Aix-en-Provence (1991), Stockholm (1995) és San Francisco (1999).

A 15. International Congress of Phonetic Sciences helyszíne Barcelona, Spanyolország, 2003. augusztus 1-től 7-ig. A résztvevők száma ezúttal is ezer körül volt, a világ több, mint 52 országából. Magyarországról öten vettek részt az eseményen, de magyarok majdnem még egyszer ennyien voltak jelen. A kongresszusra több mint 1300 előadást jelentettek be, s ebből 855-öt tartott érdemesnek a nemzetközi bíráló bizottság (amelynek magyar tagja is volt) arra, hogy szóbeli vagy poszterelőadásként, illetve szimpózium részeként bekerüljön a programba. A közel 40%-os szűrés rendkívül magas, és azt mutatja, hogy – a lehetőségekhez mérten – a szervezők igyekeztek a legnívósabb előadásjavaslatokat elfogadni. A bírálati szempontok érintették a téma újdonságát, a kutatási cél és a módszer összehangoltságát, az eredmények jelentőségét, illetőleg, hogy mekkora érdeklődésre tarthat számot az adott téma. A kongresszus témaköreibe sorolt előadások arányait az 1. ábra diagramja szemlélteti (prod.=produkció, perc.=percepció, int.=intonáció, alk.=alkalmazott fonetika, f.=fonológia, más=egyéb témakörök, ld. a továbbiakban).

A „más” kategóriába azokat a témákat soroltuk, amelyeknek aránya nem érte el a 7%-ot (bár megjegyezzük, hogy a 7% mintegy 60 előadást jelent): idegennyelv-elsajátítás (5,53%), akusztika (4,9%), szociofonetika (4,8%), anyanyelv-elsajátítás (4,1%), fonetikaoktatás (2,19%), IPA-átírás (0,8%), fonetikatörténet (0,7%). Az alkalmazott nyelvészet témakörei a következők voltak: beszédszintézis (6,08%), mesterséges beszédfelismerés (3,74%), klinikai fonetika (3,74%) és bűnügyi fonetika (1,4%).

A plenáris előadások különféle témákkal foglalkoztak, mindegyiket egy rövidebb vitaelőadás követett. Az elhangzás sorrendjében: *A beszéd temporális jellemzőinek modellálása és észlelése* (Yoshinori Sagisaka, Japan); *A beszélő, a nyelvközösség, azonosság: tapasztalati és elméleti perspektívák a szociofonetikai variációkról* (Gerard J. Docherty, Anglia); *A nyelvspecifikus fonetikai kategóriák elsajátítása csecsemőkorban* (Janet F. Werker, Kanada); *Az artikulációs programozás összefüggései a beszédprodukcióban* (Anders Löfqvist, Svédország) és *Az univerzális fonológiai jegyek akusztikai és percepció tényezői* (Kenneth N. Stevens, USA).



1. ábra

A 15. Fonetikai Világkongresszus főbb témakörei (%-ban)

A szimpóziumok egy összetett kérdéskört több oldalról mutattak be, összesen 22 témában, ilyen volt például a neurofonetika, a fonetika/fonológia interfész, az artikulációs szintézis, a fonetikai terepmunka, a folyamatos beszéd modelljei, a hangsúly és ritmus, avagy a hangrendszerek tipológiája. Mind a szekcióelőadások, mind a posztereken bemutatott kutatási eredmények rendkívül érdekes és értékes kutatásokról számoltak be (számos egyéb mellett nagy sikert aratott a latin nyelv szintéziséről, a szinkronizálás beszédtempóra gyakorolt hatásának vizsgálatáról szóló előadás vagy a beszéd utánzásával, illetve az alkohol és a prozódia összefüggéseivel foglalkozó poszter). A két hazai előadás témája egyrészt a fonológiai és fonetikai kódolás sajátosságait elemezte (Gósy Mária), másrészt a SPECO beszédtanító programot ismertette (Vicsi Klára, Szaszák György).

A kongresszus 1200 oldalt meghaladó nyomtatott kötetei valamennyi elfogadott előadást tartalmazzák; az anyag CD-n is hozzáférhető. A Barcelonán-

ban bemutatott kutatások, a problémák, megoldási javaslatok, az eredmények és az igen termékeny viták a szó szoros értelmében távlatokat nyitottak a 21. század beszéd kutatásának.

Gósy Mária

8. Beszédkommunikációs és Beszédtechnológiai Konferencia (Eurospeech '03), Genf, 2003. szeptember 1-4.

Rendező: International Speech Communication Association (ISCA)

Ezt a konferenciát immáron több mint egy évtizede, kétévenként rendezik meg. A kétéves ciklus elég idő ahhoz, hogy új eredmények szülessenek ebben a speciális tudományágban, ugyanakkor elég sűrű is ahhoz, hogy az egyre gyorsuló technikai, technológiai fejlődéssel lépést tartson. Ez a világ legjelentősebb beszédtechnológiával foglalkozó konferenciája, amit a beküldött előadások nagy száma is mutat. A nemzetközi programbizottság a beérkezett 1081 előadásból 748-at fogadott el, a szűrés tehát igen szigorú volt. Az elfogadott előadásokat 27 szekcióban szóbeli előadásként, illetve 33 poszterként mutatták be a szerzők.

A konferencia témaköreinek száma jelentős. Ezekből próbálunk meg íze-
lített adni a következő felsorolással:

- automatikus beszéd felismerés, beszélő felismerés, beszélő azonosítás;
- a beszédjel digitális feldolgozása;
- akusztikai modellezés, nyelvi modellezés;
- gépi beszéd előállítás,
- beszéd–szöveg–fordítás–beszéd rendszerek, érzelemszintézis;
- a beszéd modellezése, szegmentálás, nyelvdetekció;
- fonológia, fonetika, prozódia, érzelmek a beszédben;
- dialógusok, beszélt nyelvi dialógusrendszerek, természetes nyelvi dialógusok feldolgozása;
- beszédprodukció és fiziológia, beszéd megértés;
- beszédinterfészek tervezése, a beszédtechnológia gyakorlati felhasználásai, multimodális rendszerek;
- beszéd kódolás, beszéd tisztítás; bűnügyi fonetika;
- gépi tanulás a beszéd és a nyelv modellezéséhez; beszéd szabványok.

A konferencián több, mint ezer fő vett részt, nemcsak Európából, de főleg a távol-keleti országokból. Magyarországról három előadás hangzott el, egy az MTA Nyelvtudományi Intézetéből, kettő pedig a BME Távközlési és Média-informatikai Tanszékéről. A konferencián nagy elismerést kapott Kempen Farkas beszélőgépezének működő másolata, amelyet az MTA Nyelvtudományi Intézetének munkatársai rekonstruáltak, és amelyen 1791-es manipulációs technikával lehetett beszédhangokat, szavakat megszólaltatni.

A következő Eurospeech konferenciának Lisszabon ad otthont 2005-ben.

Olaszy Gábor

Megakadások a spontán beszédben (Diss '03) konferencia 2003. szeptember 5-8., Göteborg (Svédország)

2003. szeptember 5-8. között került megrendezésre a *Megakadások a spontán beszédben* (Disfluency of Spontaneous Speech) című nemzetközi ISCA-workshop, mint a Eurospeech konferencia társult rendezvénye. A konferenciának ebben az évben a Göteborgi Egyetem adott otthont. Sorban ez már a harmadik összejövetel ebben a témában, a legelső 1999-ben, Berkleyben (USA) szervezték, míg a másodikra az Edinburgh-i Egyetem szervezésében került sor 2001-ben.

A spontán beszéd megakadásainak (vagyis a szünetek, hezitálások, grammatikai tévesztések, újraindítások, ismétlések stb. vizsgálata) témakörében elhangzott előadások több tudomány határterületére esnek, mint például az általános nyelvészet, fonetika, pszicholingvisztika, számítógépes nyelvészet, kognitív pszichológia vagy a beszédtechnológia.

A jelen ülészszaon 19 előadás hangzott el összesen hét témakörben:

- általános vonatkozások (a megakadások kutatásának elméleti háttere és megjelenésük lehetséges okai);
- produkció és percepció (a megakadások a beszélő/hallgató szemszögéből);
- megakadásjelenségek az anyanyelvi beszédben és az idegen nyelv elsajátításában;
- számítógépes nyelvészet (a spontán beszéd gépi feldolgozásának problémái);

- ismétlések és újraindítások a különböző nyelvekben;
- fonológia és prozódia a megakadások kutatásában;
- korpuszok és annotációk.

Az előadók és a résztvevők összesen 12 országból érkeztek, Magyarországot egy előadó képviselte.

Menyhárt Krisztina

A kötet szerzői

Név	Munkahely	e-levél
Gocsál Ákos	PTE Pollack Mihály Műszaki Főiskolai Kar, Pedagógia Tanszék	gocsal@falcon.pmmfk.pte.hu
Gósy Mária	MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium	gosy@nytud.hu
Grassegger, Hans	Universität Graz, Institut für Sprachwissenschaft	hans.grassegger@kfunigraz.ac.at
Huszár Ágnes	Veszprémi Egyetem, Pedagógiai Kutatóintézet	vepi99@almos.vein.hu
Kocsor András	MTA-SZTE Mesterséges Intelligencia Tanszéki Kutatócsoport	kocsor@inf.u-szeged.hu
Markó Alexandra	ELTE Fonetikai Tanszék	markoxa@ludens.elte.hu
Menyhárt Krisztina	MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium	menyhart@nytud.hu
Navracsics Judit	Veszprémi Egyetem, Alkalmazott Nyelvészeti Tanszék	navracsj@almos.vein.hu
Nikléczy Péter	MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium	nikleczy@nytud.hu
Olaszy Gábor	MTA Nyelvtudományi Intézet, Kempelen Farkas Beszédkutató Laboratórium	olaszy@nytud.hu
Szaszák György	BME – Távközlési és Telematikai Tanszék	szaszak@alpha.ttt.bme.hu
Tóth László	MTA-SZTE Mesterséges Intelligencia Tanszéki Kutatócsoport	tothl@inf.u-szeged.hu
Varga László	ELTE Angol Nyelvészeti Tanszék	vrgl@ludens.elte.hu
Vicsi Klára	BME – Távközlési és Telematikai Tanszék	vicsi@alpha.ttt.bme.hu



