

# **BESZÉDKUTATÁS 2000**

**Beszéd és társadalom**



**MTA Nyelvtudományi Intézet  
Budapest 2000**





# **BESZÉDKUTATÁS 2000**

## **Beszéd és társadalom**

A „Beszédkutatás 2000” tudományos ülészekon  
elhangzott előadások válogatott és átdolgozott  
tanulmányai

Szerkesztette:

**Gósy Mária**

Magyar Tudományos Akadémia Nyelvtudományi Intézet  
Budapest 2000

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztő:  
Menyhárt Krisztina

ISSN 1218 - 8727  
© MTA Nyelvtudományi Intézet

Felelős kiadó: Kiefer Ferenc igazgató

## Tartalom

<b>Gósy Mária:</b> A beszédszünetek kettős funkciója	1
<b>Kovács Magdolna:</b> Beszédhangok kontextusfüggő időviszonyai	15
<b>Olaszy Gábor:</b> Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre	26
<b>Gocsál Ákos:</b> A beszéd időviszonyai különböző életkorú személyeknél	39
<b>Menyhárt Krisztina:</b> A beszéd temporális sajátosságai kétnyelvűeknél (kisiskoláskortól időskorig)	51
<b>Huszár Ágnes:</b> A versengési elv a nyelvbotlások létrejöttében	63
<b>Vékás Domokos:</b> Magánhangzó-rendszerek elemzése informatizált nyelvjárási korpuszon	75
<b>Varga László:</b> Dallamcsere az intonációs frázis belsejében	87
<b>Szépe Judit:</b> A fonológiai kontrasztfenntartás elve afáziások közlésfolyamataiban	100
<b>Szuhaj Eszter:</b> Mozgássérült óvodások beszédészlelése és beszédmegértése	111
<b>Adamikné Jászó Anna:</b> Változott-e húsz év alatt a főiskolások kiejtése és olvasása?	124
<b>Gósy Mária – Nikléczy Péter:</b> Az idő változásának és a beszéd állandóságának paradoxona	132
<b>Szarvas Máté – Fegyő Tibor – Mihajlik Péter – Tatai Péter:</b> Eredmények a magyar nyelvű nagyszótáras és kapcsoltszavas beszéd felismerésben	144
<b>Németh Géza – Zainkó Csaba:</b> Statisztikai szövegelemzés automatikus felolvasáshoz	156
<b>Olaszy Gábor – Kiss Géza – Németh Géza – Olaszi Péter:</b> Profivox: a legkorszerűbb hazai beszéd szintetizátor	167
<b>L. Aczél Petra:</b> Működő szöveg a retorikában	180
<b>A magyar kísérleti beszéd kutatás története – állandó kiállítás</b>	190



# A BESZÉDSZÜNETEK KETTŐS FUNKCIÓJA

Gósy Mária  
MTA Nyelvtudományi Intézet

## Bevezetés

Beszédünk szünetekkel tarkított. A hagyományos fonetika a beszédszünet megjelenésére két okot sorol fel: a) a lélegzetvételt (oxigén pótlása) és b) az értelmi tagolást (Szende 1976). A szakirodalom számos további szünetfajtáról is említést tesz, mint például a gondolkodási szünet vagy a hatásszünet, amelyek más-más funkciót töltenek be a beszédprodukcióban. A beszéd tartalmas, azaz értelemmel bíró jeleinek folyamatosságát egyéb jelenségek is megtörik, amelyek tradicionálisan nem tekintendők szünetnek, de a tartalmas jelek szempontjából jelkimaradások. Ilyenek a nyúlások, bizonyos ismétlések avagy az újramezdések, illetőleg a nyelvbötlések önkorrekciói (Laver 1995, 535). A szünet is lehet néma vagy kitöltött, azaz hezitálás. Újabban mindezek miatt a spontán beszéd vizsgálatában nem a *szünet* szót használják a fentebb említett 'jelkimaradások' megnevezésére, hanem a jelenségkört mint a folyamatosság megakadásait értelmezik (angol terminussal: *dysfluency*). A folyamatos beszédben olyan jelkimaradások is előfordulnak, amelyek egy adott beszéd artikulációs mozgássorának részeként funkcionálnak, s így – bár a szó szoros értelmében jelkimaradások – nem minősülnek szünetnek. Ilyenek a zöngétlen zárhangokat megelőző ún. néma fázisok (ezek a beszédhang részét képező „szünetek” időtartamban a valós szünetekkel azonos tartamúak is lehetnek). Kérdéses a junktúrák szerepe, amelyeket hol szünetként, hol mint artikulációs jelenséget tárgyalnak (Szende 1976). A beszédszüneteket értelemszerűen elsősorban a produkció szempontjából elemezték, elemzik, még akkor is, ha a háttérükben bizonyos pszicholingvisztikailag magyarázható eredetet tételeztek fel (vö. Hieke–Kowal–O'Connell 1982). Goldman-Eisler ugyan már a hatvanas években utal a szüneteknek a beszédmegértésben betöltött szerepére (1968), de a részletekbe menő vizsgálatokkal adós a szakirodalom. Hegedűs La-



jos a beszédtempó elemzések kapcsán ír a szünetről, voltaképpen az észlelés szempontjából. Kijelenti, hogy a szüneteknek fontos szerepük van a tempóélmény kialakításában (1957).

A szünet hosszú időn keresztül elfogadott definíciója módosításra szorul, mivel a konvencionális meghatározásokból hiányzik egyfelől a spontán beszédben a tervezés és a kivitelezés „paradoxonából” adódó szünet megnevezése (Gósy 1999), másfelől pedig a szünet funkciójának pontos leírása. Mindezeket figyelembe véve a módosított definíció a következőkben fogalmazható meg. A beszéd folyamatban jelentkező szünet olyan kismértékben akaratlagos beszédkimaradás, amely néma vagy jellel kitöltött, de független a beszédhang képzésétől. Funkcióját tekintve a beszédprodukciónban 1. biztosítja az artikulációt lehetővé tevő légáramot, 2. elősegíti a közlés értelmi tagolását, 3. a beszédtervezés során az ún. ellentmondások, téves utak stb. feloldására szolgál, 4. a mentális lexikonban történő keresési idő kitöltését biztosítja, illetőleg lehetőséget nyújt a nyelvi kódolás módosítására. Funkciója a beszédmegértésben: 1. az elhangzottak könnyebb feldolgozása, 2. az entrópia csökkentése és 3. a megértés és az értelmezés működési folyamatainak biztosítása.

Szükségesnek tűnik a beszédszünet minimum tartamának megállapítása, amely azonban a kezdetektől vita tárgyát képezi. Eleinte a 250 ms-ot tekintették alsó határnak, amelyet aztán csökkentettek egészen a 30 ms-os értékig. Néma szünetekként legelfogadottabbak a 130 ms, illetőleg 150 ms fölötti tartamúak (Misono–Kiritani 1990); bár újra és újra megjelenik a 200 ms körüli érték (Laver 1995). Még a nyolcvanas évek elején is módszertani nehézségek gátolták a kutatókat a szünetek határainak meghatározásában és az időtartamuk kimérésében. A technikai lehetőségek fejlődésével újabb problémák álltak elő. A jelelemzők egészen rövid tartamú jelkimaradásokat is tudtak elemezni, bőven az emberi percepciós mechanizmus határain túl. Ez felvetette a beszédszünetek percepciós vizsgálatának szükségességét. Saját vizsgálatinkban a néma szünet tartamának alsó határául azt az értéket tekintjük, amelynél az átlagos hallgatók többsége, mintegy 60%-uk már képes a szünetet észlelni. A jelen vizsgálati anyagban ez a 80 ms-os időtartam volt.



Közel ötven évvel ezelőtt Mahl volt az első, aki úttörő munkájában kétféle kitöltött szünettípust különített el (1956). Az elsőt *ah* jelenségnek, a másodikat nem-*ah* jelenségnek nevezte. Az elsőbe azok a jelenségek tartoznak, amelyeket az *ah*, *er* és *um* betűkapcsolatokkal jelenít meg írásban az angol irodalom. Ezeknek a magyar ekvivalensei leggyakrabban az *öö* és *m* hanghatású hezitálások. Mahl definíciója szerint az *ah* típusú hezitálások a bizonytalanságot tükrözik, amely után bekövetkezik a nyelvi forma megjelenése. A nem-*ah* kategória olyan jelenségeket fog össze, mint az ismétlések, a mondat cseréje, egyéb akadozások stb.

A modern beszédkutatás olyan kérdésekre igyekszik választ kapni, hogy vajon a beszéd folyamatosságát megakasztó jelenségek milyen gyakoriak, milyen arányban fordulnak elő a folyamatos artikulációban, mi idézi elő őket, melyek és mikor észlelhetők. Különböző vizsgálatok igazolták, hogy a néma szünetek megjelenése a leggyakoribb. Korábbi vizsgálataink szerint a spontán beszéd teljes időtartamának átlagosan 20%-a csöndes és kitöltött szünet, az összes többi, a folyamatosságot csökkentő jelenség ehhez az értékhez adódik hozzá. A kitöltött szünetek aránya lényegesen kisebb, 1-4% körüli érték (Gósy 1999). Misono és Kiritani párbeszédés szövegekben 32%-ban találtak szünetet, közülük 7%-nyi volt kitöltött szünet (1990). A beszédészünetek gyakorisága (száma) és hossza (tartama) több tényező függvénye. Döntően meghatározó a beszélő személy (veleszületett tulajdonságai, gyakorlottsága a beszédben, pszichés alkata stb.), továbbá a téma, a beszédhelyzet (cél, hallgatóság stb.), a beszédstílus (vö. Duez 1982) és még egyéb tényezők is. Mint a beszéd számos területén, itt is előfordulnak egészen kivételes esetek: a szinte szünettartás nélkül beszélő avagy a rendkívül sok szünetet közbeiktató beszélő vagy a csaknem folyamatosan hezitáló személy. Harminckét horvát beszélő összesen 12 párbeszédés rádióanyagában elemezték az önkorrekciót a beszéd folyamán, vagyis azokat az ismétléses jelenségeket, amelyek a beszéd folyamatosságát megakasztották (Horga 1997). Azt találták, hogy az önkorrekciók átlagosan 42 másodpercenként fordultak elő (s ehhez járultak hozzá a néma és a kitöltött beszédészünetek!).

Mindezekből nyilvánvalóan levonható az a következtetés, hogy a folyamatos beszéd nem is olyan folyamatos. Mennyit észlel a hallgató

ezekből a folyamatosságot megtörő jelenségekből? Igazolható-e kísérletileg a beszédszüneteknek a beszédmegértésben betöltött, a korábbiakban mint-egy hipotézisként ismertetett funkciója?

A következőkben bemutatandó kísérlettel a beszédszünetnek a teljes beszéd folyamatban, tehát a beszédprodukción, a beszédészlelésben és a beszédmegértésben betöltött funkcióját tárgyaljuk.

### **Anyag és módszer**

Három fiatal beszélő (egy nő és két férfi) spontán beszédéből választottunk ki egy-egy részletet a kísérletekhez. A kiválasztás szempontja mindenekelőtt a beszédtempó különbözősége volt, ami részben az artikulációs, részben a beszédtempóból adódott. Egyikük ún. gyors, másikuk átlagos, harmadikuk kissé lassú beszélő. A szövegeket magnetofonszalagra rögzítettük, elsőként a normál, másodikként a gyors, harmadikként a lassú tempójút. A CSL 4300B típusú digitális jelelemző segítségével vizsgáltuk a szövegeket: meghatároztuk a tempóértékeket, a szüneteket, a szünetek tartamát és a szavak számát. Az 1. táblázat a tesztszövegek adatait összegzi (AT = artikulációs tempó, BT = beszédtempó, a szünet a néma és a kitöltött szüneteket egyaránt jelenti).

A beszédszünetek száma mindhárom beszélőnél ugyanúgy aránylik a teljes jelidőhöz. Az időtartamuk is hasonló, az első szövegben a teljes idő 22,6%-a, a másodikban 21,3%-a, a harmadikban pedig 16,66%-a. A táblázatban a szó/percben feltüntetett tempó a beszédtempót mutatja. Az artikulációs tempó értéke az első szövegben 162,54 szó/perc, a másodikban 213,63 szó/perc, a harmadikban pedig 134,51 szó/perc.

A szünetészlelési kísérletben a kísérleti személyek (20 ép halló fiatal nő és férfi, életkoruk 22 és 36 közötti) feladata az volt, hogy az azonosított szünetet a tesztalapon egy függőleges vonallal jelöljék. Kétszer hallgathatták meg a szövegeket; nem határoztuk meg, hogy mikor kezdjék a jelölést. Elsőként a közepes, majd a gyors, végül a lassú tempójút hallgatták.



I. táblázat: A tesztszövegek mért adatai

Szöveg	Elemzett adatok					
	AT	BT	Tempó	Szünet		Jelidő
	(hang/mp)	(hang/mp)	(szó/perc)	idő (mp)	szám (db)	(mp)
I.	15,29	11,83	125,7	13,9	24	47,61
II.	15,58	12,26	168,1	8,05	18	29,77
III.	11,47	9,56	112,1	11,95	26	59,77

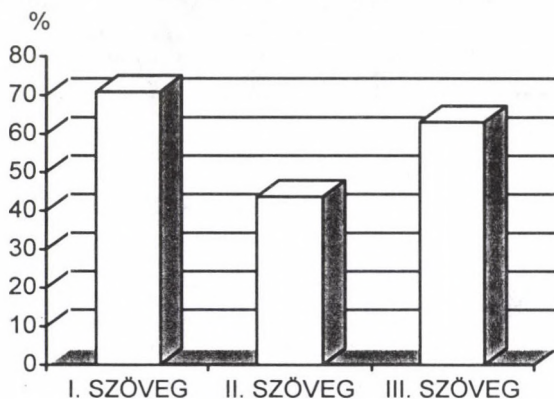
A beszédmegértést vizsgáló kísérlethez a leggyorsabb és a leglassúbb tempójú szöveget választottuk ki. A digitális jelelemző segítségével kiiktattuk az összes szünetet úgy, hogy annak hallható nyoma ne maradjon. 20-20 különböző, fiatal ép halló kísérleti személy (akik természetesen nem voltak azonosak az észlelési tesztben szereplőkkel) hallgatta le a kétféle anyagot (egyik csoport az eredeti szövegeket, a másik a szünet nélkülieket). A résztvevők életkora 22 és 35 között szórt. A kísérleti személyek feladata az volt, hogy egyszeri meghallgatás után írják le a hallottak tartalmát amilyen pontosan csak tudják.

## Eredmények

### *I. A szünetészlelési kísérlet*

Feltételeztük a korábbi hasonló vizsgálatok eredményei alapján, hogy a kísérleti személyek bizonyos százalékban lesznek képesek a szünetek azonosítására, illetőleg hogy olyan helyeken is jelölnek majd szünetet, ahol objektíve nincs jelkimaradás vagy hezitálás (Menyhárt 1998). Az eredményeket összesítettük, és a következő kérdésekre kaptunk választ: (i) mi az összefüggés a szünetek időtartama és észlelése között, (ii) befolyásolja-e az észlelési pontosságot a szöveg artikulációs, illetőleg beszédtempója, (iii) a női és férfi kísérleti személyek eltérően észlelték-e a szüneteket és (iv) mik a téves szünetészlelés sajátosságai.

A kísérleti személyek az összes szünet közel 60%-át voltak képesek helyesen azonosítani. Az egyes szövegekben megjelenő szünetek észlelése eltérést mutat. Az eltérés szignifikánsan különböző ( $F(2, 54) = 17,238, p < 0,0001$ ). Az 1. ábra a helyes szünetészlelés átlagait mutatja.



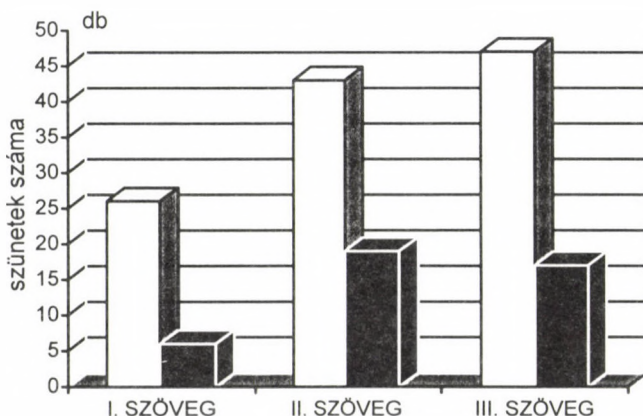
1. ábra

A különböző szövegek szüneteinek helyes észlelése

A tempó egyértelműen hatással van a szünetek percepciójára: a gyors szövegben a helyes szünetazonosítás még az 50%-ot sem éri el. Meglepő volt ugyanakkor, hogy nem az artikulációs tempóját tekintve leglassúbb beszélő szövegében azonosították legjobban a szüneteket, hanem a közepes tempójúében. Ennek okára az adatok egyértelműen rámutattak: a szünetek átlagidőtartama a közepes beszélőnél a leghosszabb, tehát a felismerésük itt a legkönnyebb. A szünetek átlagértéke 551 ms, míg a lassú beszélőnél 411,7 ms. A gyors beszélőnél ez utóbbihoz hasonló adatokat kaptunk, az átlag 413,7 ms. További ok lehet a szünetek helye: a lassú beszélő több szünetet tart szerkezet-határtól függetlenül. Ennélfogva a hallgató előfeltevése kevésbé találkozik a valósággal.

A nők és a férfiak helyes szünetazonosításaiban 10% körüli a különbség; az első szövegben az eltérés 11%, a másodikban 7%, a har-

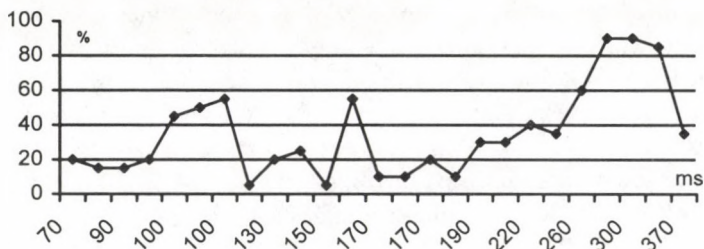
madikban 3% (mindenütt a nők észleltek több szünetet helyesen). E minimális különbségek ellenére a helyes válaszok arányába a nemek eredménye között nincs szignifikáns különbség. Különbséget találtunk ugyanakkor a téves válaszok elemzésében (2. ábra). Az eltérés a férfiak és a nők adataiban az első és a második szöveg szünetészlelésében statisztikailag is szignifikáns ( $F(1,54 = 10,205, p < 0,002)$ ). A nők több téves választ adtak: általában több szünetet észleltek, tagoltabbnak érezték a szövegeket.



2. ábra

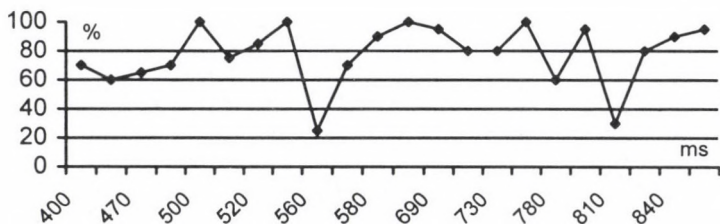
A szünetek téves azonosításai a három szövegben (a fehér oszlop a nők, a fekete a férfiak adatait mutatja)

A szünetek időtartama és észlelése igen erős korrelációt mutat, az elvártaknak megfelelően ( $r_p = 0,714, p < 0,0001$ ). Minél hosszabb a szünet, annál jobban észlelhető, bár egyéb körülmények is befolyásolják a pontos percepciót. Éppen emiatt nem szorosan egy „egyenes” mentén helyezkednek el az adatok, de erősen megközelítik azt. A 3. és 4. ábra szemlélteti a szünetek időtartamának és a helyes azonosításuknak az összefüggését.



3. ábra

A szünetek időtartamának (370 ms-ig) és észlelésének összefüggése

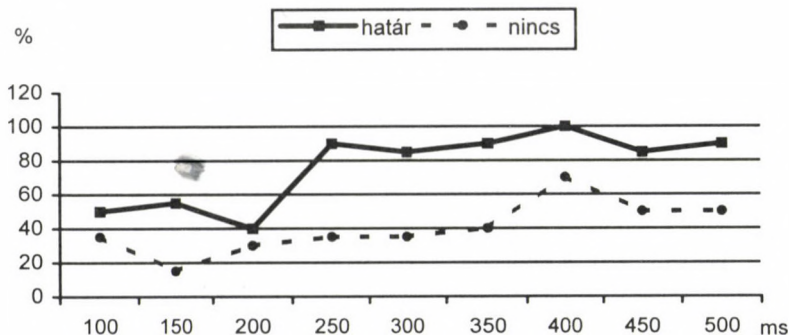


4. ábra

A szünetek időtartamának (400 és 850 ms között) és észlelésének összefüggése

A szünetek észlelését az időtartamon túl befolyásolja az is, hogy szerkezethatáron vagy attól függetlenül jelennek-e meg a szövegben. Anyagunkban azt láttuk, hogy a beszélők nagyjából azonos arányban tartottak szünetet e különböző típusú helyeken. Az észlelés azonban nem független az előfeltevésektől; a kísérleti személyek sokkal nagyobb mértékben azonosították helyesen a szerkezethatárokon megjelenő szüneteket: 75% szemben a 36,8%-kal (5. ábra).



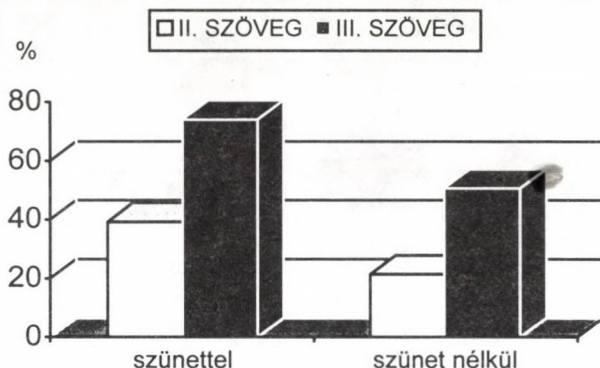


5. ábra  
Szünetek észlelése szerkezethatáron és attól függetlenül

## II. A beszédmegértési eredmények

A kísérletben résztvevők tartalomleírásaiban a szavak számát, valamint az alapgondolatok számát vettük tekintetbe. Ezeket viszonyítottuk az eredeti szövegek mért, illetőleg megállapított adataihoz. E két-féle elemzést alkalmazzák a hallott szövegek reprodukálásainak nyelvi vizsgálatában objektív adatok nyerésére. A szavak száma utal egyben a terjedelemre is, vagyis arra, hogy mi mindent képes a kísérleti személy elmondani a megértett szövegből. Másfelől jelzi azt is, hogy mennyire volt képes jól formált mondatok, illetőleg stilisztikailag igényes fogalmazás létrehozására. Ez is a szövegmegértés biztonságára és pontosságára vezethető vissza. Az alapgondolatok számát a kísérleti szövegekben található új információ alapján határoztuk meg. Minden egyes új információt alapgondolatnak tekintettünk, függetlenül attól, hogy azt az eredeti beszélő hány szóval vagy mondattal fejezte ki.

Az eredmények erősen szignifikáns különbségről tanúskodnak a szünetek jelenlététől függően (a szavak számának elemzéséből:  $t(19) = 5,936$ ,  $p < 0,0001$ ), illetőleg az alapgondolat elemzéséből:  $t(19) = 8,320$ ,  $p < 0,0001$ ). A 6. ábra a szavak számának alakulását mutatja a két szövegben.

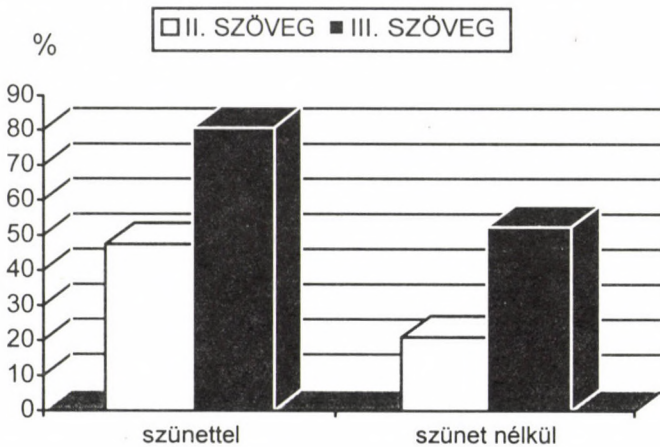


6. ábra

A szavak számának aránya a két szövegben (szünettel és a nélkül)

A szavak számának és az alap gondolatok arányának együttes vizsgálata a szövegek reprodukciójában jól jelzi a beszéd megértés szintjét nyelvtől függetlenül (vö. Wingfield–Tun 1995). A reprodukciókban jelentkező esetleges hibák további információhordozók, amelyek révén még élesebb különbségek mutatkozhatnak a feldolgozásban. Az eredeti szövegek információit tekintve nem találtunk hibát a tartalomelmondásokban akkor, ha voltak szünetek; míg jellegzetes hibák adódtak a szünet nélküli szövegek esetében. A 7. ábra az alap gondolatok megjelenésének arányát mutatja a kísérleti személyek írásaiban a szünetet tartalmazó és a nélküli szövegek alapján.

Az eredmények szembevetőek, és egyértelműen mutatják az óriási különbségeket. A szünetet tartalmazó szövegek reprodukálásában téves gondolatot alig találtunk, míg a szünet nélküliekben meglehetősen sokat. Az utóbbiakban számos többlet-információt is közöltek a résztvevők, amire a szünet nélküli szövegek lehallgatóinál alig akadt példa. Mindezek alapján kimondható, hogy a hallgatónak éppúgy szüksége van a beszéd feldolgozásában a szünetekre, mint a beszélőnek a beszédprodukciója során.



7. ábra

A felidézett alapgondolatok aránya szünetet tartalmazó és szünet nélküli szövegekben

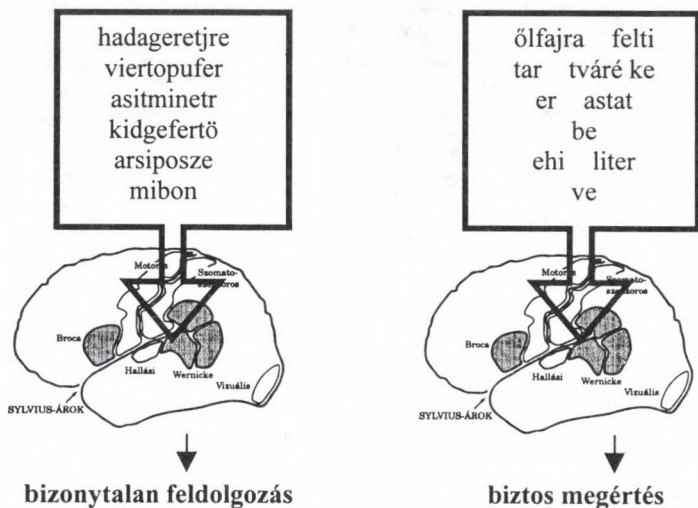
A szünetnek tehát kettős funkciója van: akár beszélők, akár hallgatók vagyunk, a (tartalmas) jelek kimaradásai nélkülözhetetlenek a kommunikációinkban. A különbség csupán a funkcióban van. A produkcióban a tervezés és a kivitelezés diszharmoniját oldják fel, a beszédmegértésben pedig csökkentik a tartalmi bizonytalanságot, továbbá a biztos követhetőséget, a pontos feldolgozást teszik lehetővé (vö. 8. ábra szemléltetését). A szünetek ekkor mintegy „tágítják” a munkamemóriát azzal, hogy több időt adnak a műveletek elvégzésére. A több idő pedig itt is jobb eredményt hoz.

### Következtetések

A beszélő közel azonos arányban tart szünetet szerkezethatáron és egyéb helyeken, a hallgató azonban mintegy kétszer annyi szünetet vár szerkezethatáron. Anyagunkban a beszélők 70 és 1200 ms közötti időtartamú szüneteket tartottak, az átlagérték jellemző a beszélőre, de



független a beszéd artikulációs tempójától. Ez a tény könnyen magyarázható azzal, hogy a beszédtervezés és a beszédkivitelezés összehangolatlansága egyénenként különböző lehet.



8. ábra

A szünetet tartalmazó és nem tartalmazó spontán beszéd feldolgozásának különbsége az agyféltekei központok szempontjából

A spontán beszédben való jártasság, a pszichés alkat, az adott beszédhelyzet, valamint a gondolat és annak nyelvi kifejezésére vonatkozó egyéni jellemzők – mint legfontosabbak – erőteljesen befolyásolják a beszédet megakasztó jelenségeket, így a szüneteket is. A kommunikációban részt vevő hallgatók valamivel egységesebbek; a nyelvi tudatosság és a nyelvi tapasztalat nagyobb mértékben idéz elő hasonló működéseket. Nem véletlen, hogy a beszédsszünetek mintegy kétharmadát észlelik az emberek, függetlenül az aktuális tartalomtól, tempótól vagy feladattól. Általánosan érvényesülő tény az is, hogy a szünet időtartama és az észlelése – a kivételektől függetlenül – igen szoros összefüggést mutat.

A hallgatók téves szünetazonosításai két okra vezethetők vissza: vagy úgy jönnek létre, hogy nem észlelik a szünetet, vagy szünetet észlelnek ott, ahol nincs. A hallgatók ugyan hasonló mértékben azonosan észlelik vagy nem észlelik a szünetet, döntéseiben azonban jellegzetes különbségeket is találunk. Az eltérések sokféleképpen magyarázhatók (egyéni észlelési működés, előjelzések dominancia, tesztelési stratégia stb.) A szünetészlelésben döntő az objektív időtartam, de számos más tényező is befolyásolja a percepciót, például az adott szemantika, esetleg a szerkezet, az asszociációs lehetőségek, az aktuális artikuláció avagy szupraszegmentális jelenségek. Érvényesülésüket minden egyes esetben külön-külön kell elemezni, valamennyi lehetőség figyelembevételével.

A hallgató, miközben az elhangzó beszédet feldolgozza, voltaképpen beleéli magát a beszélő helyzetébe (empátia), és ez teszi lehetővé számára a biztos megértést. A szünetek egyfelől ennek a beleérzésnek a megvalósítását teszik lehetővé, másfelől időt biztosítanak a percepció mechanizmus korrekciós folyamatainak működéséhez. A beszéd megértéséhez – mint láttuk – nélkülözhetetlenek a szünetek. A beszéd megértés pontossága drámaian csökken, ha kiiktatjuk a jelkimaradásokat. Noha a beszéd közben tartott szünetek egyik legfontosabb oka a tagolás, az értelmezés elősegítése lenne, éppen a spontán beszédben találjuk erre a legkevesebb példát. Természetesen minél gyakorlottabb a beszélő, annál ritkábban áll elő a tervezés és a kivitelezés diszharmoniója, így fokozottabban tud figyelni a tagolás akusztikai megvalósítására a könnyebb érthetőség érdekében. A kevésbé gyakorlott, szorongó vagy fáradt beszélő a beszédtervezés valamennyi szintjén nehézséggel találja magát szembe. Bizonytalan a mentális lexikon aktiválásában, a szemantikai és a szintaktikai egység létrehozásában, a kívánt egység kiválasztásában, így aztán a fonológiai, illetőleg a fonetikai tervezés sem tökéletes. A bizonytalanságok feloldásai a csendes vagy kitöltött szünetek, az akadozások. Az elhangzó szöveg tehát szétesik, nemhogy az elvárt tagolást nem segítik szünetek, hanem a beszédjel igen gyakori megtörései már alig teszik lehetővé a hallgatónak a megértést, az értelmezést.

### Irodalom

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

Goldman-Eisler, F.: *Psycholinguistics. Experiments in Spontaneous Speech.* Academic Press. London 1968.

Gósy Mária: *Pszicholingvisztika.* Corvina Kiadó. Budapest 1999.

Hegedűs Lajos: Beszédtempó-elemzések. *Magyar Nyelvőr* LXXXI, 1957, 223-227.

Hieke, A.E. – Kowal, S. – O'Connell, D.C.: The trouble with „articulatory” pauses. *Language and Speech* 26. 1983, 203-219.

Horga, D.: Samoispjavljanje u govornoj proizvodnji. *Suvremena lingvistika* 23/1-2. 1997, 91-105.

Laver, J.: *Principles of Phonetics.* Cambridge University Press. Cambridge 1995.

Mahl, G.F.: Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology* 53. 1956, 1-15.

Menyhárt Krisztina: Nyelvi meghatározottság a beszédzúnetek észlelésében. In: *Beszédkutatás '98.* Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1998, 47-57.

Misono, Y. – Kiritani, S.: The distribution pattern of pauses in lecture-style speech. *Logopedics and Phoniatrics* 2. 1990, 110-113.

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



# BESZÉDHANGOK KONTEXTUSFÜGGŐ IDŐVISZONYAI

**Kovács Magdolna**  
**Debreceni Egyetem, Általános és Alkalmazott**  
**Nyelvészeti Tanszék**

## **Bevezetés**

A már lassan egy évszázados hagyományokkal rendelkező eszköz-fonetikai kísérletek révén tudjuk, hogy a megnyilatkozások fizikai időtartam-viszonyai számos különféle tényező kölcsönhatásában formálódnak. A magánhangzók időtartamát egyfelől fonetikai kondicionáló tényezők – magának a magánhangzónak és az azt követő mássalhangzónak a fonetikai alkata, a hangsor terjedelme, a hangzónak a hangsor szerkezetében elfoglalt helye, a szó- és mondathangsúly stb. – alakítják. Másfelől az időtartamok jelentősen módosulhatnak a beszélő egyéni beszédjellegzetességeit tükröző és a kommunikáció körülményei szerint akaratlagosan változtatható tényezők, így a beszédtempó, a beszédstílus, az érzelmek, a társalgás átlagos hangerő szintje stb. függvényében.

A jelen dolgozat vizsgálódásának tárgyát az első csoportba tartozó kontextusfüggő időtartam-változások képezik. Kiindulópontként tekintünk át, hogy a szakirodalmi megfigyelések – Meyer–Gombocz (1909), Szabó (1938), Magdics (1966), Kassai (1979, 1982), Olaszky (1994) – alapján a magyar magánhangzók kontextusfüggő időtartam-változásaiban mely tényezők a meghatározóak, és milyen összefüggések rajzolódnak ki. Az alábbi listában a tényezőket az indukált időtartam-változás növekvő mértéke szerint rendeztük; a forrást azokban az esetekben szerepeltettük, amikor az álláspontok alapvetően eltértek.

## **1. Mássalhangzó környezet**

✓ A magánhangzók rövidebbek zöngétlen, mint zöngés obstruensek előtt. A magánhangzók a legrövidebbek a zárhangok szomszédságában, leghosszabbak a likvidák előtt.

✓ A magánhangzók időtartama egyszerű mássalhangzók előtt rövidebb, mint mássalhangzó-kapcsolatok előtt.

? Meyer–Gombocz (1909) szerint a magánhangzók a követő mássalhangzó hosszúságának hatására rövidülnek. Szabó (1938) úgy találja, hogy csak a fonológialag hosszú magánhangzók rövidülnek, a röviddek időtartama nő. Kassai (1979) álláspontja annyiban tér el, hogy a rövid magánhangzók időtartamában hangsúlyos helyzetben növekedést, hangsúlytalanban csökkenést állapít meg.

## **2. Hangsúly**

? Magdics (1966) és Olasz (1994) szerint mind a fonológialag rövid, mind a hosszú hangsúlyos magánhangzók körülbelül 20%-kal hosszabbak, mint a hangsúlytalanok. Kassai (1979) azt állapítja meg, hogy a szó hangsúlyos szótagjában a rövid/hosszú arány 19%-kal nagyobb, mint hangsúlytalanban, mivel a rövid magánhangzók nyúlnak, a hosszúak pedig változatlanok maradnak.

## **3. Hangsorban elfoglalt hely**

✓ A magánhangzók a hangsor végén a leghosszabbak.

? A hangsorvégi nyílt szótagok magánhangzója jelentős, míg a zárt szótagoké következetes, de jelentéktelen nyúlást mutat (Hockey–Fagyal 1999). Magdics (1966) és Kassai (1979) zárt szótagokat ebben a pozícióban nem vizsgálnak, Olasz (1994) pedig nem állapít meg különbséget.

? Magdics (1966) szerint a magánhangzók hangsor eleji helyzetben rövidebbek, Kassai (1979) és Olasz (1994) szerint hosszabbak, mint a hangsor belsejében.

Az alábbiakban ismertetendő kísérletben elsősorban a vitatott összefüggések értékeléséhez kívántunk adalékokkal szolgálni. Azt vizsgáltuk, hogyan, mennyire következetesen változik a fonológialag rövid magyar magánhangzók időtartama a kontextus e tényezőinek függvényében.

## **Anyag és módszer**

A kísérlet alapjául szolgáló fonetikai adatbázis – 900 mondat 1220 különböző szóalakja – egy női adatközlő (budapesti középkorú tanár) által felolvasott neutrális kijelentő mondatokból állt elő. Az adatközlő

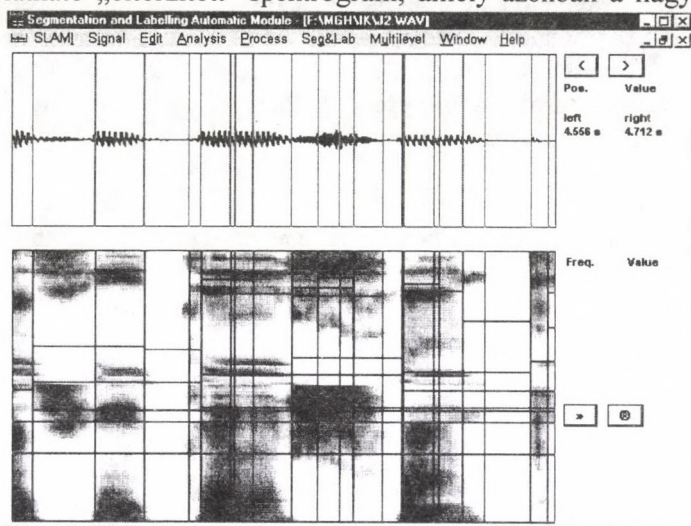


minden mondatot háromszor megismételt. Ennek célja kettős volt: egyrészt ezzel mérsékelhető a bemondás esetlegessége; másfelől csökkenthető az egyes mondatok felolvasásában óhatatlanul adódó beszédtempóbeli különbségek. A CV<sub>rövid</sub>C szekvenciák átlagos időtartama alapján becsülhető artikulációs tempó 14 beszédhang/ másodperc volt, ami a mai magyar köznyelvben normál értéknek tekinthető (vö.: Gósy 1991).

A teljes vizsgált korpusz 1600 ismételt bemondásra és környezetre átlagolt szegmentum-sorozatot tartalmazott; például a [köt] CVC kapcsolatot a *kötés*, *kötet*; a [mez] a *hómezők*, *lemezért* szavak megfelelő hangjainak átlagából állt elő. A vizsgált szekvenciák C1[i/ε/ø]C2 felépítésűek voltak, ahol C2 a) zöngés/ zöngétlen rész- vagy zárhang, illetve nazális; b) hosszú mássalhangzó, c) egyszerű mássalhangzó ↔ mássalhangzó kapcsolat (azaz a magánhangzó a szóban jobbról nyílt/zárt szótagban fordult elő); s különböztek a többségében 2-3 szótagú szavak szerkezetében elfoglalt helyük vagy mondatbeli pozíciójuk szerint (a mondat első, nem első és nem utolsó, valamint utolsó szekvenciája). A magánhangzók megválasztását az indokolta, hogy a korábbi szakirodalmi adatok alapján leginkább eltérően viselkedő hangzókat akartuk vizsgálni, így a célszótagokban a legrövidebb speciális időtartamú [i] és a leghosszabbaknak bizonyult [ε, ø] szerepeltek (Kassai 1982). A van Santen (1992) által „*single factor independence of the vowel factor*” (magánhangzó függetlenség) alapelvek nevezett összefüggés alapján azt feltételeztük, hogy a vizsgálat eredményei valamennyi fonológiailag rövid magánhangzó időtartam-változásaira jellemző összefüggéseket tárnak fel. A „magánhangzó függetlenségnek” alapelve ugyanis azt mondja ki, hogy a magánhangzók minden szövegben ugyanazon, a specifikus időtartamuknak megfelelő sorrendben fordulnak elő, ezáltal a kontextus tényezőinek hatására azonos irányú változást mutatnak.

A következőket egységes szegmentum-határok megállapításához egy fél-automatikus szegmentálási segédeszközt vettünk igénybe, a Segmentation and Labelling Automatic Module (SLAM) elnevezésű szoftvert (Cosi 1993). A program három lépcsőben dolgozza fel a 16000 Hz-en digitalizált jelet. Az első fázisban szűrési eljárásokkal igyekszik az emberi hallószerv feltételezett működésének megfelelően

átalakítani a hanghullámot. Az eljárás eredménye egy az 1. ábrán látható „eltorzított” spektrogram, amely azonban a hagyományos re-



1. ábra

Az *idősek* hangsor automatikusan kijelölt szegmentum-határa

gisztrátumoknál jobban kiemeli a szegmentumok határait. A második fázisban a SLAM megkeresi a hangsor akusztikai töréspontjait (az 1. ábra oscillogramján látható vonalak). Az eljárás, az automatikus szegmentáló programokra általában jellemző módon, a konvencionális szegmentum-határoknál több töréspontot jelez. A végső feldolgozási szakaszban a hasonló szeleteket klaszteranalízis segítségével csoportosítja, s ún. dendrogram formájában jeleníti meg. A program természetesen lehetőséget nyújt az állandó auditív ellenőrzésre is.

## Kísérleti eredmények

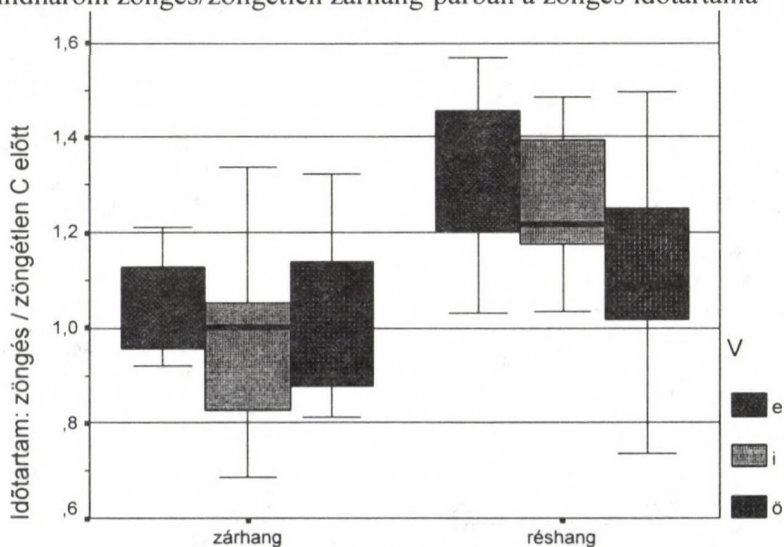
### 1. A mássalhangzó környezet hatása

A mássalhangzó környezet hatását mondat belseji 2-3 szótagú szavakban előforduló kvázi-minimális – C1-nek csak képzési módjában azonos – C1VC2 kapcsolatok alapján elemeztük, például: a *Kiszőköm ötödikén hozzád, A csőszöket kedvelik a Tiszán ↔ Kiszögez valamit*

*Áginak, Növöget a rózsá tövise; Megezem az édeskés csemegett, Min-  
dent megeszik a piszok ↔ Régen tegezett minden diákot, Nálunk  
nincs fedezet erre* mondatok félkövérrel jelölt szekvenciái.

Az összevetések alapján a követő rés- és zárhangok hatásában általános tendenciaként kitapinthatók a mássalhangzó zöngésségére és képzési módjára vonatkozó korábbi megállapítások: a magánhangzók rövidebbek zöngétlenek és zárhangok, mint zöngések, illetve réshangok előtt.

Ugyanakkor a 2. ábra, amely a zöngés/zöngétlen környezetet jellemző időtartam-arányokat foglalja össze, azt tükrözi, hogy a követő mássalhangzó zöngésségének hatása lényegesen nagyobb volt rés-, mint a zárhangok esetében. A réshangok előtti 22%-os ( $p < 0,0001$ ) időtartam-növekedéssel szemben a magánhangzók zöngés/zöngétlen zárhangok előtt átlagosan mindössze 3%-os, és jóval kevésbé következetesen érvényesülő változást mutattak. Megvizsgáltuk, hogyan alakulnak a zárhangok időtartamai ugyanezen fonetikai helyzetben. Mindhárom zöngés/zöngétlen zárhang-párban a zöngés időtartama



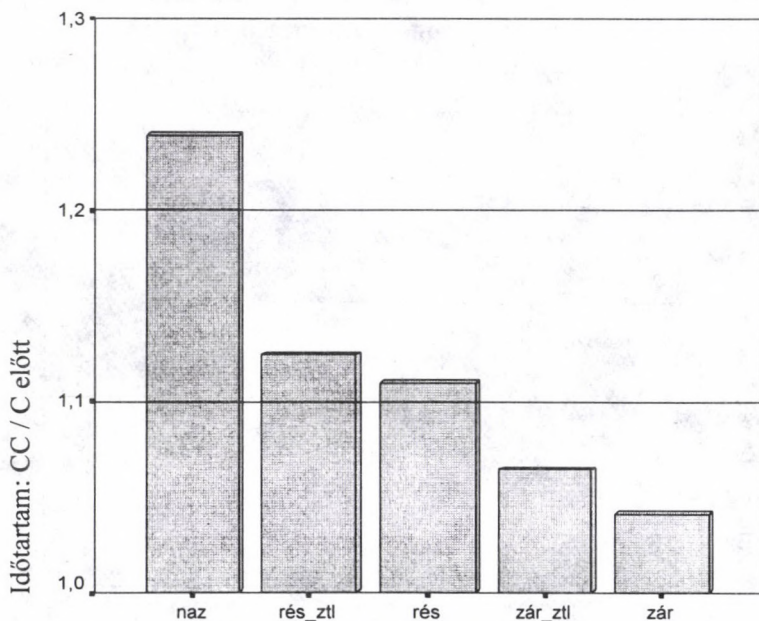
2. ábra

A követő mássalhangzó zöngésségének hatása mondat belseji helyzetben



volt a rövidebb ( $[t(59) = 4,107, p < 0,0001]$ ); az egyes mássalhangzókra kapott átlagok: [p]: 69,3 ↔ [b]: 53,16; [t]: 72,43 ↔ [d]: 62,54; [k]: 66,4 ↔ [g]: 59,08 ms. A zöngés zárhangok átlagosan 18%-os rövidüléséből azt a következtetést vonhatjuk le, hogy a hangsor belsejében a zár időtartama fontosabb produkciós korrelátuma a zárhangok zöngés/zöngétlen szembenállásának, mint a megelőző magánhangzó időtartam-változása.

A mássalhangzó-kapcsolatok időtartam-módosító hatására kapott adatokat a 3. ábra összesíti. Megállapítható, hogy ez a fonetikai kontextus is a rövid magánhangzók nyúlását váltotta ki. A hatás nazálisok előtt volt a legmarkánsabb: 24% ( $p < 0,0001$ ), ami nem tartalmazza a nazális + réshang kapcsolatok előtti pótnyúlás, például: *közönség*, adatait. Rés- és zárhangok esetében a különbség 9%-nak adódott (a kétmintás párosított t-próba adatai:  $t(54) = 3,26, p < 0,002$ ).



3. ábra

A magánhangzók időtartama mássalhangzó-kapcsolatok előtt



A követő mássalhangzó hosszúságának tulajdonítható időtartam-változás vizsgálata, a korpusz CVC ↔ CVC: párojaira vonatkozó adatok szerint a rövid magánhangzók átlagos időtartama hosszú mássalhangzók előtt nő (a kétmintás párosított t-próba szerint  $t(54) = 3,645$ ,  $p < 0,001$ ). A növekedés tehát 10%-ot tett ki, s ennek mértékében a statisztikai elemzés nem jelzett különbséget a hangsúlyos/ hangsúlytalan helyzet között.

Figyelemre méltó a rövid magánhangzóknak a hosszú mássalhangzók és mássalhangzó-kapcsolatok előtt mutatott azonos viselkedése. Feltételezhetjük, hogy az időtartam-viszonyokat a szótagstruktúra szabályozza: az indukált időtartam-változás iránya a hosszú mássalhangzóknak a szakirodalomban sok vitát kiváltó VC\$CV elemzését erősíti, vagyis azt, hogy geminátákról van szó, s így a hosszú mássalhangzó előtti szótag zárt.

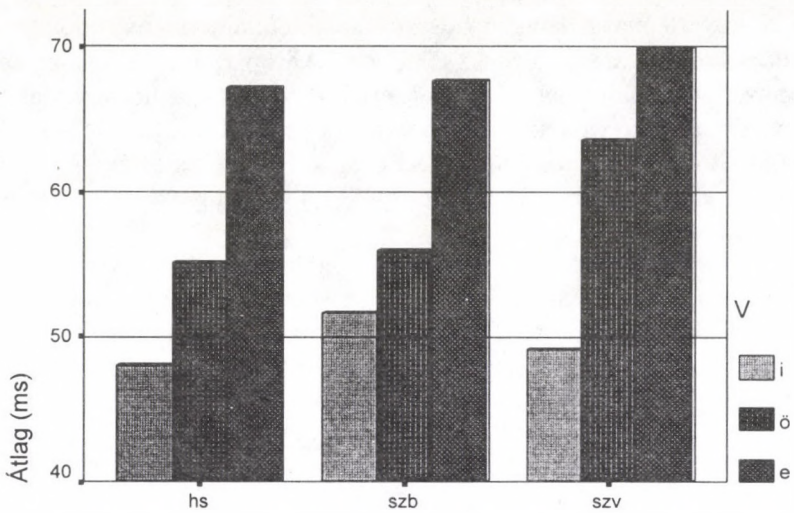
## **2. Hangsúly**

A mondat belseji 2-4 szótagú szavakban előforduló CVC szekvenciák elemzése szignifikáns különbséget jelzett a hangsúlyos és hangsúlytalan szótagok között (az egymintás t-próba eredménye:  $t(518) = 3,240$ ,  $p < 0,001$ ). A 4. ábrán bemutatott adatok tanúsága szerint a rövid magánhangzók időtartama a hangsúlyos szótagokban következetesen rövidebb. A hatás mértéke átlagosan 10%-ot tett ki.

Megvizsgáltuk, változik-e a hangsúlyos szótagbeli magánhangzó időtartama a szó szótagszámának függvényében. A hangsúlyos magánhangzók időtartama csak az egy és két szótagú szavak összevetésében volt nagyobb átlagosan 11%-kal (a kétmintás t-próba szerint  $t(28) = 4,320$ ,  $p < 0,0001$ ). A 2-4 szótagú szavak hangsúlyos magánhangzóinak időtartamát a szótagszám nem befolyásolta számottevően.

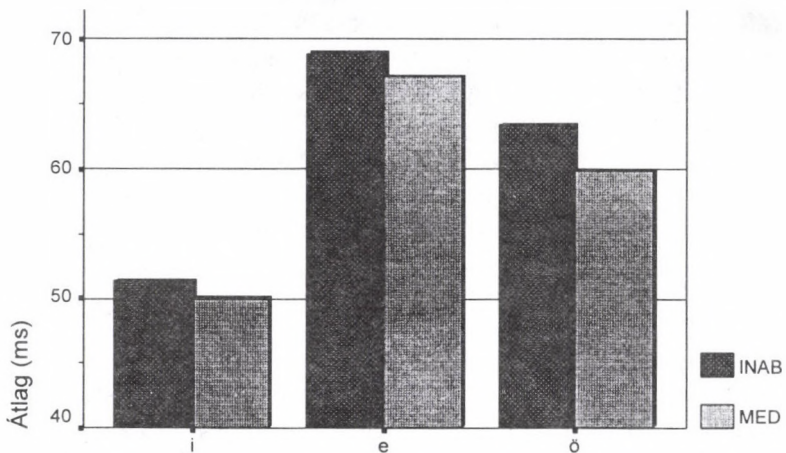
## **3. Hangsorban elfoglalt hely**

A hangsor eleji pozíció hatásának értékeléséhez összevettem a mondat első ↔ nem első és nem utolsó szavának első szótagjában ejtett magánhangzók időtartamát. A kvázi-minimális párok az 5. ábráról leolvasható szignifikáns ( $t(90) = 3,021$ ,  $p < 0,003$ ), de csekély változást



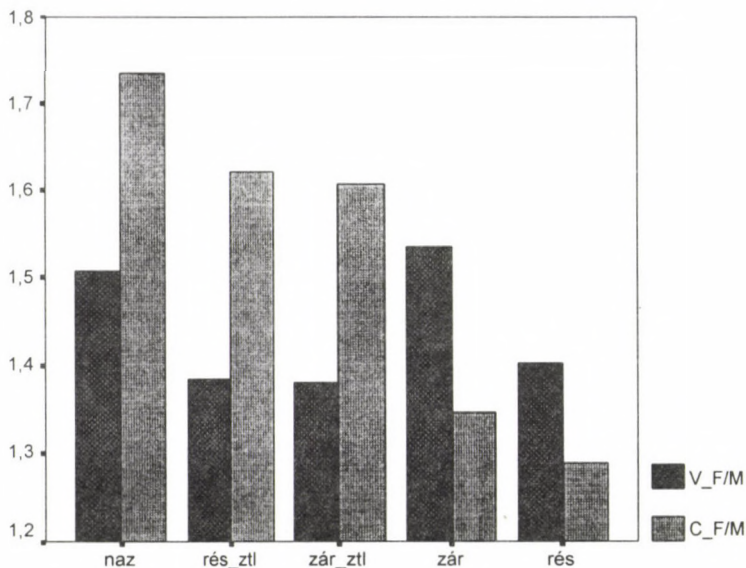
4. ábra

A magánhangzó időtartama hangsúlyos (hs), szóbelseji (szb) és szóvégi (szv) helyzetben



5. ábra

A magánhangzók átlagos időtartama a mondat első szótágjában (INAB) és a mondat belsejében (MED)



6. ábra

A magánhangzóknak (V\_F/M) és a mássalhangzóknak (C\_F/M) a mondat belseji előforduláshoz viszonyított aránya

mutatták. A hangzók időtartama átlagosan 6%-kal növekedett a mondat első szótagjában.

A hangsorvég időtartam-módosító szerepét ugyanazon szavak mondat belseji és mondat végi előfordulásai alapján elemeztük. A korpusz adatai arról tanúskodnak, hogy mondatvégi, szünet előtti helyzet jelentékenyen megnövelte a szegmentumok időtartamát. Az utolsó szótag magánhangzója a mondat belsejéhez képest 38%-os szignifikáns nyúlást mutatott. Az időtartamok növekedése jellemzően az utolsó szótagra összpontosult, de kisebb mértékben, átlagosan 7%-os arányban kiterjedt az utolsó előtti magánhangzókra is. A nyúlás utolsó szótagon belüli eloszlása elsősorban a mássalhangzó artikulációs-akusztikai alkatának a függvényében, a 6. ábrán ábrázolt arányokban alakult. Megfigyelhető, hogy a magánhangzó + rés- és zárhang kapcsolatok belső időviszonyai a mondat belseji helyzethez képest



átrendeződtek. A réshangok esetében felerősödött a zöngés/ zöngétlen mássalhangzók időtartam-különbsége; míg a zárhangok esetében megnövekedett a magánhangzók időtartamában mutatkozó eltérés jelentősége. A magánhangzók zöngés zárhangok előtti 53%-os változásával szemben zöngétlenek előtt 38%-os növekedést tapasztaltunk (az egy- mintás t-próba adatai:  $t(51) = 2,549$ ,  $p < 0,014$ ).

### **Összegzés**

A rövid magánhangzók kontextusfüggő időtartam-módosulásában a vizsgált korpusz elemzése alapján a következő fontosabb összefüggések emelhetők ki.

- Zöngés/zöngétlen zárhangok előtt a zöngések előtti növekedés mondat belsejében nem következetes; a magánhangzó és mássalhangzó időtartamok negatív korrelációja felerősödik a hangsorvégi szótagokban.
- Hangsor belseji zárt szótagokban – mind mássalhangzó-kapcsolatok, mind hosszú mássalhangzók előtt – a rövid magánhangzók időtartama növekszik.
- A szó hangsúlyos szótagjában a rövid magánhangzók rövidülnek.
- A magánhangzók hangsor végi nyúlása zárt szótagokban is igen jelentős, a mondatvég egyik akusztikai jelölőjének minősíthető.

### **Irodalom**

Cosi, P.: Segmentation and labelling automatic module. In: Proceedings of EUROSPEECH '93, 3rd European Conference on Speech Technology I. Berlin 1993, 88-91.

Gósy, M.: The perception of tempo. In: Temporal factors in speech. Szerk: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1991, 39-49.

Hockey, B. A. – Fagyal, Zs.: Phonemic length and pre-boundary lengthening: An experimental investigation on the use of durational cues in Hungarian. In: Proceedings of the XIVth ICPHS I. San Francisco 1999, 1313-1316.

Kassai Ilona: Időtartam és kvantitás a magyar nyelvben. Nyelvtudományi Értekezések 112. Akadémiai Kiadó. Budapest 1979.

Kassai Ilona: A magyar beszéd időtartamviszonyai. In: Fejezetek a magyar leíró hangtanból. Szerk.: Bolla Kálmán. Akadémiai Kiadó. Budapest 1982, 115-154.



Magdics Klára: A magyar beszédhangok időtartama. Nyelvtudományi Közlemények 68. 1966, 125-139.

Meyer, E. A. – Gombocz, Z.: Zur Phonetik der ungarischen Sprache. Uppsala 1909.

Olaszy Gábor: Hangidőtartam-módosító kísérletek a gépi beszéd ritmusának javításához. In: Beszédkutatás '94. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1994, 140-151.

Szabó P.: La quantité de la voyelle accentuée en hongrois. [Hivatkozás: Kassai Ilona (1979)]

van Santen, J. P. H.: Contextual effects on vowel duration. Speech communication 11. 1992, 513-546.

# KÍSÉRLET A MAGYAR BESZÉDHANGOK SPECIFIKUS IDŐTARTAMAINAK MEGHATÁROZÁSÁRA FOLYAMATOS BESZÉDRE

Olaszy Gábor  
MTA Nyelvtudományi Intézet

## Bevezetés

A specifikus időtartam elméleti fogalom. A szerző meghatározása szerint olyan alapidőtartamot fejez ki, amelyik a beszédképzés legsemlegesebb szintjén jellemző a hangra. Számszerű meghatározása nehéz, mivel az a kritérium, hogy legsemlegesebb szint nem fordul elő a természetes beszédképzés során. A beszédhangok időtartamának alakulását a folyamatos beszédben sok tényező befolyásolja, ilyenek a hang helyzete a hangsorban, a megelőző és a követő hang, a hangsúly, az artikulációs sebesség, a mondat típusa, a beszélő egyéni jellemzői stb. (Lazicius 1944; Magdics 1966; Kassai 1979; van Santen 1992).

A beszéd során az elhangzott közlést több tényező által kialakított komplex időszerkezet jellemzi. Ebbe a komplex időszerkezetbe beletartozik a szegmentális szintű szerkezet (mint alap), amikor csak az artikuláció befolyásolja a hangidőtartamot és beletartozik a szupra-segmentális szintű is, amikor a szegmentális szintre ráépülő összes további tényező alakítja ki a beszédhang végleges időtartamát. A hallgató ezt a végleges időtartamot hallja. Ha a szegmentális szintre jellemző időtartamok meghatározása a cél, akkor ebből a komplex időszerkezetből kell visszafejteni az egyes befolyásoló tényezők hatását, hogy a műveletsor végén megkapjuk azt az időtartamot, amelynek kialakulását már csak az artikuláció befolyásolja.

A beszédhangok időtartamával a szakirodalom gazdagon foglalkozik. A magyar beszédhangok időtartamára vonatkozó kutatások inkább csak a szó vagy mondatba ágyazott szó szintjén foglalkoztak a beszédhangok időtartamával (Magdics 1966; Kassai 1979). Magyar nyelvre folyamatos beszédre vonatkoztatott specifikus időtartamokat mindeddig nem vizsgáltak. Shiga és szerzőtársai (1988) vetik fel

ennek igényét, amikor az artikuláció hatását veszik figyelembe a beszédszintézishez történő hangidőtartam beállításnál.

A kutatás célja, hogy konkrét adatokkal jellemezzük a magyar folyamatos beszédben jelenlévő alap-időtartamokat, azaz számszerűsítsük a specifikus időtartam fogalmát. Olyan adatrendszerrel kívántunk felállítani, amelynek a felhasználásával meg lehet határozni a folyamatos beszédre jellemző specifikus hangidőtartamokat, vagyis jósolni lehet egy mondat hangjainak alapidőtartamait az artikuláció függvényében. A feltételezés az volt, hogy a specifikus időtartamok már önmagukban is reprezentálják a magyar beszédhangokra jellemző alapvető időtartam-szerkezeti képet (a magánhangzók egymáshoz viszonyított időtartam-különbségeit, a magánhangzókra jellemző időtartam-eloszlást, a mássalhangzók jellemző időtartamait, végül pedig a magánhangzók és mássalhangzók közötti időtartam arányokat).

### **Anyag és módszer**

A cél megvalósításához kialakított módszer új. Lényege az indirekt megközelítés, mely azt fejezi ki, hogy a beszédhangok időtartamát nem közvetlenül mérjük, hanem beszédszintézis és percepciók teszt kombinációjával olyan beszédet hozunk létre, amelyik jó közelítéssel reprezentálja a szegmentális szerkezetre jellemző tulajdonságokat. Ezen az adathalmazon végezzük el az időtartam-méréseket. Az így kapott hangidőtartamokat tekintjük a specifikus időtartamnak. Az indirekt mérés tehát négy részből tevődik össze: **első szakasz:** célirányos beszédszintetizálás, szegmentális szintű beszéd létrehozása; **második szakasz:** percepciók teszt és ennek alapján történő hangidőtartam beállítás; **harmadik szakasz:** a specifikus hangidőtartamok megmérése, az eredmények összesítése; **negyedik szakasz:** az eredmények ellenőrzése, összehasonlítása korábbi mérési adatokkal.

A kísérletben a célirányos beszédszintézis helyettesíti az emberi beszédet. E beszéd létrehozásához speciális, emberi hangból kialakított hullámforma-részeket használtunk. A hullámforma részletek kialakításánál (a hangfelvételnél) úgy jártunk el, hogy a később felhasználó jelrészlet építőelemekben lehetőleg csak a szegmentális szintre jellemző komponensek szerepeljenek. Ezekből a jelrészletekből adatbázist állítottunk össze. A gépi beszédet az adatbázisból kiválasztott

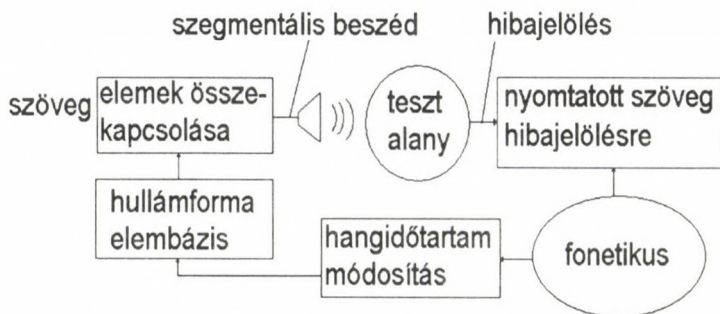


eseti elemek egymás után való kapcsolásával hozzuk létre. Így imitáljuk a szegmentális szintű beszédet. A gép a kísérlet során szöveget olvas fel.

A percepciók teszt lényege az, hogy a gép által generált beszéd hangjainak az időtartamát kell minősíteni túl hosszú, illetve túl rövid jelzővel. A kísérleti személyek döntéseiket a szövegben jelölik. Ezután a túl hosszúnak, illetve túl rövidnek minősített hangok időtartamát a gép adatbázisában korrigáljuk, így közelítünk a percepciósan helyesnek talált időtartamok felé. A percepciók teszt és a korrekciók többszöri megismétlésével fokozatosan elérjük, hogy a gép által generált beszédben minden beszédhang időtartamát elfogadhatónak minősítik a percepciók teszt résztvevői. Az így létrehozott beszéd reprezentálja a specifikus időtartamokat, amelyek kialakításában csak az artikuláció hatása szerepel. A kísérletben alkalmazott percepciók teszt zárt láncú (1. ábra), önmagába visszatérő 5-6 fordulós vizsgálat volt, amelyben minden fordulóban hallás alapján korrigáltuk a szövegben jelzett időtartam eltéréseket. A korrekciót az akusztikai adatbázis elemein végeztük el, így a következő fordulóban ugyanaz a hangsor már a korábban kijavított hangidőtartamokkal épült fel a szintézis során. Egy-egy tesztforduló általában egy hónap alatt zajlott le. Az egész percepciók teszt 8 hónapig tartott 1999-ben.

A percepciók teszthez egyrésztől speciális tartalmú 2480 mondatból álló szöveget állítottunk össze. Minden mondatban 5-10 szó szerepelt, olyanok, amelyek hangszerkezete lefedte a CVC és VCV elemeket. A tesztek első részében ennek a szövegkorpusznak a meghallgatásával végeztük a kirívóan hosszú, illetve rövidnek ítélt hangidőtartamok bejelölését és korrigálását. Ezután, a későbbi fázisokban összefüggő szövegeket hallgattak a kísérleti személyek (újságcikkek, tanulmányok stb.). Négy fő (egy nő és három férfi, életkoruk 30 és 40 év közötti) vett részt a specifikus hangidőtartamok meghatározásában. Meg kellett hallgatniuk mondatról mondatra a szintetizált mintát, és a kinyomtatott szövegben be kellett jelölni, hogy mely hangok túl hosszúak (- jellel) és melyek túl rövidek (x jellel). A mondatot többször is meg lehetett hallgatni egymás után. A tesztet egyenként végeztük. Az első harmadban a durva időtartam hibák kerültek kijavításra, a második harmadban már finomabb időtartam eltéréseket is észlel-

tek, a teszt végén pedig már 10-15 ms-os finomsággal észlelték a kissé hosszú, illetve kissé rövid időtartamokat. Ilyen alacsony érzékelési küszöbről már Huggins (1972) is közölt eredményeket. A teszt alanyok ítéleteit minden esetben fonetikus is megvizsgálta és a végső időtartam-változtatást ő határozta meg. Az időtartamok fizikai korrigálása egy korábban kifejlesztett speciális algoritmus segítségével történt (Olaszy–Olaszi 1998).



1. ábra

Kísérleti összeállítás a beszédhangok specifikus időtartamának inverz módszerrel történő meghatározására

A kísérleti beszéd-szintézis során 9 magánhangzó (á, a, o, u, ü, i, é, ö, e) és 23 mássalhangzó (b, p, d, t, g, k, gy(G), ty(T), m, n, ny(N), j, h, v, f, z, sz(s), c, zs(Z), s(S), cs(C), l, r) specifikus időtartamait vizsgáltuk. A hangok jelölésére azok betűjelét használjuk, a számítógépes táblázatokban egyedi jeleket (lásd fent a zárójelbe tett karaktereket a hang betűképe mellett) is használunk.

### A szintetikus beszéd építőelemei

Az indirekt mérési eljárásból következik, hogy a kísérlet leglényesebb eleme a célirányos beszéd-szintézissel létrehozott beszéd (hullámforma elemeket kapcsolunk egymás után és így hozzuk létre a hangsorozatot). A szintetikus beszéd tervezésekor a következő szempontokat vettük figyelembe: lehetőleg ne okozzunk torzulást a hangok



akusztikai szerkezetében, érvényesüljön a hangok előfordulási gyakorisága, a beszédépítés során felhasznált elemek lehetőleg csak a szegmentális szerkezetre jellemző tulajdonságokkal rendelkezzenek (azonos hangmagasságúak legyenek, dallam és ritmikából származó időtartam-módosulások ne legyenek bennük). Az artikulációval szemben támasztott követelményt emberi hang alapú szintézissel teljesítettük. Az alapvető cél szerint a hangok időtartamát a megelőző és a követő hang hatásának függvényében akarjuk meghatározni, mivel a folyamatos beszédben ez a hangsorozat természetes szerkezete. Ennek megfelelően hármas hangkapcsolatokban kell vizsgálni a középső hang időtartamát. A leggyakoribb hangkapcsolati szerkezeteket figyelembe véve (Szende 1976) CVC, VCV, CCV, VCC, VVC és CVV elemekben kell vizsgálni a középső hangot. A C kategóriába számítjuk a hangsorkezdő és hangsorzáró csend állapotot is.

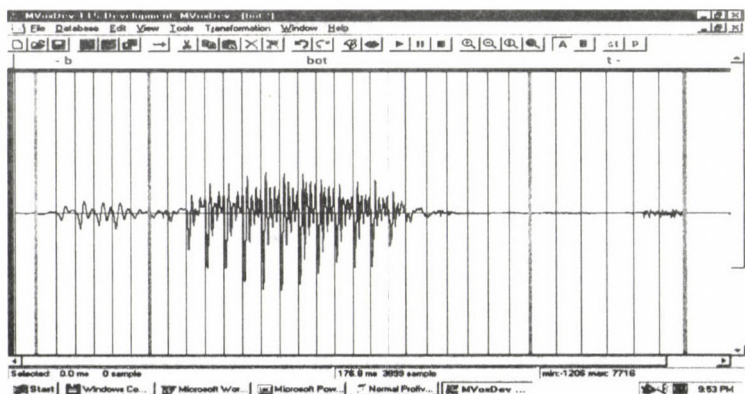
Korábbi mérések szerint a hallgatók jobban érzékenyek a magánhangzók időtartam hibáira (túl hosszú, túl rövid), mint a mássalhangzóknál generált hasonló hibákra (Kato et al. 1998). Ezért a magánhangzók vizsgálatára CVC kapcsolatokat, a többi hangra pedig két félhangból álló elemeket hoztunk létre. A CVC elemekben a C-k félhangot tartalmaznak (a mássalhangzók felénél van elvágva a hang) a magánhangzó pedig eredeti formájában marad. Így elértük azt, hogy a magánhangzók akusztikai szerkezete nem torzult (hiszen nem vágtuk el őket), továbbá, hogy minden CVC kapcsolatban szereplő magánhangzónak külön-külön kezelhető saját időtartama volt. A CVC elemeken túl – mivel a hangsorok nem kizárólag CVCVCVC.... szerkezetűek – két félhangot tartalmazó CV, VC, CC, VV félhang kapcsolatokat is alkalmaztunk építőelemként. Ezeknél az elemeknél a hang közepén történt a vágás.

### **Az akusztikai adatbázis létrehozása**

Az akusztikai adatbázisba CVC, CV, VC, CC és VV elemeket tároltunk. Az elemek kialakításához három szótagos szövegelemeket olvasott fel a bemondó (részletesen lásd Olaszgy 1999). A felolvasott szövegkorpuszból 4761 CVC, 207 CV, 207 VC, 529 CC és 81 VV hullámforma elemet hoztunk létre az adatbázisban. A 2. ábrán bemu-



tatjuk, hogy például a *bot* hangsor a **-b/|bot/t-** elemek összekapcsolásával hogyan alakul ki.



2. ábra

A *bot* hangsor -C, CVC, és C- elemei összekapcsolva (a szélesebb vonal jelzi az elemhatárokat)

A kezdő -C elem 6 periódusból áll és csak a [b] zöngéjét tartalmazza. A második CVC elem 20 periódusból áll, a [b] zárfelpattanásával kezdődik, majd a teljes [o] hangot tartalmazza és a [t] hang néma fázisával zárul. A hangsort záró C–elem a [t] néma fázisát is tartalmazza, valamint a zárfeloldódásból származó hosszabb zörejt.

A hanghatár meghatározása a kísérlet szempontjából (és minden, hangidőtartam mérésre vonatkozó kutatás szempontjából) kritikus fontosságú (van Santen 1992). A hanghatárok megjelölése ebben a kísérletben manuálisan történt, a munkát vizuális és auditív szoftver eszközök segítették. Vizuálisan az időfüggvényből, és a hangintenzitás alakulásából, auditíve pedig a hangátmenetek meghallgatásával döntöttük el, hogy hol lesz a hanghatár. A hanghatárok meghatározása erősen hangfüggő. A zár, zár-rés és réshangoknál viszonylag jól meg lehet határozni a pontos helyet, a többi mássalhangzó esetében nehezebb. A zárhang+V kapcsolatoknál a mássalhangzóhoz számítottuk a zárfelpattanást a lecsengés végéig. A réshang, zár-rés hang+V kapcsolatok-

nál a hanghatárt a magánhangzó első induló periódusára jelöltük meg. A szonoráns+V+szonoráns kapcsolatoknál a vizuális elemzésen túl erősen támaszkodtunk az auditív meghallgatás eredményére (a hangkapcsolatot meg lehetett hallgatni úgy, hogy periódusról periódusra tágítottuk a hullámforma ablakot).

A fentiekből következik, hogy a magánhangzók kezdete és vége viszonylag pontosan meghatározható zár, zárrés és réshangok szomszédságában, egyébként a kijelölés erősen személyfüggő. A kérdéses esetekben 10-20 ms-os eltérések is lehetnek egyes személyek hanghatar-jelölései között.

### Eredmények

A magánhangzókra CVC helyzetre kapott adatokat, mivel három hang kapcsolódását kell szemléltetni speciális mátrix formában szemléltetjük (ms-ban). A mátrix első oszlopában szerepelnek a magánhangzót megelőző, az első sorában az őt követő mássalhangzót jelképező karakterek, a kérdéses időtartam az oszlop és a sor kereszteződésében található. A kérdéses magánhangzó, amelyre az adatok vonatkoznak a mátrix bal felső sarkában van megadva. Példaképpen bemutatjuk az [o] hang specifikus időtartamait az 1. táblázatban. A szórás 72 ms-tól 115 ms-ig terjed. Az adatokból számított átlag 90 ms. Ha az adatokat 10 ms-os csoportokba rendezzük, akkor képet kapunk arról, hogy milyen hangkörnyezetnek milyen hatása van az [o] hang időtartamára. A 2. táblázatban bemutatjuk, hogy mely CVC kapcsolatokban jellemző a [o] hangra a 81-90 ms-os sáv. A 2. táblázat szerint az [o] hang időtartama a leghosszabb ha palatálisokhoz kapcsolódik, a legrövidebb pedig a nazálisok szomszédságában. Ez utóbbi ellentmond Magdics (1966) megállapításának. Minden magánhangzóra elvégeztük az átlagok kiszámítását is, amely a következő sorrendet adta: i(80), u,ü(86), o(90), a,e(91), ö(92); é(146), á(164). Ez a sorrend megegyezik Kassai (1979) adataival, aki hangsúlyos helyzetben mért magyar magánhangzókra a következő adatokat közölte: [i]<[u]<[y]<[o]<[E]<[□]<[O]. Más nyelvekre mért adatok is hasonló sorrendet mutatnak. O'Shaughnessy (1981) francia magánhangzókat vizsgált zárt szótagokban és úgy találta, hogy a legrövidebb

I. táblázat: Az [o] hang specifikus időtartamai CVC kapcsolatokban ms-ban

O	b	p	d	t	G	k	G	T	m	n	N	j	h	v	f	z	s	c	Z	S	C	l	r
b	88	93	84	95	93	90	93	103	83	84	94	95	95	85	94	93	90	94	94	85	83	84	94
p	88	93	83	95	92	90	92	103	82	83	93	95	95	84	93	93	90	94	94	84	83	83	93
d	86	91	82	93	91	88	91	101	81	81	92	93	93	83	91	91	88	92	92	83	81	82	92
t	84	90	80	92	89	86	89	100	79	80	90	92	92	81	90	90	87	91	91	81	80	80	90
g	87	92	83	94	92	89	92	102	82	82	93	94	94	84	92	92	89	93	93	84	82	83	93
k	79	84	75	86	84	81	84	94	74	74	85	86	86	76	84	84	81	85	85	76	74	75	85
G	90	95	85	97	94	92	94	105	84	85	95	97	97	86	95	95	92	96	96	86	85	85	95
T	99	104	95	106	104	101	104	115	94	95	105	106	106	96	105	105	101	106	106	96	95	95	105
m	79	85	75	87	84	81	84	95	74	75	85	86	86	76	85	85	81	86	86	76	75	75	85
n	90	96	86	98	95	92	95	106	85	86	96	98	98	87	96	96	93	97	97	87	86	86	96
N	94	99	90	101	99	96	99	109	89	90	100	101	101	91	100	99	96	100	100	91	89	90	100
j	80	86	76	88	85	82	85	96	75	76	86	87	87	77	86	86	82	87	87	77	76	76	86
h	90	96	86	98	95	92	95	106	85	86	96	98	98	87	96	96	93	97	97	87	86	86	96
v	88	94	84	96	93	90	93	104	83	84	94	96	96	85	94	94	91	95	95	85	84	84	94
f	80	86	76	88	85	82	85	96	75	76	86	88	88	77	86	86	83	87	87	77	76	76	86
z	91	96	87	98	96	93	96	106	86	86	97	98	98	88	96	96	93	97	97	88	86	87	97
s	87	92	83	94	92	89	92	103	82	83	93	94	94	84	93	93	89	94	94	84	83	83	93
c	93	98	88	100	97	95	97	108	87	88	98	100	100	89	98	98	95	99	99	89	88	88	98
Z	87	92	82	94	91	89	91	102	81	82	92	94	94	83	92	92	89	93	93	83	82	82	92
S	77	83	73	85	82	79	82	93	72	73	83	85	85	74	83	83	80	84	84	74	73	73	83
C	88	94	84	96	93	90	93	104	83	84	94	95	95	85	94	94	90	95	95	85	84	84	94
l	80	85	76	87	85	82	85	95	75	75	86	87	87	77	85	85	82	86	86	77	75	76	86
r	89	95	85	97	94	91	94	105	84	85	95	97	97	86	95	95	92	96	96	86	85	85	95

magánhangzó a magas nyelvállású [i] és [u] volt, utána az [E] következett, majd az alsó nyelvállású [a]. Angolra mért adatok szerint (van Santen 1992) a sorrend hasonló mind hangsúlyos, mind pedig hangsúlytalan helyzetű magánhangzó esetén. Ez a vizsgálat is megerősítette, hogy a magánhangzó időtartama és a nyelvállás közötti összefüggés a vizsgált nyelvekre általános érvényűnek látszik. Néhány ponton adataink viszont ellentmondanak a korábbi eredményeknek.



## 2. táblázat: CVC kapcsolatok, amelyekben az [o] hang 81-90 ms-os

81-90	Bob	bod	Bok	bom	bon	bov	boS	boC	bol	pob	pod	pok	pom	Pon	pov	pos
ms	poS	poC	Pol	dob	dod	dok	dom	don	dov	dos	doS	doC	dol	Tob	top	tod
	tog	tok	toG	ton	tov	tof	toz	tos	toS	tol	gob	god	gok	Gom	gon	gov
	gos	goS	goC	gol	kop	kot	kog	kok	koG	koN	koj	koh	kof	koz	kos	koc
	koZ	kor	Gob	God	Gom	Gon	Gov	GoS	GoC	Gol	mop	mot	mog	mok	moG	moN
	moj	moh	mof	moz	mos	moc	moZ	mor	nod	nom	non	nov	noS	noC	nol	Nod
	Nom	Non	NoC	Nol	job	jop	jot	jog	jok	joG	joN	joj	joh	jof	joz	jos
	joc	joZ	jor	hod	hom	hon	hov	hoS	hoC	hol	vob	vod	vom	von	vov	voS
	voC	vol	fob	fop	fot	fog	fok	foG	foN	foj	foh	fof	foz	fos	foc	foZ
	for	zod	zom	zon	zov	zoS	zoC	zol	sob	sod	sok	som	son	sov	sos	soS
	soC	sol	cod	com	con	cov	coS	coC	col	Zob	Zod	Zok	Zom	Zon	Zov	Zos
	ZoS	ZoC	Zol	Sop	Sot	Sog	SoG	SoN	Soj	Soh	Sof	Soz	Soc	SoZ	Sor	Cob
	Cod	Com	Con	Cov	CoS	CoC	Col	lop	lot	log	lok	loG	loN	loj	loh	lof
	loz	los	loc	loZ	lor	rob	rod	rom	ron	rov	roS	roC	rol			

Nem találtuk például konzekvensen hosszabbnak a magánhangzót az [l] és [r] hangok előtt és nem mutatható ki az sem, hogy a magánhangzó időtartama konzekvensen hosszabb zöngés mássalhangzó előtt, mint zöngétlen előtt, ahogy Magdics (1966) is megállapította. Az adatainkból számított eloszlásokat a mért magánhangzókra a 3. ábra mutatja.

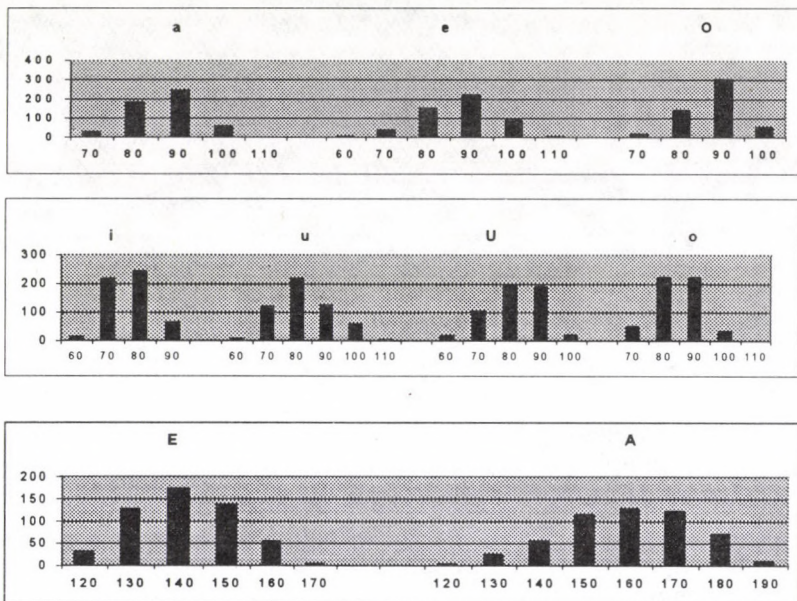
A VCV kapcsolatokban mért mássalhangzók specifikus időtartamait 23 mátrixban összesen 1863 különböző VCV hangkapcsolatra összegeztük. A 3. táblázatban bemutatjuk a [b] hang mátrixát. Az időtaramok ms-ban értendők. Itt a [b] hang időtartamának szórása sokkal kisebb, mint például a bemutatott [o] hangé volt. A legrövidebb időtartam 55 ms, a leghosszabb pedig 78 ms. A mátrixban a legtöbb hangkapcsolat (55 a 81-ből) a 61-70 ms-os sávban van. A [b] hang időtartameloszlása tehát erre a sávra koncentrálódik.

3. táblázat: A [b] hang mátrixa VCV kapcsolatokra

b	A	a	o	u	U	I	E	O	e
A	61	63	63	56	66	68	66	57	61
a	62	64	64	57	67	70	67	58	62
o	67	69	69	62	72	75	72	63	67
u	70	72	72	65	75	78	75	66	70
U	61	63	63	56	66	69	66	57	61
I	67	69	69	62	72	75	72	63	67
E	60	62	62	55	65	68	65	56	60
O	60	62	62	55	66	68	65	56	60
e	69	71	71	64	74	76	74	65	69

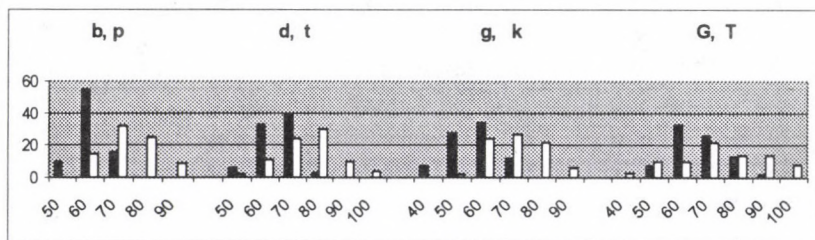
A mássalhangzók átlagidőtartamára a következő adatokat kaptuk: b(67), p(77),d(70), t(76), g(62), k(74), gy(68), ty(76), m(67), n(48), ny(66), j(59), h(62), v(61), f(85), z(68), sz(82), c(92), zs(67), s(83), cs(98), l(52), r(37). Látható, hogy például a zöngétlen zárhangok átlagosan hosszabb időtartamúak, mint a zöngések (4. ábra). Ez korrelál Magdics (1966) megállapításával is. Kassai (1979) a következő sorrendet adta meg a mássalhangzókra szám adatok nélkül: likvidák < nazálisok < zöngés zárhangok < zöngés réshangok < zöngétlen zárhangok < zöngétlen réshangok < zöngétlen zár-réshangok. Olasz (1985) későbbi vizsgálatában a következő értékeket állapította meg a fenti hangsorrendre vetítve: 45, 67, 69, 65, 117, 120, 125 ms, a jelen kísérlet eredményei pedig a következő adatokat adják: 44, 61, 66, 65, 76, 79, 95 ms. A tendenciák tehát egyeznek.

A mássalhangzók időtartamát VCC és CCV kapcsolatokban is meghatároztuk (minden esetben a középső C-re). Az eredményeket 23-23 mátrixban, 2x4761 hangkapcsolatra (9 magánhangzó és 23 mássalhangzó kombinációja) összegeztük. Példaként elemezzük a [b] hangra kapott eredményeket. VCC kapcsolatban a [b] hang legrövidebb időtartama 58 ms, a leghosszabb 95 ms.



3. ábra

A magánhangzók specifikus időtartamainak eloszlása CVC helyzetre



4. ábra

A zárhangok specifikus időtartamainak eloszlása VCV helyzetre



Az átlagidőtartama tehát hosszabb, mint a VCV kapcsolatokban szereplő [b] hangé. VCC kapcsolatban a [b] legrövidebb időtartama 52 ms, a leghosszabb 95 ms. Ezek az adatok hasonlóak, mint amilyenek a CCV helyzetben jellemezték a [b] hangot, tehát a két helyzet között nincs lényegi eltérés.

### Összefoglalás

A dolgozatban bemutatott inverz eljárás segítségével meghatároztuk a magyar beszédhangokra jellemző specifikus időtartamokat 9 magánhangzóra és 23 mássalhangzóra mint a hangok kombinációjából származtatható hármas hangkapcsolatok (összesen 16146-féle hangkapcsolat) középső hangjára. Az adatok tükrözik az artikuláció hangidőtartamokra vonatkoztatott hatását. A kapott eredményeket összehasonlítottuk korábbi kutatások eredményeivel és lényegi eltérést nem tapasztaltunk. Tehát úgy ítéljük meg, hogy a jelen eljárással meghatározott hangidőtartamok releváns adatoknak minősíthetők. Ezzel az eljárással sikerült az elméleti kategóriaként kezelt specifikus időtartamokat számszerűsíteni folyamatos beszédre. Így a kapott adatok alapján előre jóslni lehet egy mondat hangjainak specifikus időtartamát. Példaként bemutatjuk a *“Tegnap este hosszú levelet írtam a barátnőmnek Budapestre.”* példamondat hangjainak specifikus időtartamát ms-ban jelölve.

T(84)e(91)g(79)n(65)a(95)p(84) e(95)s(89)t(84)e(92) h(42)o(93)ssz(167)ú(140)  
l(58)e(95)v(37)e(73)l(57)e(85)t(85) i(140)r(76)t(70)a(90)m(63) a(83)  
b(64)a(85)r(39)á(171)t(64)n(70)ó(139)m(91)n(74)e(99)k(72)  
B(47)u(83) d(66)a(95)p(84)e(95)s(89)t(81)r(55)e(109).

A vizsgálat eredményeit beszéd szintetizáló rendszerbe is beépítettük, az eredmények felhasználásával olyan beszédet lehet előállítani, amelyik nagy részben tartalmazza a magyar beszédképzésre jellemző hangidőtartamokat, időtartam arányokat.

A bemutatott eljárás előnyei: reprodukálható, személyfüggetlen, korrekt adatokat ad, a hangidőtartamok tükrözik a nyelvre jellemző adatokat, lefedi az összes hangot, minden hangra van jellemző időtartam bármely hangkapcsolatban, statisztikai vizsgálatokra is alkalmas

(pl. arányok a hangok között, eloszlások stb.), hanggal is ellenőrizhető a mérési eredmények emberi erőforrás (bemondó) nélkül.

### **Köszönetnyilvánítás**

A szerző köszönetét fejezi ki a Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Telematikai Tanszékének, Dr. Németh Géának és Kiss Géza doktoranduszának, akik a vizsgálathoz szükséges számítástechnikai háttérrel tervezték és a szerző által kér algoritmusokat programozták.

### **Irodalom**

Huggins, A. W. F.: Just noticeable differences for segment duration in natural speech. *Journal of the Acoustic Society of America* 51, 1972, 1270-1278.

Kato, H., – Tsuzaki M. – Sagosaka Y.: Effects of phonetic quality and duration on perceptual acceptibility of Temporal changes in speech. *Proc. of the 5<sup>th</sup> International Conference on Spoken Language Processing*. Sydney 1998, 892-895.

Kassai Ilona: Időtartam és kvantitás a magyar nyelvben. *Nyelvtudományi Értekezések* 102. Budapest 1979.

Magdics Klára: A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 1966, 125-139.

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

Olaszy Gábor – Olaszi Péter: Hangidőtartamok mesterséges változtatása periódusok kivágásával, megismétlésével. *Beszéd kutatás'98*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1998, 151-162.

Olaszy Gábor: Beszédatatbázisok tervezése gépi beszédelőállításához. *Beszéd kutatás'99*. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1999, 68-89.

O'Shaughnessy, D.: A study of French vowel and consonant durations. *Journal of Phonetics* 9. 1981, 385-406.

Santen van, J. P. H.: Contextual effects on vowel duration. *Speech Communication* 11. 1992, 513-546.

Shiga, Y. – Matsuura, H. – Nitta, T.: Segmental duration control based on an articulatory model. *Proc. of the 5<sup>th</sup> International Conference on Spoken Language Processing*. Sydney 1998, 1244-1247.

**Ezt a kutatást az OTKA T 030292 is támogatta.**

# A BESZÉD IDŐVISZONYAI KÜLÖNBÖZŐ ÉLETKORÚ SZEMÉLYEKNÉL

Gocsál Ákos

Pécsi Tudományegyetem, Tanárképző Intézet

## Bevezetés

Mindannyiunk hétköznapi tapasztalata, hogy amikor valakit beszélni hallunk, akkor nem csupán a beszédben kódolt szemantikai és pragmatikai jelentéseket tudjuk megérteni, de képesek vagyunk arra is, hogy a beszélő személyéről bizonyos megállapításokat tegyünk. E jelenség háttérében az áll, hogy a beszéd létrehozása során keletkező akusztikai jelsorozat fonetikai paramétereit nagyban befolyásolják, sőt, meghatározzák a beszélő személy egyes tulajdonságai. Az utóbbi néhány évtized akusztikai és percepciós fonetikai szakirodalmában több olyan kutatási beszámolót is találunk, amely a beszélő személy sajátosságai, illetve beszédének egyes paraméterei között keres kapcsolatokat. E kutatások főbb irányai az alábbiak:

- a beszélő személy **személyazonosságának** megállapítása hangja alapján (például Gósy 1996; 1999),
- a beszélő személy **testméretei** (magasság, súly – például van Dommelen–Moxness 1995),
- **érzelmi állapota** (Scherer 1995), illetve
- **egészségi állapota** és egyes fonetikai paraméterei közötti kapcsolat megállapítása (Balázs–Gósy 1986; Campbell–Dollaghan 1995), továbbá
- a különböző **életkorokra** jellemző akusztikai sajátosságok feltárása, ideértve annak vizsgálatát is, hogy az életkor előrehaladtával az egyes fonetikai paraméterek hogyan változnak.

A jelen dolgozat ez utóbbi irányzathoz kapcsolódóan arra a kérdésre keresi a választ, hogy mi jellemző a különböző életkorú személyek spontán beszédének időszerkezetére.



### **Az életkor változásának hatásai a beszédparaméterekre**

Az e témához kapcsolódó kutatások két csoportra oszthatók. A **longitudinális** vizsgálatok esettanulmány-szerűen, egy-egy kiválasztott személy különböző időpontokban rögzített hangmintáit elemzik, és a beszédparaméterek változási tendenciáit írják le. Erre a célra általában archívumból vett, más célból rögzített beszédmintákat alkalmaznak a kutatók, mivel egyébként több évet vagy akár évtizedet is várni kellene a kutatás elvégzéséhez.

A hazai szakirodalomban Balázs (1993) tanulmánya volt az első ilyen jellegű munka. Balázs Boglárka Dajka Margit (DM) és Páger Antal (PA) egy-egy fiatal-, illetve időskori hangmintáját elemezte. Megállapította, hogy idős korokra mindkét beszélőnél jelentősen csökkent a beszédtempó értéke, nőtt a szünettartás, a beszédhangok pedig keskenyebb frekvenciasávban helyezkednek el. Különbség ott mutatkozott, hogy míg DM esetében az alaphang frekvenciája jelentősen csökkent, PA idős korában magasabb alaphangon beszélt. A beszédparaméterek változása azonban felnőttkorban lassan megy végbe. Gósy (1996) mérései alapján arra a következtetésre jutott, hogy ugyanazon (középkorú) felnőtt beszélő hat év különbséggel rögzített hangmintái azonos akusztikai paramétereket mutatnak.

Russell és munkatársai (1995) ausztrál női beszélőkkel végeztek longitudinális vizsgálatot, az alaphang változásának megállapítására. A kutatók rendelkezésére állt 28 kísérleti személy 1945-ben (ekkor 18-25 évesek), továbbá néhányuk 1981-ben rögzített hangmintája. A hangfelvételt 1993-ban megismételték. Az eredmények azt mutatták, hogy az 1945-ben felvett hangmintákon mért 229 Hz-es átlagos hangmagasság 1993-ra szignifikánsan, 181 Hz-re csökkent. Mivel az 1981-ben rögzített hangmintáknál az átlagos hangmagasság 184 Hz volt, a szerzők arra a következtetésre jutottak, hogy az alaphangmagasság csökkenésének üteme kb. 20 és 55 éves kor között gyorsabb, később mérsékeltebb.

A kutatások másik csoportjába a **keresztmetszeti** vizsgálatok tartoznak, melyek arra adnak választ, hogy egy viszonylag rövid időszakon belül a különböző életkorú személyek beszédét milyen paraméterek jellemzik. Ezekkel az adatokkal differenciáltabb képet nyerhetünk

a köznyelvi nyelvhasználat egyes jellemzőiről, például egy adott évben.

Gyermekek hangmintáival végzett keresztmetszeti kutatást Laczkó (1991), aki megállapította, hogy a 14 évesek artikulációs és beszédtempó-értékei lényegesen magasabbak voltak, mint a 4 évesek tempóadatai, azonban a 14 évesek esetében a szünettartás is nagyobb arányú volt.

Subosits (1990) 18 és 22, illetve 50 és 60 év közötti kísérleti személyekkel olvastatta fel ugyanazt a 112 szóból álló szöveget. Az idősebbek 1,29-szer több időt fordítottak a felolvasásra, és szünettartásuk időtartama a fiatalokénak 1,28-szorosa volt. Jelentős különbség adódott az artikulációs tempók között is, amelynek okát a szerző a biológiai tényezőkön kívül a felgyorsult életvitel fiatalokra gyakorolt hatásában is látja.

Walker és munkatársai (1992) 3 és 5 éves gyermekek artikulációs-tempó-értékeit összehasonlítva szignifikáns különbséget találtak, így az eredményeik a gyermeknyelv és a beszédfejlődés kutatása szempontjából is igen fontosak. Beck (1999) pedig 21 női és 35 férfi beszélő hangmintái alapján megrajzolta az életkor-alaphang grafikonokat. A grafikonok azt illusztrálják, hogy kb. 20 éves kor alatt igen széles az alaphang értékeinek szórása (2 éves kor alatt 350-450 Hz mindkét nem esetén, később egyre alacsonyabb, 20 éves kor körül a leányoknál 200-250 Hz, a fiúknál pedig 100-150 Hz). Idősebb korúaknál az alaphang magassága ezen az értéken marad, vagy valamivel alacsonyabb lesz. A grafikonon a 65 év feletti férfiak alaphangja viszont ismét magasabb értékeket vesz fel, ami más kutatási eredményekkel is egybecseng.

A longitudinális, illetve a keresztmetszeti kutatások eredményeit grafikusán (vízszintesen az életkort, függőlegesen a vizsgált beszédparamétert) ábrázolva hasonló jellegű ábrákhoz jutunk. Fontos azonban hangsúlyozni a két vizsgálati módszer közötti különbségeket: míg a longitudinális kutatás **tendenciákat**, változásokat, a keresztmetszeti kutatás a vizsgálat időpontjára jellemzően **állapotokat** tükröz. Ha az egyik módszerrel elvégeztünk egy kutatást, esetenként előfordulhat, hogy olyan következtetésekre is juthatunk, amelyekhez – elvileg – csak a másik módszer segítségével juthatnánk. Az ilyen megállapítá-



sokkal azonban óvatosan kell bánnunk. Ha például megmérjük, hogy a különböző életkorú beszélők ma milyen artikulációs tempóval beszélnek, arra következtethetnénk, hogy felnőttkorban az artikulációs tempó értéke lassan, de folyamatosan csökken. Ezt a megállapítást azonban csak akkor lehet bizonyítottnak tekinteni, ha korábban, egy másik időpontban is végeztünk volna egy hasonló, keresztmetszeti vizsgálatot. Jól látható tehát, hogy a két módszer kiegészíti egymást: a nyelvről, a nyelvhasználatról együttesen adnak pontos képet.

### **A kutatás feladata, kísérleti személyek, beszédminták**

Az itt ismertetendő keresztmetszeti kutatásban azt vizsgáltuk, hogy ma, 2000-ben, a különböző életkorú felnőtt beszélők spontán beszédének időszerkezete hogyan alakul. 37 személy beszédmintáját dolgoztuk fel (16 férfi, 21 nő, életkoruk: 17-80 év). A személyekre jellemző, hogy szellemi foglalkozásúak, a fiatalabbak gimnáziumi tanulók vagy egyetemisták, az idősebbek jelentős része diplomás, kisebb része csak érettségivel rendelkezett.

A hangfelvételek elkészítésekor a kísérleti személyek olyan témákról beszéltek, amelyekhez érzelmileg semleges viszonyultak, s amikről könnyedén tudtak beszélni. A hangfelvételek hossza 5-15 perc volt.

### **A beszédminták feldolgozása**

A rendelkezésre álló hangminták mindegyikéből véletlenszerűen kiválasztva kb. 1-1 percnyit WAV-formátumban számítógépen rögzítettünk, majd a Snack 2.0 program segítségével megmértük a megfelelő időtartamokat.

Az alábbi adatokat mértük meg:

- a hangminta teljes hossza, ms-ban
- a szünetek időtartama összesen, ms-ban
- a hezitálások időtartama összesen, ms-ban
- a nyújtások időtartama összesen, ms-ban.

Ezen kívül megszámloltuk a hangmintákban szereplő beszédhangokat is.



Az így rendelkezésre álló adatokból az alábbi újabb adatokat számítottuk ki:

- beszédtempó
- artikulációs tempó
- a szünettartás időtartama, a teljes időtartam százalékában
- a hezitálások időtartama, a teljes időtartam százalékában
- „artikulációs hatások” (az artikulációra fordított idő és a beszédminta teljes időtartamának hányadosa, 0 és 1 közötti szám)

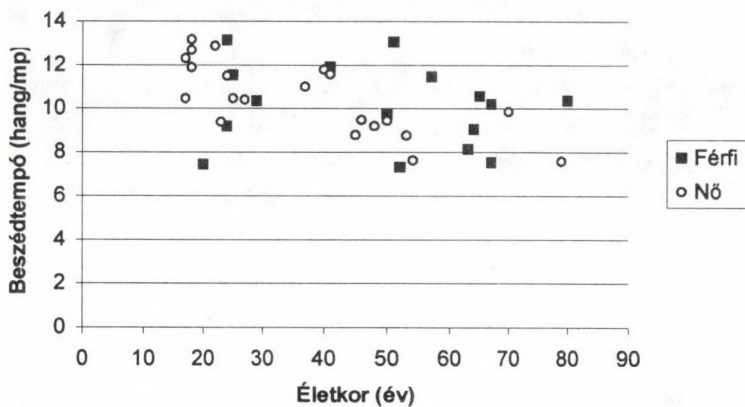
### **A kutatás eredményei**

Az 1. ábrán ábrázoltuk azt, hogy a különböző életkorú kísérleti személyek beszédtempó-értékei hogyan alakulnak. Jól megfigyelhető, hogy kb. 40-45 éves korig a beszédtempó-értékek jellemzően 10 hang/mp felett vannak, sőt, néhány egyetemista lány esetén megközelítette a 14 hang/mp-es értéket. A 45 évesnél idősebb beszélők esetén a beszédtempó értéke már gyakrabban esett 10 hang/mp alatti tartományba, több esetben 8 hang/mp alatti értéket kaptunk.

Hasonló elrendeződést mutatnak az adatok az életkor-artikulációs tempó grafikonon is. A 2. ábrán megfigyelhető, hogy kb. 45 éves kor alatt a kapott tempóértékek kevés kivétellel 14 hang/mp-nál nagyobbak. Idősebbeknél a 14 hang/mp-nél lassabb tempók lesznek jellemzők. Érdekes megfigyelni, hogy miközben az egyetemistákhoz hasonlóan a 40 éveseknél több beszélő is közel 16 hangot ejt ki másodpercenként, addig az 50 év körüli hölgyek artikulációja lényegesen lassabb, 12 hang/mp, vagy az alatti az artikulációs tempó értéke önáluk.

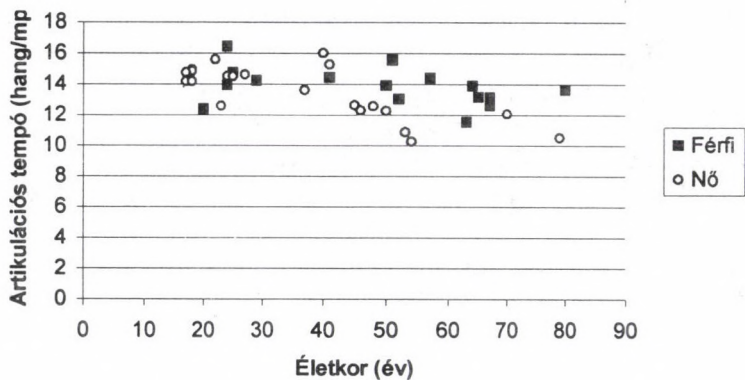
A 3. ábrán látható, hogy az egyes beszélők a beszédminta időtartamának hány százalékát töltötték csendben. Előzetesen arra számítottunk – korábbi kutatások alapján –, hogy az idősebbek lényegesen több szünetet fognak tartani, mint a fiatalabbak. A grafikon azonban véletlenszerű eloszlást mutat, 15% körüli eredményeket épp úgy találunk fiataloknál és időseknél, mint a 30% közelében adódott szünettartás-értékeket.

Hasonlóképpen véletlenszerű eloszlás figyelhető meg a 4. ábrán is, amelyen a hezitálás százalékos időtartamát ábrázoltuk az életkor



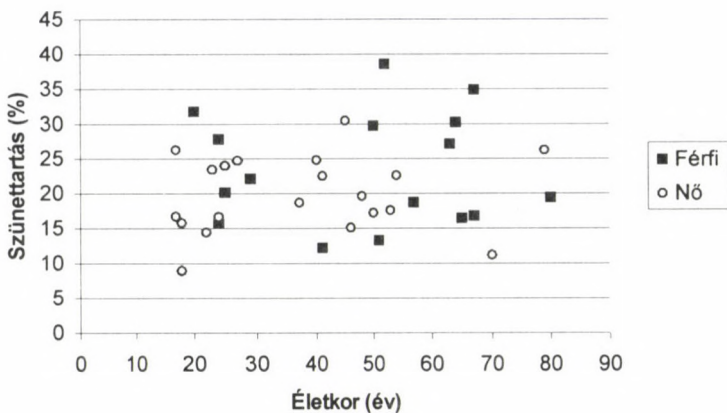
1. ábra

A beszédtempó értékei különböző életkorú személyeknél

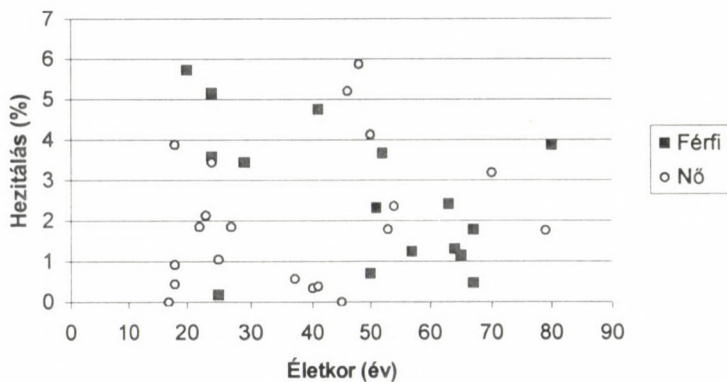


2. ábra

Az artikulációs tempó értékei különböző életkorú személyeknél

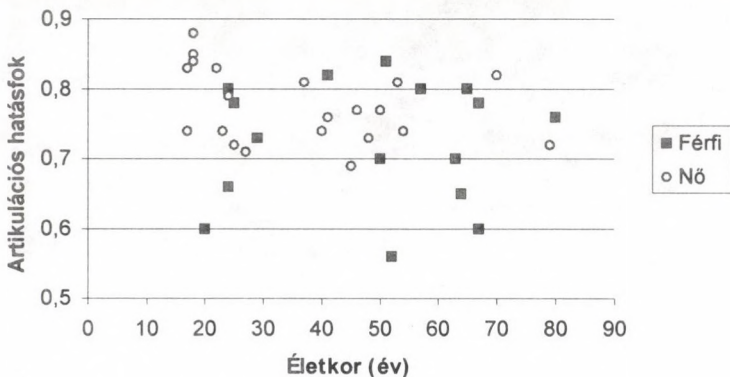


3. ábra  
A szünnettartás értékei különböző életkorú személyeknél



4. ábra  
A hezitálás értékei különböző életkorú személyeknél





5. ábra

Az artikulációs hatásfok értékei különböző életkorú személyeknél

függvényében. Ha azonban figyelembe vesszük a beszélők nemét is, akkor azt a megállapítást tehetjük, hogy a vizsgálatban részt vett idősebb férfiak egyre kevesebb hezitálással beszéltek.

Az 5. ábrán tüntettük fel az artikulációs hatásfok értékeit. (Az „artikulációs hatásfok” fogalmának használatával korántsem azt szeretnénk érzékeltetni, hogy a nagyobb hatásfokot mutató beszélők bizonyos szempontból „jobbak”, „gazdaságosabbak” – amihez hasonló megállapításokat például gépek működésének vizsgálatakor szokás tenni. Az artikulációs hatásfok értékét ugyanis leginkább a szünettartás csökkenti, amelynek – többek között – meghatározott kommunikációs és beszédproduktív funkciói is vannak. Valószínűsíthető, hogy egyéni sajátosságok, esetleg jól meghatározható stratégiák is állnak a szünettartások alkalmazása – hossza, gyakorisága stb. – hátterében. A beszédészünetek nemritkán fontos jelentéstoppletet nyújtó, szükséges részei a beszéd folyamatnak, tehát az alacsonyabb artikulációs hatásfok-értéktől nem „rosszabb” valakinek a beszédképzése. E megállapítások természetesen a beszédproduktív szempontból egészséges beszélőkre vonatkoznak.) A grafikonról leolvasható, hogy nincsenek életkorra jellemző eltérések az artikulációs hatásfokban. Így az a következtetés vonható le, hogy bár jelentős eltéréseket találtunk az be-

szédtempó és az artikulációs tempó mérésekor a különböző életkorú személyeknél – azaz az idősebbek lassabban beszélnek –, az életkor nem befolyásolja az időkihasználást a beszéd során. Erre a 3. ábra alapján is könnyen következtethetünk. Amit azonban csak az 5. ábra illusztrál, az a férfiak és a nők hatásfokértékei közötti különbség. Jól megfigyelhető, hogy a nők esetében életkortól függetlenül a hatásfok értéke szinte minden esetben 0,7 és 0,85 közé esik. A férfiaknál viszont jóval szélesebb tartományban helyezkednek el ezek az értékek, az intervallum két határa náluk 0,55 és 0,85. A nők tehát az artikulációs hatásfok szempontjából jóval koherensebb csoportot alkotnak, mint a férfiak. E jelenségre jelenleg nem tudunk magyarázatot adni. Feltételezhető, hogy a beszédproduktions folyamatok során a férfiak és a nők helyenként vagy más stratégiákat alkalmaznak, vagy eltérő sebességgel hajtják végre bizonyos mentális műveleteket – de ezekre a kérdésekre csak részletesebb pszicholingvisztikai kísérletekkel lehet válaszolni.

### A kapott eredmények statisztikai elemzése

A grafikonokon ábrázolt adatokat statisztikailag is feldolgoztuk. A statisztikai feldolgozás során azt vizsgáltuk, hogy az egyes beszédparaméterek értékei és a kísérleti személyek életkora között milyen szintű korreláció áll fenn. A statisztikai elemzés eredményeit az 1. táblázat tartalmazza.

1. táblázat: A korrelációs együtthatók és a szignifikanciaszintek

	BESZÉD.T.	ARTIKUL.T.	SZÜNET	HEZITÁLÁS	HATÁSFOK
ÉLETKOR	-0,487	-0,542	0,181	0,036	-0,244
p	0,002	0,001	0,283	0,832	0,146

További szignifikáns korrelációk:

- beszédtempó – art. tempó:  $p < 0,001$
- beszédtempó – szünettartás:  $p < 0,001$  (negatív korreláció)
- beszédtempó – hatásfok:  $p < 0,001$
- artikulációs tempó – hatásfok:  $p < 0,041$
- szünettartás – hatásfok:  $p < 0,001$  (negatív korreláció)



A statisztikai elemzésből tehát kiderül, hogy a tempógrafikonokon ábrázolt adatok elrendeződése nem a véletlen műve, azaz **az életkor szignifikáns módon befolyásolja a beszédtempó és az artikulációs tempó alakulását**. A korrelációs együttható negatív előjele arra utal, hogy nagyobb életkorok esetén alacsonyabbak a tempóértékek. Feltételezhetően egy szélesebb populáció vizsgálatakor mindkét esetben még erősebb, 0,001-nél kisebb szignifikancia-szintet kapnánk.

Sem a szünettartás, sem a hezitálás értékei nem mutattak szignifikáns összefüggést az életkorral, tehát adataink alapján azt mondhatjuk, hogy **a spontán beszédben tapasztalt szünettartások és hezitálások mértéke nem jellemzője a különböző korosztályoknak**. A 3. és a 4. ábrán látható jelentős szórás más tényezőknek tudható be; nagy valószínűséggel az egyéni sajátosságok, esetleg a beszéd témája vagy körülményei miatt adódnak a grafikonon megfigyelhető különbségek. Az artikulációs hatásfok esetén sem találtunk szignifikáns összefüggést az életkorral.

A korrelációs együtthatók és a szignifikanciaszintek kiszámításával statisztikailag is alátámasztottuk az ábrák vizuális elemzése alapján megfogalmazott megállapításainkat. Megjegyezzük, hogy a hezitálás értékei egyik vizsgált beszédparaméterrel sem mutattak összefüggést, tehát a hezitálás hátterében meghúzódó okokat másutt kell keresnünk.

### **Lehetséges magyarázatok, további kérdések**

1. Elsősorban a **szervezet lassú fáradása, gerontológiai jellegű változásai** lehetnek a hátterében annak, hogy az idősebbek lassabban beszélnek. E folyamatokat, az időskorúak hangképzésének artikulációs sajátosságait Balázs (1993) részletesen ismertette. Kevés információ áll azonban rendelkezésünkre arról, hogy a beszédprodukción pszicholingvisztikai folyamatai hogyan alakulnak időseknél. Hipotetikusan ugyanis felvethető, hogy idős korra egyes pszichés működések is lassulnak, tehát ha például a mentális lexikon elérése lassul, akkor ez a „végeredményben”, tehát a beszéd tempóértékeiben is tükröződik. Erre a kérdésre is pszicholingvisztikai kísérletekkel kaphatunk választ.

2. A fiataloknál tapasztalt gyors beszéd- és artikulációs tempó rájuk jellemző **generációs sajátosság** is (vö.: Subosits 1990). Több kutatás ismeretes, amely a köznyelvi beszéd tempójának gyorsulását



mutatta ki (például Kassai 1993). Ezek az eredmények azt bizonyítják, hogy például a néhány évtizeddel ezelőtti húsz évesek valamivel lassabban beszéltek, mint a maiak. A mai idősebb generáció tagjainak tehát már a fiatalkori beszédtempói is alacsonyabbak voltak, mint a mai fiatalokéi. Feltételezhető tehát, hogy a fiatalok a saját generációjukra jellemző beszédtempót (és még sok más beszédjellemzőt) mintegy „magukkal viszik” idősebb korukra, és így a tempólassulás náluk más érteken fog realizálódni.

\*

A mérések során találtunk egy olyan tényezőt – az életkort –, amely szignifikánsan differenciálja a nyelvi közösség tagjainak egyes beszédparamétereit. Korábbi kutatások a mai köznyelvi beszéd átlagos artikulációs tempójára 14 hang/mp körüli értéket adnak meg (az itt feldolgozott beszédminták alapján átlagos artikulációs tempóra 13,6 hang/mp, átlagos beszédtempóra pedig 10,3 hang/mp adódott). Eredményeink szerint a kb. 40 éves kor alattiak beszédére ennél gyorsabb, az idősebbekére ennél lassabb tempóértékek a jellemzők. Úgy véljük, az ilyen jellegű kutatásokkal árnyaltabb képet alkothatunk mai köznyelvi beszédünkről.

### Irodalom

Balázs Boglárka: Az időskori hangképzés jellemzői. In: Beszédkutatás '93. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest 1993, 156-166.

Balázs Boglárka – Gósy Mária: Hangszalagcsomós betegek beszédének akusztikai vizsgálata. Fül-orr-gégegyógyászat, 32. 1986, 162-169.

Beck, J. M.: Organic variation of the vocal tract. In: The Handbook of Phonetic Sciences. Eds.: Hardcastle, W. J. – Laver, J. Blackwell. Oxford–Malden 1999, 256-299.

Campbell, F. T. – Dollaghan, C. A.: Speaking rate, articulatory speed, and linguistic processing in children and adolescents with severe traumatic brain injury. Speech and Hearing Research Vol. 38. Issue 4. 1995, 864-876.

van Dommelen, W. A. – Moxness, B.H.: Acoustic parameters in speaker height and weight identification: Sex-specific behaviour. Language and Speech 38 (3). 1995, 267-287.

Gósy Mária: A beszéd akusztikai szerkezetének állandóságáról. In: Nyelv, nyelvész, társadalom. Emlékkönyv Szépe György 65. születésnapjára

barátaiktól, kollégáitól, tanítványaitól. II. Szerk.: Terts István. Keraban Könyvkiadó. JPTE. Pécs 1996, 66-75.

Gósy Mária: Az egyéni hangszínezet és a beszélő felismerésének kísérleti-fonetikai megközelítése. Magyar Nyelvőr 123. 1999, 424-438.

Kassai Ilona: Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? In: Beszédkutatás '93. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest 1993, 62-69.

Laczkó, M.: The interrelation of articulation rate and pauses in children's speech. In: Temporal Factors in Speech. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1991, 139-151.

Russell, A. – Penny, L. – Pemberton, C.: Speaking fundamental frequency changes over time in women: A longitudinal study. Journal of Speech and Hearing Research Vol. 38 Issue 1. 1995, 101-110.

Scherer, P: How emotion is expressed in speech and singing. In: Proceedings of the XIIIth International Congress of Phonetic Sciences. Eds.: Elenius, K. – Branderud, P. Vol. 3. Stockholm 1995. 90-96.

Subosits István: Az életkor hatása a beszédtempó alakulására. In: Magyar Fonetikai Füzetek 22. Szerk.: Szende Tamás. MTA Nyelvtudományi Intézet. Budapest 1990, 131-132.

Walker, J. F. – Archibald, L. – Cherniak, S. R. – Fish, V. G.: Articulation rate in 3- and 5-year-old children. Journal of Speech and Hearing Research Vol. 35. Issue 1. 1992, 4-14.

A szerző köszönetet mond Hartung Bernadett, Jóna Lilla, Pál Mária, Pesti Mariann és Ziegler Teréz egyetemi hallgatóknak a beszédminták összegyűjtésében nyújtott segítségükért.

# A BESZÉD TEMPORÁLIS SAJÁTOSSÁGAI KÉTNyelvűEKNEél (KISISKOLÁSKORTÓL IDŐSKORIG)

**Menyhárt Krisztina**  
**MTA Nyelvtudományi Intézet**

## **Bevezetés**

Annak ellenére, hogy a kétnyelvűség kutatása már közel egy évszázados történelemmel rendelkezik, még mindig igen nehéz feladatot jelent annak a meghatározása, ki tekinthető kétnyelvűnek. Az egymásnak gyakran ellentmondó elméletek és irányzatok útvesztőjében komoly gondot jelent nemcsak magának a kétnyelvűségnek mint jelenségnek a vizsgálata, hanem az is, hogyan ítéljük meg annak pozitív és negatív hatását, hogyan viszonyuljunk az életkori tényezőhöz, a két nyelv elsajátításának idejéhez és módjához, illetőleg a kulturális és társadalmi hatásokhoz.

Már a korai kutatások (például Ronjat 1913; Leopold 1939–1947) is kísérletet tettek arra, hogy valamilyen módon rendszerezzék mind a kétnyelvű személyek, mind pedig a kétnyelvűség különböző típusait. A kezdeti szakaszban igen nagy hatást gyakorolt az ún. kettős monolingvis elmélet (Navracsics 1999, 15), amely a „tökéletes”, vagyis mind a két nyelvet anyanyelvi szinten beszélő kétnyelvűt tartotta egyedül elfogadhatónak. Számos félreértésre adott okot Lambert (1975, idézi Bartha 1999, 184-186) elmélete a két nyelv tudásának mértékéről. Ő vezette be a kiegyensúlyozott (balansz) és az egyenlőtlen (domináns) kétnyelvűség fogalmát. Nála azonban a „kiegyenlített” nem a két nyelv ismeretének szintjére, hanem csupán a két nyelv használatbeli egyensúlyára vonatkozott. A nyelvi kompetenciával foglalkozó elméletekben szintén találkozhatunk végletes megközelítésekkel: a minimalista elmélet szerint elég, ha a kétnyelvű személy az egyik nyelven jól tud, míg a másikat épp csak valamennyire, míg a maximalista elmélet szerint mind a két nyelv grammatikai rendszerének tökéletesen birtokában kell lennie (Navracsics 1999, 15).



A kétnyelvűség kutatásának pszicholingvisztikai megközelítése új szempontokkal és újabb vitatható állításokkal bővítette eddigi ismereteinket. Eleve vita tárgyát képezi a kétnyelvűek mentális lexikonának és nyelvi rendszereinek felépítése (egy közös vagy két szeparált rendszer létezik-e), illetve a kétnyelvű agy lateralizációjának kérdése is (Bartha 1999, 177). A legerjedtebb felfogás (Grosjean 1997, 109) pszicholingvisztikai szempontból megkülönböztet koordinált, összetett és szubordinált kétnyelvűeket. Koordinált kétnyelvű az, akinek mentális lexikonában a két nyelv szavai teljes mértékben elkülönülnek egymástól, így mindegyik lexikai egység mellé sajátos jelentés rendelődik. (Kérdéses azonban, hogy ez lehetséges-e egyáltalán.) Az összetett kétnyelvű személynél már egy közös nyelvi rendszert találunk, amelyben a két nyelv ekvivalens szavai közös, összekapcsolt jelentéssel rendelkeznek. A szubordinált kétnyelvűség esetében a kevésbé ismert nyelv szavainak értelmezése a jobban tudott nyelven keresztül történik, vagyis az agyban tulajdonképpen fordítás zajlik.

A kétnyelvűség kialakulásának ideje, vagyis az életkori tényező vizsgálata is fontos adatokkal szolgálhat. Ebből a szempontból megkülönböztetünk korai vagy kései, illetve gyermekkori, serdülő és felnőttkori kétnyelvűséget. A gyermekkori kétnyelvűség lehet egyidejű (szimultán), amikor is a gyermek egyszerre sajátítja el mind a két nyelvet, illetve egymás utáni (szukcesszív/konszekutív), amikor a második nyelv elsajátítása az első nyelv nyelvi rendszerének megszilárdulása után kezdődik meg (Hoffmann 1991; Bartha 1999). Természetesen a serdülőkori és a felnőttkori kétnyelvűség csak konszekutív típusú lehet. Ekkor a második biológiai sorompó gyakorol döntő hatást leginkább a beszédprodukción (vö. akcentus) szintjén.

A kétnyelvűség meghatározásánál a legcélravezetőbb több tényező is figyelembe venni. Ezek közül a leglényesebbek: (i) a nyelvi kompetencia, (ii) a pszicholingvisztikai szempont, (iii) a nyelvek hétköznapi használata, (iv) a nyelvek elsajátításának ideje és módja, (v) kulturális jártasság, illetve (vi) érzelmi kötődés.

A felsorolt tényezők akkor különösen fontosak, ha beszédvizsgálatokat végzünk kétnyelvűek körében, hiszen ekkor feltételezhetjük, hogy a résztvevők mind a két nyelven tudnak kommunikálni (beszélni és érteni). A kétnyelvűek beszédét sok szempontból vizsgálhatjuk,

ezek közül a **beszédtempó** számos fontos adattal szolgálhat mind az egy nyelvűek, mind pedig a két nyelvűek beszédének időviszonyairól. A beszéd időszerkezetének megismerése nemcsak a fonetika egyik alapfeladata, segítségével új ismereteket szerezhetünk a mentális folyamatok működéséről is. Így a két nyelvűségről is gazdagíthatjuk ismereteinket, megtudhatjuk, hogyan viszonyul a két nyelv egymáshoz, milyen a nyelvismeret szintje, hogyan szerveződnek a mentális lexikonok (a szavak előhívása is bizonyos időintervallumban történik, vö. lexikális hozzáférés), hogyan kapcsolódnak egymáshoz a grammatikai struktúrák és ez mind hogyan hat ki a beszédprodukcóra. A beszédtempó olyan egyéni sajátosság, amit a hétköznapi beszélő nem nagyon tud változtatni, éppen ezért igen alkalmas arra, hogy új adatokkal gazdagítsa a mentális folyamatok időbeli működéséről szerzett ismereteinket.

A beszéd időszerkezete az életkorral is változik (Gósy 1988). Ezért egy longitudinális vizsgálat pontosabb képet adhat a beszéd és az életkor kapcsolatáról a két nyelvű személyeknél is. Ismert, hogy a beszédtempó gyermekkorban lassúbb, az aktív felnőtt lakosságnál egyre gyorsuló tendenciát mutat (Kassai 1993), majd időskorban újra lassul. Kérdéses azonban, hogy ez a folyamat ugyanígy zajlik-e a két nyelvű személyek esetében is, illetőleg egyformán érinti-e mind a két nyelvet.

Megfelelő elemzések elvégzésével a beszédtempó a beszédprodukciónak a folyamatának temporális szervezéséről nyújt felvilágosítást. Nem véletlen, hogy a beszédprodukciónak a modellek felépítéséhez is felhasználták a temporális szerveződésről nyert adatokat (vö. Goldman-Eisler 1968).

A jelen tanulmány témája a beszédtempó vizsgálata magyar–bolgár két nyelvű személyek különböző életkorú csoportjainál – kisiskolásoknál, felnőtteknél és időseknél. Az elvégzett kísérlettel egyrészt arra kerestük a választ, hogyan változik az életkor függvényében a két nyelvűek magyar, illetve bolgár beszéde, másrészt pedig azt vizsgáltuk, utal-e a beszédtempó a nyelvdominanciára.

### **Anyag és módszer**

Kísérletünkben 15 magyar–bolgár két nyelvű személy (7 nő és 8 férfi) vett részt (1. táblázat). Közülük öten gyermekek, a budapesti



bolgár iskola 3. és 4. osztályos diákjai, öten aktívan dolgozó felnőttek, öten pedig idősek. Közülük nyolcan felsőfokú, ketten középfokú végzettséggel rendelkeznek.

1. táblázat: A résztvevők adatai

Csoport	Nem		Életkor	Átlagéletkor
kisiskolások	2 leány	3 fiú	9–11 év	10 év
felnőttek	2 nő	3 férfi	21–56 év	32 év
idősek	3 nő	2 férfi	63–83 év	73 év

Spontán, illetőleg irányított spontán beszéd felvételek készültek a kísérlet személyekkel (a felnőtteknél interjúk, a gyermekeknél képsorozatról mesélés formájában) mind bolgár, mind magyar nyelven. A közléseket magnetofonra rögzítettük, a felvett anyagokból a jelen vizsgálathoz személyenként és nyelvenként 2-3 percet választottunk ki. A szövegeket IPA-átírással, fonetikus lejegyeztük, majd folyamatos akusztikus és vizuális megerősítés mellett, CSL 4300B digitális jelelemzővel feldolgoztuk. Megállapítottuk az artikulációs tempó értékeit, amely az artikulációra fordított időt jelenti, majd ennek és a szünetek időtartamának függvényében megkaptuk a beszédtempó értékeit is, amelyeket beszédhang/s-ban adtunk meg (h/s).

A kísérletben részt vett korcsoportok rendkívül heterogének voltak mind a származás, mind pedig a nyelvtudás szempontjából, másrésről viszont éppen ezáltal rendkívül jól reprezentálták a magyarországi bolgár közösséget. Az időskorú csoport tagjai 50-60 éve kerültek Magyarországra, minden nyelvi készségnek (beszéd, értés, írás, olvasás) mindkét nyelven birtokában vannak, de magyar nyelven igen erős akcentussal és számos nyelvtani hibával beszélnek. A felnőtt csoport tagjai már Magyarországon születtek, a két nyelvet vagy egyszerre, vagy kisgyermekkorban egymás után sajátították el, így azokat igen magas szinten, gyakorlatilag akcentus nélkül beszélik, és a mindennapi élet folyamán használják. A gyermekeknél, annak ellenére, hogy bolgár tannyelvű iskolába járnak, a bolgár nyelv háttérbe szorult (enyhe akcentus, grammatikai hibák), ők már szívesebben beszélnek magyarul.

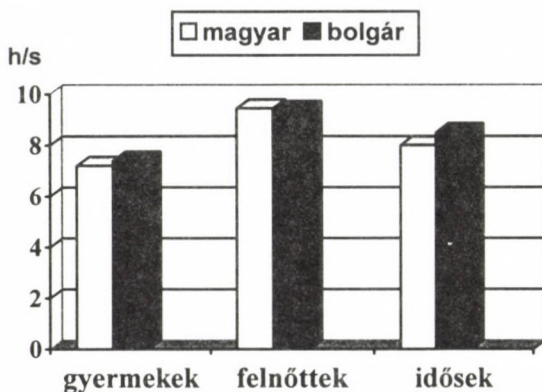


A kísérleti csoportok heterogén felépítése nagyon jól szemlélteti azokat a definíciós problémákat, amelyekkel a bevezetésben foglalkoztunk.

### Kísérleti eredmények

A beszéd megértésében és sebességének megítélésében az artikulációs sebességnek és szüneteknek az együttes hatása játszik döntő szerepet. Ezért az elvégzett kísérletek eredményei közül mindenekelőtt a beszédtempóval foglalkoztunk.

Az 1. ábrán láthatók mindhárom korcsoport beszédtempójának összesített eredményei.



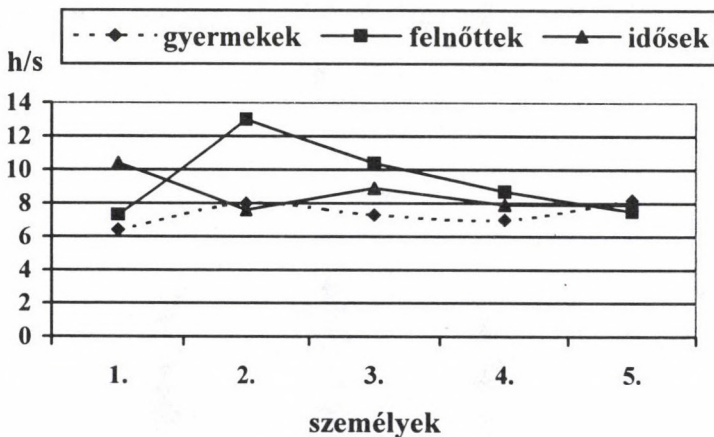
1. ábra

A beszédtempó összesített eredményei

A kapott adatok a kétnyelvűeknél igazolják azt a hipotézist, amely szerint a gyermekek beszédtempója a leglassúbb (7,2 h/s a magyar és 7,4 h/s a bolgár nyelvre), leggyorsabban a felnőttek beszélnek (9,4 h/s – magyarul; 9,38 h/s – bolgárul), és az idős korcsoport beszédtempója lassul (8 h/s a magyar és 8,5 h/s a bolgár nyelv tempója). Másrészről viszont a két nyelv beszédtempója között mind a három korcsoportnál csak kis különbséget regisztráltunk. Ez valószínűleg annak

bizonyítéka, hogy a részt vevő személyek a különbségek ellenére két-nyelvűeknek tekinthetők.

A vizsgált korosztályok beszédtempójának egyéni különbségei fontos információkkal gazdagíthatják ismerteinket a kétnyelvűek mentális folyamatainak működéséről. A három csoport eredményeit bolgár, illetve magyar nyelven a 2. és a 3. ábra mutatja.

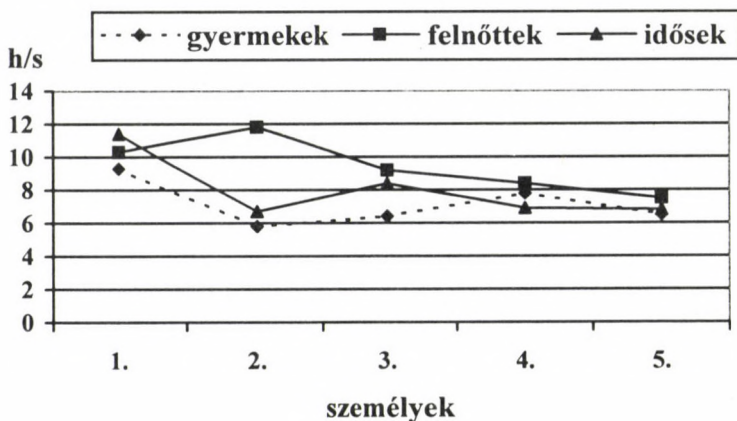


2. ábra

A három korcsoport beszédtempója bolgár nyelven

A legnagyobb egyéni különbségeket a bolgár nyelv beszédtempójában a felnőtt csoportnál találjuk – 7,3 h/s-tól 13 h/s-ig szórt az eredményük (ez az 5,6 h/s 45 szó különbséget jelent percenként). Ennek oka valószínűleg az életkorban (a legidősebb résztvevő 35 évvel volt idősebb a legfiatalabbnál), a foglalkozásban és a nyelvhasználatban keresendő. Legkisebb szórást a gyermekeknél tapasztaltunk – 6,5 h/s-tól 8,2 h/s-ig (1,7 h/s vagyis 14 szó/perc), náluk ez az azonos életkorral és életmóddal magyarázható. Időskorra a bolgár nyelv beszédtempója újra kiegyensúlyozódik – az idősök eredményei 7,6 h/s és 10,4 h/s közé esnek (2,8 h/s: 22 szó/perc).

A 3. ábrán láthatók a három vizsgált korcsoport beszédtempójának egyéni értékei magyar nyelvre.



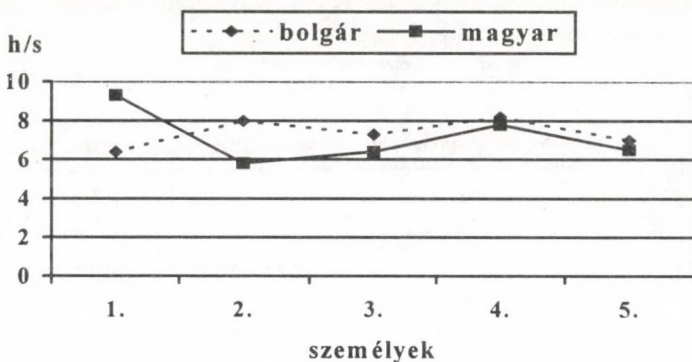
3. ábra

A három korcsoport beszédtempója magyar nyelven

A magyar nyelvre kapott adatok szórási arányai másképpen alakultak. Csökkentek a felnőttek közötti különbségek (7,5 h/s és 11, 8 h/s közé esnek – a különbség 4,3 h/s vagyis 34 szó/s), viszont növekedtek az egyéni eltérések mind a gyermekek, mind az idősek között. Legnagyobb a szórás az idősek adataiban – 6,7 h/s-tól 11,4 h/s-ig (4,7 h/s vagyis 38 szó/perc). A gyermekek adatainak szórása kétszeresére növekedett – 5,8 h/s-tól 9,3 h/s-ig (a különbség 3,5 h/s: 28 szó/perc). Az okok valószínűleg ugyancsak a foglalkozási, életmódbeli és nyelvhasználatbeli különbségekben keresendők.

Ha összevetjük a három korcsoporton belül a két nyelv egyéni beszédtempójának értékeit (4, 5, 6. ábra), akkor a nyelvdominancia alakulására, illetve állapotára vonatkozóan fogalmazhatunk meg munkahipotézist.

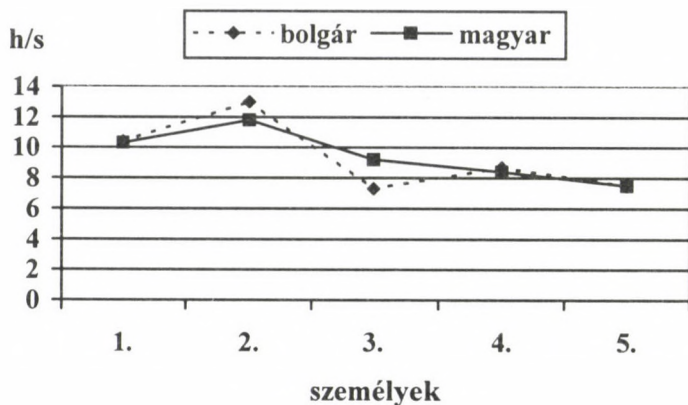




4. ábra  
A gyermekek beszédtempója bolgár és magyar nyelven

A kapott adatok szerint az öt gyermek közül négynek a bolgár beszédtempója gyorsabb, mint a magyar. Közülük viszont háromnál a két érték igen közel esik egymáshoz, a különbség tehát nem számottevő. Nagyobb eltérést az 1. és a 2. gyermeknél tapasztalhatunk, náluk valószínűleg már a domináns nyelv hatását jelzik az eredmények (az 1. gyermek inkább magyar, a 2. inkább bolgár domináns). A 2. számú gyermek magyar beszédében egyértelműen hallható volt kismértékű akcentus.

Az 5. ábrán a felnőtt korcsoport beszédtempóinak egyéni értékeit tüntettük fel. Mint azt a bevezetőben kiemeltük, a felnőtt résztvevők mindegyike igen magas szinten ismeri és használja mind a két nyelvet. Ez a kapott eredményekben is tükröződik. A két nyelv beszédtempójának értékei gyakorlatilag itt kiegyenlítődnek, a különbségek elhanyagolhatók. Az öt résztvevő közül háromnál a két érték szinte azonos, egyénél a bolgár tempó valamivel gyorsabb.

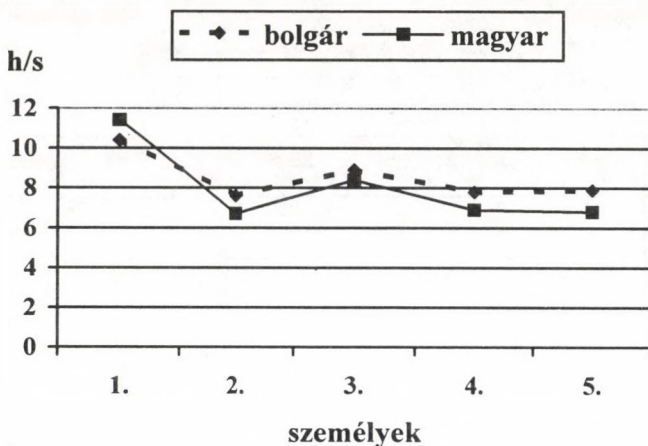


5. ábra

A felnőttek beszédtempója bolgár és magyar nyelven

Az egyetlen résztvevő, akinél számottevőbb, 2 h/s körüli különbséget találunk (ennyivel gyorsabb a magyar beszéde), foglalkozási okokból az utóbbi években kiesett a bolgár közösség életéből, így nála ez az eredmény a magyar környezet erős hatásának tulajdonítható. Ez a tény jól mutatja a bilingvis személyek két nyelvének dinamikus egyensúlyát, amely mindig a domináns nyelvi környezet hatásaira reagál.

Az idős korcsoport eredményeit a 6. ábra szemlélteti. Az öt résztvevő közül négyenél egyértelműen a bolgár nyelv beszédtempója volt gyorsabb. A kísérleti személyek életének ismeretében ez az adat nem meglepő, hiszen mindnyájan felnőtt korukban tanultak meg magyarul. A különbség mindazonáltal nem kiugróan nagy, ami a kétnyelvűség tényét támasztja alá. A bolgár nyelv előtérbe kerülését az is elősegíti, hogy az idős személyek már kikerültek az aktív életből, kapcsolataik már szinte csak a bolgár közösségre szorítkoznak, ahol természetesen a bolgár nyelvet használják.



6. ábra

Az idősek beszédtempója bolgár és magyar nyelven

Az egyetlen résztvevő, akinél a magyar beszéd bizonyult gyorsabbnak, bizonyos értelemben kivételes eset. Egyrészt a nyelvcsajátítás folyamata nála másképp zajlott (a két nyelvet egyszerre sajátította el, de később elfelejtett magyarul, majd felnőttkorban újra megtanult). Ennek eredményeként nála figyelhető meg a legkisebb mértékű akcentus, grammatikai hibákat szinte alig követ el magyar beszédében. Másrészt, ebben a csoportban ő volt a legfiatalabb, így beszédtempóját az is befolyásolta, hogy csak nemrég került ki a magyar nyelvű közezből.

Megvizsgáltuk a nemek eredményeit is (2. táblázat).

2. táblázat: A két nem beszédtempó-eredményeinek összevetése (h/s)

Nem nyelv	Gyermekek		Felnőttek		Idősek		Átlag	
	bolgár	magyar	bolgár	magyar	bolgár	magyar	bolgár	magyar
nők	6,7	7,9	9,6	9,4	8,6	8,3	8,3	8,5
férfiak	7,8	6,7	9,3	9,5	8,4	7,7	8,5	8



A két nem adatainak összevetése során nem találtunk jelentős eltérést. Ez arra utal, hogy a beszédtempó olyan egyéni jellegzetesség, amely (leszámítva az anyanyelv-elsajátítás korai szakaszát) inkább az életkortól és nem annyira a beszélő nemétől függ.

### **Következtetések**

Kiinduló hipotézisünkben arra kerestük a választ, hogy hogyan hat az életkor változása a kétnyelvűek beszédének időviszonyaira, illetőleg arra is, hogy a beszédtempó utalhat-e a nyelvdominanciára.

Az eredmények alapján igazolódott az életkor hatása a beszédtempóra – mindkét nyelv esetében a gyermekek beszéde volt a leglassúbb, valamivel gyorsabb volt az idős korcsoport beszéd, míg leggyorsabbnak a felnőttek bizonyult. Az életkori hatás elsődleges tényező, amely mindkét nyelvre kimutatható, és amelyet a kétnyelvűség ténye nem befolyásol.

A kétnyelvűek különböző korcsoportjainak bolgár és magyar beszédtempója között az összesített eredményekben gyakorlatilag nem találtunk különbséget. Ez a tény a kétnyelvűség egyik bizonyítéka, amelyet a csoportok nyelvileg heterogén összetétele sem módosít. Nem mutattunk ki eltéréseket a nők és a férfiak között sem.

A nyelvdominancia kérdésére az egyéni különbségek elemzésével kaphattunk választ. A legkisebb különbséget a két nyelv beszédtempójában az aktívan dolgozó felnőtt csoportnál találtunk, a gyermekeknél és az időseknél az eltérések nagyobbak voltak. Az okok valószínűleg a külső tényezőkben keresendők: az életmód, a foglalkozás és a nyelvhasználat különbözőségeiben.

A hallható akcentus összefüggésben van a beszédtempó adataival is. A különbségek nem kiugróan nagyok (egy ilyen eredmény kétségessé tenné résztvevők kétnyelvűségét), de a meglévő eltérések jól tükrözik a beszédtempó és a nyelvdominancia összefüggését.

Az elvégzett kísérletek új adatokkal gazdagították ismereteinket a kétnyelvűség működési mechanizmusáról. Sajnálatos tény, hogy a kétnyelvűséggel foglalkozó munkák általában csak a gyermekkori, esetleg a fiatal felnőttkori kétnyelvűséget vizsgálják pszicholingvisztikai szempontból. Az időskorú kétnyelvűek mentális folyamatainak elem-

zése számos új információval szolgálhat úgy a kétnyelvűség, mint az öregedési folyamatok nyelvi vonatkozásáról is.

### Irodalom

Bartha Csilla: A kétnyelvűség alapkérdései. Nemzeti Tankönyvkiadó. Budapest 1999.

Goldman-Eisler, F.: Psycholinguistics: Experiments in Spontaneous Speech. Academic Press. London 1968.

Gósy Mária: Tempóészlelés és beszédmegértés. In: Tanulmányok a beszéd időviszonyairól. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest 1988, 87-123.

Grosjean, F.: Individual bilingualism. In: Applied Linguistic Studies in Central Europe. I. Eds.: Lengyel, Zs. – Navracsics, J. – Simon, O. Veszprémi Egyetem. Veszprém 1997, 103-114.

Hoffmann, Ch.: An Introduction to Bilingualism. Longman. London 1991.

Kassai Ilona: Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? In: Beszédkutatás '93. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest 1993, 62-70.

Leopold, W.: Speech development of a bilingual child: a linguist's record. I. The vocabulary growth in the first two years. II. Sound learning in the first two years. III. Grammar and general problems in the first two years. IV. Diary from age two. Northwestern University Press. Evanston 1939, 1947, 1949a, 1949b.

Navracsics Judit: A kétnyelvű gyermek. Corvina. Budapest 1999.

Ronjat, J.: Le developoment du language observé, chez un enfant bilingue. Champion. Paris 1913.

**A tanulmány a Magyar Nemzeti és Etnikai Kisebbségekért Közalapítvány szíves támogatásával készült.**

# A VERSENGÉSI ELV A NYELVBOTLÁSOK LÉTREJÖTTÉBEN

Huszár Ágnes

Pécsi Tudományegyetem BTK Alkalmazott Nyelvészeti  
Doktori Program

## Bevezetés

Nyelvbotlásoknak a beszélő normatudatának meg nem felelő, attól eltérő, alkalmilag előforduló alakokat nevezzük. A nyelvbotlást a beszélő sokszor azonnal javítja is, hiszen feltűnik neki, vagy hallgatója figyelmezteti rá, hogy „a létrehozott szekvencia nem adekvát a mikrotervezési szándékkal” (Gósy 1999, 50). Létrejöttük mechanizmusát a témakör első kutatói a nyelvi elemek küzdelmével magyarázták. Wundt és Freud egyébként egymástól eléggé eltérő magyarázataiban közös az a feltevés, hogy alternatív beszédproduktions tervek „harca” hozza létre a nyelvbotlást. „A versengési hipotézis lényege az, hogy két, alternatív módon lehet ugyanazt a dolgot kimondani, vagy két, meglehetősen eltérő értelmezési nehézséget jelentő alternatív üzenetünk van, melyek mindketten ugyanazt a mentális és/vagy artikulációs helyet akarják elfoglalni. Freud és Meringer, bár az egymással versengő üzenetformák forrására és természetére nézve igen eltérő felfogást hirdettek, mindketten a versengési elvre hivatkoztak.,, (Garrett 1989, 307). Garrett megállapítása helytálló. Meringer, Wundt és Freud konkrét példák sorával próbálja bizonyítani, hogy a szócserek esetén küzdelem zajlik a két szó között a „tudat előterébe” kerülésért. Meringer még az egyszerű sorrendi hibák, az anticipációk és a perszeverációk esetében is úgy jellemzi a jelenség létrejöttét, hogy az „erősebb intenzitású” nyelvi elem kiszorítja helyéről a gyengébbet. Freud úgy magyarázza az általa elszólásnak (Versprechen) nevezett jelenséget, hogy valamely tudati (és nyelvi elem) kimondási tilalom, elfojtás alá kerül. Az elfojtás tökéletlen működése következtében azonban ilyen elemek is kimondódhatnak, leíródhatnak, habár töredékes módon, például kontaminációk elemeként.



A sorrendi hibák létrejöttének mechanizmusát az elmúlt évtizedekben lefolytatott kísérletek (magyarra lásd Huszár 1998) eredményeinek tükrében ma már másként látjuk. Van azonban egy elszólástípus, a szócsere, amelynek létrejöttében tagadhatatlanul fellelhető a vételkedés mozzanata. Ezt kívánom bizonyítani spontán magyar nyelvű beszédből származó nyelvbtlások segítségével.

### **A szócserek**

Már Meringer, a téma első kutatója, és egy 8500 adatból álló német nyelvű korpusz összeállítója is megkülönböztetett a szócsereknél:

1. hangzási hasonlóságon alapuló cserét,
2. fogalmi összefüggésen alapuló cserét, mégpedig
  - a. szinonim jelentésű,
  - b. antonim jelentésű szó fellépését a mondani szándékozott szó helyett (Meringer 1908, 46).

Létrejöttük okát ő abban látja, hogy több szó „küzd” azért, hogy a tudatküszöb fölé kerüljön. A nyelvbtlások, leglátványosabban pedig a szócserek esetében a kimondani szándékozott szó csatát vesz. A kimondani szándékozott szó helyére egy vele asszociációs kapcsolatban álló kerül. Így végeredményben a célszóhoz hangzásában hasonló vagy vele jelentésében rokon, szinonim vagy antonim jelentésű szó hangzik el.

Meringer, Wundt és Freud elképzelése a nyelvbtlások létrejöttéről egy nagyon egyszerű elméletet implikál. A beszéd folyamat az ő felfogásuk szerint – Meringer többször is használt metaforája szerint – egy sínpályákkal behálózott területhez hasonlít, amelyben a rossz váltóállítás okoz zavarokat azzal, hogy a nem megfelelő szerelvényt indítja el. Sem Meringer, sem Freud nem kételkedik abban, hogy beszéd közben mintegy „hangosan felolvassuk” a gondolatainkat. E mögött az elképzelés mögött mindössze egy szintet tartalmazó, „földszintes” modell áll.

### **A beszédproduktív folyamat szakaszai**

A gondolattól a beszédig vezető folyamat túlegyszerűsített elképzelését Lashley 1951-es, alapvető jelentőségű cikkében cáfolta meg. Bebizonyította, hogy a beszédproduktív szükségzerűen több,

egymást követő szakaszból áll. Példái a spoonerizmusok közül kerülnek ki, amelyekben az egymást követő szavak kezdő hangzói cserélnek helyet, például *font pordítva*. Ez a fajta elszólás a nevét egy W. A. Spooner nevű teológiai tanárról kapta, akinek mulatságos nyelvbotlásai közismertek voltak. Csak egy példát említünk Spooner tiszteletes gyöngyszemei közül: *The Lord is a shoving leopard to his flock* 'Az Úr nyájának lökdösődő leopárdja' *The Lord is a loving shepard to his flock* 'Az Úr nyájának szerető pásztorja' helyett. Lashley legfontosabb felismerése az volt, hogy a spontán beszéd és a gépirás sorrendi hibáiért az asszociációs kapcsolatok nem tehetők felelőssé. Bebizonyította, hogy szükségszerűen léteznie kell egy, a sorrendiséget és fonetikai reprezentációkat tartalmazó előkészítési fázisnak ahhoz, hogy ennek elemei térbeli és időbeli sorrendjükben helyet cserélhessenek.

A beszédprodukciónak szakaszosságára vonatkozó korai sejtéseket nem is oly régen cáfolhatatlan, természettudományos bizonyítékokkal támasztották alá. Turenout és munkatársai híres kísérletükben képeket mutattak a kísérletben részt vevő személyeknek. Az agykéreg elektromos aktivitásának mérése azt bizonyította, hogy a kép megpillantásától a rá vonatkozó nyelvi elem bizonyos morfológiai-szintaktikai jellemzőjének (nyelvtani nem) meghatározásáig 370 ms telt el. Ekkor a keresett szó még csak absztrakt entitásként jelent meg. További 40 ms-ra van szükség ahhoz, hogy a kísérleti személy fel tudja idézni a szókezdő beszédhangot, azaz a szó fonetikai reprezentációiban is rendelkezésére álljon.

Ezeket a felismeréseket is figyelembe véve egy adekvát beszédprodukciónak a következő részegységek működéséről kell számot adnia:

◆ **A konceptus szintje** (gondolat)

Itt zajlik a nyelvi közlés előkészítése. Még nem dőlt el, melyik konkrét nyelven fog a beszélő megszólalni. Az elemek elrendezése nem lineáris.

◆ **A konkrét nyelvi feldolgozás szintje.** Ezen a szinten zajlik a lexémakeresés, a szintaktikai, morfológiai szabályok alkalmazása, az egyeztetési szabályok végrehajtása. Az elemek hierarchikus elrendezésben kapcsolódnak egymáshoz.



- ◆ Az **artikuláció előkészítése**. Itt zajlik a már előhívott elemek lineáris sorba rendezése, a zöngésülés, az összeolvadás előkészítése.
- ◆ **Artikuláció**: az előkészített közlés megvalósul. Ennek a szakasznak a megléte nem feltétlenül kötelező, az írás esetében például el is marad.

### **Moduláris modellek**

A gondolatól, szándéktól a hangzó beszédig tartó folyamatot a kognitív paradigmába tartozó ún. moduláris beszédproduktions modellek legalább két szakasszá bontják fel.

Az autonómnak is nevezett modellek az elemek lineáris sorrendjét teszik meg legfőbb szerveződési elvüknek. Több, egymástól viszonylag függetlenül működő szakaszból, ún. modulból állnak. Ezek szigorú egymásutániségben lépnek működésbe, kiindulópontjuk az előző modul kimenete. Párhuzamos működésre nincs lehetőség, mivel az aktivációs hatás csak előrefelé terjed. Korrekcióra egy konzervatívan autonóm modellben tulajdonképpen nincs lehetőség. Az újabb, dinamikus modellek ezzel szemben, ha korlátozott mértékben is, megengedik az egyes szakaszok párhuzamos, automatikus működését. Ennek következtében tartalmazznak korrekciós lehetőségeket. A korai modellek közül azért idézzük fel Garrettét, mert modellje adekvátságának bizonyítására a nyelvbtlásokat is felhasználta.

Garrett egy 1976-ban közölt tanulmányában a beszédproduktio folyamatában két szakaszt (modult) különít el:

- ◆ a funkcionális szinten a mondat szintaktikai struktúrája jön létre,
- ◆ a pozicionális szinten a szintaktikai szerkezet váza fonológiai, prozódiai jegyekkel szerelkezik fel.

A két szintnek másfajta elsólási hiba felel meg. A szótévesztést Garrett a funkcionális szint hibájának tartja. Ilyenkor a szándékolt szó helyett egy másik, hangzásban esetleg nagyon is eltérő szó valósul meg. A szófaj általában megörzödik, tehát főnév helyére főnév, ige helyére ige kerül. A pozicionális szint hibái a fonémátévesztések. Ezek létrejöttében a szófaj nem játszik szerepet. A közvetlen nyelvi környezet hatása érhető tetten ennek az elsólásnak a létrejöttében. Garrett megemlíti még a szerinte ritkán előforduló ún. kevert cseréket



(*stranding exchanges, mixed errors*), melyek esetében mind a funkcionális, mind a pozicionális szint szerepet játszik.

Látszólag hasonlóan érvel Stemberger is. Ő 1985-ös tanulmányában (idézi Berg 1988, 26) szintén két modulhoz (tervezési és megvalósítási) rendel hozzá elszólástípusokat. A tervezési szakaszban jönnek létre az azonos szóosztályba tartozó lexikai elemek cseréjén alapuló elvétések. Például: *The list was 'nt in the word The word was 'nt in the list* helyett. A megvalósítási szakaszban létrejövő hibáknál viszont más-más szóosztályba tartozó elemek cserélnek helyet. Például *Did a lot of people different see it Did a lot of different people see it* helyett. Stemberger tehát mindkét szakasz jellegzetes elszólásaként sorrendi hibát említ. Ezek a példák önmagukban még nem bizonyítják a két szakasz független működését, hiszen mindkét esetben már meglevő reprezentációk cserélnek helyet.

### **Levelt modellje**

Az autonóm elméletek közül kidolgozottságával és magyarázó erejével kiemelkedik Levelt 1989-es modellje. Nagy figyelmet fordít a korai kognitív modellek által eddig elhanyagolt konceptus-szinten zajló folyamatokra. A modell jellemzője, hogy a konceptus szintet két szakaszra bontja, a makro- és a mikrotervezési szakaszra. A levelti modell a beszédtervezés következő szakaszait különíti el:

A **konceptuális tervező** a pragmatikai és szemantikai terület felelőse. Előhívja a hosszú távú memóriából a helyzethez illő ismereteket, így egy diskurzusmodellt hoz létre. A felügyelete alá tartozó tevékenységek:

**makrotervezés** – létrehozza a kommunikatív célok és alcélok szekvenciáját,

**mikrotervezés** – az üzenet egyes elemeit a kommunikatív célnak megfelelő helyükre teszi (topikalizáció).

Kimenet: propozícióformába rendezett preverbális üzenet, bármely konkrét nyelvi közlés alapjául szolgálhat.

Ezzel átlépünk a konkrét nyelvi feldolgozás területére. Az **átalakító**, melynek központi mozgatója, motorja a mentális lexikon, grammatikai kódolással olyan felszíni szerkezet hoz létre, amely szemantikai viszonyokat modellál szintaktikai funkciók segítségével. Az utolsó

modul, az **artikulációs tervező** feladata a beszéd motoros megvalósításának előkészítése. A levelti modell jellemzője, hogy benne csak az első fázisok, valamint az ellenőrző folyamatok kívánnak meg irányított figyelmet, a többiek lényegében automatikusan működnek. Ez magyarázza azt, hogy bár az egyes modulok rá vannak utalva az előzők végtermékére, de egyazon időben, egymástól függetlenül is működnek.

A beszédprodukción rendszer, az átalakító központi mozgatója a mentális lexikon. Ez tartalmazza azokat az információkat, amelyeknek segítségével a „preverbális üzenetből” mindenki számára érthető közlés lesz. A lexikon egyes elemei a következő tulajdonságokkal rendelkeznek:

- ◆ a szemantikai tulajdonságok a konceptuális feltételeknek felelnek meg, hiszen csak így választódnak ki;
- ◆ a szintaktikai tulajdonságok szintaktikai kategóriáknak felelnek meg;
- ◆ a morfológiai tulajdonságok olyan kategóriáknak felelnek meg, mint a flexiószerű végződések, pre- és szuffixumok;
- ◆ a fonológiai információk a szó szótagszerkezetére és az egyes fonémákra vonatkoznak.

Lényeges tulajdonsága a levelti modellnek az, hogy a lexikon nem egy helyen tárolja az adott nyelvi elemre vonatkozó szemantikai-szintaktikai és a fonológiai információkat. A szintaktikai, grammatikai és szemantikai információkat egy szemantikai vezérlésű lexikon tárolja. Ennek alapegysége a **lemma** (Levelt 1989, 6, 187). A morfológiai és fonológiai információt ezzel szemben egy formalexikon tárolja. Ennek alapegysége a **lexéma**. A két egység közötti kapcsolatot egy **mutató** (*lexical pointer*) hozza létre. A beszédprodukción folyamán először a lemma kerül elő, ezt követi a lexéma megkeresése.

Levelt és munkatársai kísérleti módon is bizonyították azt, hogy a lemma-szint elérése nem esik egybe a szó konkrét alakjának felidézésével. A kísérleti személyeknek képeket mutattak, majd olyan szavak írott képét, melyek szemantikai vagy hangzásbeli hasonlóságban álltak a képen ábrázolt dolgokat. Az eredmény az volt, hogy a szemantikai kapcsolat nem gyorsította meg a célzott szó megtalálását, mivel mind a kép, mind a szó a lemma keresésére irányult. Ezzel



szemben a hangzásbeli hasonlóság meggyorsította a feladat megoldását. Ez azzal magyarázható, hogy a hasonló hangzású szó megkönnyítette a lemmáról a lexémára való átlépést. (Levelt et al. 1991.)

Ezt a sorrendet és a két feladat elvégzésére fordított külön-külön időtartamot igazolja egyébként a Turennout és kollégái által végzett, már idézett vizsgálat is.

A mentális lexikon ilyenfajta megosztottsága és a két rész viszonylagos önállósága jól magyarázza az ún. *nyelvemen van (tip of the tongue)* jelenséget. Ebben az esetben pontosan tudjuk, „mire gondolunk” – a lemma már megvan – a szót mégsem tudjuk kimondani. Töredékes információink azonban annak hangzó alakjáról is lehetnek, például *két szótagból áll, d-vel kezdődik*. Néha szerencsénk van, és beszélgetőtársunk kiségit bennünket, kimondja a számunkra csak töredékesen megjelenő lexémát.

A mentális lexikonban való keresés egyébként ritkán hoz felszínre csupán lexémákat. Ez csak abban a ritka és mesterséges helyzetben fordul elő, amikor az embernek olyanféle kérdésekre kell választ adni, hogyan is hívják azt a vörösesbarna vagy szürke bundás, erdei fákon élő kis rágcsálót, vagy felszólítják, hogy nevezze meg a képen látható tárgyat vagy élőlényt. Erre a kérdésre felidéződik bennünk a lemma, de ha a kérdező hozzáteszi, hogy a keresett szó öt beszédhangból (betűből) áll és *m*-mel kezdődik, a lexémát is azonnal fel tudjuk idézni.

A természetes beszédhelyzetben ez nem így van. A beszédtevékenység során nemcsak lemmákat, lexémákat emelünk ki a lexikonból, hanem velük együtt szintaktikai struktúrákat is. A mentális lexikonban, mégpedig annak mindkét részében nemcsak lemmák, illetve lexémák, hanem azoknak szemantikai jegyekkel, szintaktikai struktúrákkal, hangtani jellegzetességekkel jellemezhető holdudvara is jelen van. Ezekből a szintaktikai struktúrákból a közlési szándéknak megfelelő átalakítások – topikalizáció, mozgatási, törlési szabályok végrehajtása – után áll elő a közlésre szánt mondat.

### **A lemmakeresés során létrejövő nyelvbtlások**

Nyelvbtlásokat létrehozó defektusok a beszédprodukciónak mechanizmusának több szakaszán is bekövetkezhetnek. Ha már a lemma-



keresés fázisában, a mentális lexikon szemantikailag irányított részében való kereséskor bekövetkezik a defektus, a hibás útra terelt folyamat végén megjelenő lexéma hangalakjában általában nem is hasonlít a szándékozottra, mint például az egy meteorológiai jelentésben elhangzott *vékonyul az ég* – *vékonyul a felhőzet* helyett. Ezek a hibák általában az asszociációs pályák működését bizonyítják: antonimák, szinonimák, metonimikus, rész-egész kapcsolatok aktiválódnak. Antoníma lép a célszó helyére a következő példákban: *a jobboldal vörös sárkánya* – *a baloldal vörös sárkánya* helyett, *vandálok késői ősei* – *leszármazottai* helyett. Szinonima választódik ki ezekben az esetekben: *szeptás* – *hívó* helyett, *a merénylet végrehajtása* – *kísérlet végrehajtása* helyett, *foltot vethet* – *árnyékot vethet* helyett. Rész-egész viszonyban álló elemek cserélnek helyet ebben a mondatban: *a ravaszt nem hivatásos vadász fogja* – *a puskát nem hivatásos vadász fogja* helyett.

A lemmakeresés szakaszában bekövetkező elvétések közé tartozik az is, amikor végeredményben egy szó saját két alternatív, egymást kizáró vonzatstruktúrájának elemei egy szintagmában lépnek fel. Például *átlépte a Rubiconon* – *átlépte a Rubicont* helyett, *visszagondolják azt* – *visszagondolnak arra* helyett.

A mentális lexikon az egyes szavak mellett tartalmazza a több elemből álló szókapcsolatokat, az ún. frazeologizmusokat is. Ezek egy része mögött elhalványult köznyelvi kép áll. A lemmakeresés defektusai közé tartozik ezeknek a frazeologizmusoknak hibás felidézése is. Gyakran előfordul az, hogy a hasonló jelentésű frazeologizmusok elemei összekeverednek. A végeredmény képzavar lesz. Néhány példa: az *elfogyott a cérna* az *elszakadt a cérna* és az *elfogyott a türelem* keveredéseként jött létre. A *tücsköt bogarat kiabál* a *kígyót békát kiabál* és a *tücsköt bogarat összehord*, a *végét rúgja* pedig az *utolsókat rúgja* és a *végét járja* elemeiből állt össze.

### **A lexémakeresés során létrejövő nyelvbtlások**

Ha a működési zavar egyértelműen a következő stádiumban, a formalexikonban, a lexémakeresés során lép fel, a végeredmény a szándékozotthoz hangzásában hasonló lexéma lesz. Ilyenek a *megromlott a viszály köztük* – *viszony* helyett, *kábítószerek-kereskedő kutyák* –

*kábítószer-kereső kutyák* helyett, *értelmetlenül néz* – *értetlenül néz* helyett. Ezek a lexémák hangalakjukban alig térnek el, de jelentésükben sem állnak messze egymástól, egyazon szemantikai mezőhöz tartoznak.

A lexématalálás zavarai különösen gyakran érintik a tulajdonneveket, azon belül pedig a személyneveket. Ennek oka a tulajdonnevek szemantikai rigiditásának is nevezett sajátosságában kereshető. A tulajdonnevek egy része csak referenciális kapcsolatban áll az általa jelölt dologgal vagy személlyel, s nem tartozik bele egy sokféle kapcsolatot implikáló szemantikai mezőbe. A referenciális kapcsolat létrejöttét és felidézését pedig zavarják a homonimák. (Nem könnyű eligazodni a sok gyakori családnevű és keresztnévű ember között.) Nem csoda hát, hogy a tulajdonnevek gyakran válnak elszólások alapjává. Így lesz *Szent Pál*-ból *Szent Péter*, *Ádler*-ből *Áder*, *Endrényi*-ből *Endrei*.

A lexémakeresés zavarai okozzák a malapropizmusok létrejöttét is. Ez az elszólástípus Sheridannek egy 1775-ben, Londonban bemutatott darabjának egyik szereplőjéről, Miss Malapropról kapta a nevét. Ez a hölgy főleg műveltségzavakat kever össze egymással, így például *superstitious*-t mond *superfluous* helyett. A malapropizmusokra az a jellemző, hogy olyan, hangzásukban egymáshoz nagyon hasonló lexémák lépnek egymás helyére, melyeknek jelentésánál semmi közük nincs egymáshoz. Így lesz a *rabiátus*-ból *rabinátus*, az *interpellál*-ból *introvertál*, a *restaurál*-ból *retusál*. Az összetételi tagok egyikének rosszul tudása hozza létre a *humánteológia*-t a *humánételológia* helyett, a *garzonréteget* az *ózonréteg*, a *mágnáskártyá*-t a *mágneskártya* helyett.

A lemma- és lexémakeresés fázisainak közös hibája okozza a szóalakvegyüléseket, kontaminációkat. Esetükben a szemantikai mezőben egy helyett két lemma választódik ki, például *politikus* és *pszichológus*. Ezek a formalexikonban keresve több lexéma töredékeit tartalmazó hibrid alakot emelnek ki: *polipszichológus*. Hasonló mechanizmus alapján lesz a *balszerencse* és a *szerencsétlenség* szerencsétlen egyesüléséből *balszerencsétlenség*, a *tudálékoskodás* és az *álokoskodás*-éből *tudálokoskodás*. A *kirendez* példa esetében az *elrendez*



igéje és a *kifizet* igekötője cseréltek helyet. Ez a fajta kontamináció a nyelv paradigmaticus tengelye körül szerveződő kapcsolat terméke.

Ebbe a csoportba tartozik egyébként a freudi elszólások egy része is. Freud leír egy esetet, mikor egyik nőbetege arra a kérdésre, hol található a háza, ezt válaszolta: *An der Berglende*. A szándékolt válasz ez lett volna: *An der Berglehne* 'a hegyoldalon'. A szóösszetétel második helyére azonban a *Lende* 'ágyék' szó furakodott be, az analitikus magyarázat szerint a szexuális tartalmú elfojtás tökéletlen működésének következményeként (Freud 1991, 60).

Előfordulnak egyébként a szintagmatikus tengely körül szerveződő nyelvbtlások, szóösszevonások is. Ilyen az *Arany születésében írta* a helyett, hogy *Arany születésének évében írta*, az *összefogható* az *összefüggésbe hozható*, a *szervezőinek* a *szervezetei vezetőinek* helyett.

Néha a kontamináció az egész mondatra kiterjed, például *Szinetárra jár a neve*. A beszélő szándéka szerint valami ilyesmit mondott volna: *Azon jár az eszem, nem Szinetár-e a keresett név*.

A spontán beszédből gyűjtött elszólások között vannak olyanok is szép számmal, amelyeknek létrejöttében több ok is szerepet játszhat. Jó példa erre az a válasz, amelyet egy televíziós vetélkedő egyik versenyzője mondott az idézett vers címét találgatva: *Április*. A helyes válasz *Áprily Lajos Március* című verse lett volna.

### **A nyelvbtlások okai**

Felvetődik a kérdés, miért, milyen okból jönnek létre nyelvbtlások. Tudomásul kell vennünk, hogy ezek a kis defektusok a mindennapi beszédtevékenység elkerülhetetlen, kevés zavart okozó velejárói. Sokszor maga a beszélő, sőt hallgatója sem veszi észre a spontán beszédben átlagosan 10-15 percenként elkövetett apróbb bakikat, elvételeket. Fáradság vagy stressz, alkohol vagy drog hatására megnövekszik a hibázások száma. Beszélni tanuló kisgyermeknek gyakrabban hibáznak, mint a felnőttek. Az agyi sérülésekre visszavezethető betegségeknek, az afáziáknak pedig szimptomái az nyelvbtlásokhoz hasonló hibázások.



A beszédtevékenység gyorsaságához képest csekély számú hibát általában az okozza, hogy az egymást követő szakaszok (modulok) kapcsolódásában zökkenő következik be.

A legfőbb külső ok a szöveggörnyezet hatása. Az adott elemet megelőző szövegrészben előforduló szavak, szókapcsolatok nagyobb valószínűséggel lépnek a kimondani szándékozott szavak helyére, mint azok, amelyek nem szerepelnek a szöveggörnyezetben.

A másik fontos ok a nyelvi elemek gyakorisága. A gyakrabban előforduló nyelvi elemeknek nagyobb esélyük van arra, hogy ritkább társaik helyére kerüljenek, mint fordítva. A *telekspekuláns* minden bizonnyal gyakoribb a mai szóhasználatban, mint a *kamatspekuláns*. Egy viszonylag hosszú televíziós beszélgetés során a riporter következetesen *telekspekuláció*-ról beszélt, amellyel egy újság vádolta beszélgetőtársát. A bemutatott újságkivágás bizonyította, hogy az újság *kamatspekuláns*-nak nevezte az urat. Sem a riporter, sem maga az érintett megkérdezett nem vette észre a nyelvbtlást, úgy beszélgettek, mintha a szándék és a végeredmény nem különbözött volna egymástól.

A legritkábban előforduló, de néha igazán meghökkentő hatást kiváltó ok a nyelven kívüli környezet lehet. Az ember beszéd közben szemével pásztázza környezetét, s néha előfordul, hogy a látott tárgyak megnevezése nyelvbtlásként beleszövődik a beszédébe. A *szerelemi együttlég*y – mondta egy orvos egy rádióműsorban. Rögtön ki is javította magát, *természetesen a szerelemi együttlétre gondoltam*, s rögtön okát is adta nyelvbtlása okának: *csak ez az átkozott légy állandóan itt röpköd az orrom előtt*.

### Irodalom

Berg, T.: Die Abbildung des Sprachproduktionsprozesses in einem Aktivationsflussmodell. Niemeyer. Tübingen 1988.

Freud, S.: A mindennapi élet pszichopatológiája. Cserépfalvi. Budapest 1991.

Garrett, M. F.: Syntactic processes in sentence production. In: Wales, R. J.–Walker, E. C. (eds.): New approaches to language mechanisms. North Holland. Amsterdam 1976, 231-256.

Garrett, M. F.: A feldolgozási szint a mondatprodukción. In: Pléh Csaba (szerk.): A beszédmegértés és a beszédprodukción pszichológiája. Szöveg-

gyűjtemény. Tankönyvkiadó. Budapest 1989, 275-322. (első megjelenés: 1980)

Garrett, M. F.: Processes in language production. In: Newmeyer, F., (ed.): Linguistics: The Cambridge Survey. Vol. III. Language. Psychological and Biological Aspects. CUP. Cambridge 1988, 6-96.

Gósy Mária: Pszicholingvisztika. Corvina. Budapest 1999.

Huszár Ágnes: Az elszólásvizsgálat egy lehetséges módja. In: Gósy Mária (szerk.): Beszédkutatás '98. MTA Nyelvtudományi Intézete. Budapest 1998, 21-34.

Lashley, K. S.: The Problem of Serial Order in Behaviour. In: Saporta, S. (ed.): Psycholinguistics. A Book of Reading. Holt, Rinehet and Winston. New York 1961, 180-197.

Levelt, W. J. M.: Speaking. From Intention to Articulation, MIT Press. Cambridge, Massachusetts 1989.

Levelt, W. J. M. – Schiefers, H. – Vorberg, D. – Meyer D. – Pechmann T. – Havinga, J.: The time course of lexical access in speech production. A study of picture naming. Psychological Review 98. 1991, 122-142.

Meringer, R.: Aus dem Leben der Sprache. Versprechen, Kindersprache, Nachahmungstrieb. Behr's Verlag. Berlin 1908.

Turenout, van M. – Hagoort, P.– Brown, C. M.: Brain activity during speaking: From syntax to phonology in 40 milliseconds. Science vol. 280, April, 1998, 572-574.

# MAGÁNHANGZÓ-RENDSZEREK ELEMZÉSE INFORMATIZÁLT NYELVJÁRÁSI KORPUSZON

Vékás Domokos  
ELTE Fonetikai Tanszék

## Bevezetés

Dialektológiai célú informatizálással régen megálmodott lehetőségeket teremthetünk meg, új kutatási eszközöket hozhatunk létre. A multimédiás archívumokba rendezett hanganyagok a képernyőn megjelenő térképlapokról gyorsan elérhetőkké és könnyen megszólaltathatókká válnak, továbbá azonnal rendelkezésünkre állnak fonetikai (akusztikai) elemzésekhez is. E lehetőségek és egyben feladatok újában alapot teremtenek fonetika és dialektológia szoros együttműködéséhez külföldön és itthon egyaránt (pl. Bruce et al. 1999; Vékás 1999; Juhász és mtsai 2000).

Fölmerül azonban a kérdés: a hagyományos papír-nyelvatlaszok és nyelvföldrajzi szótárak meg nem szólaló és akusztikailag nem vizsgálható, ám nagy mennyiséget kitevő adatai — megfelelő számítógépes rögzítés és rendszerbe illesztés után — felhasználhatók-e hangtani kérdések megválaszolásában?

A továbbiakban a még csak részben megjelentetett *A romániai magyar nyelvjárások atlasza* két informatizált kötetét korpusznak tekintve kíséreljük meg a 136 kutatópont magyar magánhangzó-rendszerét kvantitatív módon vizsgálni. A településenként közel kétezer magánhangzónyi adatállomány informatizálása és feldolgozása a fejlesztés alatt álló Bihalbocs számítógépes program segítségével történt. Az atlasz a nyelvterület egy jelentős és változatos részét fogja át, a tíz magyar nyelvjárás régióból hat teljesen vagy részben beletartozik. Hagyományos (papír) változatának kiadása jelenleg is folyik Juhász Dezső irányításával, elkészültekor tizenegy kötetével az egyik legnagyobb nyelvjárás adattár lesz (bővebben lásd pl. Murádin–Juhász 1995, 17-18; Juhász 1997).

Vizsgálódásunk során központi feladatnak tekintjük a rövid-hosszú szembenállás szilárdságának feltárását a fonetikus lejegyzések által



hordozott információ számbavételével és értelmezésével. Ezzel összefüggésben bemutatjuk az alsóbb nyelvállású (rövid és hosszú) magánhangzó-minőségek együtt járását, rámutatva a félhosszúként lejegyzett magánhangzók és a diftongusok szerepére is.

### **A lejegyzések pontossági szintje és használhatósága**

A gyűjtést és fonetikus lejegyzést minden kutatóponton ugyanaz a nyelvész, Murádin László végezte, ezért az adatok egybevetésre kiemelkedően alkalmasak.

Ha informatizált adatok állnak rendelkezésünkre, könnyen kiszámoltathatjuk az egyes hangzóárnyalatok arányát, szótagolási „szabályok” formalizálásával pedig (a megvalósíthatatlanságig hosszadalmas manuális munkát megtakarítva) informatikai úton juthatunk szótagviszonyokra lebontott adatokhoz.

Lejegyzett adatok informatizálása során megkerülhetetlen a lejegyzési rendszerek (elsősorban az IPA és az egyezményes hangjelölés) dialektológiai használhatóságának a kérdése. Nagyban egyszerűsítene ugyanis a feladatot, ha csak a hagyományos vonulat alapján működtetnénk az informatikai rendszerünket, a másik legfeljebb úgy jöhetne szóba, mint az eredményeknek (állítható konverzió utáni) egyik prezentációs változata, elsősorban nemzetközi fórumokon. Számos konkrét kérdés azonban, például az *á* mellékjelezése, egyszerre jelent fonetikai és informatizálási problémát. Ha az *á* (morfológiai, toldalékolási okokból) velárisnak minősül, zártabbnak jelölt változata az *a*-t közelíti. Márpedig az *á* IPA-szimbólumának [a:] a besorolása egyértelműen palatális; ha mellékjellel egy enyhén magasabb nyelvhátra utalunk, az eredmény az *e* felé mutat, nem az *a* felé. Az ilyen problémák egyelőre megoldatlanok, ezért is használunk egyezményes szimbólumokat az informatizálási munkafolyamatokban.

Informatikai szempontból a finoman mellékjelezett lejegyzések azért értékesek, mert nem sematikus fonetikus átírás eredményei. A magyar hangtani-nyelvjárástani hagyományt éppen az jellemzi, hogy a nyelvi valóságot a mozgás és változás sokféleségében ragadja meg, az éppen elhangzó megvalósulást rögzíti, illetve több rokon fonetikai alakot dokumentál. A lejegyzés így kevesebb interpretációt, és ezért kevesebb önkényes értékelést tartalmaz, kevesebb elméletfüggő meg-

oldással jár, ezáltal pedig alkalmasabb az igényes feldolgozásra. Egy település nyelvváltozatáról így nem azt tudjuk meg (elsősorban), hogy funkcionálisan hány magánhangzót állít szembe, hanem például azt, hogy mekkora a köznyelvi *e*-nél egy-két árnyalattal nyíltabb minőségűnek ítélt *e*-féleségek (palatális, illabiális, alsó nyelvállású vagy még nyíltabb magánhangzók) aránya. Márpedig feltételezhetjük, hogy minél gyakrabban észlel a lejegyző nyíltabb változatot a különböző beszélőközösségekben, annál nyíltabb *e* hangtípus jellegzetes megvalósulása a szóban forgó nyelvjárásokban. Ha a viszonyítási alapul szolgáló köznyelvi (pontosabban: köznyelvinek tartott) hangzominőségek spektrális jellegzetességeit ismertnek vesszük, akkor az egyes hangárnyalatok gyakorisága alapján megjósolható az eltérések mértéke, a várható formánsértékeket pedig akár számszerűen is megbecsülhetnénk. A lejegyzett változatok arányából tehát az egyes hangtípusok minőségére következtethetünk: mennyiségi adatokból becült minőségi adatokhoz juthatunk. A továbbiakban a vizsgált nyelvváltozatokat egy-egy magánhangzó kvantitatív módon becült nyíltsági mértéke szerint (várható formánsértékek szerint) fogjuk sorrendezni.

### **Egy klasszikus kérdés: van-e „zárt e” a magyarban?**

Induljunk ki egy klasszikus, magánhangzó-rendszert (és a különböző hangtáblázatokkal szembesülő kisiskolások és egyetemi hallgatók tömegeit) érintő kérdésből: van-e „zárt e” a magyarban? A klasszikus válaszoknak két fajtáját ismerjük: igen, mert a nyelvterület nagyobbik részén a beszélők jelentős része, főleg az idősebbek, megkülönböztetnek egy nyíltabb és egy zártabb változatot; illetve: nem, mert a legtöbb beszélő, főleg a fiatalabbak, nemigen érzékelnek ilyen különbséget.

Egy árnyaltabb válasz megfogalmazásához figyelembe vehető, hogy korpuszunk azon nyelvjárásaiban, ahol lényegében csak egy *e*-árnyalat van, annak minősége nem zártabb, mint a viszonyítási minőség. Márpedig a köznyelvi *e* minősége is inkább nyíltabbnak tekinthető: kontrasztív vizsgálatokban magasabb F1- és alacsonyabb F2-értékeket mutat, mint például a középső és alsó nyelvállású illabiális palatális egymástól megkülönböztető francia vagy olasz (toszán)



nyíltabb változata (pl. Herman 1984). Ha nincs is meg a „zárt e”, a rövid magánhangzók rendszerében a helye megvan: a hangminőségbeli különbség *e* és *i* között nem két szomszédos magánhangzóra vall, és az *ür* megmaradni látszik: az *e* nem záródik, nem nyomul a szabaddá vált akusztikai-percepció terület felé.

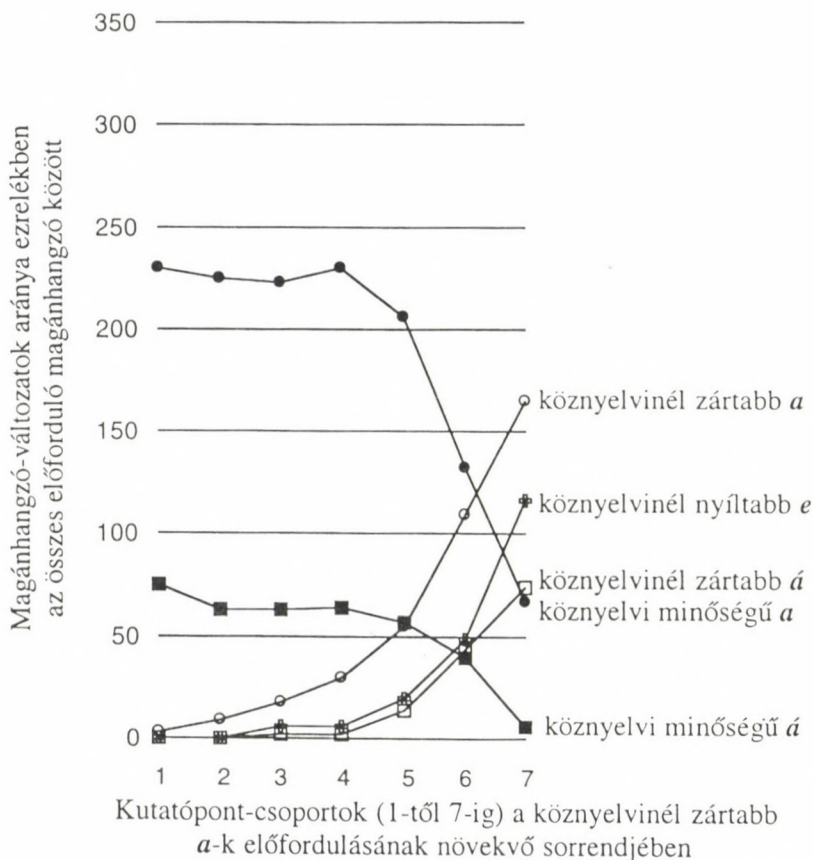
A fenti gondolatmenet azt feltételezi, hogy a rövid magánhangzók rendszere önálló, és a hosszú magánhangzók nem játszanak jelentős szerepet a rövid minőségek akusztikai-percepció elhelyezkedésében. Valójában a rövid *e* nem csak az esetleg meglévő középső nyelvállás rövid szomszédjával, de például a hosszú *é*-vel is versengésben áll, különbségükben a minőség releváns szerepet játszik (a köznyelv vonatkozásában vö. pl. Kovács 1998). A „zárt e” helyzete talán éppen ott tekinthető bizonytalanabbnak, ahol — a rövid-hosszú szembenállás felbomlásának folyamatában — például *e* és *é* elkülönülésében is egyre nagyobb szerepet játszik a minőség, az időtartam rovására.

### **Hogyan függ össze rövid és hosszú magánhangzók minősége?**

A magánhangzó-kvantitás megléte klasszikus értelmezésben azt jelenti, hogy minőségükben azonosnak tekinthető magánhangzók pusztán időtartamuk alapján különbözhetnek. Ez csak részben áll a magyarra. Gyakran előfordul, hogy egy nyelv történetében a rövid és hosszú vokálisok szembenállását minőségi különbségek erősítik: ilyenkor az oppozíció fonetikailag akár stabilabbnak is tekinthető. A rövid-hosszú megkülönböztetés ugyanakkor törékennyé is válhat: az elkülönülő minőségek önmagukban is biztosíthatják a különbségtételt. A legjobban ismert példát a latin vokalizmus története szolgáltatja, ahol szempontunkból legalább három szakaszt szokás elkülöníteni: az első (latin) tisztán kvantitáson alapul, a másodikban (preromán) együtt él a kvantitás és a minőség, a harmadik (román) szakaszt kizárólag az utóbbi jellemzi (Hagège—Haudricourt 1978, 53-54). E szakaszolási kísérlet azt szemlélteti, hogy hosszúságok és minőségek történetileg változó súllyal vesznek részt a magánhangzó-rendszerek kialakításában. Engstrand (1986) eszközfonetikai vizsgálataira támaszkodva kifejezetten hosszúsági/spektrális kontinuumról ír, amely-



nek különböző pontjain helyezhetők el az egyes nyelvek<sup>1</sup>. Spontán beszéd vizsgálata is hasonló eredményeket mutat (Engstrand—Krull 1994). A továbbiakban feltételezzük, hogy a különböző magyar nyelvváltozatok is e fonetikai kontinuum egy-egy (nem feltétlenül azonos) pontjához tartoznak.



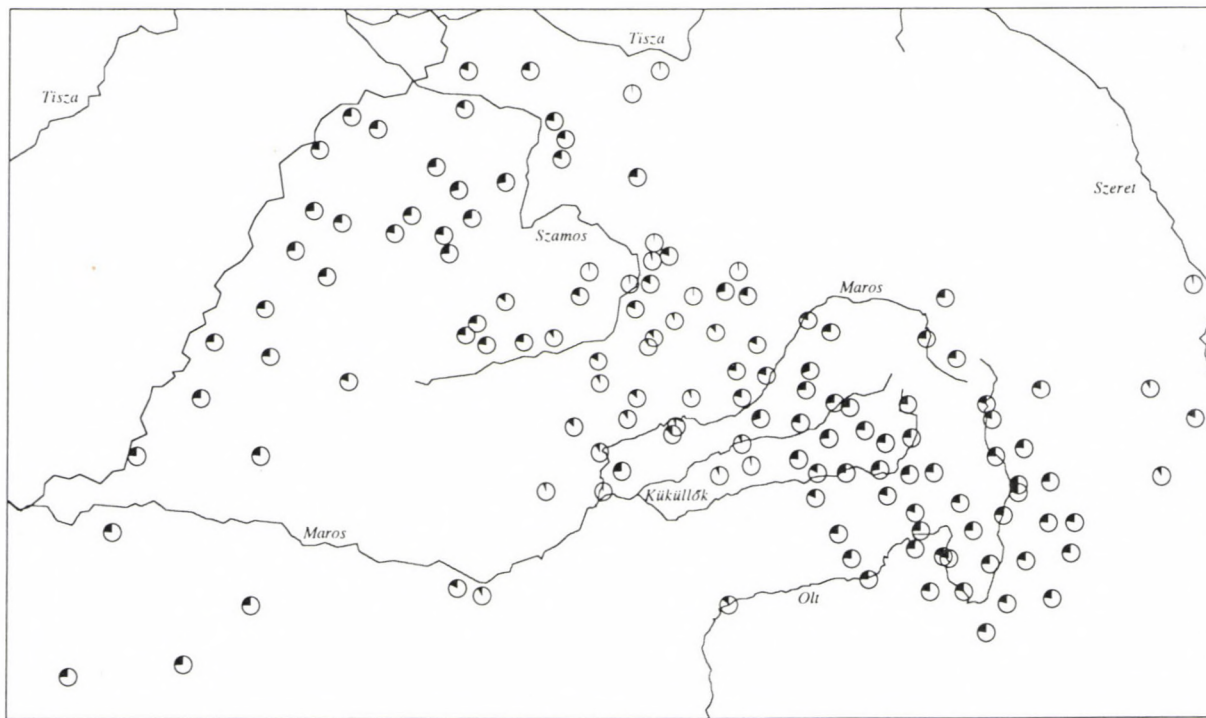
1. ábra: egyes magánhangzó-változatok együtt járása

<sup>1</sup> "... a phonetic quantity typology with Finnish and Czech near one end of a duration vs. spectrum continuum, and with Swedish near the other end."

Korpuszunkat elemezve az alacsonyabb nyelvállású magánhangzó-minőségek összefüggéseit fedhetjük fel. Az 1. ábra 7 kutatópont-csoporthoz tartozó, magánhangzó-mennyiségekre vonatkozó átlagértékeket mutat: a 136 települést a zártabbnak minősített *a*-k számaránya szerint sorrendeztük, és emelkedő sorrendben 7, egyenként mintegy 20 településből álló csoportot határoltunk el. A zártabb *a*-k előfordulásának növekedésével együtt jár a nyíltabb *e*-k és a zártabbnak jelzett *á*-k mennyiségi növekedése. Statisztikailag e három mennyiség (és vonatkozó minőség) egymással szignifikáns mértékű pozitív korrelációban áll (a korrelációs együttható *á* és *e* vonatkozásában  $r = 0,953$ ; *a* és *á*:  $r = 0,788$ ; *a* és *e*:  $r = 0,781$ ). Az ábrán látható a köznyelvi minőségű *a* és *á* változatok ellentétes irányú, tehát csökkenő mennyiségi alakulása is. Az egyes lejegyzett árnyalatok arányából a várható jellegzetes minőségekre következtethetünk: balról jobbra haladva fokozatosan egyre zártabb *a*-kat, egyre nyíltabb *e*-ket és egyre inkább az *a* irányába mozduló *á*-kat feltételezhetünk.

A vizsgált nyelvjárásokban tehát az *e* minőségét (nyíltsági árnyalatát) jó eséllyel megjósolhatjuk az *a* minőségéből, de az *á* minőségéből is: a hosszú *á* is részt vesz az alacsonyabb nyelvállású magánhangzók minőség-átrendeződésében, amelynek lényege egy optimális hangminőségbeli távolság fenntartása lehet. Ez a most nagy mintán grafikusán is bemutatott és statisztikailag is igazolt összefüggés nem ismeretlen a hagyományos dialektológiában sem, ám nevet nem kapott: az egyezményes hangjelölés mögötti fonetikai szemlélettel inkább csak az artikulációs terminológiával kifejezhető tendenciák képezhetők el (például: nyíltabbá válás, labializáció). Itt azonban olyan jelenséggel van dolgunk, ahol nyíltabbá válás és zártabbá válás együtt jelentkezik. A hagyományosan különböző irányúnak tekintett eltérések közös lényegét viszont jól megragadhatjuk egy egyszerű beszédakusztikai elem, a második formáns (F2) segítségével. Az alacsonyabb nyelvállású magánhangzók (*e*, *á*, *a*) bármelyikének F2 süllyedése/emelkedése a másik két magánhangzó azonos irányú és hasonló mértékű F2-változásával jár együtt: F2-eltolódásról beszélhetünk (amit még kísérleti úton is pontosítani szükséges).

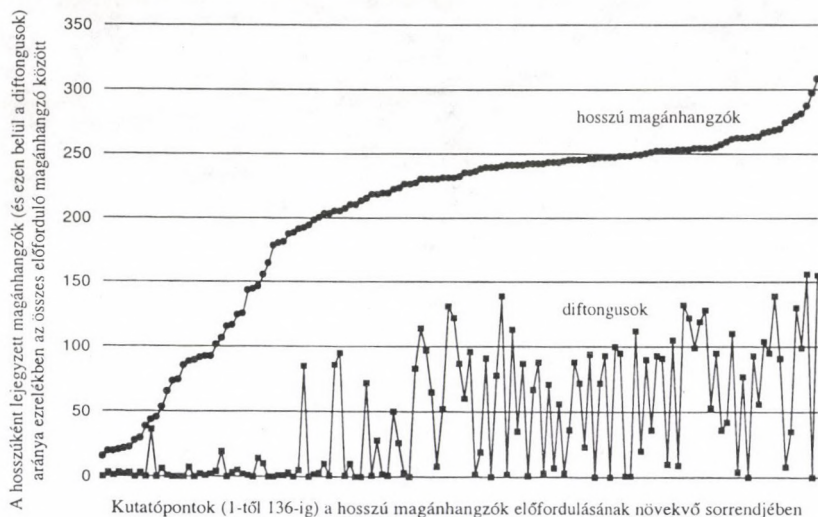
Megjegyzendő, hogy veláris irányban a tendencia a középső nyelvállás magasságában megáll: az *o* nemigen záródik záródó *a* hatá-



2. ábra  
 A hosszúként lejegyzett magánhangzók aránya (fekete köröckek)  
 az összes előforduló magánhangzó között, 136 kutatóponton



sára, így több településen — szintén a mennyiségi adatokból következtetve — *a* és *o* között nem éles a határ.



3. ábra

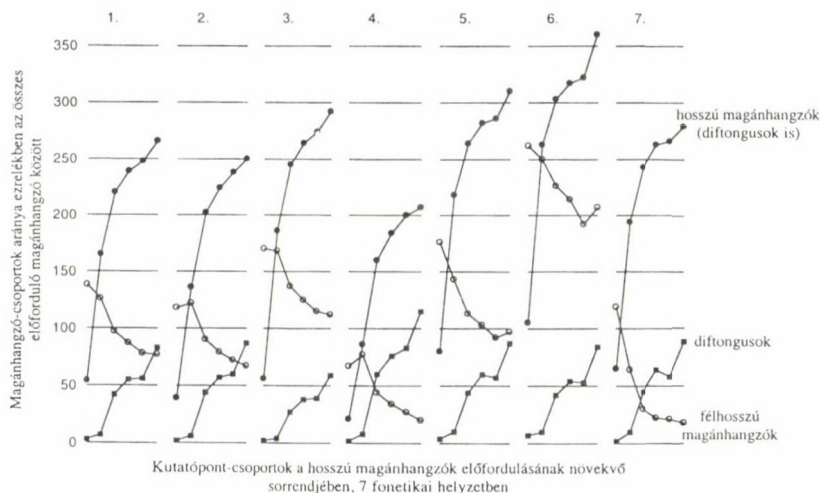
Hosszúként lejegyzett magánhangzók  
(és ezen belül a diftongusok) aránya kutatópontonként

### Mennyire nyugszik szilárd fonetikai alapokon a magánhangzó-kvantitás?

Láttuk, hogy az alacsonyabb nyelvállású magánhangzók — tekintet nélkül időtartamukra — egymáshoz képest optimálisan elkülönülő minőségűek. E megfigyelés érvényessége igen változatos nyelvjárásokra terjed ki, tehát a magyar magánhangzórendszer(ek) lényeges tulajdonságáról van szó: a magyar nem lehet túl közel a korábban idézett hosszúsági/spektrális kontinuum hosszúságot maximalizáló végéhez. Az informatizált nyelvjárási korpusz lehetőséget ad az időtartamon alapuló magánhangzó-kvantitás fonetikai szilárdságának tesztelésére is. Nem a szembenállás pusztá fonológiai tényét vizsgáljuk (ami szinte minden magyar nyelvváltozatra igaz), hanem a fonetikai hátteret. Egy ilyen oppozíciót akkor tekinthetünk ugyanis szilárdnak, ha rövid és hosszú megvalósulásból is kellő mennyiséget talál-

lunk, és ha a félhosszúnak ítélt magánhangzók száma alacsony, ugyanis ezek is homályosíthatják, tompíthatják a szembenállást, amennyiben a rövid és a hosszú kategória időtartam szerinti nagyobb mértékű átfedését jelzik.

Általánosságban az mondható el (2. és 3. ábra), hogy a vizsgált települések legnagyobb részében korpuszunk 20% feletti (nemritkán 25% feletti) fonetikailag hosszú magánhangzó-arányt mutat, de számos kutatóponthoz jóval kisebb, 11 esetben pedig 5% alatti értékek tartoznak. Utóbbiaknál a rövid-hosszú szembenállás megbomlásának többé-kevésbé előrehaladott állapotát feltételezhetjük. Nem szerepel az ábrán, de említést érdemel, hogy a rövidülési tendenciával szemben legellenállóbb magánhangzó az *á*, amelynek inherens hosszúsága és intenzitása jól ismert a legnagyobb.



4. ábra

Hosszú és félhosszú magánhangzók, valamint diftongusok aránya  
 1. minden helyzetben, 2. hangsúlytalan szótagban, 3. hangsúlytalan, zárt sz.;  
 4. hangsúlytalan, nyílt sz.; 5. hangsúlyos szótagban; 6. hangsúlyos, zárt sz.;  
 7. hangsúlyos, nyílt sz.

A 4. ábra hét fonetikai helyzetet jelenít meg (az első rész minden helyzetet összesít). A 136 települést a hosszúnak minősített magánhangzók számaránya szerint sorrendeztük, és emelkedő sorrendben 6, egyenként mintegy 22 településből álló csoportot határoltunk el. Mind a hét fonetikai helyzetben e csoportokra jellemző, magánhangzó-kategóriák számarányára vonatkozó átlagértékeket olvashatunk le. Jól látható, hogy az egyes helyzeteket bemutató részabrákon belül balról jobbra haladva, a hosszú magánhangzók arányának növekedésével a félhosszú magánhangzók aránya csökken, a diftongusok (a hosszú magánhangzók egy részhalmaza) előfordulása nő. Statisztikailag a negatív ( $r = -0,596$ ) és a pozitív ( $r = 0,533$ ) korreláció szoros összefüggést mutat.

Kiemelendő, hogy a hangsúlyos és a hangsúlytalan helyzetre egyaránt igaz: zárt szótagban található arányosan a legtöbb hosszú realizáció, akár kicsi, akár nagy a hosszú vokálisok abszolút száma, jóllehet más nyelvekben a fonetikailag (nem fonológiai) hosszú magánhangzók inkább nyílt szótagban tekinthetők természetesebbnek. Az eltérés azzal függhet össze, hogy a magyar vokalikus rendszer alapvetően mégiscsak az időtartam szerinti szembenállásra épül: a hosszúság arányos (és inkább az univerzálisnak tekinthető tendenciákkal szembenemő) disztribúciója eleve kizárja a szótagviszonyok szerinti reinterpretálásának lehetőségét.

A diftongusok jelenléte összefügg a hosszú és félhosszú magánhangzók számával, végső soron az oppozíció stabilitásával: minél több a diftongus, annál több hosszú magánhangzóra számíthatunk. Úgy tűnik, a diftongusoknak kérdésünk szempontjából stabilizáló szerepük van, esetleg azért is, mert a két artikulációs célpontot is megvalósító kettőshangzók kevésbé lehetnek hajlamosak a rövidülésre, és e tulajdonságuk (időzítesi jellegzetességük) áttevődhet más hosszú magánhangzók produkciójára is.

### **Mi várható a köznyelvben?**

A nyelvjárási tendenciák alapján a köznyelv változási irányaira is hipotéziseket állíthatunk fel. Egyrészt a köznyelvben nem tartunk számon diftongusokat, amelyek szilárdítanak a vizsgált szembenállást.



Másrészt az *e* minősége más magyar nyelvváltozatokhoz képest nem túl nyílt, márpedig nyíltabb *e*-t nagyobb számban tartalmazó nyelvváltozatokban rendre nagy mennyiségű (19% feletti) hosszúnak lejegyzett magánhangzót, és feltevésünkkel összhangban az időtartamon alapuló szilárdabb szembenállást találunk. E nyelvjárásokban nyíltabb *e*-vel a „zárt *e*” nagyobb vitalitása járhat együtt: eggyel több magánhangzó-minőség jelentkezik egy rendszeren belül, s több magánhangzó nyilván nehezebben különíthető el pusztán minőségi alapon. Így a tényleges időtartam-különbségek szerepe is nagyobb lehet, a rövid-hosszú szembenállás stabilabb marad. Az *e* nagyobb nyíltsága sem érvényes azonban a köznyelvre, amely ezért a rövid-hosszú magánhangzó-szembenállás lassú felbomlásának folyamatában inkább innovatív nyelvváltozatnak minősül.

Vizsgálatunkban azt szerettük volna példázni, hogy nagyobb mennyiségű informatizált nyelvjárási adat kvantitatív elemzése elméleti fonetikai vonatkozásban is figyelmet érdemelhet. Még nagyobb korpuszon (és lehetőleg kísérleti fonetikai ellenőrzéssel) minden bizonnyal további összefüggések is feltárhatók vagy pontosabban megismerhetők lesznek.

### Irodalom

Bruce, G. – Elert, C.-C. – Engstrand, O. – Wretling, P.: Phonetics and phonology of the Swedish dialects - a project presentation and a database demonstrator. Proceedings of the 14<sup>th</sup> ICPhS. Vol. I. San Francisco 1999, 321-325.

Engstrand, O.: Durational correlates of quantity and sentence stress: A cross-language study of Swedish, Finnish and Czech. UCLA Working Papers in Phonetics 63. 1986, 1-25.

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

Hagège, C. – Haudricourt, A.: La phonologie panchronique. PUF. Paris 1978.

Herman, J.: Phonétique et phonologie du français contemporain. Tankönyvkiadó. Budapest 1984.

Juhász Dezső: A romániai magyar nyelvjárások atlaszának kiadástörténeti, nyelvi és településtörténeti tanulságaiból. *Magyar Nyelv* 2. 1997, 199-207.

Juhász Dezső – Radványi Péter – Vékás Domokos: Egy dialektológiai adatbázis és elektronikus prezentációs rendszer alapvetése. Magyar Nyelv 1. 2000, 47-58.

Kovács Magdolna: A spektrális minőség és az időtartam szerepe a magánhangzók percepciójában. In: Beszédkutatás '98. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1998, 35-46.

Murádin László (gyűjt.) – Juhász Dezső (szerk.): A romániai magyar nyelvjárások atlasza. I. kötet. Magyar Nyelvtudományi Társaság. Budapest 1995.

Vékás Domokos: Informatikai lehetőségek a dialektológiában különös tekintettel a fonetikai szempontokra. Kézirat. ELTE. Budapest 1999. <http://www.btk.elte.hu/fonetika/p/kozle.html>

# DALLAMCSERE AZ INTONÁCIÓS FRÁZIS BELSEJÉBEN

Varga László  
Eötvös Loránd Tudományegyetem

## Első megközelítés

A magyarban mind a spontán beszédben, mind a felolvasásban elég gyakori a dallamcsere jelensége. Első megközelítésben ehhez az szükséges, hogy a kiindulásul szolgáló intonációs frázis (IF) utolsó három dallama féleső, féleső és eső dallam legyen. A dallamcsereben a második féleső dallamot magas szinttartóra cseréljük. A jelenség részletesebb megbeszélése előtt először mutassuk be a benne szereplő dallamokat.

A féleső és az eső dallam hasonlít egymáshoz: mindkettő elől esik, azaz többszótagú dallamhordozó esetén az első és második szótag között viszonylag nagy esés van. Bolinger osztályozásában mind a féleső, mind az eső dallamot A-profilnak hívhatjuk (vö. Bolinger 1986, 1989). A különbség köztük az, hogy a féleső dallam észrevehetően a beszélői hangterjedelem alsó szintje *fölött* ér véget, míg az eső dallam vége leér a beszélői hangterjedelem alsó szintjére; ld. (1), (2), vö. Varga (1993, 1994). A hallgató általában rövid idő alatt beméri, hogy a beszélő hangterjedelmének alsó szintje hol húzódik, és így a féleső és eső dallamokat viszonylag nagy biztonsággal meg tudja különböztetni (Pierrehumbert 1980). Az eső dallam ún. terminális dallam, tehát egy intonációs frázis utolsó dallama. Ezért utána intonációs frázishatár áll. Az intonációs frázishatár jele a függőleges vonal [ | ]. A féleső dallam csak akkor terminális, ha szünet követi. Ha nincs utána szünet, akkor a féleső dallam az intonációs frázis belsejében, a skálában helyezkedik el, és természetesen nincs utána intonációs frázishatár. A (3) ábra a magas szinttartó dallamot mutatja be. Akárcsak az eső dallam, ez is terminális dallam, tehát a vége egyúttal az intonációs frázis vége is.



(1) *Féleső:*

| 'Angéla néni |

An

géla néni

---

(2) *Eső:*

| 'Angéla néni |

An

gél<sup>e</sup> la né<sub>ni</sub>

---

(3) *Magas szinttartó:*

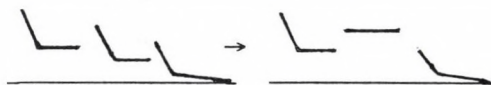
| <sup>-</sup>Angéla néni |

Angéla néni

---

Amikor tehát az intonációs frázis utolsó három dallama féleső + féleső + eső, akkor a beszélő megteheti, hogy az utolsó előtti (féleső) dallamot magas szinttartó dallamra cseréli; vagyis féleső, magas szinttartó, eső dallamsorozatot állít elő. Bolinger terminusait használva, az eredeti A+A+A profilsorból A+B+A profilok sora jön létre. Ezt a jelenséget nevezem dallamcserének. Mindezt sematikus dallamgörbék segítségével a (4), valóságos példával pedig az (5) szemlélteti:

(4) A    A    A            A    B    A



(5) A kiinduló („közönséges”) változat:

a. | Mert 'nem volt 'elegendő 'bizonyítékom. |

nem  
 e  
 Mert volt bi  
 legendő  
 z<sup>o</sup>nyíté kom

---

A dallamcsereélt változat:

b. | Mert 'nem volt 'elegendő | 'bizonyítékom. |

nem elegendő  
 bi  
 Mert volt  
 z<sup>o</sup>nyí tékom

---

Mivel a magas szinttartó dallam terminális dallam, utána új intonációs frázis kezdődik; ezért van IF határ az *elegendő* szó után. Amikor tehát egy féleső dallamot magas szinttartóra cserélünk, az intonációs frázist két intonációs frázisra bontjuk szét. Az újonnan létrejövő, másodlagos intonációs frázisoknak nem kell a Selkirk-féle szintaktiko-szemantikai feltételnek („Sense Unit Condition on Intonational Phrasing”, Selkirk 1986, 286) megfelelniük.

A dallamcsere jelentésbeli változással jár együtt: a létrejövő új dallamsorozat azt fejezi ki, hogy a beszélő a mondat tartalmát magától értetődőnek (nyilvánvalónak, természetesnek, könnyen kitalálhatónak stb.) tartja. A dallamsorozat egyúttal a beszélő intellektuális vagy hivatali fölényét is érzékelteti, és így egyszerre leereszkedő és kategorikus attitűdöt is kifejezhet, de nem barátságatlan. Ezért gyakori a csoportoknak szóló, kisgyermeknek

adott tanítói magyarázatokban és előadásokban, tornagyakorlatok vezényleésében, de párbeszédekben elhangzó egyéni védekezésekben vagy tiltakozásokban, meg nem értett mondatok megismétlésében sem ritka. Ezt a komplex attitűdöt a „tanítós” címkével foglalhatjuk össze.

A dallamcsere jelenségét tehát, első megközelítésben, a következő szabállyal írhatjuk le:

(6) *Dallamcsere (első közelítés):*

'σ... =opc=> ˘σ... | /'σ... \_\_\_ ˘σ... |

ahol σ = szótag,

... = egy vagy több szótagból álló füzér, főhangsúly és IF-határ nélkül.

Feltétel: „tanítós” attitűd.

Lássunk néhány további példát. Az átiratok aláhúzott szakaszai illusztrálják a dallamcserét. A (7) példa az Iskolarádió „Törd a fejed” c. adásából származik. Egy tanárnő olvassa fel megjegyzéseit a beküldött munkák helyesírási problémáival kapcsolatban.

(7) | ˘Pedig | ugye ˘tanultátok, hogy a -'val, -'vel ˘rag ˘véje |  
 a ˘mássalhangzóval ˘végződő ˘szavaknál | ˘általában |  
 ˘áthasonul | ˘azzá a hanggá, | ami a ˘szó ˘végén | ˘van. |

A következő példák „A hát gyógytornája” c. kazettáról valók.

(8) | ˘Ellazítunk, | a ˘lábat ˘talpra | ˘visszahúzzuk. |

(9) ... | az ˘ujjak a ˘plafon felé mutatnak, | a ˘két tenyér ˘egymás felé  
 néz, | és ˘zárjuk ˘ökölbe a | ˘kezeket. |

Az utolsó példák családi párbeszéd-részletek:



(10) | 'Angolul drágám 'nem lehet 'megtanulni ˘tökéletesen |  
`sohasem. |

(11) | 'Iszom egy kis ˘teát, kisleány | `süteménnyel. |

(12) | 'Tudod ^mit? | 'Felveszem az ˘egészet | `magnóra. |

A (13)-as párbeszéd harmadik megnyilatkozása szintén informális, családi párbeszédből való, és azt illusztrálja, hogy amikor egy meg nem értett vagy nem jól hallott közlésünket újra megismételjük, dallamcserehez folyamodhatunk.

(13) A: | 'Lemegyek a 'Jánoshoz a `lottózóba. |

B: | ^Hogy mondod? |

A: | 'Lemegyek a ˘Jánoshoz a | `lottózóba. |

### Kiterjesztés

További példák azonban arról győznek meg minket, hogy a dallamcsere első definíciója, a (6), nem eléggé általános. Így például nem alkalmazható az olyan esetekre, mint a (14c), holott a dallamcsere ott is lehetséges.

(14) *A kiinduló („közönséges”) változat:*

a. | 'Ráborította a 'hamutartót a 'frissen festett `asztalra. |

Rá

borította a ha fris  
mutartót a sen festett asz  
ta | r a

---

*A dallamcserélt változatok:*

b. | 'Ráborította a 'hamutartót a -frissen festett | `asztalra. |

Rá  
                  ha          frissen festett  
borította a                                  asz  
                  mutartót a  

---

  ta l<sub>ra</sub>

c. | 'Ráborította a -hamutartót a | 'frissen festett `asztalra. |

Rá                  hamutartót a  
                                  fris  
borította a                                  asz  
                                  sen festett  

---

  ta l<sub>ra</sub>

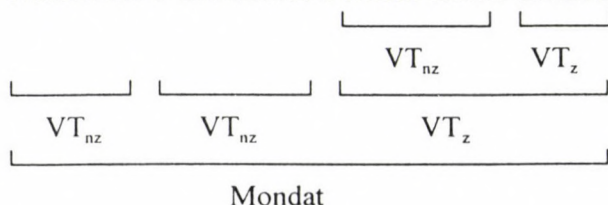
A korábbi példák és a (14b) alapján azt hihettük, hogy dallamcsere csak akkor lehetséges, ha a cserében érintett féleső dallam közvetlenül egy eső dallam előtt áll. Ezt a részigazságot ragadta meg a (6) szabály. A (14c)-ben azonban más a helyzet: az érintett dallam (a *hamutartót a* szótagfüzére) egy féleső dallam előtt található, és a dallamcsere mégis lehetséges.

Ezért a (6) helyett egy olyan általánosabb dallamcsere szabályra van szükségünk, amelynek érvénye nemcsak a korábbi példákra és a (14b)-re, hanem a (14c)-re is kiterjed. Ehhez fel kell elevenítenünk a váltótömbök fogalmát (vö. Varga 1993, 1994). A váltótömbök (VT-k) olyan főhangsúlyos szintaktikai konstituensek, amelyek a mondat közvetlen vagy közvetett összetevői, és amelyek között komplementáris viszony feszül: a nyitott (nem-záró) VT-k folytatást, befejezést kívánnak, és ezt a záró VT adja meg.

A dallamcsere valamennyi esetéről számot adhatunk, ha

elismerjük, hogy a dallamcsere a mondat utolsó (záró) váltótömbje előtt történik, bármilyen mélyre is van az a váltótömb ágyazva. A (14) mondatban például két konstituens tekinthető záró váltótömbnek, egy nagyobb, ez a *frissen festett asztalra*, és egy kisebb, amely a nagyobbiknak része, ez az *asztalra* szó; ld. (15). A *z* index jelentése 'záró', az *nz* jelentése 'nem záró'.

(15) Ráborította a hamutartót a frissen festett asztalra.



Eszerint a dallamcsere fogalmát újra kell definiálnunk, és a (6) szabályt ki kell terjesztenünk. Az új, kiterjesztett szabályt a (16) tartalmazza.

(16) *Dallamcsere (kiterjesztett)*:

'σ... =opc=> ˘σ... | /['σ... \_\_\_[\_z(Y) ˘σ... | ]]

ahol σ = szótag,

Y = a záró VT-nek az eső dallam előtti anyaga,

... = egy vagy több szótagból álló fűzér, főhangsúly és IF-határ nélkül.

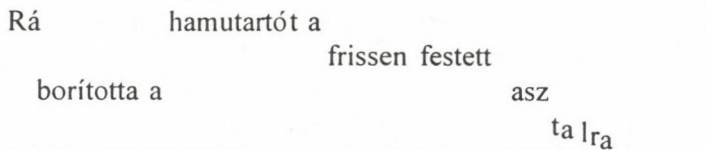
Feltétel: „tanítós” attitűd.

A kiterjesztett szabály alapján már a (14c)-t is elő tudjuk állítani. Az új szabály szerint ugyanis a (14) mondatban dallamcsere nemcsak a *fris-* szótagon lehetséges (az *asztalra* záró-VT előtt), hanem a *ha-* szótagon is (a *frissen festett asztalra* záró-VT előtt), ahogy a (14c) mutatja. Az új szabály azt is megmagyarázza, hogy esetenként hogyan lehetséges egy mondatban több dallamcsere. Így például a



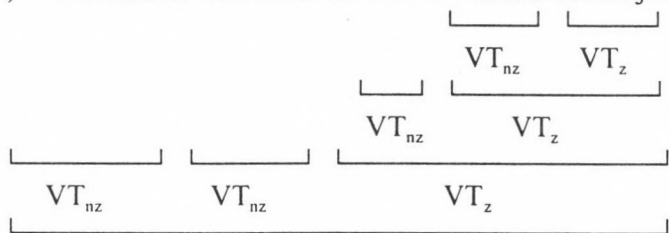
(14d)-ben kettő is előfordul:

(14)d. | 'Ráborította a 'hamutartót a | 'frissen festett | 'asztalra. |



Végezetül vizsgáljuk meg a (17) mondat összes dallamcsere lehetőségét.

(17) Mindenkinék elmesélné az utolsó balatoni kalandját.



Mondat

A (17) mondatban három konstituens minősül záró váltótömbnek: *kalandját*, *balatoni kalandját*, és *az utolsó balatoni kalandját*. A lehetséges realizációkat a (18) mutatja. A kiinduló változat (18a) helyébe öt dallamcsereált változat léphet. Ezek közül a (18e)-ben egyszerre kettő, a (18f)-ben egyszerre három dallamcsereált hajtunk végre:

(18) A kiinduló („közönséges”) változat:

a. | 'Mindenkinék 'elmesélné az 'utolsó 'balatoni 'kalandját. |

Min  
 el  
 denkinek u  
 mesélné az ba  
 tolsó ka  
 latoni  
 la ndj á t

---

*A dallamcserélt változatok:*

b. | 'Mindenkinek 'elmesélné az 'utolsó -balatoni | `kalandját. |

Min  
 el  
 denkinek u balatoni  
 mesélné az ka  
 tolsó  
 la ndj á t

---

c. | 'Mindenkinek 'elmesélné az -utolsó | 'balatoni `kalandját. |

Min  
 el utolsó  
 denkinek ba  
 mesélné az ka  
 latoni  
 la ndj á t

---

d. | 'Mindenkinek  $\bar{e}$ lmesélné az | 'utolsó 'balatoni 'kalandját. |

Min                elmesélné az  
   u  
           denkinek   ba  
   tolsó   ka  
   latoni  
   landjút

---

e. | 'Mindenkinek  $\bar{e}$ lmesélné az | 'utolsó  $\bar{e}$ balatoni | 'kalandját. |

Min                elmesélné az  
   u                balatoni  
           denkinek   ka  
   tolsó  
   landjút

---

f. | 'Mindenkinek  $\bar{e}$ lmesélné az |  $\bar{e}$ utolsó |  $\bar{e}$ balatoni | 'kalandját. |

Min                elmesélné az  
   utolsó  
           denkinek   balatoni  
   ka  
   landjút

---

### Három megjegyzés

Végezetül három megjegyzés kívánczik még témánkhoz. Az első a következő. Mint láttuk, a dallamcsere eredményeként olyan dallamképlet jön létre, amelyben egy magas szinttartó dallamot találunk féleső dallamok között, illetve féleső és eső dallam között. Nem szabad azonban azt hinnünk, hogy minden ilyen dallamképlet



dallamcsere eredménye. A magyar hanglejtés mindezidáig legrészletesebb fonetikai leírásában, Fónagy Iván és Magdics Klára „A magyar beszéd dallama” c. 1967-es művében számos példát találunk a szóban forgó konfigurációra. A (19) példa a Fónagy-Magdics könyv 131. oldaláról, a (20) példa a 135. oldalról való (mindkettő az én jelöléssel):

(19) | 'Behozta az ̀abroszt | és megterített. |

(20) | Ha 'minden leckét ̀megírtatok, | 'elmehettek ̀moziba. |

Ezekben a példákban azonban a szinttartó dallam egy tagmondat végét jelöli, tehát szintaktikai oka van: grammatikai határt jelez, és nem dallamcsere eredménye. Az ilyeneket a felszíni hasonlóság ellenére sem szabad a dallamcsere által létrehozott dallamképlettel összekeverni. A Fónagy--Magdics kötetben egyébként valóságos dallamcserére is találunk egy példát, a 28. oldal utolsó előtti kottájában, amit megint a saját jelöléssel közlök:

(21) | 'Szorgalmasan ̀kefélgette a | ̀nadrágját. |

Ez valódi dallamcsere, de a Fónagy–Magdics kötet nem tud a dallamcsere jelenségéről és ebből a szempontból nem is kommentálja a példát.

A második megjegyzésem az, hogy a dallamcsere révén előálló jellegzetes dallamszekvenciát és az általa kifejezett „tanítós” attitűdöt nem csak dallamcsere útján lehet elérni. Elérhető ez ún. dallambeékelés révén is (vö. Varga 2000). Ennek röviden az a lényege, hogy egy ha egy megnyilatkozás záró váltótömbje eső dallamot tartalmaz, és ez előtt a váltótömb előtt egy fésűs dallamot hordozó beszédszakasz áll, akkor a fésűs dallamú beszédszakasz mellékhangsúlyos szótagján magas szinttartó dallamot indíthatunk, miközben ez a mellékhangsúlyos szótag főhangsúlyossá lép elő. Ezt szemlélteti a (22b) példa. A (22a) *-gyek* szótagján ritmikai mellékhangsúly van, ez lesz a (22b)-ben a dallambeékelés helye.

(22) *A közönséges változat:*

a. | 'Elme.gyék a `könyvtárba. |

El  
                                könyv  
                    megyek a  
  t á r b a

---

*A dallambeékelte változat:*

b. | 'Elme<sup>-</sup>gyék a | `könyvtárba. |

El gyék a  
                                könyv  
                    me  
  t á r b a

---

A dallambeékelés kimerítő ismertetése azonban meghaladná ennek az előadásnak a kereteit.

Végezetül megjegyzem, hogy az itt bemutatott attitudinális jelentés (tehát a jobb híján „tanító” attitűdnek nevezett sajátos jelentéstartalom) kifejezése szempontjából az angol nyelvben, úgy tűnik, csak a magas szinttartó és az eső dallam releváns. A kettő kombinációja az angolban is létezik, és hasonló attitűdtartalmakat képes kifejezni. A Bolinger által B + A szekvenciának nevezett dallamképlet a megnyilatkozás végén nem annyira informatív, mint inkább a beszélő domináns helyzetét tükröző, kommentáló, „tudó megjegyzés” (‘knowing remark’) kísérője (Bolinger 1986, 287, 296).

### Irodalom

Bolinger, D. L.: Intonation and Its Parts. Edward Arnold. London 1986.

Bolinger, D. L. : Intonation and Its Uses. Edward Arnold. London 1989.

Fónagy Iván – Madgics Klára: A magyar beszéd dallama. Akadémiai Kiadó. Budapest 1967.

Pierrehumbert, J. B.: The Phonetics and Phonology of English Intonation. PhD dissertation, MIT 1980.

Selkirk, E. O.: Phonology and Syntax. The Relation between Sound and Structure. The MIT Press. Cambridge, Mass. & London 1986.

Varga László: A magyar beszéddallamok fonológiai, szemantikai és szintaktikai vonatkozásai. Nyelvtudományi Értekezések 135, 1993.

Varga László: A hanglejtés. In: Kiefer Ferenc (szerk.): Strukturális magyar nyelvtan 2. Fonológia. Akadémiai Kiadó. Budapest 1994, 486-549.

Varga László: A magyar mellékhangsúly fonológiai státusáról. Magyar Nyelvőr 124, 2000, 91-108.



# A FONOLÓGIAI KONTRASZTFENNTARTÁS ELVE AFÁZIÁSOK KÖZLÉSFOLYAMATAIBAN

Szépe Judit  
MTA Nyelvtudományi Intézet

A nyelvészeti megközelítések az afáziás jelenségek lényegét nem a beszéd fizikai–fiziológiai természetében (de gyakran még csak nem is a beszédet létrehozó kognitív mechanizmusban) látják megragadhatónak, hanem annak a jelrendszernek a sajátosságaiiban, amelyet a tevékenység kivitelezéséhez az afáziás közlő használ. Ettől válik relevánsá egy nyelvészeti megközelítés számára az afáziás jelenség, miközben a szintén a beszédtevékenységet érintő hangképzési rendellenességek (selypítés, pöszeség) – nem.

Figyeljük meg, hogy azok a grammatikai modellek, amelyek megkülönböztetik az eltérő szintű mögöttes reprezentációkat a tényleges kivitelezés által létrehozott alakoktól, utóbbiakat kizárólag az alsóbb nyelvi szintek körében tartják értelmezhetőnek. Vannak fonetikai implementációs szabályok, de nincsenek morfológiaiak vagy szintaktikaiak. A modelleknek egyetlen nyelvi szintje állhat közvetlen kapcsolatban a tényleges kivitelezéssel: a fonológiai szint.

Ez nyilván abból adódik, hogy kizárólag a fonológiai szint absztrakt rendszeregységei feleltethetők meg közvetlenül fizikai entitásoknak. Sem a morfológiai, sem a szintaktikai kategóriáknak nincs velük egyegyértelmű megfelelésben álló, saját fizikai ekvivalensük.

Valószínűleg az afáziakutatásban is ezért merül fel élesebben az a kérdés, hogy a beszédhangokat érintő jelenségek az absztrakt struktúra vagy a ténylegesen kivitelezett alak szintjén értelmezendők, illetőleg hogy e distinkció releváns-e a jelenségeknek klinikai tipológiához való köthetősége szempontjából.

A több évtizedes kérdésben az afázia nyelvészeti kutatását illetően, hogy t.i. osztályozhatók-e az afáziás beszédjelenségek valamely elmélet szerveződési szintjei, kategóriái és szabályai alapján, a legtöbb elemzés vagy megpróbál a saját vizsgálataira vonatkozóan állást fog-

lalni, vagy eleve a besorolhatóságból indul ki. Abban a kérdésben, hogy milyen elemzési szinthez tartozónak tekintendők a beszédhang-  
tévesztések afáziás jelenségei, erősen megoszlanak a vélemények. Tartsuk szem előtt, hogy a most bemutatandó állítások a szóalak, ille-  
tőleg frázis nagyságrendjére kiterjedő vizsgálatok adatainak kifejezet-  
ten fonológiai alapú elemzésére épülnek, ezért bizonyos jelenségek  
értelmezéséhez megkerülhetetlenek a mentális lexikon belső rendjé-  
nek, hozzáférési mechanizmusainak egyes kérdései.

A korai szakirodalmi megállapítások szerint azoknak az afáziások-  
nak, akiknek a beszédét nem érinti motoros kivitelezési károsodás,  
anyanyelvük minden fonémája az adott nyelvben fonotaktikailag meg-  
engedett valamennyi kontextusban rendelkezésre áll. A károsodások  
nem bontják meg a beszélő fonémarendszerének integritását, s a jelen-  
ségek között még olyan összefüggés sem mutatható ki, amely arra utal-  
na, hogy a károsodott fonémarealizációk a fonémarendszernek vala-  
mely alrendszeréhez volnának kapcsolhatók (Marcie 1972). Mind-  
azonáltal már a korai megfigyelések is jellemzőnek találták a fonéma-  
realizációk afáziás károsodására a magánhangzókhoz a mássalhang-  
zókhoz képest kisebb mértékű sérülékenységét, a mássalhangzó-kap-  
csolatok egyszerűsítését, a réshangok és az affrikáták erősebb érintett-  
ségét, valamint a helyettesítések gyakoribb előfordulását szó elején  
(Shankweiler–Harris 1966). Ennek a felfogásnak egy markánsabb  
megfogalmazású változata szerint még abban sincs különbség, sem az  
egy (motoros kivitelezési károsodásuktól mentes) afáziatípusok  
között, sem afáziások és ép beszélők megnyilvánulásai között, hogy  
mely fonémák mely pozíciókban milyen gyakorisággal fordulnak elő,  
továbbá a beszédhang-helyettesítések típusai sem térnek el szignifi-  
kánsan az egyes klinikai típusokban. Azt látjuk ugyanis, hogy bármely  
afáziatípus hasonló módon befolyásolja a beszélő fonémarealizációit.  
Jellemző az egymástól egy disztinktív jegyben eltérő fonémák realizá-  
ciója közötti helyettesítés, általában a pár jelöletlen tagjának irányá-  
ban, például zöngés→zöngétlen irányban (Blumstein 1973).

Az ezzel szemben álló megközelítések eltéréseket találnak az  
egy típusokban előforduló helyettesítések között abban a tekintet-  
ben, hogy fonetikai/kivitelezésbeli avagy fonológiai/tervezésbeli az  
adott afáziás jelenség. A Lecours–Lhermitte (1969) által bevezetett



distinkció későbbi értelmezései arra a fonológiai alapfeltevésre épülnek, hogy egy szó mély (standard generatív értelemben: lexikonbeli) fonológiai reprezentációja lehet kevésbé specifikus, mint felszíni fonológiai reprezentációja, amely utóbbi közvetlenül alapját képezi a fonetikai reprezentációnak. Az absztraktabb ábrázolásban nincsenek vagy nincsenek teljesen specifikálva azok az összetevők és elemek, amelyek két szomszédos szint között működő szabályokkal levezethetők. A vizsgálatoknak az a kiindulása, hogy a reprezentációknak van pszichológiai realitásuk a beszédmechanizmusban: más-más szintű reprezentációk relevánsak a beszélő tervezési és kivitelezési folyamataiban, s ez a különbség az afáziás jelenségekben is megmutatkozik. A tervezésbeli károsodások a mély és a felszíni fonológiai reprezentáció között, a kivitelezésbeliek a felszíni fonológiai reprezentáció és a fonetikai realizáció között mennek végbe.

Ez a megközelítés tehát az afáziás beszélő mentális grammatikájában a szóalak lexikonbeli fonológiai reprezentációját érintetlennek, az ép beszélőével megegyezőnek tekinti. Fonetikai/kivitelezésbeli afáziás folyamatok esetén a szóalak felszíni fonológiai reprezentációja ugyancsak megegyezik az ép beszélőnél feltett reprezentációval, fonológiai/tervezésbeli folyamatok esetén viszont eltér attól. A kivitelezésben (nem motoros károsodás miatt) fellépő beszédhang-tévesztésekben a helyettesítő és a helyettesített elemek között artikulációs összetevőik tekintetében a különbség minimális, a helyettesítés iránya pedig a kisebb energiaráfordítást igénylő elem felé halad. Ebben a tévesztési típusban kimutatható a közvetlen hangtani kontextus erős hatása (Trost–Cantor 1974; Blumstein 1981). A tervezésbeli tévesztések esetén a szóalakban nem a megfelelő fonémák és/vagy nem a megfelelő helyre választódnak ki. Ez azt jelenti, hogy a szóalakban (i) két, egymástól több jegyben különböző fonéma „helyet cserél” (*sün* [nyʃ]; *Péterrel* [te:per:el]); vagy (ii) egy fonéma a szóalakban nemcsak a reprezentáció által determinált helyén jelenik meg, hanem más szegmentumok helyén is (*te leszel a főszerkesztő* [se reser a fõ:sersesø:]); vagy (iii) olyan fonéma is megjelenik, amely megváltoztatja a szó CV-szerkezetét (*beledolgoztam* [belgedolgoztam]). Vagyis a jelenség nem az aktuális kivitelezést, hanem a felszíni fonológiai



reprezentációt érinti (Nespoulous–Borell 1979; Nespoulous et al. 1982, 1983, 1984).

Ez a megközelítés két ponton vet fel problémát. (a) Az elemzések azon az alapon határozzák meg, mely szint lehet érintett valamely tévesztési folyamatban, hogy felteszik a kisebb energiaráfordítás elvének érvényesülését a fonetikai kivitelezésben. De csakis ott. Álláspontjuk szerint fonetikailag kivitelezni egy zöngés hangot lehet bonyolultabb, mint egy zöngétlent, de fonológiai zöngésnek tervezni – nem. Az elv következetes alkalmazása oda vezet, hogy minden olyan afáziás hangtévesztést, amely a szó tényleges megvalósulásában a kisebb energiaráfordítást igénylő artikuláció felé mutat, kivitelezésbeli jelenségnek kell tekinteni. Következésképp a nagyobb energiaigényű változat választásának szükségképpen tervezési műveletnek kell lennie. Pedig éppen az a feltevés, hogy a tervezés számára irreleváns, mekkora az artikulációhoz szükséges energia, érvényteleníti azt az elvárást, hogy a tervezés soha ne a kevésbé komplex alakot válassza. (b) A tévesztéseknek a reprezentációs szintek átmenetei közé helyezése azzal az érveléssel, hogy az egyre felszínközelebbi szintek egyre magasabb specifikációját adják meg egy szó fonológiai formájának, valamint annak a feltevése, hogy a mély fonológiai reprezentáció a helyettesítési folyamatokban nem sérül, ahhoz a következtetéshez vezet, hogy a helyettesítések megragadhatók az adott szint specifikációs folyamatainak elmaradásában. Ez pedig azzal a következménnyel jár, hogy elemezhetetlenné válnak olyan afáziás szóalakok, amelyek nem vezethetők le specifikációk elmaradása útján egyetlen mély fonológiai reprezentációból sem (*száj* [te:k]), hacsak nem tesszük fel, hogy a szó mély fonológiai reprezentációja kizárólag a szó CV-szerkezetét írja le, ami *contradictio in re* volna. A szótalálási zavarok afáziás közelítési stratégiái viszont éppen arra utalnak, hogy amennyiben a beszélő egyáltalán hozzá tud férni mentális lexikonja valamely egységéhez, az mindenképp fog tartalmazni (legalább az első szótagjában) a szónak megfelelő, magasan specifikált jegyeket. Magyar példák: *köt* [kom kø kø køm kom køm]; *lapátot* [la labda:val labda:t]. Viszont nem őrzi meg szükségképpen az adott szó CV-szerkezetét: *Kriszta* [kirista]; *sün* [jynt]; *előadás* [ɛlø:dada:f]; *kiskecske* [kiʃkece]; *pillanatig* [pilint pilintig].

Egy másik megközelítés szerint lehetséges, hogy (motoros kivitelezésbeli probléma nélkül is) valamely fonéma kiesett az afáziás beszélő rendszeréből. Crystal (1980) éppen annak alapján különíti el a fonológiai/nem fonológiai afáziás károsodást, hogy hiányzik-e a beszélő rendszeréből valamely fonéma, pontosabban neutralizálódik-e a realizációkban két fonéma kontrasztja vagy nem. Ha az afáziás közlésben egy fonéma realizációja torzul, de még megfeleltethető az adott fonémának, a károsodás fonetikai szintűnek minősül. Am ha egy fonémarealizáció torzulása olyan nagy mértékű, hogy egybeesik egy másik fonéma realizációjával, akkor a károsodás fonológiai, hiszen megszűnt egy kontraszt. A megközelítés a következő miatt nem tartható. E felfogás szerint addig, amíg egy rendszerben minden kontraszt megvan, a károsodás nem fonológiai. (De emlékezzünk vissza: ha egy fonéma kiesik, az már igen.) Következésképp, ha egy afáziás közlésben [b]-ből [p]-lesz, [p]-ből pedig [t], akkor a folyamat nem érintette (t.i. funkcionálisan) a beszélő fonémarendszerét. A megközelítés figyelmen kívül hagyja azt, hogy egy rendszerben kizárólag akkor maradhat fenn ugyanolyan számú oppozíció, ha a változás lefut a teljes rendszeren. Minden egyéb esetben a kontrasztok közül egy biztosan elvész. A teljes lefutás pedig vagy úgy történhet, hogy az utolsó állomás visszatér az elsőbe (ilyen jelenségre pedig nem találunk példát a szakirodalomban), vagy úgy, hogy az utolsó állomáson új fonéma jön létre. Ebben az utóbbi esetben viszont a fonológiai rendszer máris érintve van. Azzal kell tehát számolnunk, hogy egy afáziás  $X \rightarrow Y$  folyamat hamarabb megáll, mint hogy minden érinthető elemen lefuthatna. Ahol megáll, ott felfüggesztődik egy kontraszt. Így, a megközelítés logikája szerint, azt a típust kellene fonológiai szintűnek tekinteni, ha például egy szóban [d]-ből [b] lesz, és ez egyúttal egybeesik a szóban ugyancsak előforduló, tényleges [b]-vel, például *dob* [bob]. Ettől eltérőnek pedig azt az esetet kellene tekinteni, amikor egy szóban  $X$ -ből  $Y$  fonémarealizáció lesz, mire  $Y$ -ból  $Z$ . Az előző célszó példájánál maradva: *dob* [bot]. Miután a kontraszt fennmaradt, a jelenség nem érintette a beszélő fonémarendszerét. Am a különbségtétel ellentmondáshoz vezet.

Az afáziás nyelvi jelenségek sokszínűsége és gyakori variabilitása alapján feltehetjük, hogy a károsodások összetételében és belső



rendszerviszonyai tekintetében nem határozzák meg egyszer s mindenkorra az afáziás beszélő által alkalmazott fonológiai rendszert. E rendszer, a maga szabályaival és oppozícióival, akár közlésegyeségenként is más és más lehet. Máskülönben nem volna lehetséges, hogy egy szónak ugyanazt az elemét ugyanabban a szövegkörnyezetben a közlő egyszer így, másszor úgy helyettesítse, például: *ceruza* [tʃeruza], [teruza]; *sál* [ʃa:j], [ʃa:r]. De egy közlésegyeségen belül egy rendszerállapot érvényesül, tehát egy adott rendszerállapot összefüggései egy közlésegyeség nagyságrendjében vizsgálhatók. Továbbá kommunikációs alapelveként tegyük fel, hogy a károsodott rendszernek is tartalmaznia kell az elemek megkülönböztetésére (a fonológiai kontraszt fenntartására) való törekvést. Ennek megfelelően a *dob* [bob] példa nem lehet rendszert érintő helyettesítés, hiszen a tervezés nem törekszik a tévesztés nyomán létrejött kontrasztmegszűnést kompenzálni. Tehát a tervezés számára rendelkezésre álló rendszerben megvolt a kontraszt, amely csak a kivitelezés során függesztődött fel. Ezzel szemben a *dob* [bot] példában a [d]→[b] helyettesítés miatt felfüggesztődő különbség nem vész el, hiszen a [b]→[t] helyettesítés helyreállítja a kontrasztot. Miért nem tekinthető ez a művelet a fonetikai kivitelezés programjának? Egyfelől azért, mert nincs olyan elmélet, amely a kontrasztfenntartást ne a rendszerhez, hanem a fonetikai realizációhoz kapcsolná, másfelől pedig azért, mert a folyamat jobbról balra is működik, például *dolgozom* [voldozom]: d→v, még mielőtt g→d, illetőleg *jövőre* [jønø:ve]: v→n, még mielőtt r→v. Márpedig ha egy korábbi, kontrasztfenntartó helyettesítés már akkor bekövetkezik, amikor a létrejöttét motiváló, későbbi kontrasztvesztés még nem is történt meg, akkor annak a választásnak korábban kellett megtörténnie, mint a fonetikai kivitelezés.

A jelenségek értelmezésére más megoldás is kínálkozik. Szende (1997, 153-163) Globális Programozás Elve (GPP) alapján feltehető a lexikonbeli fonológiai reprezentációban (FR) foglalt információk előhívásának két, egymásra épülő műveleti szintje. (1) A globális előhívás szintje (GE), amelyben a fonológiai programozás az FR által tartalmazott információk közül a szótagszámot, a szóhatárokat, a hangsúlymintákat, a fonotaktikai elrendezést hívja elő, illetőleg a felhasználandó megkülönböztető jegyeket, amelyeknek sorrendezéséről



azonban csak kitüntetett pozíciókra (a szóhatárokon lévő szótagokra) vonatkozóan hív elő információkat. (2) A finomprogramozás szintje (FP), ahol a programozás a GE által a lexikai egység egészére egyszerűen aktivizált jegyek tagolási pozícióknkénti rendezésére vonatkozó információkat hívja elő: a nem kitüntetett pozíciók itt kapják meg teljes jegyspecifikációjukat. (1) bemenetét képezi (2)-nek, amely pedig a fonetikai kivitelezés kiindulását alkotja.

A GPP szerint a lexikonbeli FR által előírt (nem kitüntetett) tagolási pozíciók GE szintjén nincsenek lezárva, vagyis még nem rendelkeznek teljes specifikációjú jegyfeltöltéssel. A jegyek, amelyek GE-ben pontos pozicionálási utasítások nélkül, a teljes szóalakra (vagy a közlésfolyamatban: szekvenciára) vannak aktivizálva, FP bemenetként ezen a műveleti szinten nyerik el pozíciójukat, illetőleg a pozíciók teljes jegyfeltöltésüket.

A finomprogramozásbeli tévesztések kiváltói az FP bemenetét képező lezáratlan tagolási pozíciók és a teljes szóalakra aktivizált jegyek: a műveleti hiba valamely jegyet más pozícióba helyez, mint ahogyan azt FR előírja, s ezzel egyszerűsítést, információcsökkentést hajt végre. Ha (a) valamely lezáratlan pozíció a teljes specifikációhoz nem a lexikonban előírt saját, elkülönítő jegyét/jegyeit kapja meg, hanem valamely másik (kitüntetett) pozíció(k)ét, akkor a programozás az egyszerűsítést úgy hajtja végre, hogy csökkenti az elemszámát annak a halmaznak, amely tartalmazza a fonológiai reprezentációban szükséges és lehetségesen választandó összes jegyet; ha pedig (b)  $X$  jegy, amelynek FR előírása szerint  $A$  (lezáratlan) pozícióba kellene kerülnie,  $B$  pozícióban jelenik meg, miközben  $Y$  jegy, amelynek  $B$  (lezáratlan) pozícióban kellene megjelennie,  $A$  pozícióba kerül, sorrendezésbeli műveleti hiba történt. Ebben az esetben tehát az egyszerűsítés egy sorrendezésbeli információ hiányából áll.

E műveleti hibák hagyományos elnevezései (szegmentális szintű jelenségekre vonatkoztatva): (a) anticipáció és perszeveráció (b) metatézis. De az (a)-n belüli hagyományos terminológiai különbségtétel csak jelenség szintű lehet. Csak annyit tud megjelölni, hogy az elvárt-hoz képest milyen irányban történt az egyszerűsítő jegypozicionálás. Ezt az irányt pedig az határozza meg, hogy egymáshoz képest hol helyezkedett el a feltöltendő pozíció és az, amelyikkel közös jegyet

kapott. A *távfütés* [ta:f:i:tø:f] adat sorbarende­zési műveleti hibájánál kerek­ség tekintetében FR az /y:/ pozí­ciójára [+kerek<sub>mgh</sub>], az /e:/ pozí­ciójára pedig [–kerek<sub>mgh</sub>] jegyértéket írt elő. A kerek­ség tekintetében GE-ben mindkét pozí­ció specifikálatlan maradt, FP-ben pedig a műveleti hiba következtében a két jegyérték egymás helyére pozicionálódott. Az *összkomfort* [øspøŋkørt] adatban sem /k/, sem /f/ pozí­ciója nem kapott teljes specifikációt a labialitásra vonatkozóan, a műveleti hiba pedig egymás pozí­ciójába helyezte az eltérő értékeket. /f/ pozí­ciója ettől eltérő mértékben volt lezártatlan, hiszen folyamatos­ság szempontjából sem volt specifikálva, és a /k/ pozí­ciójában lévő specifikációnak megfelelő [+hátsó, –foly] jegyek /f/ pozí­ciójában is /k/-t eredményeztek.

A GPP az FP-műveleteket a fonetikai kivitelezés fizikai jelenségétől, a koartikulációtól elsősorban az anticipáció jelensége alapján különíti el: ha egy korábban megjelenő egységre hatással van egy későbbi, amely a korábban létrejövő egység megjelenésekor fizikailag még nem volt jelen, akkor azt annak az egységnek nem a fizikai, hanem mentális jelenléte hozhatta létre. A jelenség tehát csak fonéma szintű, nem pedig beszédhang szintű értelmezést kaphat.

A már korábban bemutatott adatok is az afáziás FP-beli tévesztések hasonló működésére utalnak. Egyfelől az anticipáció jelenségei (például *Szabadság Hotel* [həbətʃ:a:k hotel]), másfelől pedig a fonológiai kontrasztfenntartás elve az afáziás közlésekben megerősítik a FP műveleti hibáinak nem fizikai motiváltságát: a *célkitűzést* [ke:lkyø:ze:ʃt] adatban a [+kerek<sub>mgh</sub>] jegy, amely /y:/ pozí­cióján kívül /i/ pozí­ciójában is megjelenik, illetőleg a [–felső], amely /e:/ pozí­cióján kívül /y:/ pozí­ciójában is megjelenik, illetve a [+zöng] és [+foly (+érd)] jegyek, amelyek /z/ pozí­cióján kívül /t/-n is megjelennek; továbbá /k/ megjelenése /ts/ helyén számos jegy pozicionálódására utalnak egy feltehetően csak annyiban specifikált tagolási helyen, hogy [+obstr, –zöng, –foly]. A kontrasztfenntartást elve mindenképp előtt olyan afáziás adatköz­lőknél vizsgálható, akiknél FR-nek GE szintű információi szinte minden esetben elérhetők, s a tévesztések kizárólag FP szinten jelentkeznek. A stratégia a következő. Ha egy szóalakon vagy szekvencián belül  $F_1$  fonéma realizációja egybeesik  $F_2$  realizációjával, akkor a kontraszt fenntartása érdekében  $F_2$  fonéma realizációja átmegy  $F_3$ -éba.



Vegyük a következő adatokat: *dob* [bot]: d→b, b→t; *nagy nevek* [naɟ vɛlɛk]: n→v, v→l; *szervíz* [ʃɛlvis]: z→s, s→f; *nyugdíjban van* [ɟugdida bā]: v→b, b→d; *családi dolgot* [tsada:ʃi dɔlɡot]: l→d, d→j; *lakásom* [vala:ʃom] k→l, l→v; *családi képeket* [tsala:di tʃe:pɛtʃɛt]: k→tʃ, tʃ→ts; *gomb* [dɔŋɡ] b→g, g→d; *jogsegélyszolgálat* [jokʃɛby:ʃolga:lat]: ʃ→s, s→f; *vállalatnál dolgozom* [va:lalatna:l vɔldozom] g→d, d→v; *azt mondja a párttitkár* [as moŋʃa a pa:rtista:r]: k→t, t→s; *jövőre ünnepeljük* [jønø:ve yn:ɛpɛj:yk]: r→v, v→n; *klasszikus zene* [tlaksikuz:ɛnɛ] s→k (s:→ks), k→t (kl→tl).

A •-gal jelölt adatokban a művelet jobbról balra megy végbe, tehát ha a szóalakban vagy szekvenciában egy később következő fonéma realizációja esne egybe egy hamarabb megjelenő fonéma realizációjával, akkor a korábbi fonémarealizáción már akkor végbemegy a kontrasztfenntartó helyettesítés, mielőtt még a később következőnek alkalma lett volna a korábbival helyettesítődni (hiszen a később következő akkor fizikailag még meg sem jelent). Ez, hasonlóan az anticiáció jelenségéhez, azt támasztja alá, hogy az FP-beli műveleti hibák nem kivitelezési hibák, hanem már a fonetikai implementáció előtt, FP szintjén eldőlt egy fonémának egy másik helyett való választása.

Ez a jelenség újra csak megerősíti azt a kiindulást, hogy azok a nem motoros kivitelezésbeli afáziás jelenségek is, amelyek a szóalak FR-jének egyedül FP-beli tévesztéseit mutatják, fonéma szinten történőnek tekintendők.

Mindezek alapján az afáziás beszédprodukciókban a beszédhang-helyettesítések két alaptípusának elkülönítése adódik.

(1) Azok a helyettesítések, amelyekre jellemző, hogy:

(a) a helyettesített beszédhang helyén kivétel nélkül más (de következetesen ugyanaz a másik) beszédhang áll,

(b) a helyettesített beszédhang soha nem jelenik meg valamely más beszédhang helyettesítőjeként,

(c) előfordulhat, hogy a helyettesítő beszédhang nem áll egyegyértelmű megfelelésben a beszélő anyanyelvének egyetlen fonémájával sem,

(d) a helyettesítés fellépése nem fonetikai–fonológiai környezeti hatástól függ;



(2) azok a helyettesítések, amelyekre jellemző, hogy:

(a) az adott beszédhang helyettesítése valamely más beszédhanggal nem kivétel nélküli, és a helyettesítés nem szükségképpen történik mindig ugyanazzal a helyettesítővel,

(b) a helyettesített beszédhang megjelenik vagy valamely más beszédhang helyettesítőjeként is, vagy érintetlenül (= helyettesítési folyamatban részt nem vevőként),

(c) a helyettesítő beszédhang mindig egyegyértelmű megfelelésben áll a beszélő anyanyelvének valamely fonémájával,

(d) fellépésük, valamint a helyettesítő kiválasztása, amennyiben motivált, fonetikai–fonológiai környezeti motivációjú.

Az (1)-ben leírt helyettesítéstípus megjelenhet nem afáziás patológiás beszédben is (2) előfordulása, illetve bármely más nyelvi szint sérülésére utaló jelenség nélkül.

A (2)-ben leírt típus megjelenhet önmagában vagy valamely más nyelvi szint sérülésére utaló jelenségekkel együtt, az (1)-ben leírt típus előfordulása nélkül, vagy azzal együtt.

(1a,b,d) és (2a,b,d) alapján, ha ugyanabban az afáziás beszédben előfordulnak mind (1), mind (2) típusú tévesztések, azokat egymástól elkülöníthetőnek tekintem. Közülük pedig a (2) típusba tartozókat tartom fonetikai–fonológiai alapon értelmezendőnek.

(1c) és (2c) alapján kimondjuk, hogy a csak (2)-t tartalmazó afáziás beszéd a beszélő anyanyelvének minden fonémáját és csak azokat foglalja magában.

A (2)-ben leírt tévesztésekben fellépő információ elakadását nem az információ kieséséből, hanem hozzáféréseinek átmeneti gátoltságából eredőnek tekintem: a helyettesített elem vagy annak valamely összetevője nem hiányzik a beszélő fonológiai rendszeréből, hanem az adott közlésben nem vagy nem a megfelelő időzítéssel hívható elő. Ezt alátámasztja még egy megfigyelés. Nem minden esetre jellemző, hogy az információ nem érkezik be. Ez pedig csak akkor lehetséges, ha az adott elemről szóló információ nem törlődött.

## Irodalom

- Blumstein, S.: A phonological investigation of aphasic speech. Mouton. The Hague 1973.
- Blumstein, S.: Phonological aspects of aphasia. Sarno, M. T. (ed.): *Acquired Aphasia*. Academic Press. New York 1981.
- Crystal, D.: *Introduction to language pathology*. Arnold. London 1980.
- Lecours, A. R. – Lhermitte, F.: Phonemic paraphasias: Linguistic structures and tentative hypotheses. *Cortex* 5, 1969, 193-228.
- Marcie, P.: Quelques remarques sur l'étude phonétique de l'aphasie motrice. Hécaen, H. (éd.): *Neurolinguistique et neuropsychologie*. *Langages* 25, 1972, 29-36.
- Nespoulous, J.-L. – Borell, A.: A propos des perturbations phonétiques et/ou phonémiques dans le discours aphasique: réflexions sur quelques données dans la littérature aphasologique. *La Linguistique* 15, 1979, 133-146.
- Nespoulous, J.-L. – Joannette, Y. – Béland, R. – Caplan, D. – Lecours, A. R.: Phonologic disturbances in aphasia: Is there a "markedness effect" in aphasic phonemic errors? Rose, F. C. (ed.): *Advances in Neurology* 42. *Progress in Aphasiology*, 203-214. Raven Press. New York 1984.
- Nespoulous, J.-L. – Lecours, A. R. – Joannette, Y.: Stabilité et instabilité des déviations phonétiques et/ou phonémiques des aphasiques. Insuffisance d'un modèle statique d'analyse. *La Linguistique* 18, 1982, 85-96.
- Nespoulous, J.-L. – Lecours, A. R. – Joannette, Y.: La dichotomie "phonétique-phonémique" a-t-elle une valeur nosologique? Messerli, P. – Laurel, P. Nespoulous, J.-L. (éds): *Neuropsychologie de l'expression orale*, 71-91. CNRS. Paris 1983.
- Shankweiler, D. – Harris, K. S.: An experimental approach to the problem of articulation in aphasia. *Cortex* 2, 1966, 277-292.
- Szende T.: *Alapalak és lazítási folyamatok*. *Linguistica, Series A, Studia et Dissertationes* 22. MTA Nyelvtudományi Intézet. Budapest 1997.
- Trost, J. E. – Cantor, G. J.: Apraxia of speech in patients with Broca's aphasia: A study of phoneme production accuracy and error patterns. *Brain and Language* 1, 1974, 63-79.

# MOZGÁSSÉRÜLT GYERMEKEK BESZÉDÉSZLELÉSE ÉS BESZÉDMEGÉRTÉSE

Szuhaj Eszter

Mozgássérültek Pető András Nevelőképző és Nevelőintéze-  
tének Gyakorló Közoktatási Intézménye, Budapest

## Bevezetés

Az anyanyelv-elsajátítás során a kisgyermek mintegy három év alatt alapfokon birtokba veszi azokat a nyelvi eszközöket, amelyek segítségével beszél és a hozzá intézett közléseket megérti. Azoknál a gyermekeknél, akiknél valamilyen – a beszédtanulást érintő – organikus zavar áll fenn, az anyanyelv-elsajátítás nagyon nehéz feladat. A folyamat lassú, és az organikus zavartól függően a gyermek még iskoláskorára sem éri el az elvárt szintet. Ahhoz, hogy az ilyen problémával küszködőknek segíteni tudjunk, szükség van arra, hogy pontosan ismerjük az anyanyelvi állapotot egy adott életkorban, és a lehetőségek szerint objektív adatok alapján tudjuk az épen fejlődőkhöz hasonlítani a kapott teljesítményt. Különösen nehéz ez akkor, ha közvetlenül nem vizsgálható folyamatokról van szó, mint amilyenek a beszédészlelés és a beszédmegértés (Gósy 1999).

A jelen tanulmányban mozgássérült óvodások anyanyelvi szintjének vizsgálati eredményeit mutatjuk be a beszédfeldolgozás szempontjából. Ismereteink szerint ez az első, tudományosan megalapozott, egységes diagnosztikát alkalmazó kutatás ezen a területen. Az alábbiakban huszonhat, részben ép értelmű, részben értelmi fogyatékos mozgássérült gyermek beszédpercepciós teljesítményét tárgyaljuk. A tanulmány kiindulása az ép értelmű és az enyhe értelmi fogyatékos, súlyosan mozgássérült tanulók beszédfeldolgozási szintjének összehasonlítása, majd pedig az enyhe értelmi fogyatékos, súlyosan mozgássérült tanulók egyéves fejlődési eredményeinek tárgyalása következik.



### **A kísérletben részt vevő gyermekek**

A vizsgált gyermekek életkora 7 és 11 év közötti, közülük 12 leány, 14 fiú, valamennyien a Pető Intézet tanulói. Központi idegrendszeri károsodással élnek, mozgássérülésük kihat mind a négy végtagjuk mozgására. Diagnózisként tetraparesis spastica, illetőleg athetosis jelölhető meg, amelyek oka születés előtti, alatti vagy közvetlen a születés utáni agyi sérülés. Mozdásállapotuk súlyos, többségük gurulva, kúszva képes az önálló helyváltoztatásra. Ülésbiztonságuk kialakulóban van, kapaszkodás, támaszkodás nélkül nem tudnak ülni, de asztalnál támaszkodva, vagy megkapaszkodva biztonságosan ülnek. Felállni, állni, járni segítség, segédeszköz nélkül nem képesek. A legjobban járó gyerek is két botot használ, hamar elfárad. Az önellátásban az egészségügyi szükségleteiket jelzik, segítséggel el tudják látni magukat. A gyermekek többségének beszéde nem tiszta, kiejtésük néhány esetben kevésbé érthető.

A 26 gyermek közül tizet általános iskolába javasoltak, az ép értelmű mozgássérültek osztályaiba, tizenhatot pedig kisegítő iskolába, az enyhe értelmi fogyatékos mozgássérültek osztályaiba.

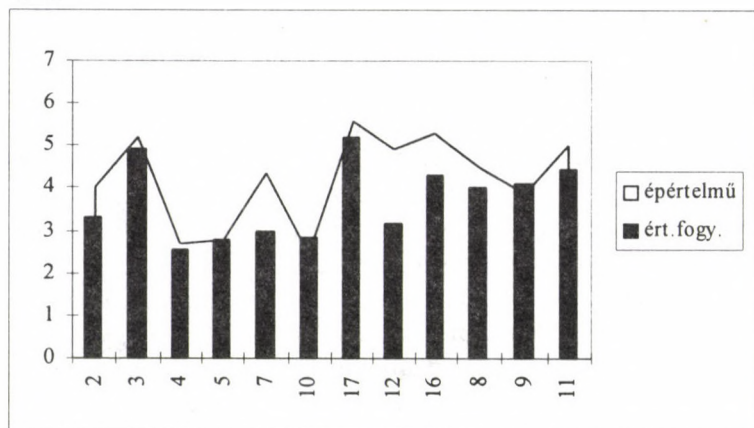
### **Anyag és módszer**

A gyermekek beszédészlelését és beszédmegértését iskolába kerülésükkor a GMP-diagnosztikával mértük fel (Gósy 1995a). A diagnosztika egyes résztesztjei a közvetlenül nem tanulmányozható percepciós folyamatok szintjéről és működéséről adnak felvilágosítást. A korspecifikus életkori adatok alapján lehetőség van a mozgássérült gyermekek teljesítményének objektív összevetésére ép társaikhoz képest. A tesztelést egyénileg végeztük valamennyi gyermekkel, a teszt felvételét csak annyiban módosítottuk, amennyire azt a speciális csoport megkívánta.

### **Eredmények**

A súlyosan mozgássérült 10 ép értelmű, és 16 enyhe értelmi fogyatékos gyermek átlageredményeit összesíti az 1. ábra. A vízszintes tengelyen az észlelési és a megértési résztesztek sorszámai láthatók. A függőleges tengely az életkort mutatja, amelyre a vizsgált gyermekek

adatait átszámítottuk. Így azonnal látható az elmaradásuk mértéke is. A **beszédészlelést** méri a 2, 3, 4, 5, 7, 10, 17-es részteszt, a **beszédmegértést** a 12, 16, az **emlékezeti működést** a 8, 9 és a **szókincsaktiválást** a 11-es részteszt.



1. ábra

Súlyosan mozgássérült, ép értelmű és enyhe értelmi fogyatékos tanulók beszédpercepció átlageredményei

A grafikonon az összefüggő, világosabb színű területtel az ép értelmű, súlyosan mozgássérült gyermekek által az egyes résztesztben elért átlageredményeket jelöltük. A fekete oszlopok az értelmi fogyatékos, súlyosan mozgássérültek átlageredményeit mutatják.

Megállapítható, hogy a 7 éves 3 hónapos átlagéletkorú tanulóknak az egyes résztesztben mutatott átlageredményei súlyos elmaradást mutatnak az ebben az életkorban elvárható képest. Néhol „csak” másfél-két éves, másutt több, mint 4 éves az elmaradás. Várhatóan az értelmi szint eltérései jelentkeznek a beszédfeldolgozási teljesítményben. A beszédészlelésben legnagyobb az elmaradás. A beszédmegértésben a szövegértésnél is és a mondatértésnél is jelentős különbség tapasztalható. Hasonló a helyzet a vizuális észlelés tekintetében is, ugyanakkor az emlékezeti működésekben és a szókincs aktivizálásában a két csoport között nincs jelentős teljesítménybeli különbség.



A mérések eredményeiből egyértelműen megállapítható a súlyosan mozgássérült gyermekek beszédészlelési és beszédmegértési zavara, illetőleg elmaradása. A vizsgált gyermekek a korábbi fejlesztések során nem részesültek célzott beszédészlelési és beszédmegértési terápiában. Felmerül a kérdés, hogy vajon az organikus sérülés kizárja-e a gyermekek beszédészlelésének és beszédmegértésének fejlődését, fejlesztését?

Erre a kérdésre kerestük választ egy fejlesztő program egyéves alkalmazásával. A fejlesztő program a GMP-diagnosztikához kapcsolódó fejlesztő könyvek felhasználásával készült egy 14 fős gyermekcsoport részére (Gósy 1994; 1995b; Porkolábné Balogh 1987). A 14 súlyosan mozgássérült, enyhén értelmi fogyatékos gyermekből 3 nem beszélő, 11 pedig a már bemutatott 26 gyermek közül került ki. A 11 súlyosan mozgássérült, enyhe értelmi fogyatékos tanulót a tanév elején (1999 szeptemberében) vizsgáltuk meg GMP-diagnosztikával, és a kapott eredmények alapján állítottuk össze a fejlesztő programot. Az egyéves fejlesztés után, 2000 szeptemberében kontroll vizsgálatot végeztünk, amelynek adatait összevetettük az egy évvel korábbi GMP-diagnosztikai vizsgálat eredményeivel.

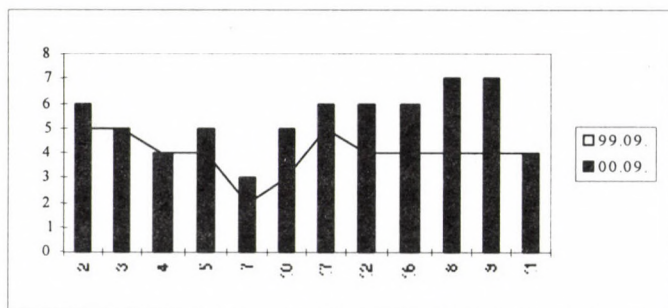
### **A fejlesztés eredményei**

A Pető Intézetben a tanulók többsége heti bentlakó. A tanév során hétfő reggeltől péntek délutánig hat konduktor foglalkozik a gyermekcsoporttal. A konduktív pedagógia jellemzője, hogy a fejlesztés csoportban, de egyénre szabottan történik. Az egyszerre együtt dolgozó 2-4 konduktor közül a nap egyes szakaszaiban váltakozva egy konduktor irányítja a csoportot, a többiek segédkeznek. Valamennyi kolléga alkalmazta az év elején összeállított fejlesztő programot. A program év közbeni módosítását – az aktuális eredményeknek megfelelően – közösen végeztük. A fejlesztésbe bevontuk a végzős főiskolai végzős hallgatókat, valamint a szülőket is, így a gyermekekkel foglalkozó valamennyi felnőtt alkalmazta a fejlesztő eljárásokat.

A 2. ábra grafikonján egy közel 8 éves kisfiú eredményei láthatók, az összefüggő halvány terület mutatja az 1999. szeptemberi, a fekete oszlopok pedig a 2000. szeptemberi adatokat. Pontosán 12 hónap,



azaz egy év volt a fejlesztés időtartama. A függőleges tengelyen az évek (életkor, illetve az életkori szint) száma olvasható, a vízszintes tengelyen pedig az egyes GMP-diagnosztikai résztesztek sorszámát látható (mint a megelőző ábrában).

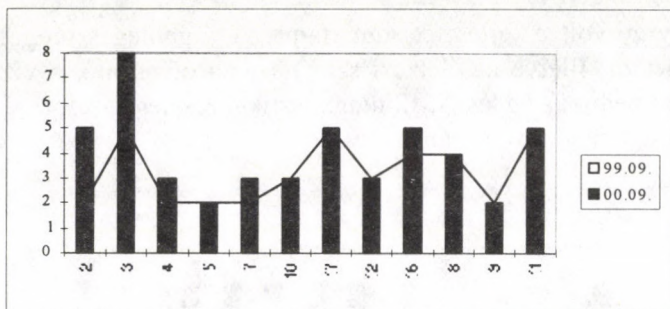


2. ábra

A. P. 7 éves 9 hónapos fiú fejlődése

A. P. a bemutatott 12 részképesség-terület közül kilencben a vizsgált időszakban több, mint egy évnyi fejlődést mutat. A további célirányos fejlesztés hatására elképzelhető, hogy egyre inkább behozza többéves elmaradását. Beszédészlelése – ezen belül is leginkább a vizuális észlelése –, valamint a szókincsaktiválási képessége lassúbb fejlődést mutat. A. P. kezessége kialakult, jobbkezes, bal szemes. Az irányokat ismeri, ritmusészlelése jó. Rajzolási szintje: 7 év a Goodinough-féle rajzvizsgálattal mérve.

A 3. ábra egy 8 éves leány fejlődési adatait tartalmazza.

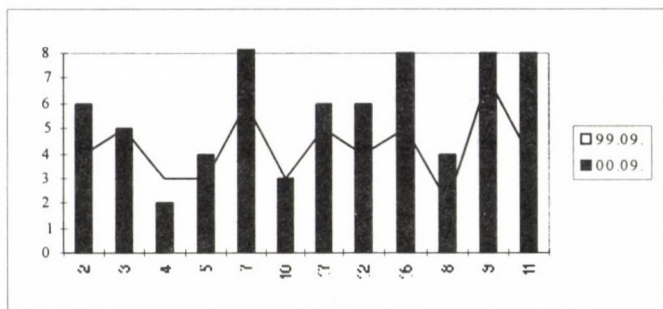


3. ábra  
B.V. 8 éves leány fejlődése

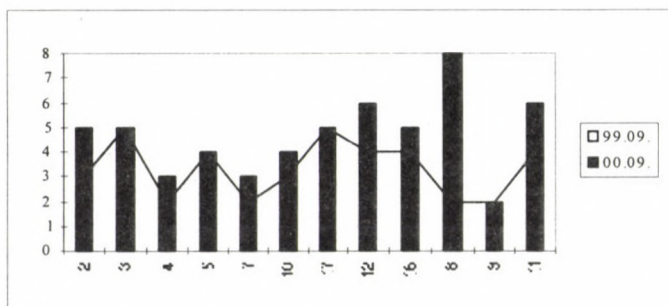
B.V. részképességeinek fejlődési üteme nem kiegyenlített. A beszédészlelés akusztikai szintjén nagy fejlődést mértünk, ezen kívül vizuális észlelése, mondatértése és szókincsaktiválási képessége fejlődött az egy év alatt egy évnnyit. A többi területen fejlődést nem mutatott. Kezessége kialakult, jobbkezes, jobb szemes. Az irányokat ismeri, ritmusészlelése jó. Rajzolási szintje 4 éves 6 hónapos szintnek felel meg.

A 4. és 5. ábra 8 év körüli gyermekek adatait szemlélteti. C. S. szintén kilenc területen mutat egy évnnyinél nagyobb mértékű fejlődést a fejlesztés egy éve alatt. Részképességei nem kiegyenlítetten fejlődnek, de remélhető, hogy több területen egyre inkább behozza lemaradását, és csakhamar eléri az életkorának megfelelő szintet. C. S. kezessége kialakult, jobbkezes, jobb szemes. Az irányokat téveszti, ritmusészlelése jó. Rajzolási szintje: 7 éves szintnek megfelelő.

E. T. diagnózisa nem egyértelmű. Tetraparesis spastica diagnózishoz kisagyi tünetek társulnak, nem kizárt, hogy állapota progrediál. A kislány fejlődése több területen meghaladja az egy év alatt elvárt mértéket. Legtöbbet rövid idejű verbális memóriája, szókincsaktiválási képessége és szövegértése fejlődött.



4. ábra  
C. S. 8 éves fiú fejlődése



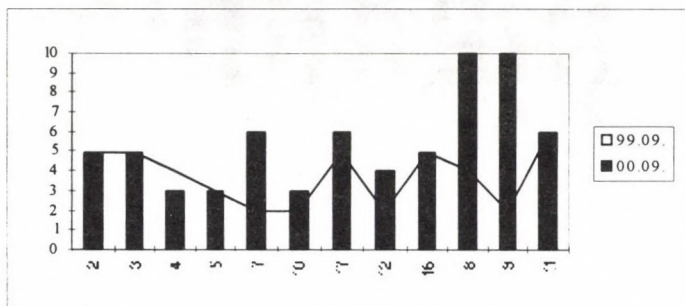
5. ábra  
E. T. 8 éves 3 hónapos leány fejlődése

E. T. kezessége nem kialakult, a vizsgálatok alapján jobbkezes, a tanév során jobboldali dominanciáját igyekeztünk kialakítani, illetőleg erősíteni. Jobb szemes. Az irányokat téveszti, ritmusészlelése gyenge. Rajzolni nem tud.

A 6. ábrán szereplő F. G. tetraparesis spastica diagnózisához microcephalia társul. 9 éves 3 hónapos korában került iskolába, addig otthonában nevelkedett. Fejlődése nem egyenletes. A vizuális észlelés és az emlékezeti működések területén nagyon sokat fejlődött (5-7 évnőt az egy év alatt!), ugyanakkor több területen – a célzott fejlesztés



ellenére sem mutatkozik eredmény. Ennek oka valószínűleg a súlyos és halmozott sérülése, valamint az életkora. Kezessége kialakult, balkezes, bal szemes. Az irányokat téveszti, ritmusészlelése gyenge. Rajzolsági szintje: 4 év 9 hónapnak megfelelő.

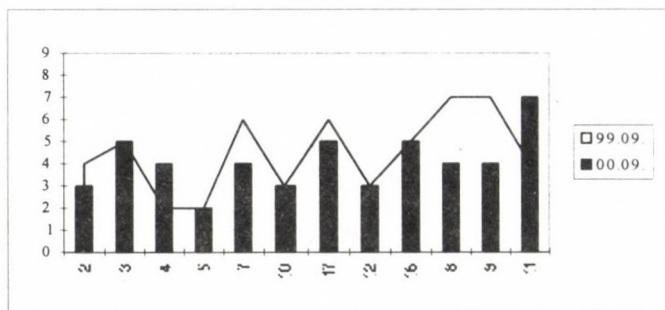


6. ábra

F. G. 10 éves 3 hónapos leány fejlődése

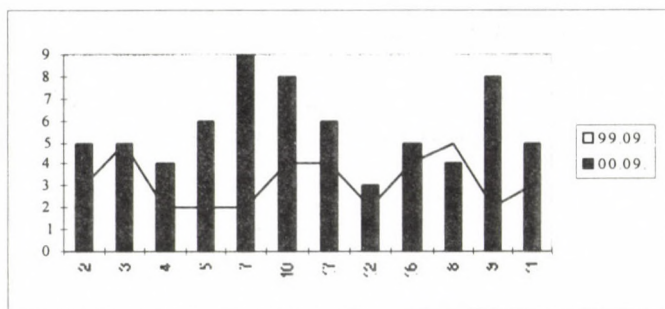
K. H.-nak a testvére is kisegítő iskolába jár. A kislány egyedül a szókincsaktiválás területén mutat fejlődést (vö. 7. ábra). Kezessége kialakult, balkezes, bal szemes. Az irányokat ismeri, ritmusészlelése gyenge. Rajzolsági szintje 4 év 3 hónapos szintnek megfelelő.

M. B. állandóan hurutos kislány, több műtétet követően sem szűnik ez a problémája. Beszéde nehezen érthető. 10 területen fejlődött egy évnél vagy annál többet az egy év alatt. Legnagyobb mértékben vizuális és szeriális észlelése, valamint rövid idejű vizuális memóriája fejlődött (8. ábra). M. B. kezessége kialakult, jobbkezes, jobb szemes. Az irányokat téveszti, ritmusészlelése gyenge. Rajzolsági szintje 4 év 3 hónapos szintnek felel meg.



7. ábra

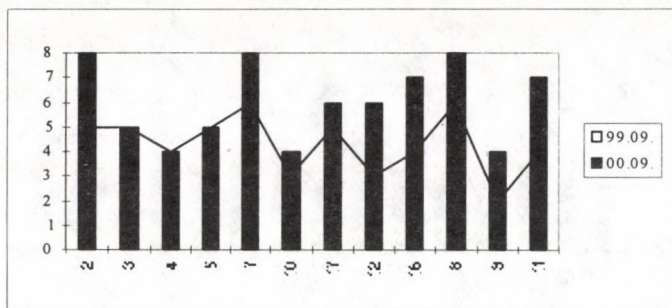
K. H. 8 éves 6 hónapos leány fejlődése



8. ábra

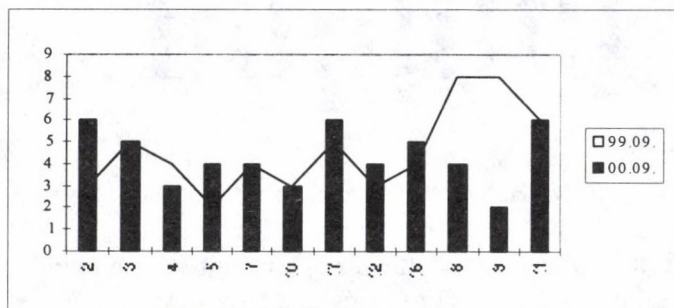
M. B. 8 éves 6 hónapos fiú fejlődése

M. O. 8 területen fejlődött az egy év alatt egy évnyinél többet. Életkori szintjét a 2-3 évnyi fejlődéssel a beszédészlelés akusztikai és vizuális észlelési területén, valamint a rövid idejű verbális memória területén érte el (9. ábra). Fejlődési ütemét ismerve remélhetjük, hogy egyre több területen behozza lemaradását. M. O. kezessége kialakult, jobbkezes, bal szemes. Az irányokat ismeri, ritmusészlelése jó. Rajzolás szintje 5 évesnek megfelelő.



9. ábra  
M. O. 8 éves fiú fejlődése

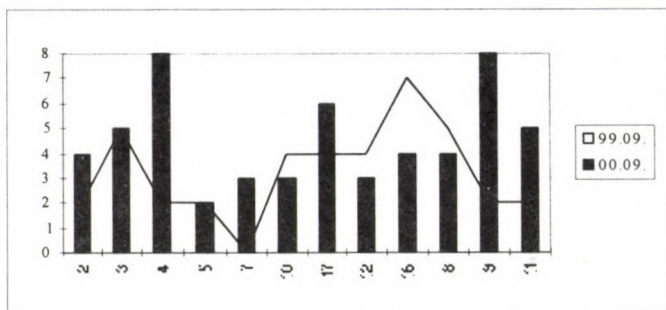
N. E. érzelmileg labilis kislány, teljesítménye változó. Az emlékezeti működések területén mutatkozó visszaesés valószínűleg ennek tudható be. Kiugró teljesítménynövekedés nem jellemzi (10. ábra). Kezessége kialakult, balkezes, bal szemes. Az irányokat téveszti, ritmusészlelése jó. Rajzolás szintje 4 év 3 hónapnak megfelelő.



10. ábra  
N.E. 9 éves leány fejlődése

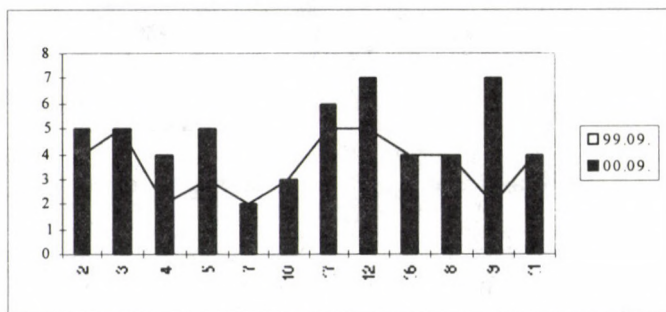
R. G. tetraparesis spastica diagnózisához kisagyi tünetek társulnak. Fejlődése két-három területen ugrásszerű, más területeken az egy évvel ezelőttihez képest gyengébben teljesített (11. ábra).





11. ábra  
R. G. 8 éves fiú fejlődése

R. G. kezessége nem kialakult, a vizsgálatok alapján jobbkezes, a tanév során jobboldali dominanciáját igyekeztünk kialakítani, illetőleg erősíteni. Jobb szemes. Az irányokat téveszti, ritmusészlelése gyenge. Rajzolási szintje 4 év 9 hónapnak megfelelő.

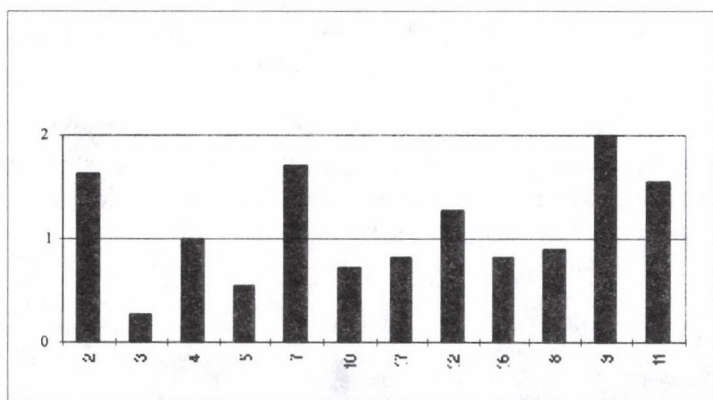


12. ábra  
Z. T. 7 éves 9 hónapos fiú fejlődése

Z. T. több területen mutat egy évnyi, vagy annál nagyobb fejlődést a vizsgált időszakban (12. ábra). Rövid idejű vizuális memóriája fejlődött a legtöbbet. Balkezes, bal szemes. Az irányokat ismeri, ritmusészlelése jó. Rajzolási szintje 5 év 3 hónapnak megfelelő.

## Összegzés

A 11 tanuló egyéves fejlődési eredményeit megismerve látható, hogy súlyos organikus sérülésük ellenére is fejleszthetők, fejlődésük a GMP-diagnosztikával objektíven mérhető. A diagnosztikához kapcsolódó fejlesztő könyvek felhasználásával készült fejlesztő program folyamatos alkalmazása eredményesnek bizonyult. Érdemes megnézni, hogy a gyermekcsoportnak az egyes részteszteken mért átlagos fejlődése mely területeken volt a legnagyobb (13. ábra).



13. ábra

Átlagos fejlődés az egyes részképeségek területén

A gyermekcsoport minden területen fejlődött. Az adott időszakban elvárhatónál nagyobb mértékű volt a fejlődés a rövid idejű vizuális memória (2 évnyi!), a vizuális észlelés, a szókincsaktiválás, a beszédészlelés akusztikai szintje és a szövegértés területén. Az organikus sérülés tehát nem zárja ki, csak jelentősen megnehezíti a beszédészlelés és a beszédmegértés fejlődését. A fejlesztés tervezéséhez, eredményeinek méréséhez kiváló eszköz a pedagógus kezében a GMP-diagnosztika.

## Irodalom

Gósy Mária: A beszédészlelés és a beszédmegértés fejlesztése óvodásoknak. Nikol. Budapest 1994.

Gósy Mária: GMP-diagnosztika. Nikol. Budapest 1995a.

Gósy Mária: A beszédészlelési és beszédmegértési folyamat zavarai és terápiája. Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola. Budapest 1995b.

Gósy Mária: Pszicholingvisztika. Corvina Kiadó. Budapest 1999.

Porkolábné dr. Balogh Katalin: Készségfejlesztő eljárások tanulási zavarral küzdő kisiskolásoknak. Iskolapszichológia sorozat 4. füzet, ELTE. Budapest 1987.

### **Köszönetnyilvánítás**

A szerző ez úton fejezi ki szívből jövő köszönetét az alábbiakban felsorolt kollégáinak, akik munkájukkal segítették a gyermekek fejlesztését: Bánki Gáborné, Nagyné Divinyi Éva, Paulin Márta, Siverné Radnóti Emőke, valamint Kocsis Katalin.



# VÁLTOZOTT-E HÚSZ ÉV ALATT A FŐISKOLÁSOK KIEJTÉSE ÉS OLVASÁSA?

Adamikné Jászó Anna  
ELTE Tanárképző Főiskolai Kar

## Bevezetés

Egy 1978 szeptemberében végzett felmérést 1989 novemberében, majd 2000 februárjában megismételtem. Az első két felmérést publikáltam, közöltem bennük a szakirodalmi háttérrel, s hasonló célú felmérések eredményeit, egészen egy 1891-ben Aradon végzett vizsgálattól kezdve; ezeket most nem ismételtem meg (Adamik 1979; Jászó 1994).

## Anyag és módszer

1978-ban 24 tanítóképzős hallgató, 1989-ben 18 tanítóképzős hallgató, majd 2000-ben 30 tanárképzős nem magyar szakos hallgató beszédét vizsgáltam meg. A felmérés a következőképpen zajlott le: néhány bemutatkozó mondatot, majd egy kb. 1,5 perces szöveget vettem magnetofonra, Sütő András Jön az erdő! című írásának első bekezdését (Anyám könnyű álmot ígér. Kriterion, Kolozsvár 1972). Az előző két tanulmányban nem tartottam szükségesnek Sütő közismert sorait közölni, mivel az olvasás nem okozott gondot; a legutóbbi felmérésben azonban nagyon sok volt az olvasási hiba, ezért közlöm a felolvasásra szánt szöveget:

*„Az elemi iskolában osztálytársaimnál hamarabb tanultam meg a betűvetést. A sikongó palatábla fölé hajolva, ebéd közben is mámorosan róttam a szavakat. Úgy éreztem: mindahánnyal egy-egy titkot csípek fölön. Még nem tudtam, mit kezdek majd velük, merre s mi végett indítok rohamot. De hogy leírtam a félelmetes szót: báró, máris többnek éreztem magam, majdnem egyenrangúnak a betűk rácsai mögé vetett fogalommal. A világ birtokbavételének illúzióját éltem át: szavaim egyes csapatának eszméltető, titkos erejét. Más örökségre – semmilyen ingó és ingatlan vagyongra –*

*nem számíthattam. Az ősök tankönyvbe foglalt testamentumaira sem. Mivel sáfárkodtam volna? A magaméihoz tehát, közeli és távoli rokonoknál, új szavakat eprésztem. Degenyegesek, fazekas vándorszékelyek szekere körül ólálkodva egy-egy friss jelzőt, igét, ikerszót dugdostam a nyelvem alá, s rohantam szívdobogva, mintha kancsót loptam volna: iafia, iafia! Anyám nevetett: pap lesz belőlem. Kérkedtem valósággal, mi mindennek tudom az elfelejtett – vagy soha meg sem tanult? – nevét.”*

A szöveg nem volt különösebben nehéz, az ismeretlen szavakat mindhárom alkalommal megbeszéltük. Minden tanulóról kiállítottam egy diagnosztikai lapot, amelyen az alábbi rovatok szerepelnek: *személyi adatok; artikulációs hibák: a hangképzés hibái, a beszédhangok hibái, a beszédfolyamat hibái (hadarás, dadogás, pattogás, lepegés); helyesejtési hibák: a beszédhangok hibái, a hangok egymásra hatásából származó hibák (időtartam, hangkapcsolatok); a szupraszegmentumok hibái (hangsúly, hanglejtés, tempó, szünet, junktúra).* A diagnosztikai lapra minden hibásan ejtett helyet feljegyeztem, majd a kapott adatokat összesítettem. Ily módon arról kaptam képet, hogy egy bizonyos kiejtési hiba a tanulók hány százalékának beszédében fordul elő. Természetesen, egy tanuló kiejtése több szempontból is kifogásolható. Meg kell említenem, hogy a felmérést eredetileg pedagógiai szándékkal készítettem: egyrészt saját hallásomat kívántam vele edzeni, másrészt segítségével a tanulók fejlesztését jobban meg lehetett tervezni. A tanulók megkapták a diagnosztikai lapokat, s visszahallgathatták a felvételeket: sokan csak ekkor észlelték hibáikat, hiszen belső hallásuk, percepció bázisuk saját beszédüknek megfelelően működik. Megjegyzem, ilyen módszert jó légkörben lehet csak alkalmazni, s a bizalom, az együttműködési szándék mindig megvolt. A mérés tehát spontán szövegből, felolvasásból és utólagos vizsgálatból állt (a beszédhibás tanítójelölteket logopédus is megvizsgálta, majd külön kezelte; logopédiai kezelésre a tanárképzőn is van lehetőség).

A spontán beszédek és a felolvasások meghallgatása után kialakult összbnyomásról a következőket lehet elmondani. Alig volt olyan hallgató, aki érthetően mondta volna a nevét, a szalagon sokukét egyáltalán nem lehet érteni. Fiúk és lányok egyaránt igen zártan beszélnek, s felkapják a mondatvég dallamát, vagyis kunkorítanak. Feltűnő

volt, hogy olvasáskor magasabb hangfekvésben beszéltek, ezt az előző felmérésekben nem tapasztaltam. Két hallgató beszédében érezhető volt némi regionális köznyelvi jelleg, egyikük – egy debreceni fiú – hátrább képezte a [t] és a [d] hangokat, másikuk – egy lány – olykor dunántúlias nyílt [e] hangokat ejtett; ezeket a változatokat természetesen nem tekintettem hibának.

### Kísérleti eredmények

A táblázatban a számok a százalékot jelentik, a csillag pedig az enyhébb eseteket jelzi.

Hibatípusok	1978	1989	2000
<b>Artikulációs hibák</b>			
[ʃ]-hangok			
lateralis	4,1%	–	–
interdentalis	4,1%	5,5%	3,3%
stridens	4,1%	–	–
stridens*	37,5%	72,2%	56,6%
[s]-hangok			
lateralis	4,1%	–	–
interdentalis	4,1%	5,5%	–
addentalis	–	–	3,3%
interdentalis*	4,1%	–	–
stridens*	–	–	56,6%
[r]-hangok			
uvuláris	4,1%	–	3,3%
renyhe képzés*	33,3%	61,1%	43,3%
dadogás	–	–	3,3%
hadarás	12,5%	–	–
lepeggés	8,3%	–	3,3%
pattogás	–	–	10%
<b>Helyesejtési hibák</b>			
nazális magánhangzók	12,5%	33,3%	6,6%
[ɔ] (illabiális)	8,3%	16,6%	3,3%
[ɛ] (labializált)	12,5%	5,5%	16%
[o](feljebb képzett)	8,3%	–	–
[ɔ] (zártabb képzésű)	12,5%	–	–



[k] (aspirált)	4,1%	–	–
<b>A beszédhangok egymásra hatásából származó hibák</b>			
magánhangzó-kivetés	29,3%	50%	–
mássalhangzó-kivetés	4,1%	50%	–
magánhangzó-rövidítés:			
– mindent röviden ejt	37,5%	–	46,6%
– az [a:] és az [e:] is röviden ejti	12,5%	72,2%	33,3%
– enyhén röviden ejt	54,1%	27,7%	20%
mássalhangzó-rövidítés	41,6%	11,1%	–
mássalhangzó-nyújtás	50%	33,3%	
<b>A szupraszegmentumok hibái</b>			
– nincs mondathangsúly	16,6%	5,5%	40%
– nincs szólamhangsúly	33,3%	5,5%	16,6%
– túlhangsúlyozás	12,5%	–	–
– névutót, kötőszót hangsúlyoz	4,1%	–	3,3%
– szóösszetételt rosszul hangsúlyoz	8,3%	16,6%	23,3%
– a mondat végén lebegteti a dallamot	25%	27,7%	3,3%
– a mondat végén felemeli a dallamot	20,8%	5,5%	6,6%
– a mondat végén erősen leejti a dallamot	8,3%	5,5%	10%
a tagolás hiánya	8,3%	–	6,6%
a kiegészítendő kérdés dallama nem megfelelő	4,1%	22,2%	3,3%
nem tart szünetet a beékelésnél	16,6%	55,5%	23,3%
túlságosan gyors tempó	33,3%	27,7%	40%
a természetesnél magasabb hangfekvés	12,5%	–	86,6%

Úgy látszik, bele kell törödnünk abba, hogy a tanító- és a tanárjelöltek között mindig előfordulnak logopédiai esetek; az első felméréskor volt egy laterális szigmatizáló, az utolsó felméréskor pedig egy súlyos klónusos dadogó. Biztosan nem lesz tanár, de azért az iskolai gyakorlaton át kell esnie. A logopédusok nem szokták megkülönböztetni a kiejtési hibákat a beszédhibáktól, s mindig beszédhibáról be-

szélnek, a beszédművelő tanárok viszont különbséget tesznek a beszédhibák és a kiejtési hibák között. Kiejtési hibának a normától való kisebb fokú eltéréseket nevezzük (Vértes O. 1956), a modorosságokat vagy az olyan susogást és sziszegést, amely inkább modorosság, illetőleg a zárt szájú beszéd következménye: a beszélő helyesen ejti a hangot, ha odafigyel.

A legfeltűnőbb kiejtési hiba a magánhangzók röviden ejtése (Gósy 1997), az arányok eltolódtak a 20 évvel ezelőtti állapothoz képest: másfélszeresére nőtt a mindent röviden ejtők száma, s háromszorosára azoknak a száma, akik az [a:] és az [e:] magánhangzókat is röviden ejtik. Ez már lényegében pattogó beszédet eredményez, az igazán pattogó beszélő semmit nem hangsúlyoz. Úgy tűnik fel, hogy a mai fiatalok beszédében nincsenek szóvégi hosszú magánhangzók, sem toldaléktalan, sem toldalékos szavakban. A harmadik felmérés megerősítette mindazt, amit 1989-ben megállapítottam: hallgatók beszéde monoton, nem bánnak jól a hangsúlyokkal és a szünetekkel. A résztvevőknek mintegy a fele nagyon gyorsan beszél, spontán beszédük még gyorsabb. A régebbi felmérésekben nem okozott gondot a szóhatár, mindig a megfelelő helyen volt. A mostani olvasók sokszor eltolták a szóhatárt a névelő és a szó esetén, efféléket olvastak: *a zelemi, a zösök, a zelfelejtett*.

Az 1989-es felmérés után óvatosan állapítottam meg, hogy talán új jelenség a [b] mássalhangzó intervokális helyzetben, velárisok közötti spirantizálódása: *szívdo**β**ogva*. Arra gyanakodtam, hogy talán a gyors beszéd miatt lazul, eltolódik az explozíva képzése. A [v] helyett is a bilabilális spiráns lehetett hallani: *szafakat, hajol**β**a*, sőt a *mivel* is inkább *mifel*nek hallatszott. Az ideji felmérés igazolta feltevésemet: mindegyik hallgató beszédében érzékelhető volt a bilabiális spiráns. Az *egyenrangú* első szótagjának a mássalhangzója is sokszor elmosódottan hangzott. A *degenyegesek*ben zöngés mediopalatális spiráns hallatszott a [g] helyett. Nem egy-két esetben, hanem a beszélőknek mintegy fele így ejtette. Lehetséges, hogy elkezdődött egy második hangeltolódás? Lehetséges, hogy korosztályhoz és társadalmi réteghez lehet kötni? Érdeemes lenne ezt a jelenséget szociolingvisztikai módszerekkel is megvizsgálni.



1978-ban mindenki jól olvasott, 1989-ben már sok olvasási hibát vettem észre, a legutóbbi felmérés az olvasás szempontjából pedig tragikusnak mondható. Mindössze két hallgató olvasott hibátlanul, négyöt apró tévesztésekkel, olyanokkal, amelyek mindig előfordulnak. A többség azonban sokat hibázott, három főiskolás olvasása pedig alsó tagozatos szinten volt. Az is árulkodó tény, hogy olvasáskor sokan megemelték hangjukat, magasabb hangfekvésre váltottak, mint a régi elemi iskolások vagy a keveset olvasó öregek. A bizonytalan olvasást jellemzik a félreolvasások, az újrakezdések, a betoldások, a hezitálások, a mást olvasások.

A félreolvasások a szóvégekre, illetőleg a szavak második felére jellemzők: *mit kezdek* > *mit kezdjek* (2), *mindahánnyal* > *mindhánnyal*, *többnek* > *többet*, *rácsai* > *rácsi*, *illúzióját* > *illúzijóját*, *szavaim* > *szavam*, *mi mindennek* > *mi mindnek*, *mi mindent* (2), *magaméihoz* > *magaméhoz*, *ólálkodva* > *ólálkodra*, *tudom* > *tudtam*, *eszméltető* > *eszmélet*, *eszméletétől*, *rokonoknál* > *rokonnál*, *betük* > *betű*, *közeli* > *közel*. A *testamentumaira* komoly próbatétel volt, négyen többszöri nekirugaszkodásra tudták kibetűzni, jobbik esetben lelassítottak és szótagoltak a szó közepén. A félreolvasások oka lehet az a fajta találgató olvasás, amelyre a szóképes előprogramok idején szoktatták rá a gyermekeket (ez a korosztály még elsős korában így „tanult” olvasni), de lehet a szó, illetőleg a grammatikai alak nem ismerete is. Érdekes a *hamarább* > *hamarabb* (6) félreolvasás, a hallgatók valószínűleg a *hamarabb* alakváltozatot ismerik, s ezt olvassák bele a megadott változatba. Sok tanuló az alakváltozatokat mereven ítéli meg, csak a párok egyik tagját tartja helyesnek (Szabó 1994). Az időtartam-problémák nagy része is ilyen: úgy olvasnak, ahogy a percpációs bázisuk diktálja, s nem úgy, ahogy látják a szót (diktálásakor is úgy írnak, ahogyan hallanak, hiába nyújtja meg a diktáló tisztességesen a hangokat).

Sok volt az újrakezdés, a következő szavakban: *magaméihoz*, *szívdobogva*, *loptam volna*, *mindahánnyal*, *dugdostam*, *degenyegesek*, *kancsót*, *illúzióját*, *elfelejtett*, *egyenrangúnak*, *nevét*. Betoldások is voltak szép számmal: *sáfárkodtam* > *sáfárkodhattam*, *számítottam* > *számíhattam*, *vagy soha* > *vagy talán soha*, *mit kezdek velük* > *mit kezdek majd velük*, *kérkedtem valósággal* > *kérkedtem a valósággal*.



Egy-egy szokatlan – vagy tán soha nem látott – szót hosszabb (szinte alapevegőnyi) hezitációs szünet előzött meg: □ *De hogy, pala* □ *tábla*, □ *eprézsstem, mit* □ *kezdek*, □ *kancsót*, □ *vándorszékelyek*.

A más olvasások oka a szó nem ismerése, illetőleg a fentebb említett találgató olvasás: meglátja az első szótagot, majd tetszés szerint kiegészíti: *mindahánnyal* > *mindannyian*, *degenyegesek* > *degenyese*, *meg sem tanult* > *meg nem tanult*, *ősök* > *örök*, *vándorszékelyek* > *vándorszínészek*. Volt olyan, aki a *sáfárkodtam volna* összetett igealakot kérdő intonációval olvasta, vagy a szót nem ismerte, vagy a feltételes múlt volt számára szokatlan. Ez történt a *nyelvem alá* szerkezettel is.

A bemutatott olvasási hibák tulajdonképpen nem meglepőek, különösen nem azok a gyakorlat és egyik-másik olvasástanítási program ismeretében. Tulajdonképpen megjósolhatóak, ha nem megfelelő a tankönyv. Nem véletlen, hogy a Gósy Máriával és Lénárd Andrással közösen írt ábécénkben szótagoltatunk a pontos olvasás elsajátíttatása végett. Az első három osztály számára pedig olyan olvasólapokat szerkesztettünk, amelyeken azonos kezdetű – különböző végű szavak, toldalékos hosszú szavak, magánhangzó-torlódásos szavak sorjáznak szép számmal, az egyéb struktúrák mellett. A szöveg megértése alapvetően az olvasási technikától függ: ha nem tudják dekódolni a szavakat, nem is érthetik őket. Az a tény azonban lesújtó, hogy a főiskolásoktól mennyire távol áll egy metaforikus szöveg, mennyire távol áll a szavak mögötti világ. Érettségi vizsga, Ady, Babits, Nagy László után.

### **Következtetések**

Változott-e húsz év alatt a főiskolások kiejtése? A feltett kérdésre igennel válaszolhatunk. Gyorsabb a beszéd, rövidek a magánhangzók és gyakoriak az enyhébb artikulációs hibák. A kiejtés ilyen jellegű változásai már húsz évvel ezelőtt elkezdődtek, az ezredfordulóra felerősödtek. Felmérésem erre a felerősödére hívja fel a figyelmet. Az olvasás színvonala nagyon sokat romlott, az olvasási hibák nemcsak a felolvasók gyakorlatlanságát jelzik, hanem azt is, hogy távol állnak a mívés szövegtől és a szöveg mögött meghúzódó valóságtól: nemzeti múltunktól. Elszomorító az a tény, hogy pedagógusok lesznek, s hibáik ily módon újratermelődnek.

### Irodalom

Adamik Tamásné – Cs. Nagy Lajos: Tanítóképzős hallgatók kiejtésének vizsgálata. A Budapesti Tanítóképző Főiskola Kiadványa 4. Tudományos Közlemények I. 1979, 80-90.

Adamikné Jászó Anna – Gósy Mária – Lénárd András: A mesék csodái. ABC és olvasókönyv. Dinasztia Kiadó. Budapest 1996.

Gósy Mária: Semleges magánhangzók a magyar beszédben. Magyar Nyelvőr 121. 1997, 9-19.

Jászó Anna: A kiejtési norma érvényesülése főiskolai hallgatók beszédében. In: A magyar nyelvi norma érvényesülése napjaink nyelvhasználatában. Szerk.: Kemény Gábor – Kardos Tamás. MTA Nyelvtudományi Intézet. 1994, 169-176.

Kontra Miklós: Az élőnyelvi vizsgálatok tárgyáról és módszereiről. In: Szabó Géza (szerk.): Dialektológiai szimpózium. MTA Veszprémi Akadémiai Bizottsága. 1960, 131-138.

Szabó Kálmán: A többalakú szavak a szakmunkástanulók szóhasználatában. Magyar Nyelvőr 118. 1994, 57-64.

Vértes O. András: A beszédhibákról. Magyar Nyelvőr 80. 1956. 343-354.

# AZ IDŐ VÁLTOZÁSÁNAK ÉS A BESZÉD ÁLLANDÓSÁGÁNAK PARADOXONA

Gósy Mária – Nikléczy Péter  
MTA Nyelvtudományi Intézet

## Bevezetés

A beszélő személy felismerésének kutatásában számos olyan kérdés is felmerül, amelyekre pontos válaszokat a tudomány még nem adott. Ha feltételezzük, hogy a beszéd alapján, a beszélő akaratlagos együttműködésével meghatároztuk mindazokat a paramétereket, amelyek a személy felismerését kétséget kizáróan lehetővé teszik, szembe kerülünk az idő változásának problémájával. Mit jelent ez?

Múlik az idő, változik az ember, s ebben a változásban óhatatlanul részt vesznek a beszélő szervek, az artikulációs mozgások, a teljes beszédprodukciónak. Az első kérdés az, hogy vajon mennyi idő elteltével kell a beszéd észlelhető változását szükségszerűen tekintetbe vennünk. Egy év vagy tíz év múltán? Minden ember esetében azonos mértékű a változás és hasonló annak következménye? A gyakorlati alkalmazás is számos kérdésre vár feleletet a tudománytól. Például, ha beszédminta alapján azonosítjuk a beszélőt telefonon keresztül, akkor milyen gyakorisággal kell a személytől új beszédmintát rögzíteni a biztos felismeréshez. Az alkalmazott tudomány az ilyen és hasonló kérdésekre ma csupán hozzávetőleges válaszokat tud adni. Ismereteink és hiányaink az idő változásával és annak a beszédre gyakorolt hatásával kapcsolatban a következőkben foglalhatók össze.

a) A beszédképzés a személy életkorának előrehaladtával változik. Nem tudjuk azonban, hogy ez a változás a beszédnek milyen paramétereit érinti.

b) Ismert, hogy a beszédképzés egyéni eltérései csak részben akaratlagos működések eredményei (Nikléczy 1996). Nincsenek sok beszélőre kiterjedően vizsgálatok azonban arra vonatkozóan, hogy ugyanazon személy esetében mekkora idő elteltével tapasztalható lényeges változás a beszédben.



c) Kísérletek igazolták, hogy a fiatalok és az idősek között a beszéd több területén tapasztalható eltérés (Gósy 2000). Például a beszédtempóban (Gocsál 2000), az alaphangmagasság alakulásában (Balázs 1993), sőt az artikulációban is, de ez utóbbi még nincsen pontosan összegezve. Rögtön felmerül ugyanakkor a kérdés: honnan számítjuk az időskort, illetőleg az öregkort. Mettől és meddig beszélhetünk fiatalokról és középkorúakról? Az években meghatározott időszakaszok jellemzőek-e a beszédre is?

Tapasztalatból tudjuk, hogy könnyen felismerjük a beszédük alapján évek elteltével is azokat a személyeket, akikkel folyamatos kapcsolatban vagyunk (például telefonon át). Ha azonban két-három évtizedig nem beszéltünk az ismerősünkkel, akkor pusztán a beszéde alapján, nem fogjuk felismerni. Mindkét tény igaz, ami viszont tökéletes ellentmondás. Egyfelől ugyanis azt állítjuk, hogy a beszéd nem (nagyon) változik, hiszen felismerjük a beszélőt, másfelől pedig ugyancsak állítjuk, hogy oly mértékű a változás, hogy nem ismerjük fel a beszélőt. Melyik tapasztalati tény igaz, illetőleg ha mindkettő bizonyítható, akkor e paradoxon hogyan oldható fel?

A kérdés megválaszolására kísérletsorozatot terveztünk. Arra kerestünk választ, hogy 25 év elteltével milyen azonosságok és változások tapasztalhatók ugyanazon beszélők beszédprodukcójában.

### **Anyag, módszer, kísérleti személyek**

A vizsgálatssorozat az tette lehetővé, hogy (i) birtokunkban volt három beszélő két és fél évtizeddel ezelőtt rögzített beszédanyaga, (ii) ugyanezek a beszélők jelenleg is rendelkezésre álltak és (iii) ugyanazon berendezésekkel volt megismételhető a hangfelvétel, mint amelyeket huszonöt évvel ezelőtt használtak. A hangrögzítés körülményei – beleértve a felvétel helyiségét, a csendes szobát – megegyeztek, ezáltal a változók számát gyakorlatilag az időre lehetett szűkíteni, vagyis a beszélő személyek életkorának változására.

Az említett három személlyel végeztük el a kísérleteket, két férfivel és egy nővel. Az első felvételnél (1975-ben) a férfiak 30 és 32 évesek, a nő 23 éves volt; a második felvételnél (2000-ben) a két férfi 55 és 57, a nő pedig 48 éves. A hangrögzítés mindkét esetben Sennheiser MKH 815T típusú mikrofonnal, Beag gyártmányú előerősítővel

történt Studer A80-as stúdiómagnetofonra (Agfa 525-ös hangszalagot használtunk). Mindkét beszédanyagot a CSL 4300B típusú digitális jelfeldolgozóval elemeztük 2000-ben (különböző beállításokban).

A vizsgálat tárgya egyetlen magánhangzó, az [e] volt, tekintettel arra, hogy az eredetileg rögzített 23 egy és két szótagú szóban, valamint a 7 mondatban ennek a beszédhangnak az előfordulása volt a leggyakoribb. Személyenként 35 [e] magánhangzót elemeztünk egy adott életkorban, az első és a második anyag tehát összesen 70 magánhangzót jelentett. Minden egyes esetben vizsgáltunk 7 formánst, a frekvenciahelyüket, valamennyinek a sáv szélességét és időtartamát. Felhasználtuk a formánsgörbe analízis, a teljesítményspektrum és az LPC-elemzések lehetőségét. Ily módon összesen 3150 adatot kaptunk. Ezen felül elemeztük az átlagos alaphangmagasságot, és további statisztikai vizsgálatokat végeztünk.

### **Eredmények**

Említettük, hogy a szakirodalomban több ízben megtalálható és adatolt tény az alaphangmagasság csökkenése az életkor előrehaladtával nőknél és emelkedése időskorban férfiaknál. Nincs egyértelmű válasz azonban arra vonatkozóan, hogy a változások pontosan hány éves korban következnek be, hiszen az egyén egészségi állapota, beszédhasználati sajátosságai alapvetően befolyásolják ennek a változásnak a létrejöttét (sőt, egyéb tényezők is tekintetbe veendők, például a dohányzás). Nem tudjuk, hogy vajon a fiatalabb és az idősebb felnőttek között van-e és ha igen, milyen mértékű változás az alaphangmagasság alakulásában. Saját eredményeinket a három személy esetében az 1. táblázat mutatja. Az adatok két beszélőnél az alaphangmagasság meglehetősen nagy mértékű csökkenését mutatják (az egyik férfi és a nő), ugyanakkor a másik férfi beszélő F0-értéke gyakorlatilag változatlan. Tekintettel arra, hogy hármójuk közül egyik sem dohányzik (és soha nem is dohányzott), az alaphang mélyülését, illetőleg változatlanágát egyéni sajátosságként kell értelmeznünk.

Az átlagosan 41 Hz és 44 Hz különbség egyértelműen olyan változás, amely a percepcióban is jelentkezik. Nézzük meg a hangszalagműködés frekvencia- és intenzitásállandóságát! Ezek az értékek arra



utalnak, hogy a beszélő zöngéje beszéd közben akusztikailag megnyire változékony.

1. táblázat: Az alaphang jellemzőinek statisztikai változása (F0 = alaphangmagasság, I = intenzitás)

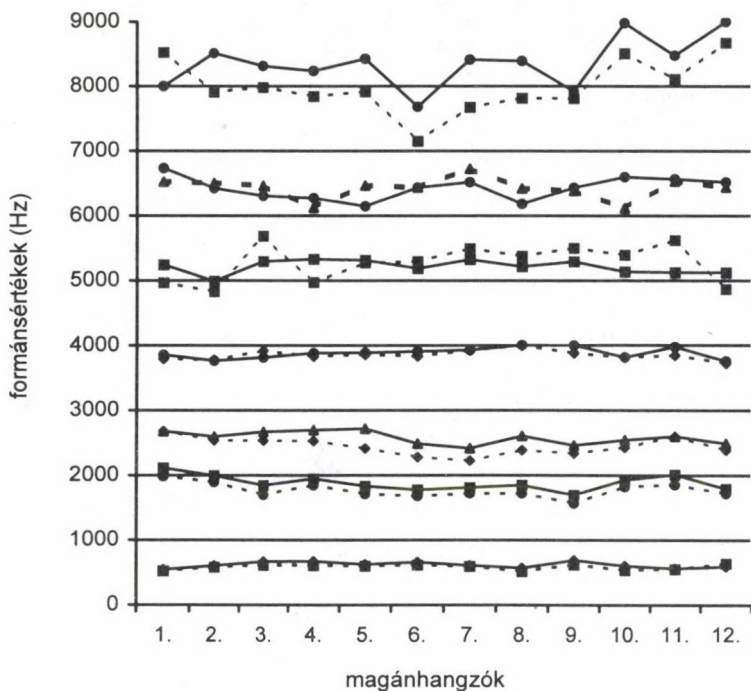
Változók	Idő	Férfi1	Férfi2	Nő
F0 (Hz)	múlt	116	150	185
	jelen	114	106	144
F0-állandóság (%)	múlt	1,13	1,28	1,41
	jelen	1,6	1,07	1,08
I-állandóság (dB)	múlt	1,26	1,11	0,83
	jelen	0,89	0,75	0,77

Az adatok szerint az idősebb korban mélyebben beszélők alaphang-magasságának frekvenciája sokkal állandóbb, mint a változatlan F0-értéken beszélőé. Feltehető, hogy valamiféle kiegyenlítődés következik be, s ha ez igaz, akkor kimondható, hogy az alaphang-magasság mélyülése fordítottan arányos a frekvenciaállandóságával. Az intenzitás állandósága nem mutat jelentős eltérést a női beszélőnél, a két férfi esetében azonos mértékű a változás. Feltételezhető, hogy itt a beszédhasználati gyakoriság kap szerepet (a női beszélő rendszeresen előad, a férfiak ritkábban).

A magánhangzók vizsgálatokor szokásosan az első három, ritkábban az első négy formáns frekvenciáját határozzák meg. Ennek az az oka, hogy az F1 és az F2 egyértelműen meghatározzák a magyar magánhangzók minőségét, szükségtelen tehát a magasabb indexszámú összetevők elemzése. (Korábban technikailag sem volt megoldható a felsőbb formánsok pontos mérése.) Tudjuk, hogy az alaphang változása nem befolyásolja az alsóbb formánsok elhelyezkedését. A magánhangzó minőségét – mint említettük – az első két formáns frekvenciahelye biztosítja, hiszen ha itt egy adott értéknél nagyobbak a változások, akkor az eredeti magánhangzó más hangminőségbe fog át-



csapni. Ennélfogva a feltételezett változások csak a felsőbb formánsok elhelyezkedésében várhatók. Elemzéseink e feltételezést egyértelműen igazolták. Az 1. ábra szemléltető jellegű, mivel csak 12 magánhangzót ábrázol egyetlen beszélő ejtésében (az egyik férfi elemzett hangjainak ún. tiszta fázisáról nyert adatokat mutatja).



1. ábra  
Elemzett magánhangzók formánsértékei  
(folyamatos vonal = 1975, szaggatott vonal = 2000)

Az ábrán jól látható, hogy a formánsok különféleképpen változtatták értéküket a 25 év elteltével. Az összes beszélő valamennyi magán-

hangzóját tekintetbe véve, a következő, százalékban kifejezett átlagos változások adódtak. Az első formánsok különbsége -0,6%, az F2-é +1,2%, az F3 +7,5%, az F4 -2,6%, az F5 -7%, az F6 +3,9% és az F7-nél -6%.

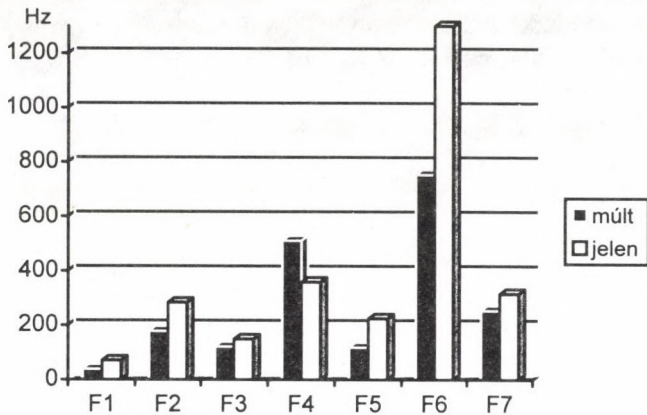
A teljes anyag vizsgálata alapján az alábbi következtetések fogalmazhatók meg.

1. Az első két formáns értékei 25 év elteltével csak kismértékű, nem szignifikáns eltérést mutatnak.
2. Hasonlóan stabil az F4 értéke.
3. Az F3, az F5, F6 és F7 frekvenciája ugyanakkor különbséget mutat az első és a második felvétel tekintetében.

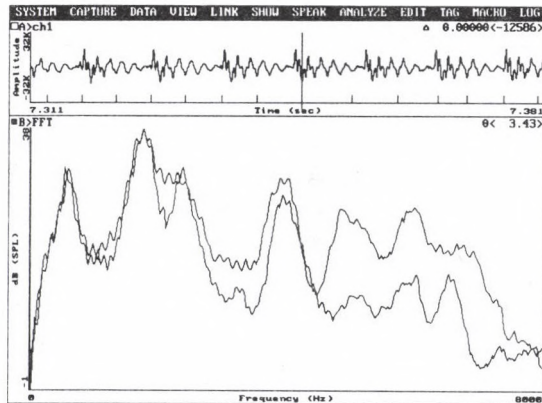
A hangszínképeken vizuálisan is jelentkeznek a fent említettek; nem csupán a tiszta fázisban mért értékek, de az ugyanazon kontextusban megjelenő magánhangzók formánshelyei egészen hasonló lefutásúak. Ez azt jelenti, hogy az első két formáns frekvenciatartományában lényegében nincs különbség, a felsőben viszont határozottan van.

A formánsok sáv szélességének elemzése ugyancsak valamennyi beszélő esetében azonos eredményre vezetett. Az egyes formánsok sáv szélességében eltéréseket találtunk a múltban és a most rögzített artikuláció következményét tekintve. Az F4 kivételével valamennyi formáns sáv szélessége megnövekedett, e növekedés mértéke azonban formánsonként változó. E változások mértéke átlagosan 34%. Egyetlen olyan formánst találtunk, amelynek a sáv szélessége csökkent, ez a negyedik formáns. Már a formánsok frekvenciájának elemzésekor is tapasztaltuk, hogy az F4 értéke a felsőbb indexszámú formánsokkéhoz képest kisebb eltérést mutat. A jelen adatok arra utalnak, hogy a nagymértékű stabilitás mellett az F4 sáv szélessége is másként változik, mint az összes többi formánsé, hiszen szűkülte valamennyi beszélő esetében. E szűkülés mértéke átlagosan 28%. A 2. ábra a formáns-sáv szélességek alakulását mutatja.

Elemeztük az [e] magánhangzó összesített FFT energiagörbéit (3. ábra). Ez a fajta elemzés is megerősítette a formánsok már elvégzett vizsgálataival kapott eredményeket.



2. ábra  
A formáns-sávszélesség változása 25 év elteltével (összesítés)

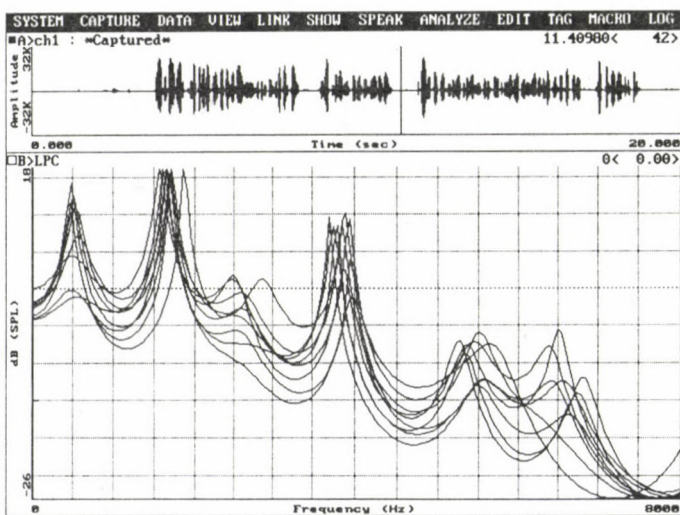


3. ábra  
Az [e] magánhangzó összesített FFT energiagörbéi az egyik férfi beszélő artikulációja alapján (a felső vonal a 25 évvel ezelőtti, az alsó a jelen ejtés alapján kapott görbét mutatja)



Az energiagörbék egybeesnek az első két formáns területén, ugyanakkor jellegzetes különbségeket mutatnak a felsőbb formánsok frekvenciatartományában. Az ábra az egyik férfi beszélő elemzett magánhangzóinak az összes előfordulásban mért energiagörbéit mutatja. További vizsgálatot igényel, hogy vajon a felsőbb formánsok között – több beszédhang elemzésbe vonása esetén – tapasztalunk-e szignifikáns, illetőleg a beszélő személyre jellemző objektíven mérhető adateltérést.

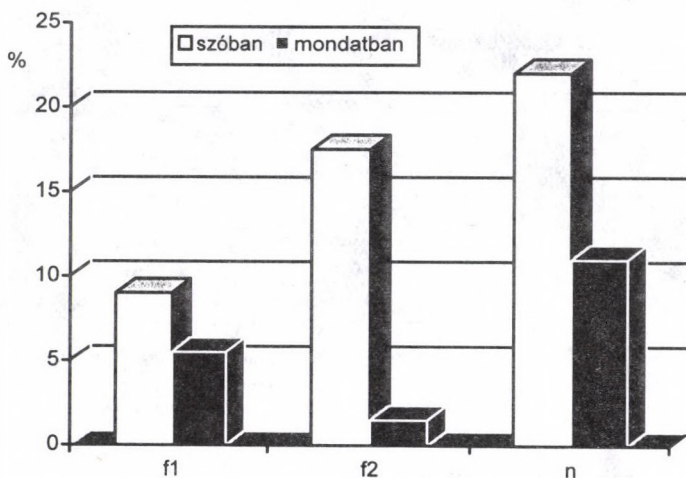
Az LPC-analízis ismételten megerősítette a korábbiakban kapott eredményeket. Az ábrán jól látható, hogy az F1-re és az F2-re kapott értékek csaknem egybeesnek; hasonló az F4-re kapott elemzési eredmény is. Az F3 és a felsőbb formánsok mért értékei azonban jellegzetesen szórnak, amit a görbék vonulatai jól szemléltetnek (4. ábra).



4. ábra

Az [e] magánhangzó összesített LPC-rezonanciagörbéi

Megvizsgáltuk, hogy a kísérleti személyek ejtésében a magánhangzók időtartama változott-e két és fél évtized alatt. Valamennyi előfordulás mért értékét összegeztük, és azt találtuk, hogy szinte minden esetben nyúlás történt (vö. 5. ábra). Az időtartamok nyúlásait százalékban fejeztük ki és ábrázoltuk, mivel így a különböző beszélők adatai összevethetőek. Jól látszik, hogy a három beszélő eltér a magánhangzók nyújtása tekintetében, de annak függvényében is, hogy izolált szavakat vagy mondatokat ejt-e. Legnagyobb mértékű az időtartam változása a női beszélő ejtésében, mind szóban, mind mondatban. Az izolált szavakban előforduló elemzett magánhangzók nála hosszabbodnak meg legjobban. A férfiak produkciójában ugyancsak nagyobb mértékű a magánhangzó nyúlása, ha mondatban fordul elő. Feltűnő, hogy az egyik férfi a mondatbeli ejtéshez képest alig nyújtja a magánhangzókat az izolált szavakban.



5. ábra

Az [e] magánhangzók időtartamának változása az egyes beszélők ejtésében (f=férfi, n=nő)

Az egyes beszélők átlagadatai – függetlenül a beszédhang tágabb kontextusától – a két férfinél igen hasonló, hozzájuk képest a női beszélőnél szignifikáns különbséget mutatnak. A kapott adatokból két következtetés vonható le. 1. Az életkor előrehaladtával az artikulációs mozgások lassúbbodnak, s ez a beszédhang időtartamának nyúlásában ölt testet. 2. Az időtartam-változások aktuális mértéke nagyobb individuális különbségeket mutat, amelyek kevésbé előre jelezhetőek.

### **Összegzés és kitekintés**

A tárgyalt kísérletsorozat különlegességét az adja, hogy nemcsak a részt vevő személyek, hanem a beszéd rögzítésének körülményei és berendezései is ugyanazok. Ezáltal az idő múlásának következményeit vizsgálva erősen le lehetett szűkíteni a változókat. Az eredmények igazolták a „paradoxonnak” tekintett jelenséget: az emberi beszéd változik is, és állandó is hosszabb idő, jelen esetben 25 év elteltével.

Hogyan oldható fel tehát a paradoxon? Mi állandó és mi változó a beszéd akusztikai szerkezetében? Elemzéseinkben erősen leszűkítettük a lehetséges változókat, amikor egyetlen magánhangzó különféle előfordulásainak akusztikai szerkezetét vizsgálatuk a 25 évvel ezelőtti és a mostani artikulációnak megfelelően. Eredményeink a következőkben foglalhatók össze:

1. Az életkor hatással van az alaphangmagasság alakulására, ennek mértéke azonban egyénenként változó. A mintegy 25-50 év és a 30-55 év közötti beszélőknél részben erőteljes csökkenést, részben változatlanságot tapasztaltunk.
2. A magánhangzók első két formánsának elhelyezkedése gyakorlatilag nem változik. Ezen formánsok stabilitása a hang minőségének biztosítója, itt tehát jelentős változást nem is vártunk.
3. A negyedik formáns állandóságot mutat mind a három beszélőnél, az értékek hasonlósága 25 év elteltével is igen nagy, gyakran egybe is esnek.
4. Az F5, F6 és F7 formánsok helye jellegzetesen megváltozik, magasabb és alacsonyabb frekvenciák irányába egyaránt történik eltolódás. Az eltolódások mértéke 3-10% körüli, azonban sem a mérték, sem az irány nem köthető a beszélő személyhez.



5. A formáns-sávszélességek két és fél évtized elteltével átlagosan 20-40%-kal növekednek. Az F4 esetében éppen ellenkezőleg, szűkülés tapasztalható, kivétel nélkül minden helyzetben és minden beszélőnél.
6. Az elemzett magánhangzók időtartama idősebb korban egyértelműen megnyúlik. Ennek az időtartam-növekedésnek a sajátosságait azonban csak nagyobb anyag vizsgálatával lehet egyértelműen leírni.

Noha a kutatás tovább folytatódik, a kapott adathalmaz mennyisége és a kialakult eredmények alapján kétséget kizáróan megfogalmazható a következő gondolat. A beszélő életkorának változása kihat a beszédprodukcijára, az artikulációs gesztusaira, s ezáltal beszédének hangzására. A változások mértéke ugyanakkor mégis lehetővé teszi a beszélő felismerését, ha az invariáns elemek dominánsan érvényesülnek. A kérdés ezek után az, hogy mit jelent az **invariáns elemek domináns érvényesülése**. Ember és ember közötti kapcsolatban ez azt jelenti, hogy az állandó ingerhatás biztosított. Ha a beszélő személlyel folyamatos a kapcsolatunk, s a beszédében bekövetkező változásokat az agyunk állandóan regisztrálja, s az adott személy beszédéről kialakított ún. neurális spektrogramot ennek megfelelően folyamatosan módosítja, akkor az invariáns elemek dominánssá válnak, s a beszélő személyt hosszú évek elteltével is felismerjük a beszéde alapján. Ha nincs mód a neurális spektrogram módosítására, mivel nélkülözzük az ingert (ti. nem találkozunk a személlyel, nem beszélgetünk vele), akkor hosszabb idő elteltével a beszéde alapján nem, illetve nehezen fogjuk őt felismerni. Az invariáns elemek ugyanis nem válhattak dominánssá az észlelésünkben.

További, rendszeres kutatásoknak kell számot adniuk az invariáns elemekről, s arról, hogy ezek milyen összefüggésben képesek az egyén beszédét oly mértékben jellemezni, hogy a beszélő személy felismerhető legyen.

### Irodalom

Balázs Boglárka: Hangképzési zavarok időskorban. In: Beszédkutatás 1993. Szerk.: Gósy Mária – Siptár Péter. MTA Nyelvtudományi Intézet. Budapest 1993, 156-166.

Gocsál Ákos: A beszéd időviszonyai különböző életkorú személyeknél. In: Beszédkutatás 2000. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 2000.

Gósy Mária: Állandóság és változás a beszédben. Magyar Nyelv XCVI, 2000/1, 1-14.

Nikléczy Péter: Beszélő személy azonosítása szűk frekvenciás szavak alapján. In: Beszédkutatás '96. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1996, 20-31.

**A kutatás a T 025965 sz. OTKA keretében készült.**

# EREDMÉNYEK A MAGYAR NYELVŰ NAGYSZÓTÁRAS ÉS KAPCSOLT-SZAVAS GÉPI BESZÉDFELISMERÉSBEN

**Szarvas Máté – Fegyő Tibor – Mihajlik Péter – Tatai Péter**  
**Budapesti Műszaki és Gazdaságtudományi Egyetem,**  
**Távközlési és Telematikai Tanszék**

## **Bevezetés**

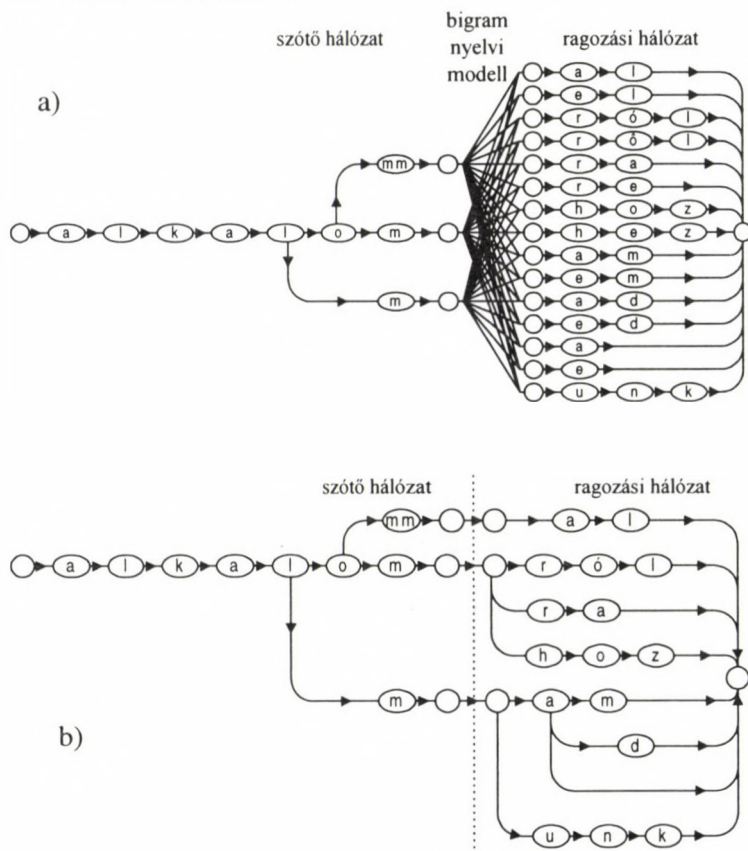
A számítógépek egyre szorosabb integrálódása a mindennapi életbe folyamatosan növeli az igényt a billentyűzetnél természetesebb adatbeviteli eszközök iránt. Mivel az ember számára a beszéd az egyik legtermészetesebb kommunikációs forma, világszerte folyamatosan növekszik az emberi beszédet mint számítógép bemenetet hasznosító technológiák kutatása. Ebben a dolgozatban röviden ismertetjük a magyar nyelvű gépi beszédfelismerés terén a BME TSP laboratóriumában elért legújabb eredményeket. A magyar nyelv esetében a legnagyobb nehézséget a szókészlet mérete okozza. Ennek kezelésére bevezetjük a sztochasztikus morfológiai elemzés módszerét, amely a helyesírás ellenőrzésben már bevált módszert terjeszti ki a nem-determinisztikus megfigyelés sorozatok esetére. A gyakorlati eredmények közül egy beszélőfüggetlen kapcsolt szavas számjegyfelismerő, valamint egy 2000 szavas városnév felismerő rendszerrel végzett kísérleteket mutatjuk be. Ennek során ismertetjük a kísérletek során használt jelreprezentációt, a használt akusztikus modellek szerkezetét, valamint az elért felismerési pontosságot.

## **Morféma alapú szókészlet ábrázolás**

Ebben a fejezetben két alternatív megoldást mutatunk be az agglutináló nyelvek szókészletének ábrázolására. Mindkét módszer morfémákat használ alapelemként, de míg az első alapvetően csak szabványos statisztikus nyelv-modellezési technikákat használ, a második lehetővé teszi az adott nyelv morfológiájáról rendelkezésre álló nyelvészeti tudás explicit beépítését is. Ennek megfelelően az



elsőként bemutatandó módszer fő erénye, hogy egyszerűen és gyorsan alkalmazható tetszőleges nyelvre, a második viszont sokkal „erősebb” modellt eredményez, és alkalmazása a beszédfelismerő rendszerekben várhatóan kevesebb felismerési hibát okoz.



1. ábra

Ragozási információ ábrázolása bigram nyelvi modellel (a), illetve morfológiai elemzővel (b).

### 1. A morféma bigram modell

A szókészlet ábrázolása morfémák használatával a legkézen-

fekvőbb módon úgy történik, hogy létrehozunk két részlistát. Egyikben a felismerési feladathoz kapcsolódó szavak szótári és alternatív tövei, a másikban pedig az összes lehetséges toldalék szerepel. A ragozott alakokat ezek után a szótó és a toldaléklista bigram nyelvi modellel történő összekötése ábrázolja. Egy ilyen reprezentációt az 1.a ábrán mutatunk be. Ebben a példában csak egyetlen szó, de annak mindhárom lehetséges tőváltozata szerepel a szótó listán. A toldalék listát 8 különböző toldalék minden lehetséges alakja alkotja. (Természetesen a toldaléklista közel sem kimerítő a példában, hiszen legalább 1000 különböző toldalék-kombináció kapcsolható ehhez a szóhoz.) Mint ismeretes, a magyar nyelvben a régi eredetű toldalékoknak több különböző alakja van, de ezek közül általában csak egy kapcsolható egy-egy adott szóhoz. Mivel a toldaléklista általános, az összes toldalék összes alakjának szerepelnie kell rajta. A megengedett kapcsolódásokat az egyes szótövek (tulajdonképpen tőváltozatok) és a toldalékok között a bigram nyelvi modell (sztochasztikus átmenet mátrix) fogja megadni.

Nyilvánvaló, hogy ez az ábrázolási módszer a szótó listán szereplő összes szó összes ragozott alakját magában foglalja. A módszer nagy előnye, hogy a toldalékok csak egyszer szerepelnek a modellben, míg a teljes szavas ábrázolási módszer esetén minden toldalék annyiszor szerepelne, ahány szót a tőlista tartalmaz. Ezenfelül a tövek is annyiszor szerepelnének, ahány lehetséges toldalék-kombináció kapcsolódhat hozzájuk. A nagymértékű memóriamegtakarítás mellett a módszer erőssége, hogy létrehozásához csak a lehetséges szótöveket és a toldalékokat kell összegyűjteni, míg a bigram átmenet mátrix becslése szabványosnak tekinthető automatikus módszerekkel történhet.

A módszer hátránya is az automatikus becslési eljárásból ered. A bigram átmenet mátrix pontos becslése a gyakorlatban ugyanis nehéz feladatnak bizonyul. Ennek az az oka, hogy véges méretű szövegadatbázis használata esetén mindig lesznek olyan ragozott szóalakok, amelyek nem szerepelnek a tanító adatbázisban, de a rendszer használata során előfordulhatnak. A hiányzó ragozott alakokhoz a becsülő eljárás nulla valószínűséget becsülne, és ezért a rendszer nem lenne képes az adott szóalak felismerésére. Az ilyen esetek szokásos

kezelése úgynevezett „simító” eljárások használatával történik. Ezek lényege, hogy a tanító adatbázisban nem szereplő kombinációkhoz is nullától eltérő (kis) valószínűséget rendelnek.

Az ilyen algoritmusok azonban nem képesek az adathiány, illetve a nyelvi szabályok miatt hiányzó ragozott alakok megkülönböztetésére. Az eredmény az, hogy a simított nyelvi modell a nyelvileg helytelen ragozott alakokhoz is pozitív valószínűséget rendel, azaz például a helyes „alkalm-*ad*” alak mellett a helytelen „alkalm-*ed*” alakhoz is.

## 2. A sztochasztikus morfológiai analízis használata

A fentiekben vázolt megoldás alternatívája lehet a helyesírás ellenőrzésben már bevált (Prószéky–Kis 1999) morfológiai elemző használata.

Mindezidáig azonban a morfológiai elemzők főként írott nyelvi rendszerekben – mint például lemmatizáció, fordítás támogatás, helyesírás ellenőrzés – kerültek hasznosításra. Ennek következtében a számítógépes morfológia kialakult módszerei arra a feltevésre épülnek, hogy az elemzendő szavak szöveges formában vannak ábrázolva.

Pontosabban, a jelenlegi módszerek feltételezik, hogy az egyes szimbólumokhoz kapcsolódó megfigyelések determinisztikusak és ismertek az elemző rendszer fejlesztése során. Például a *bot* szó mindig a *b*, *o* és *t* megfigyelések sorozataként jelenik meg. A beszéd-felismerés során azonban az alapmegfigyelések a jelfeldolgozó előrendszer által kiszámított lényegvektorok, ezek pedig a gyakorlatban minden esetben eltérőek, még akkor is, ha ugyanaz a beszélő ejti ki ugyanazt a hangot kétszer egymás után.

Mivel a hagyományos módszerek nem képesek az ilyen sztochasztikus megfigyelés sorozatok kezelésére, a morfológiai elemzőket idáig főként a felismerés utófeldolgozási lépéseként javasolták alkalmazni (Prószéky–Kis 1999), miután egy „előfelismerő” már meghatározott egy megközelítőleg pontos fonéma sorozatot.

Ez a paradigma azonban nem teszi lehetővé, hogy az „előfelismerő” hozzáférjen a nyelvi információhoz, ezért az első lépésben meghatározott fonéma sorozat minősége nagyon alacsony lehet, erősen lecsökkentve annak az esélyét, hogy a második lépés során a morfoló-



giai elemző elfogadható szóalakokat találjon.

Ebben a részben egy javaslatot mutatunk be a hagyományos – egészen pontosan az ún. két szintű (Koskenniemi 1983) – morfológiai elemzők sztochasztikus megfigyeléssorozatok kezelésére történő kiterjesztésére. A javasolt módszert *sztochasztikus morfológiának*, az elemzőt pedig *sztochasztikus morfológiai elemzőnek* fogjuk hívni. A következőkben felvázoljuk a sztochasztikus morfológiai elemzés alapelveit és megmutatjuk, hogyan használható az elemző a beszédfelismerésben.

### 3. Formalizáció

A morfológiai elemző alapfeladata, hogy a bemeneti szót a felépítő morfémák sorozatává konvertálja. Bár számos implementációs alternatíva kínálkozik, a következőkben a morfológiai elemzőt alapvetően egy véges fordítónak fogjuk tekinteni, a „két szintű morfológiában” (Koskenniemi 1983) használt modellnek megfelelően.

A véges fordító egyesével olvassa a bemeneti megfigyelés sorozatot, és minden olvasott megfigyeléshez kiír egy kimeneti szimbólumot, miközben az állapotgép következő állapotába lép. Mind a bemeneti, mind a kimeneti szimbólum lehet az üres megfigyelés is. A kimenetet általában a bemeneti szó morfémáinak kanonikus alakjai adják, esetlegesen az adott morféma nyelvtani szerepére utaló címkéssel ellátva. Ha az állapotgép az utolsó bemeneti szimbólum feldolgozása után elfogadó állapotban áll meg, akkor a bemenő sorozat elemzése sikeres volt, és a kimeneti sorozat tartalmazza az elemzés eredményét. Ha a gép nem elfogadó állapotban áll meg, akkor a sorozat elemzése sikertelen volt (nem része annak a nyelvnek, amit az automata feldolgoz). Ennek a modellnek az az erőssége, hogy hatékony módon képes ábrázolni a nyelv szókészletét.

Másrészről viszont a modell egyik korlátja az, hogy egy-egy állapot átmenet csak akkor tehető meg, ha pontosan az előírt szimbólumot olvassa a gép a bemenetről. Ez a tulajdonság megakadályozza a való életben (és a beszédfelismerésben) oly gyakori zajos bemeneti sorozatok elemzését.

A fentiekben leírt determinisztikus morfológiai elemzőt a működé-

sének alapját adó véges állapotú gép rejtett Markov-modellre történő cserélésével alakíthatjuk sztochasztikussá. Az állapot átmenetek a Markov modell esetében is függenek a bemenő szimbólumoktól, de ez a függőség nem determinisztikus, hanem sztochasztikus. Tetszőleges bemenő szimbólum esetén bármely átmenet megtehető bizonyos valószínűséggel, de az átmenet valószínűsége függ a bemenő szimbólumtól. (A modell paramétereinek speciális megválasztásával az elemző működése a determinisztikus elemzőnek megfelelővé tehető.) A sztochasztikus működés következménye, hogy az elemző bármely bemeneti sorozathoz több elemzést is elfogadhat (a választott állapot átmenetek függvényében). A gyakorlati alkalmazások szempontjából a legvalószínűbb (vagy a legvalószínűbb  $N$ ) elemzés meghatározása a feladat, ami a jól ismert Viterbi algoritmussal végezhető el.

Az ily módon létrehozott elemző a determinisztikus elemzőhöz hasonlóan hatékonyan reprezentálja az adott szókészletet. A nyelvtanilag helytelen szóalakokhoz tartozó állapotsorozatok továbbra is 0 valószínűségűek, mivel az állapotgép topológiája nem változott meg. A sztochasztikus elemző esetében azonban már nem követelmény, hogy minden szó egy előre meghatározott sorozattal legyen ábrázolva a bemeneten. Ez a tulajdonság a sztochasztikus morfológiai elemzőt ideális eszközzé teszi az agglutináló nyelvek változatos szóalakjainak beszédfelismerő rendszerekben történő ábrázolására.

Az „alkalom” szó 8 ragozott alakjának sztochasztikus morfológiai elemzővel történő ábrázolására az 1. b ábra mutat egy példát.

### **Felismerési kísérletek**

A mai napig viszonylag kevés kutatás folyt a magyar nyelvű beszélőfüggetlen beszédfelismerés területén, és kísérleti eredmények csak fonéma, illetve félszótag felismerésről kerültek publikálásra (Vicsi-Vig 1995; Kocsor-Kuba-Tóth 1999; Kocsor et al. 2000). A gyakorlatiasabb eredményeket megcélzó kísérletek egyik legnagyobb akadálya a megfelelő beszédadatbázis hiánya volt. A közelmúltban azonban két adatbázis is hozzáférhetővé vált. A kisebb méretű izolált szavas számjegyeket (Kocsor-Kuba-Tóth 1999), a nagyobb méretű pedig kapcsolt szavas számsorozatokat és folyamatosan kiejtett



olvasott mondatokat (Vicsi–Vig 1997) tartalmaz. Bár a fenti adatbázisok felhasználásával végzett kutatás még mindig csak kezdeti stádiumban van, ebben a részben ismertetjük az első kísérleti eredményeinket, amelyeket majd összehasonlítási alapként lehet használni a későbbi eredményekhez.

### 1. Jelfeldolgozás

Ebben a részben a jelfeldolgozási modul beállításait ismertetjük, amelyek az összes kísérlet során azonosak voltak. A jelfeldolgozási lépés során a mintavételezett beszédet egy vektor-sorozattá konvertáltuk, melynek minden eleme a beszéd spektrális jellemzőit írja le egy kis ablak – vagy a szokásos terminológiát használva „keret” – idejére. A beszéd felvétel során a mintavételi frekvencia eredetileg 20 kHz volt. Ezután 10 milliszekundumos időközönként egy 13 dimenziós, mel-skála szerinti kepsztrum vektort számítottunk ki, 25 milliszekundum hosszúságú Hamming ablakkal ablakozott beszéd szegmens használatával.

Ezek mellett a statikus együtthatók mellett az úgynevezett „delta,” vagy más szóval differenciális együtthatókat is használtuk. A delta együtthatókat a regressziós módszerrel számítottuk,  $\pm 2$  keretnyi adat felhasználásával (például Rabiner–Juang 1993).

Bizonyos kísérletekben a kepsztrum-átlag kompenzáció (cepstral mean compensation, CMC) nevű módszert is alkalmaztuk a csatorna és beszélő különbségek kompenzálására. A CMC-t a bemondás szintjén alkalmaztuk, azaz először parametrizáltuk a teljes bemondást, majd meghatároztuk a vektorok átlagát és kivontuk a sorozat minden vektorából. A számítást végző szoftver korlátai miatt mind a „csönd,” mind az „aktív” kereteket felhasználtuk a kepsztrum-átlag számításához.

### 2. Eredmények az OASIS adatbázis használatával

Az OASIS adatbázis (Kocsor–Kuba–Tóth 1999) egy 26 beszélős izolált szavas számadatbázis. Minden beszélő kétszer olvassa fel azt a 26 szót, amelyek kombinálásával 100 millióig előállíthatók a számok.

A tanítás során környezetfüggetlen Markov-modelleket tanítottunk



be az adatbázisban előforduló összes fonémára a FlexiVoice programcsomag (Szarvas–Fegyő–Mihajlik–Tatai 2000) használatával. A modellek balról-jobbra modellek voltak, fonémánként három állapottal. Minden állapothoz egy 8 komponensű Gaussi-keverék megfigyelési sűrűségfüggvény tartozott. A komponens eloszlások diagonális kovarianciamátrixúak voltak.

### 2.1 Fonémaosztályozási kísérletek

A fonémaosztályozási feladatban a cél az volt, hogy minden fonémát felismerjünk, (a való élethez képest) azzal a könnyítéssel, hogy használhattuk a kézzel hozzájuk rendelt fonémahatárokat. A korlátozott szókészlet miatt csak 28 különböző fonéma szerepelt az adatbázisban. Monofón fonéma modellek és kepsztrális kompenzáció használatával 7,47%-os felismerési hibaarányt értünk el.

### 2.2 Izolált szavas számjegy felismerés

A számjegy felismerési kísérletek során egyáltalán nem használtuk a bemondásokhoz rendelt címkéket. A kezdet-vég detektálást a szó modellek elé és után kapcsolt opcionális csönd modell használatával végeztük. A szó modelleket az előző részben ismertetett kontextus független fonéma modellekből építettük. Az elért felismerési hibaarány 0,37% volt. A kísérletek részletei a (Szarvas–Fegyő–Mihajlik–Tatai 2000) publikációban olvashatóak.

## 3. Eredmények a BABEL adatbázis használatával

A BABEL adatbázis a mai napig a legnagyobb jó minőségű, magyar nyelvű beszédatadtbázis, ami kutatási célokra rendelkezésre áll. Az adatbázis három különböző részből áll: izolált és kapcsolt szavas számjegy bemondásokból, CVC (mássalhangzó-magánhangzó-mássalhangzó) szótagokból, valamint folyamatos olvasott beszédből (Vicsi–Vig 1997). A folytonos részben a bemondások némelyike suttogó hangú, de ezeket nem használtuk fel a jelenlegi kísérletekben.

Mind az olvasott mondatokat, mind a számjegy sorozatokat oly módon tervezték, hogy jól lefedjék a magyar nyelvben előforduló hang-kombinációkat. Az adatbázis egy kis része fonémákra van szegmentálva és meg van címkézve. A rendelkezésünkre álló mintá-

ban összesen 5 beszélő hangja tartalmazott fonéma szintű címkéket, összesen 400 másodpercnyi beszédben. (A beszédfelismerő fejlesztés során legalább több órányi fonéma szinten címkézett beszédet szokás használni.)

Az adatbázisban összesen 30 beszélő (15 férfi és 15 nő) hangja és körülbelül 2000 mondat és 14 000 kapcsolt szavas számjegysorozat szerepel. Azonban az adatbázis fejlesztői csak az adatok felét (15 beszélő hangját) bocsátották a rendelkezésünkre, ami csökkenti a beszült fonéma modellek robusztusságát. Az összes kísérletben 13 beszélő hangját használtuk tanításra, a tesztelő halmazt pedig a fennmaradó felvételek képezték.

### 3.1 A modellek betanítása

Mivel az adatbázisnak csak egy kis része tartalmaz fonéma szintű címkéket, a tanítást két lépésben hajtottuk végre. Az első lépésben kezdeti modelleket tanítottunk be a rendelkezésre álló fonéma címkék felhasználásával, majd ezeknek a modelleknek és a FlexiScribe (Szarvas–Fegyő–Mihajlik–Tatai 2000) program felhasználásával automatikusan felcímkéztük a teljes tanító halmazt. A második lépésben az ily módon gyártott címkékkel és a teljes tanítóhalmazzal folytattuk a modellek tanítását. A modellek balról-jobbra modellek voltak, fonémánként 3 állapottal. A megfigyelési valószínűségeket minden állapotban egy 10 komponensű diagonális kovariancia mátrixú Gaussi-keverék eloszlással modelleztük. A magyar nyelv 64 fonémája közül csak a magánhangzókat és a rövid mássalhangzókat használtuk (39 fonémát összesen), mivel a rövid és hosszú mássalhangzók főként csak az időtartamban különböznek, a használt Markov-modellek viszont nem modellezik az időtartamot.

### 3.2 Fonémaosztályozási kísérletek

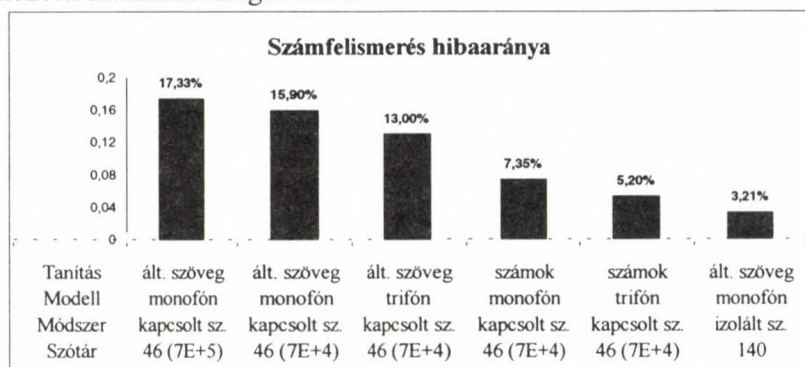
A fonémaosztályozási kísérletek során a tanítás első lépésében gyártott automatikus címkéket használtuk. A fonémaosztályozás hibaránya 46,93% volt, azaz körülbelül ötszöröse az OASIS adatbázison elért hibarányoknak. A nagy eltérés legvalószínűbb magyarázata az, hogy a BABEL adatbázis fonetikailag gazdag, míg az OASIS adatbázisban csak kevesebb mint 30 szó szerepel, erősen csökkentve az



előforduló fonetikai kontextusok számát. A kísérletekben használt monofón modellek könnyebben tudták modellezni az OASIS adatbázisban előforduló néhány allofónt, mint a BABEL adatbázisban előforduló nagy számú allofónt. Egy másik ok, hogy a BABEL feladatban 39, míg az OASIS feladatban csak 28 fonéma szerepelt.

### 3.3 Kapcsolt szavas számfelismerés

A BABEL adatbázis kapcsolt szavas számokat tartalmazó részén is végeztünk kezdeti felismerési kísérleteket. Az adatbázisnak ez a része 1 és 10 000 közötti számokat tartalmaz, összesen 8939 bemondást. Ezekben a kísérletekben a szám modelleket kontextusfüggő illetve kontextus független fonéma modellekből építettük fel. A rendszer-szótár 46 számalakot tartalmazott (például *egy, egyes, ..., tíz, tizen, tizes, ... ezres*), és a nyelvtani hálózat képes volt az összes 1 és 10 000 közötti számalak elfogadására.



2. ábra

Kapcsolt szavas számsorozat felismerési hibaaránya különböző beállítások mellett

Az első kísérletnél a nyelvtan által elfogadott alakok száma közel 700 000 volt, amit a többi kísérletnél a nyelvtan javításával 70 000-re csökkentettünk, mivel az előzőben szerepeltek nem használt alakok is. A kísérleteket elvégeztük általános olvasott szöveggel, és olvasott számokkal betanított beszédhang modellekkel is. A különböző modelltípus és tanítóadat kombinációk melletti felismerési hibaará-



nyokat a 2. ábra ábrázolja. Összehasonlításként az utolsó kísérletben izolált szavas felismerési tesztet is elvégeztünk, ahol a szótárban csak a valóban előforduló számok szerepeltek.

Az ábrából kitűnik, hogy nagyon fontos a feladatnak megfelelő tanítóadatok, illetve a kontextusfüggő fonéma modellek használata. Az ábrán szereplő számok a számsorozat hibák arányát jelzik, de a legtöbb hiba esetében a többjegyű számnak csak egyetlen számjegye volt félreismerve.

### 3.4 Városnév felismerés

Ez a kísérlet egy nagyszótáros felismerési probléma volt általános irodai környezetben. A feladat a 2000 legnagyobb magyar város nevének felismerése volt. A felismeréshez a kiejtési szótárt a (Mihajlik–Tatai 2000) hivatkozásban ismertetett algoritmussal készítettük automatikusan. A szó modelleket ebben az esetben is a BABEL fonémaosztályozási feladat fonéma modelljeivel építettük. A tanítóhalmaztól eltérően a tesztelőhalmazt egy zajos számítógép-laboratóriumban vettük fel, miközben a számítógépek működtek, és a bent lévő emberek végezték a napi munkájukat. A használt mikrofon is eltérő volt a tanító adatbázis felvételéhez használttól. Az akusztikus modellek nem lettek az eltérő körülményekhez adaptálva. A tesztelőhalmazban két férfi hang volt és összesen 2550 bemondás, szünetekkel az egyes szavak között. Az elért hibaarány 14,58% volt. A hibaarány csökkentésére irányuló munka jelenleg is folyik az irodai környezetnek megfelelő tanító adatbázis és kontextusfüggő fonéma modellek fejlesztésével.

### Összefoglalás

A dolgozatban ismertettük a BME TSP laboratóriumban a gépi beszéd felismerés területén elért legújabb eredményeket. Két morféma alapú szókészlet ábrázolási módszert javasoltunk és bemutattuk mindkettő előnyeit és hátrányait. Egy teljesen automatikus kiejtésmodellező rendszer egy külön dolgozatban kerül bemutatásra. Ezeket a módszereket a FlexiVoice nevű beszéd felismerő programrendszerben valósítottuk meg. Végül bemutattuk a programrendszer különböző beszélőfüggetlen felismerési feladatokon végzett kísérleteinek eredményeit. Az eredmények jónak tekinthetők az izolált szavas számjegy-

felismerési feladaton, de a nagyszótáras felismerés esetében további javítás szükséges a gyakorlati alkalmazhatósághoz. A jelenlegi kutatásaink nagy pontosságú kontextusfüggő fonéma modellek kifejlesztésére és a magyar morfológia sztochasztikus reprezentációjának kialakítására irányulnak.

### **Köszönetnyilvánítás**

A dolgozatban ismertetett kutatások a Magyar Távközlési Vállalat, illetve a japán Nippon Telephone and Telegraph Corporation támogatásával folytak. A támogatásért köszönetünket fejezzük ki. Köszönjük továbbá, hogy a BME Akusztikai Kutatólaboratóriuma rendelkezésünkre bocsátotta a BABEL beszédadatbázis egy részét, az MTA Mesterséges Intelligencia Kutatócsoportja pedig az OASIS adatbázist.

### **Irodalom**

Kocsor, A. – Kuba A. – Tóth, L.: An overview of the OASIS speech recognition project. ICAI'99. 1999. Megjelenés alatt.

Kocsor, A. – Tóth, L. – Kuba, L. – Kovács, K. – Jelasity, M. – Gyimóthy, T. – Csirik, J.: A comparative study of several feature transformation and learning methods for phoneme classification. International Journal of Speech Technology. Megjelenés várhatóan 2000 decemberében.

Koskenniemi, K.: Two-level morphology. University of Helsinki. Helsinki 1983.

Mihajlik Péter – Tatai Péter: Automatikus fonetikus átírás magyar nyelvű beszéd felismeréshez. Beszédkutatás 2000. Megjelenés alatt.

Prószték Gábor – Kis Balázs: Számítógéppel emberi nyelven. Szak kiadó. Bicske 1999.

Rabiner, L. R. – Juang, B.-H.: Fundamentals of Speech Recognition. PTR Prentice Hall. Englewood Cliffs 1993.

Szarvas, M. – Fegyő, T. – Mihajlik, P. – Tatai, P.: Automatic recognition of hungarian: Theory and practice. International Journal of Speech Technology. Megjelenés várhatóan 2000 decemberében.

Vicsi, K. – Vig, A.: Text independent neural network/rule based hybrid, continuous speech recognition. EUROSPEECH '95. Madrid 1995. 2201-04.

Vicsi, K. – Vig, A.: Babel—a multi-lingual database. Technical report. <http://www.ttt.bme.hu/speech/database.htm>. „György Békésy” Acoustics Research Laboratory of the Budapest University of Technology and Economics. Budapest 1997.



# STATISZTIKAI SZÖVEGELEMZÉS AUTOMATIKUS FELOLVASÁSHOZ

Németh Géza – Zainkó Csaba  
Budapesti Műszaki és Gazdaságtudományi Egyetem,  
Távokzlési és Telematikai Tanszék

## Bevezetés

A statisztikai szövegelemzés segít megismerni a vizsgált írott nyelv olyan tulajdonságait, amelyek a nyelvtani szabályokból korlátozottan vagy egyáltalán nem vezethetők le. Ilyen irányú kutatásokat már végeztek a magyar nyelvre, de általában más célokkal, például címszavak előállítására céljából (Pajzs–Váradi 1997; Pajzs–Kiss 2000). Mi elsősorban szószintű vizsgálatokat végeztünk, mivel az ékezetesítés problémájának megoldásához a magyar nyelv ilyen jellegű ismeretére volt szükségünk.

Az elemzés során gyakorisági listákat készítettünk, amelyek felhasználhatóak a mesterséges beszédfelismerésben, a mesterséges beszédkeltésben vagy egyéb, a nyelvvel összefüggő eljárásokban. A statisztika a beszédfeldolgozásnál használt különböző szótárak készítésekor nyújt segítséget, vagy a beszéd szintetizátorok optimalizálásakor, de több nyelv statisztikájának ismeretében a nyelvdetekció is könnyen megvalósítható.

A magyarban, mivel ragozó nyelv, rengeteg szóalakot lehet alkotni. Egyes vélemények szerint ez eléri a több milliárdot is, ami azt jelenti, hogy a gépi beszédfelismerésben nem lehetséges minden szóalakot egyenként betanítani és felismertetni. A kutatásban azt is érdemes megvizsgálni, hogy az elméleti számon túlmenően valóságos környezetben hány különböző ragozott vagy ragozatlan szóalak fordul elő. A jelen tanulmányban a szótövek meghatározásával nem foglalkoztunk.

## Korpuszok gyűjtése

A szövegtörzsek mérete meghatározza, hogy a belőlük kapott statisztikák mennyire általános érvényűek, illetve mennyire csak az



adott korpusz sajátosságait mutatják. Ahhoz, hogy az írott magyar nyelvre jellemző statisztikákat kapjunk, nem elég kis méretű szövegeket vizsgálunk, hanem több millió szavas korpuszok szükségesek. A korpuszok vizsgálatánál figyelembe kell venni azt is, hogy milyen tematikájú szövegeket tartalmaznak. Szűk tematika esetén a korpuszból csak az adott területre jellemző adatokat nyerhetünk, de nem tudunk következtetni a nyelv összes sajátosságára. Speciális felhasználás esetén azonban előállhat az az eset is, hogy célirányosan a nyelv egy adott területének tulajdonságait vizsgáljuk, és ilyenkor arra kell nagy figyelmet fordítani a szövegek gyűjtésekor, hogy a témához oda nem tartozó szövegek ne kerüljenek bele ebbe a speciális korpuszba.

A statisztikai feldolgozáshoz csak olyan szövegek használhatóak fel, amelyek elektronikus formában léteznek. A szöveggörpuszok gyűjtésekor elsősorban az interneten elérhető szövegeket használtuk fel, mert ezek széles tematikájú dokumentumok, és a vizsgálat szempontjából megfelelően nagy mennyiségben elérhetőek.

### **A statisztikai elemzés minőségét befolyásoló tényezők**

Írott szövegben is gyakran szerepelnek elírások, amelyek rontják a statisztika minőségét. Ez azt eredményezi, hogy olyan szóalakokat is figyelembe veszünk, amelyek a magyar nyelvben nem is léteznek, így a különböző szóalakok számát az elírások előfordulásának gyakoriságától függően felülbecsüljük.

A helyesírási hibák az elírásokhoz hasonlóak, de létezik egy olyan kategóriájuk is, amelyik nem a szöveg írójának hibájából keletkezik. Még mostanában írott szövegek esetében is előfordul, hogy az *i* és *í* betűt nem használja a gépelő, mivel a billentyűzetén nem található meg. Hasonló probléma az is, hogy egyes számítógépes programok rosszul kezelhetik az ékezetes betűket, így elveszhetnek ékezetek a betűkről. Ilyen okokból leggyakrabban az *ő* és az *ű* betűk módosulnak.

Problémát okoznak továbbá azok az idegen nyelvű szövegrészletek is, amelyek a korpuszba bekerülnek. Azokat a szövegrészleteket, amelyek összefüggően idegen nyelvűek, mondat szintű nyelvdetekcióval ki lehet szűrni, de azokat a szavakat, amelyek a magyar szövegbe beágyazva csak elvétve fordulnak elő, nem lehet

biztosan megtalálni. A magyar szövegekben ilyen szavak leggyakrabban a külföldi tulajdonnevek. Ezek rendszerint a hírekben fordulnak elő, különösen a sporthírekben. A túl sok idegen szó miatt, a korpusz elveszítheti azokat a tulajdonságokat, amelyek csak az adott nyelvre lennének jellemzőek.

A szövegek tördelése során tördelő-szerkesztők, szövegírók elválasztásokat alkalmaznak a kellemesebb vizuális hatás, illetve a jobb olvashatóság elérése érdekében. Ezek az elválasztójelek azonban a szövegek újratördelése során is bent maradhatnak a mondatokban, de már nem sorvégi pozícióban. Ez azt eredményezi, hogy a szó formailag kötőjeles szó lesz, amely újabb téves szóalakokhoz vezet.

A statisztika eredményeit az is befolyásolja, hogyan határozzuk meg, hogy a szövegekben mit tekintünk szónak, és mik azok a karaktersorozatok, amelyeket nem veszünk figyelembe. A jelen vizsgálatnál azokat a karaktersorozatok tekintettük szavaknak, amelyek csak betűket tartalmaztak. A német eredetű szavak egy részét kizártuk azzal a megkötéssel, hogy a szavakban nem szerepelhetnek  $\beta$  és  $\alpha$  betűk. Ezen szigorú megkötések miatt elveszthetünk néhány értelmes szóalapot, azonban a nyelvben értelmetlen alakokat kiszűrjük.

### Korpuszok tulajdonságai

- **nytud:** A MTA Nyelvtudományi Intézetének Magyar Nemzeti Szövegtára. Itt a kiindulási állapot nem a szöveg volt, mivel már a gyakorisági lista állt rendelkezésünkre. Ilyenkor azonban nincs lehetőség arra, hogy szövegkörnyezetben vizsgáljuk a szavakat. Ezzel az adatbázissal elsősorban a szógyakorisággal összefüggő sajátosságokat vizsgáltuk. *A korpusz összesen (ö): 20 805 975 szót tartalmazott, és a különböző szóalakok száma (k): 691 159 volt.*
- **mek:** A Magyar Elektronikus Könyvtár gyűjteményeiből több mint 2000 magyar nyelvű dokumentumot használtunk fel. A kiválasztott szövegek tematikája nagyon vegyes, irodalmi műveket, tudományos cikkeket, jogi és köznapis írásokat is találunk közöttük. Összehasonlítva a *nytud* korpuszsal, észrevehető, hogy a majd harmadnyi méret ellenére a különböző szóalakok száma nagy, amely a széles tematikával magyarázható.  
*ö: 6 799 701 szó, k: 522 432 szó*



- **mn:** A Magyar Nemzet 10 ezer, az interneten is megtalálható cikkét tartalmazza, amelyek 2000. április és október között jelentek meg. Ez a korpusz már nem általános tematikájú, főként sajtónyelvezetű írásokat tartalmaz. Fontos azonban, hogy a korpuszban szereplő dokumentumok frissek, a jelenkor nyelvezetét tükrözik. Az eredmények kiértékelésekor figyelmet kell arra fordítani, hogy a vizsgált időszakba belesik a 2000. évi Sydney-ben megrendezett olimpia is. Az arról szóló beszámolók sok idegen nevet, illetve a köznyelvben ritkábban használt, sporttal kapcsolatos szót is tartalmaznak.  
ö: 4 373 412 szó, k: 345 657 szó.
- **hvg:** A Heti világgazdaság 4000 cikke az elmúlt 3,5 évben jelent meg. A heti megjelenés miatt a cikkek nagyobb terjedelműek voltak.  
ö: 4 091 732 szó, k: 311 578 szó.
- **mh:** A Magyar Hírlap megközelítőleg 8 ezer cikkét szintén az internetről gyűjtöttük. A cikkek a 2000. év első 4 hónapjából valók.  
ö: 2 054 777 szó, k: 196 965 szó.

### Egyesített korpuszok

Az előzőekben tárgyalt korpuszok függetlenek egymástól, ezért lehetőség volt arra, hogy egyesítsük tartalmukat. A 1. táblázatban láthatók a vizsgált korpuszok adatai. Az egyesítés során az összes szó

1. táblázat: Egyesített korpuszok méretei

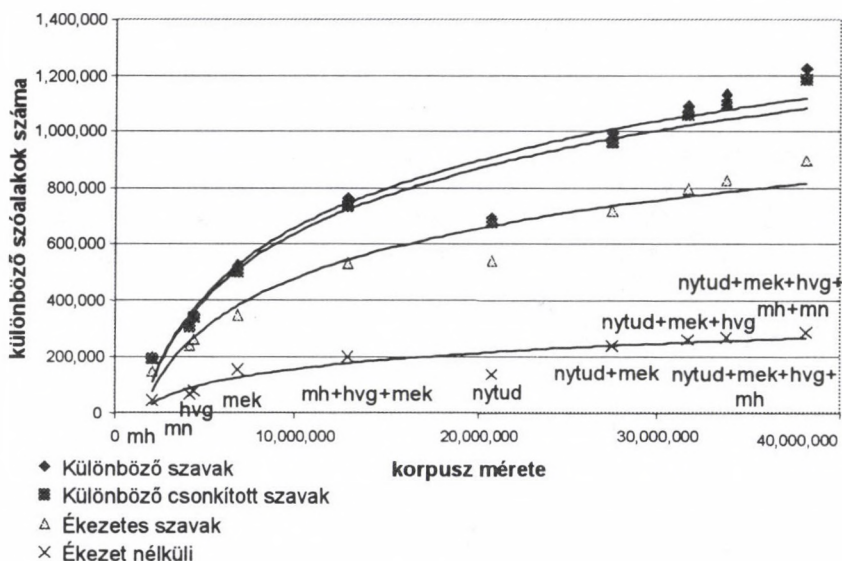
rövidítés	szavak száma	különböző szavak száma
mh+hvg+mek	12 946 210	761 255
nytud+mek	27 605 673	991 546
nytud+mek+hvg	31 697 487	1 090 916
nytud+mek+hvg+mh	33 752 182	1 129 811
nytud+mek+hvg+mh+mn	38 125 594	1 225 101



száma összedődött, de a különböző szavak száma már nem, és így a különböző szavak számának növekedéséből meg lehetett állapítani, hogy a két vagy több korpusznak mekkora volt az a részhalmaza, amely mindegyik korpuszban megtalálható. Például a *nytud+mek* egyesített korpuszhoz hozzáadtuk még a *hvg* korpuszt, ami 306 ezer különböző szóalakat tartalmazott. Az így keletkezett korpusznál csak 100 ezerrel nőtt a különböző alakok száma, tehát mintegy 200 ezer alak mindkét korpuszban megtalálható. A legnagyobb korpusz, amit összeállítottunk, 38 millió szóalakat tartalmaz, ezekből több mint 1,2 millió szó különböző.

### Statistikák

Az 1. ábrán láthatók az önálló és az egyesített korpuszok legfontosabb tulajdonságai. A függőleges tengely a különböző szóalalakok számát, amíg a vízszintes tengely a különböző korpuszok



1. ábra

Önálló és egyesített korpuszok összehasonlítása

méretét jelzi. A korpuszok nevei rövidítve az ábra alján találhatóak, a leírásnak, illetve a fenti táblázatnak megfelelően.

A legfelső pontsor a korpuszokban előforduló különböző szóalakok számát jelöli. A szorosan alatta található négyzetek azt mutatják, hogy hány különböző szóalakot különböztethetünk meg, amennyiben az ékezetes szavakat úgy csonkoljuk, hogy az ékezetes betűk helyére az ékezet nélküli alakot írjuk. Például az *ágyat* szó csonkolt alakja az *agyat* szó. Végül a két alsó sorozat azt mutatja, hogy hány különböző ékezetes, illetve ékezet nélküli szóalak található a korpuszokban. Minden egyes korpuszhoz a négy egymás alatt elhelyezkedő pont tartozik, és a korpuszok rövidítései a pontok alatt szerepelnek. A 4 laposodó görbe a pontsorozatok logaritmikus közelítései, amelyek a változások trendjét jelenítik meg.

A legfelső görbe tehát a korpuszokban található különböző szavak számát adja meg a korpusz méretének függvényében. Látható, hogy kis korpuszméret esetén még meredeken emelkedik a különböző szavak száma, de ahogy egyre több szöveget vizsgálunk, egyre gyakrabban fordulnak elő azonos szavak, így a különböző szavak számának növekedése lelassul. A görbéhez tartozó pontsorozatban egy törést találhatunk, melyet a *nytud* okoz. A korpusz több mint 20 milliós szószámához képest a különböző szavak száma csak 700 ezer körül van, pedig a közelítő görbe már 900 ezret indokolna. A kapott érték valószínűleg azért kisebb a vártnál, mert az ábrán szereplő többi korpusz széles tematikából kisebb mennyiséget tartalmaz. A *nytud*-ban azonban szűkebb tematikából nagyobb mennyiségű szöveg található. Ez a kisebb tematika viszont kisebb változatosságot biztosít, ez okozza a pontsorozat „beesését”.

A közelítő görbe egyenletéből megállapítható, hogy az 1% változási sebesség – ami azt jelenti, hogy átlagosan minden egyes új különböző szóhoz a korpusz méretét 100 szóval kell megnövelnünk – 31 millió szónál található. Az 1%-os határ viszont már 310 millió szónál van, ami már nemzetközi összehasonlításban is rendkívül nagy korpuszt jelent és elérése rendkívül munkaigényes feladat.

Fentről a második görbe szorosán simul a különböző szavak görbéjéhez. Ez azt jelenti, hogy a korpusz gyakorlatilag olyan szavakkal bővül, amelyeknek a „csonkolt” (ékezet nélküli) alakja már



eddig is szerepelt a korpuszban. Ennek az ékezetesítésben van fontos szerepe, amelyet a következő fejezetben részletesebben vizsgálunk.

Az alsó két görbéből az ékezet nélküli szavak számát figyelve látható, hogy a görbe 30 milliós korpusznál szinte már vízszintes, a 400 ezres határt elméletileg 200 millió szónál érné el, ha ezt a tendenciát tartaná.

### Ékezetesítés

A szavak előfordulásának gyakorisága felhasználható ékezet nélkül írott szövegek ékezetesítésére. Tökéletes visszaállításra nincs lehetőség, de a statisztikából kinyert adatok segítségével az esetek nagy részében jól becsülhető az eredeti alak. Ez a fajta ékezetesítés elsősorban olyan helyen használható, ahol a sebesség lényeges, nincs idő és erőforrás nyelvtani vagy egyéb elemzésekre.

Első lépésben az ékezetes alakról leszedjük az ékezeteket, így megkapjuk azt az alakot, amellyel az ékezetesítés során találkozni fogunk.

A statisztika készítésekor minden csonkolt alakhoz feljegyezzük, hogy milyen ékezetes alakok tartoznak hozzá, és azok milyen gyakran fordulnak elő. Azokban az esetekben fordulhat elő pontatlanság, ahol egy ékezet nélküli alakhoz több ékezetes alak tartozik, vagy csak egy ékezetes alak található, de az ékezet nélküli alak

2. táblázat: Azonos csonkolt alakkal rendelkező szavak száma

2	29076	2,7561%
3	1609	0,1525%
4	285	0,0270%
5	19	0,0018%
6	1	0,0001%
7	1	0,0001%
8	0	0,0000%
9	1	0,0001%

is értelmes. Ilyen például a *meg* és *még* szó, ahol mindkét alak értelmes és a 2 alakú szavakhoz tartoznak. A 2. táblázatban az szerepel, hogy hány olyan ékezet nélküli alak van a korpuszban, amelyek ebben az értelemben „kétesek”. Az abszolút előfordulás (középső oszlop) mellett, a szélső oszlop azt mutatja, hogy az összesen 1 054 983 különböző csonkolt szóalakot tartalmazó korpusz hány százalékát teszi ki a 2, 3 stb. kétes alakot tartalmazó szó. A következő példán egy olyan kétes szó látható, amelynek 4 alakja van és összesen 2062-szer fordultak elő a teljes korpuszban.



Példa:	arat	árát	árat	arát
	5,19%	52,67%	41,10%	0,05%

A táblázat utolsó sorában látható, hogy egy darab 9 alakos szó is előfordult a korpuszban. Ez a nyelvben azonban nem létezik, csak a már említett elírások, karakterkészlet-hibák miatt keletkezett a valóságosnál több alakja.

Ékezetesítéskor, az ékezet nélküli szöveget szavakra bontjuk, majd az ékezet nélküli alakhoz tartozó leggyakoribbat írjuk az ékezetesítendő szó helyére. Ha például azt találjuk a szövegben, hogy *arat*, akkor az ékezetesítő ezt a szót az *árát* alakkal helyettesíti, ami sok szöveget vizsgálva az esetek 52,67%-ban szolgáltat jó megoldást.

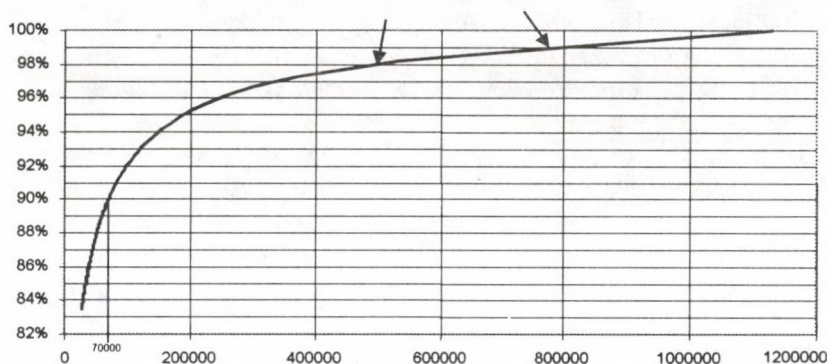
A 2. táblázatból az is látható, hogy a szavak 3%-ánál elégtelen ez a módszer, csak ilyen arányban fordulnak elő olyan szavak, amelyeknek több lehetséges alakja is van. A helyzet azonban rosszabb ennél, mivel a 3%-ban előfordulnak gyakori szavak is, tehát ez a megoldás nem nyújt ebben a formában 97%-os biztonságot, hanem néhány százalékkal kevesebbet.

Ezekben az esetekben – mint például a *meg-még* pár, ami a leggyakoribb kétes pár – további elemzésekre van szükség, amelyek lehetnek statisztikai vagy egyéb, például nyelvtani elemzések.

### A statisztikák minősége

A kapott eredmények minőségére az egyes adatok konvergenciájából következtethetünk. Megvizsgáltuk, hogy a leggyakoribb 20 szó a korpusz mekkora hányadát fedi le. A 22,8%-os arány hasonlít a más forrásból meghatározott 24,5%-hoz (Pajzs-Váradi 1997). A 20 leggyakoribb szóból 18 megegyezett egy korábbi kutatásunk során előállított listával (Németh et al. 1999). A 2. ábrán látható, hogy egy adott korpuszban előforduló leggyakoribb szavak mekkora mértékben fedik le az eredeti korpuszt. Az ábrán függőleges vonallal bejelölt 70 000 leggyakoribb szó már a korpusz összes szavának 90%-át lefedi. Ha tehát a korpusz első 70 000 leggyakoribb szavát használjuk fel valamilyen eljárásban (például kötött szótáras beszédészintézisben, beszéd felismerésnél vagy ékezetesítés során),

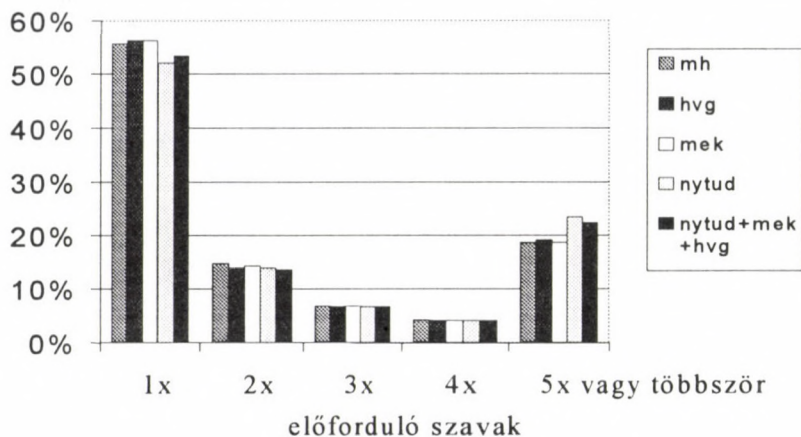
akkor a korpuszban lévő szövegben átlagosan csak minden tizedik esetén fordul elő lefedettségi hiba.



A görbe laposodásából az is kiolvasható viszont, hogy a lefedettség növelésére egyre több szót kell felhasználnunk. Míg az előző példánál 70 ezer szóval 90%-ra sikerült jutni, addig a 98%-ról a 99%-ra kerüléshez – az ábrán a két nyíl jelöli az értékeket – hozzávetőlegesen 300 ezer szóval többet kell felhasználnunk. Tehát egy viszonylag nagy lefedettséghez kevés szó is elegendő, de a nagyon pontos lefedéshez sokszorta több különböző szóra van szükségünk.

A korpusz mérete elegendő, ha a leggyakoribb szavakkal összefüggő eljárásokat használunk, mivel a megtalált kevés szó is jó lefedettséget biztosít. Azonban, ha arra keressük a választ, hogy mennyire vagyunk attól a gyakorlati korláttól, amit az összes elképzelhető használt szóalakra becsülünk, azt kell megállapítani, hogy korpuszaink mérete még rendkívül kicsi. Vannak becslések, amelyek 500 millió és 1 milliárd közé teszik a nyelvtanilag lehetséges szóalakok számát (Prószyński 1993).

A 3. ábra azt szemlélteti, hogy a különböző korpuszokban előforduló szavak hány százaléka szerepel csak egyszer, kétszer,



3. ábra  
A szavak gyakoriságának eloszlása

háromszor, négyszer, illetve ötször vagy annál többször. Ebből a diagramból leolvasható, hogy a korpuszokban szereplő szavak száma még nem elég nagy, mivel a különböző szavak átlagosan 55%-ban csak egyszer fordulnak elő. A korpuszra viszont csak akkor mondhatjuk, hogy közelít a gyakorlati határhoz, ha a grafikon jellege megfordul és a csak egyszer előforduló szavak száma már nem jelentős a korpusz méretéhez képest.

Az eloszlás majdnem független a korpusz méretétől, mert az alig 2 milliós *mh* korpusznak is lényegében ugyanaz az eloszlása, mint a 31 milliós *nytud+mek+hvg* korpusznak. Ebből arra következtetünk, hogy a megfelelő korpusz mérete több százmilliós lehet.

### Összehasonlítás más nyelvekkel

Azért, hogy a kapott eredményeinket össze lehessen hasonlítani más nyelvekkel, elvégeztük a magyar nyelvre kidolgozott vizsgálatainkat két, régi és modern szövegeket egyaránt tartalmazó angol nyelvű korpuszra is (Morgan). Brit és amerikai angolra azt az



eredményt kaptuk, hogy mindkét tízmilliós korpuszban közelítőleg 80 ezer különböző szó szerepel, a közös szóalakok száma 47 ezer.

Az angol és a magyar nyelv közti különbséget jól jelzi, hogy a mi eredményeink szerint egy tízmillió szavas magyar szövegben már 700 ezer különböző alak szerepel az angol 80 ezerhez képest.

A British National Corpus (BNC) mérete 100 millió szó, ami abszolút értékben is jóval nagyobb, mint a saját korpuszaink. Ha még azt is figyelembe vesszük, hogy az angol nyelvben kb. tízszer kevesebb különböző szó van, akkor a magyar nyelvű korpuszoknak 800 millió – egymilliárd szavasnak kell lennie. Ebben az esetben lennének a gyakoriságok a BNC-hez hasonló pontosságúak.

### **Köszönetnyilvánítás**

Szeretnénk megköszönni az MTA Nyelvtudományi Intézetének és Váradi Tamásnak, hogy a Magyar Nemzeti Szövegtár 1999-es változatának gyakorisági listáját a rendelkezésünkre bocsátotta.

### **Irodalom**

British National Corpus (BNC). [http://www.elda.fr/cata/text\\_det.html](http://www.elda.fr/cata/text_det.html)

HVG Online. <http://www.hvg.hu>

Magyar Elektronikus Könyvtár. <http://www.mek.iif.hu>

Magyar Hírlap Online. <http://www.magyarhirlap.hu>

Magyar Nemzet Online. <http://www.magyarnemzet.com>

Morgan, E. L.: Alex Catalogue of Electronic Texts.

<http://www.infomotions.com/alex/downloads/>

MTA Nyelvtudományi Intézet, Magyar Nemzeti Szövegtár.

<http://bobac.nyttud.hu/del>

Németh, G. – Zainkó, Cs. – Olaszy, G. – Prószycki, G.: Problems of creating a flexible e-mail reader for hungarian. In: EUROSPEECH '99 ESCA Budapest 1999, 939-942.

Pajzs Júlia – Kiss Gabriella: A nagyszótári korpusz címszavainak előállításáról. 2000, <http://nws.iif.hu/NwScd/docs/eloadas/80/index.htm>

Pajzs Júlia – Váradi Tamás: A magyar irodalmi és köznyelvi nagyszótárának korpusza a HUNGARNET közösség számára. 1997,

<http://www.iif.hu/rendezvenyek/networkshop/97/tartalom/NWS/3/22/index.htm>

Prószycki Gábor: Nyelvművelés számítógéppel? (A számítógépes helyesírás-ellenőrzés új útjai) Magyar Nyelvőr 117 1993, 509-511.

[http://www.morphologic.hu/h\\_pgp5.htm](http://www.morphologic.hu/h_pgp5.htm)

# **PROFIVOX – A LEGKORSZERŰBB HAZAI BESZÉDSZINTETIZÁTOR ÉS SZÖVEGFELOLVASÓ**

**Olaszy Gábor\*, Németh Géza, Olaszi Péter, Kiss Géza**

**\*MTA Nyelvtudományi Intézet**

**Budapesti Műszaki és Gazdaságtudományi Egyetem**

**Távközlési és Telematikai Tanszék**

## **Bevezetés**

Az elmúlt húsz év során a beszédszintézis tudománya és technikája nagy fejlődésen ment keresztül. Ez főleg az automatikus szövegfelolvasókra vonatkozik, hiszen ez az a szolgáltatás, amelyre egyre több gyakorlati alkalmazást fejlesztenének, ha lenne megfelelő beszédminőséget adó beszédszintetizátor. A beszéd mesterséges előállítása bonyolult, hosszú időt, összetett szakképzettséget igénylő feladat. Jó beszédszintetizátort csak speciális tudással rendelkező szakembergárda képes megalkotni, olyan, amelyikben beszédakusztikával, nyelvészetrel, ezen belül fonetikával foglalkozó szakemberek, valamint a számítástechnikában jól képzett mérnökök vesznek részt. A fejlesztés másik fontos szempontja, a nyelvfüggőség. Magyar nyelvű igényes beszédszintetizátort csak saját, jelen esetben magyar szakemberekkel lehet fejleszteni. Magyarországon – immár 20 éve – az MTA Nyelvtudományi Intézetének Kempelen Farkas Beszédkutató Laboratóriuma foglalkozik beszédszintézis kutatással a BME Távközlési és Telematikai Tanszékével szoros kapcsolatban, ahol számos rendszert valósítottak már meg. Az elmúlt tíz év során voltak ugyan más próbálkozások is magyar beszédszintetizátorok fejlesztésére, de egyik sem járt igazi sikerrel (nagyon rossz minőségű, nehezen érthető beszédet tudtak csak előállítani). A jelen tanulmányban ismertetett beszédelőállító rendszer az első olyan hazai beszédszintetizátor, amelyik teljesíti azt a három alapkövetelményt, amellyel egy korszerű beszédszintetizátort jellemezni lehet.

Az első, hogy emberi hangszínezettel rendelkező hangot állítson elő, amely tiszta, érthető kiejtést hordoz, a szintetizátor beszéde dalla-



mos, és ritmikailag is változatos. Más szóval a beszéd első hallásra is megérthető, és a hosszabb szövegek hallgatása sem fárasztja a hallgatót.

A második tulajdonság, hogy illeszkedjen a korszerű, általános technikai háttérhez, ezen megszólaltatható legyen. A mai követelmény az, hogy csak szoftver eszközökkel állítsuk elő a beszédet és azt hangkártyán lehessen megszólaltatni.

A harmadik fontos jellemző, hogy jól kiépített háttér-támogatás (szoftver eszkörendszer) álljon rendelkezésre a fejlesztéshez, módosításokhoz.

A tanulmányban az első feltétel teljesítéséhez, vagyis a szép, érthető beszéd előállításának megvalósításához szükséges fonetikai adatok és szabályok rendszerbe foglalt összességét tárgyaljuk. Nem tárgyaljuk a szöveg-beszédhang átalakításának részleteit (Olaszy et al. 1992), csak az akusztikai megvalósítást mutatjuk be.

### **A szintetizátor általános felépítése**

A szintetizátor elkülönített blokkokból áll, amelyek között az adatáramlást a keretrendszer biztosítja (Olaszy et al. 2000). Így lehetett elérni azt, hogy a különböző feldolgozási szintek helyes működését külön-külön is, csoportokban is és teljes egészében is ellenőrizni lehet a fejlesztés során. Fonetikai szempontból a szintetizátor három részből áll, a szöveg–hang meghatározóból, a szegmentális, valamint a szuprasegmentális szintű feldolgozóból.

**A szöveg–hang átalakító** részben készítjük elő a dallam, a ritmus és az intenzitás szerkezet kialakításához szükséges főbb adatokat.

**A szegmentális modul** az elem összekapcsolóból és a hangelemtárból áll. A hangelemtár elemeivel párhuzamosan több információt tárolunk. Ilyenek a specifikus időtartamok (jelen kötetben Olaszy 2000), valamint azok az információk, amelyekre szükség van a precíz beszédépítéshez. Ezek különböző szintűek. A szövegből vezetjük le és tároljuk a hangokat, azok fő tulajdonságait, a szótagok elhelyezkedését a szón belül, a szó szótagszámát, a szó elhelyezkedését a szövegben (első/utolsó, vessző előtti, frázishatáron van stb.), továbbá a prozódiai jeleket (frázis határok, dallam és hangsúly jelek, mondat-



jellemzők stb.). Frázisnak tekintjük az új dallamvonulattal megvalósuló szövegrészt.



1. ábra

A szintetizátor előkészítő részegységei

Hangszinten tároljuk a hangok képzési tulajdonságait, valamint a hanghoz tartozó belső időstruktúrát periódus jelzők segítségével (zöngéseknél a periódusok számát és határait, zöngétleneknél 10 ms-os osztást férfi, 5ms-os osztást női hangnál), továbbá a hanghatárokat és az adatbázisban tárolt elemek határát. A fenti információk nagy része pontosan kinyerhető a kiindulási adatokból (szöveg és elembázis), egy részüket azonban nyelvészeti és fonetikai vizsgálatok eredményeiből kell származtatni. A prozódia előrejelző jelek meghatározásához csak a bemenő szöveg áll rendelkezésre. A szövegelemzés szintje határozza meg, hogy mennyire pontosan lehet a mondatdallamot, a hangsúlyozást, a beszédsebesség változását meghatározni. Ezen a téren az optimális megoldást a szöveg- és azon belül a mondat szintű szemantikai és szintaktikai elemzés jelentené, azonban ilyen komplex tudományos vizsgálatokat a magyar nyelvre algoritmizálható formában még nem folytattak. Az itt ismertetett rendszerben két kompromisszumos megoldást alkalmazunk a prozódia előrejelzésére. Mindkettő csak az adott mondat szintjére vonatkoztatva végez vizsgálatot, mondatok közötti összefüggéseket nem tudunk megadni. A beszéddallam vonatkozásában az egyszerűbb megoldás statisztikai jellegű, melynek során jóslani próbáljuk a mondat teljes dallamát (az esetleges részdallamok összekapcsolását), valamint a szó szintű hangsúlyozást. A megoldás

lényege, hogy bizonyos szövegelemekhez (vessző, névelő, pozitív, illetve negatív fogalmat hordozó szó, szóegyüttes, frázishatárt magában hordozó szó stb.) hozzárendeljük a megfelelő információt. Ezen kívül jósoljuk a beszédszüneteket a mondaton belül, valamint a beszédsebesség változását is (például, gyorsabb a közlés elején, lassabb az utolsó szóban, ahogy Magdics (1966) is megállapította). A komplexebb megoldás a mondat célirányos szintaktikai elemzésén alapszik (Koutny–Olaszy 2000), amely lényegesen jobb eredményt ad, azonban megvalósítása sokkal bonyolultabb, mint az előző változaté.

A fentiekben összefoglalt sokrétű előkészítő eljárás és tárolt információhalmaz ad módot arra, hogy a magasabb szinteken működő szabályokat a szintézis során elég részletesen lehessen kialakítani ahhoz, hogy a felépített beszéd jó minőségű legyen.

### **Hangelemtár**

A szintetizátor alaphangzását (szegmentális alap) a hangelemtár elemeinek formája és belső akusztikai minősége határozza meg. A hangelemtár elemeit kapcsoljuk egymás után, így hozzuk létre a folyamatos beszédet.

Formánsszintetizátoroknál az építőelemek formája paramétermező, a belső akusztikai minőségüket pedig a formánsszintetizátor paramétereinek a száma, valamint a paraméter-értékek helyes megválasztása (amely a szintetizátor tervezőjétől is erősen függ) határozza meg. Ezek a szintetizátorok szinte mindig egy speciális, robotos hangszínezettel rendelkeznek, nem tudják megvalósítani az emberi hangszínezetet. A fentiekben kívül ennek az is oka, hogy a formánsszintetizátoroknál a hang gerjesztésére szolgáló jel (zöngé, illetve zörej) túl szabályos, nem valósítja meg az emberi hangképzéskor keletkező változatos és kváziperiodikus zöngé jellemzőit, illetve a zörejek szerkezetét. Ennek a problémának a megoldására alkalmaznak a korszerű szintetizátorokban emberi hangból kivágott hullámforma-részleteket a hangelembázis elemeire. A jelen szintetizátor is ilyenekből építkezik. Ezen elemek belső akusztikai minőségét alapvetően az emberi hangot felépítő részegységek alakítják ki, tehát a szintetizátor megszólalásakor mindenképpen érezni lehet, hogy a hang nem for-



mánsszintetizátorból ered. Az így szintetizált beszéd végleges hangzását azonban erősen befolyásolja, hogy milyen módon készítettük elő az elemek konkrét megvalósítását, mennyire felkészült (felkészített) bemondó olvasta fel az elkészítésre kialakított szöveget, milyen szöveget alakítottunk ki a felolvasáshoz stb. A későbbi beszédépítés során ugyanis az elemek összekapcsolási pontjain a lehető legpontosabban biztosítani kell az akusztikai paraméterek értékeinek időbeli folyamatosságát, tehát, hogy a formánsok értékeiben ne legyenek ugrások, az alaphang közel ugyanolyan értékű legyen két összekapcsolt elemben, a jelamplitudó ne legyen kiugróan magas az egyik elemben, kiugróan alacsony pedig a másokban. Ilyen soktényezős követelményrendszer szerint készítettük el a hangelemtár elemeit (Olaszy 1999). A hangelemtár kialakítása a specifikus hangidőtartamok beállításával fejeződik be. A szegmentális szinten készített hangsorban (az elemek össze vannak kapcsolva) tehát a hangidőtartamok kiegyensúlyozottak lesznek, nem lesznek túl hosszú, illetve feltűnően rövid hangok, továbbá, az egyes magánhangzókra, valamint mássalhangzókra jellemző hangidőtartamok is megvalósulnak és a hangok közötti időtartam-arányok (például a magánhangzók és mássalhangzók között) is a nyelvi normának megfelelőek lesznek. A végleges hangidőtartamokat a szupraszegmentális részben alakítjuk ki.

A hangelemtár elemeinek intenzitásviszonyait is ki kell egyenlíteni, hiszen a bemondás során a bemondó hangereje – még a leggondosabb kiejtés során is – változik. Ez azt eredményezi, hogy ugyanazon hangok egyes elemekben intenzívebbek, más elemekben halkabbak lesznek. Az általunk kidolgozott hangintenzitást kiegyenlítő eljárás lényege a következő. 1. Megmérjük, az adatbázis elemeiben minden hang átlagolt intenzitását. 2. A kapott eredményeket a hangzóssági sorrendnek és az arra jellemző dB skálának feleltetjük meg (Olaszy 1989). Ebből meghatározzuk azokat a szorzókat, amiket alkalmazni kell, hogy a helyes hangzóssági sorrendet megvalósítsuk. 3. A hangok intenzitását (szorzással, illetve osztással) a hangzóssági sorrendnek megfelelő arányokra állítjuk be. 4. Az elemhatárokon az intenzitáskülönbségeket interpolációval kiegyenlítjük



Az eljárás eredménye, hogy a szintetizált beszédben a hangzás egyenletesebb, simább lesz. A végleges hangintenzitásokat a szupraszegmentális részben alakítjuk ki.

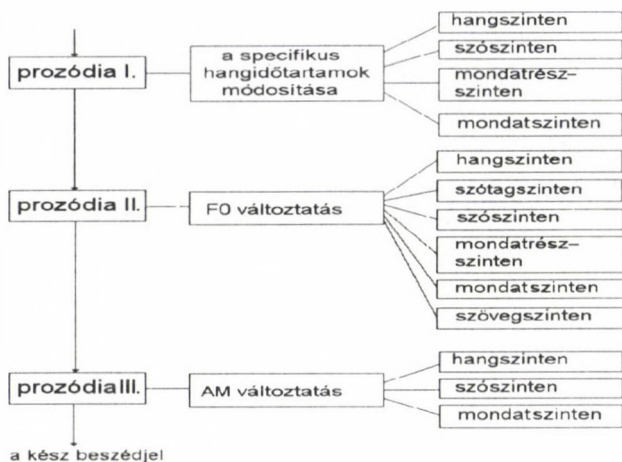
### **A prozódiai szerkezet kialakítása**

A beszéd végleges akusztikai szerkezetének felépítése során kialakítjuk a ritmikát, vagyis a hangsor meghatározott pontjain néhány százalékkal felgyorsítjuk, illetve lelassítjuk a beszédsebességet, a szegmentális szintű beszédre ráépítjük a dallamot, beállítjuk minden szóra a rá kijelölt hangsúlykategóriát, és végül a hangintenzitást is úgy módosítjuk, hogy közelítsük a természetes beszédre jellemző intenzitáviszonyokat. Mindezeket a szupraszegmentális modulban végezzük el (2. ábra).

### **A végleges hangidőtartamok beállítása**

A szupraszegmentális szintre jellemző korrekt időszerkezet a jó hangzású beszéd elsődlegesen legfontosabb követelménye. Hiába valósítunk meg megfelelő intonációt és intenzitásszerkezetet, ha a hangidőtartamok helyenként túl hosszúak, helyenként pedig túl rövidek ("sántít" a beszéd). A természetes beszédben az időszerkezet kialakítása automatikusan, sokszintű szabályrendszer működésének eredményeképpen jön létre. A beszédépítéshez minél részletesebben fel kell tárnunk ezeket a szabályokat és belőlük olyan algoritmizálható szabályrendszert kell kialakítani, amellyel jól lehet közelíteni a valós helyzetet. Ezt a munkát segíti a szintézis lehetősége is azzal, hogy meg tudjuk hallgatni a végeredményt és el tudjuk dönteni, hogy az adott szabály kielégítő eredményt ad-e, vagy esetleg módosítani kell rajta.

A hangidőtartamok végleges, szupraszegmentális szintű beállítása úgy történik, hogy egy szorzófaktort alakítunk ki, amellyel megszorozzuk a specifikus időtartamot. A szorzófaktor 0,5-2,5 közötti értéket vehet fel. A szorzófaktor kialakításához négy szintről használunk fel információt. Minden szintre külön szabályokat dolgoztunk ki.



2. ábra

A szupraszegmentális modul feldolgozási egységei

**Hangszinten** a magánhangzók specifikus időtartamainak hosszúsági sorrendje adja a kiindulási alapot. A legrövidebb hang az [i], a leg-hosszabb az [ø]. Az időtartammódosító szabályban figyelembe vesz-szük az adott magánhangzó elhelyzkedését az időtatamra számított hangsorrendben. Például, ha az [i]-t hosszabbítani kell egy adott sza-bály szerint (pl. a mondat végi szóban), akkor azt nagyobb mértékkel hosszítjuk, mint az [ø]-t hasonló helyzetben.

**Szószinten** figyelembe vesszük, hogy a szó hány szótagos, majd ennek függvényében azt, hogy a hang a szón belül hol helyezkedik el, és hogy milyen a hangkörnyezete. Továbbá figyeljük azt is, hogy a szón belül előfordul-e ugyanaz a magánhangzó többször (például: *elfe-lejtette*). A szó szintű feldolgozás részét képezi a szóhatáron lévő han-gok figyelése is (mi volt az utolsó hang és mi a következő szó első hangja). Például a *jobb bútor* szövegrészben az összevont [b] hang időtartamát nem a [b:] -re jellemző időtartammal kell beállítani, hanem annál rövidebbel. Ha a szóbelseji [b:] időtartamát például 200%-os ér-tékkel jelöljük, akkor a fenti esetre 150%-os időtartamot kell beállítani. Ezt a rövidülést hangsorokon végzett mérési eredményeink, valamint a szintézisből kapott hangsorok meghallgatásos teszthei alapján alakítot-



tuk ki. Ha 200%-os értékkel valósítjuk meg ezt a [b] hangot, akkor az elhangzásban a dadogásra emlékeztető jelenséget hallunk, mert a [b] zöngeszakasza túl hosszú. A természetes beszéd során az ilyen és ehhez hasonló szabályok automatikusan működnek.

**Szövegrész**-szinten a szövegben kijelölt valós szünetek adják az információt az időtartam módosításhoz. A mondat belsejében tartott szünetek előtt kissé megnyújtjuk a hangidőtartamot, a szünetek utáni folytatáskor pedig rövidebbre állítjuk be, tehát gyorsítjuk a beszédet.

A **mondat** szintjén a mondat első és utolsó szavát külön kezeljük. Az előbbiben rövidítjük, az utóbbiban nyújtjuk az időtartamokat.

A végleges hangidőtartam-beállítás gyakorlatilag úgy történik, hogy a közlés minden hangjának specifikus időtartamát megszorozzuk a fenti négy szintre kidolgozott részletes szabályok alkalmazásából **összegzett szorzófaktorral**. A kapott végleges időszerkezet már ritmikus beszédet reprezentál.

### **Az alapfrekvencia-változások beállítása**

Az alapfrekvencia változtatása adja a leglátványosabb színezést a beszédben, ez képviseli a beszéd hangszerelését (Szende 1995). Lényegesnek tartjuk kiemelni, hogy az alapfrekvencia-változáson itt nemcsak a beszéddallamot értjük, hanem minden változást, amely a beszédképzés során valamilyen szabályszerűséghez kapcsolható. Itt modellezzük tehát a hang szintű változásokat is, a hangsúlykategóriák alapfrekvencia vonatkozású részeit és természetesen a mondat dallamot is. Az alapfrekvencia változtatása kihat az időstruktúrára is, mivel a hangperiódusokat rövidítjük, illetve hosszabbítjuk. A korábbi szinten beállított időtartamok megtartására idővetemítést kell alkalmazni.

Az alapfrekvencia változtatáshoz az információt hat szintre végzett elemzés eredményei alapján kialakított szabályokból gyűjtjük össze. Ezek a szintek a következők: hang, szótag, szó, mondatrész, mondat, szöveg (2. ábra). Ez azt jelenti, hogy vannak szabályok, amelyeket adott hangokra alkalmazunk, vannak olyanok, amelyek adott szótaghoz (első, utolsó, utolsó előtti) kötöttek, egyeseket meghatározott szavakra (tartalom, hely szerint) alkalmazunk, majd következnek a magasabb szintű szabályok, amelyek szövegrészre, illetve a teljes mondatra



vonatkoznak. A legmagasabb szintet képviselik a mondatok között működő törvényszerűségeket leíró szabályok. A komplex alapfrekvencia-struktúra felépítését a legmagasabb szintről kezdve, mindig az alacsonyabb szintre visszalépve végezzük. Ez úgy zajlik, hogy a magasabb szint struktúrájára ráépítjük az alacsonyabb szintű változásokat, majd a feldolgozás végén a hangok szintjén számítjuk ki a hangra vonatkozó végleges alapfrekvenciát. Ezt állítjuk be.

**Szövegszinten** az egymás után következő mondat(ok) típusa határozza meg az egyes mondatokra kidolgozott dallamgörbe kezdőpontját és végpontját (hány Hz-en kezdődik a mondat dallama és hányon fejeződik be, hogy dallamszerkezetiileg jól illeszkedjen a megelőző, illetve a következő mondathoz). Szabályokat dolgoztunk ki arra, hogy milyen összefüggések vannak a kijelentő, a kérdő, a felszólító és a felkiáltó mondatok dallamszerkezete között (Olaszy 2000).

A **mondat** szintjén az adott mondatra vonatkozó általános dallamstruktúrát valósítjuk meg. Ezt hosszabb, illetve összetett mondatok esetében a belső frázisokra alkalmazott egyedi dallamstruktúrák egymás után való kapcsolásával hozunk létre. A mondatdallam felépítésére Varga (1994) elméleti struktúrái alapján emelkedő, enyhén emelkedő, eső, enyhén eső és lebegő dallamelemeket alakítottunk ki konkrét adatokkal (1. táblázat), amelyeket szavak, vagy több szóból álló szövegrészek dallamának közelítésére használunk. A táblázat elemeiben az alapfrekvencia kezdő és végpontjait jelöltük %-ban a változás a két pont között lineáris. A konkrét Hz-érték kiszámítása a szintetizátorban beállított kiindulási alapfrekvencia-értékből történik.

1. táblázat: A mondatdallam kialakítására meghatározott lineáris dallamelemek és kezdőpontjaik, illetve végpontjaik %-ban kifejezve

Dallamelemek	1.	2.	3.	4.	5.	6.
1. eső	100-85	95-80	95-70	90-70	70-65	100-70
2. enyhén eső	100-95	95-90	95-85	85-75	80-75	75-70
3. szinttartó	100	95	90	85	80	75
4. enyhén emelkedő	95-100	90-95	85-90	80-85	80-95	80-100
5. emelkedő	90-100	85-100	75-90	70-80	70-100	75-100

A dallamelemek meredeksége a megvalósulás során attól függ, hogy milyen hosszúságú beszédre szre vonatkoztatjuk a változást. Ebből belátható, hogy az 1. táblázat elemeivel végtelen számú különböző dallammenetet lehet összeállítani. Egy természetes ejtésből adódó dallamformát tehát törtvonalas közelítéssel írunk le. Az egyes elemekre a táblázat sor-oszlop számával lehet hivatkozni a szintézis során. Látható, hogy az elemek változatos összekapcsolásával szinte bármilyen egyszerű dallamforma létrehozható. A dallamelemek kiválasztását, összekapcsolásuk szabályait a szövegben elhelyezett prozódiai jelzők vezérlik. Az alábbiakban néhány példamondatot mutatunk be dallamforma előrejelzőkkel. A //jel az új dallamegység kezdetét jelenti rövid szünet tartással, a / jel ugyanezt szünet tartás nélkül. A számok az 1. táblázat sor és oszlop számai által jelölt dallammenetet jelentik.

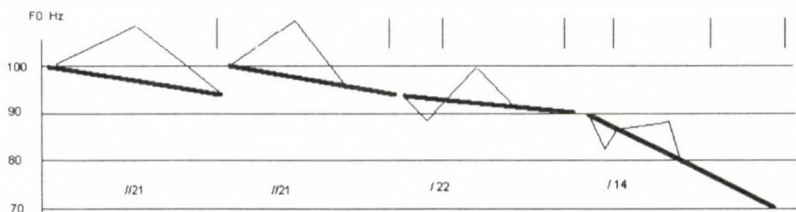
//21Holnap //21megírom /22a levelet /14a külföldi partnernek. (1)  
 //11A tervezett tárgyalás után //21 levelet írok /11a külföldi partnernek. (2)  
 //11Magyarországon /44novemberben //11esik az eső. (3)  
 //14 A külföldi partnernek is írok levelet. (4)  
 //21István bekapcsolta /41a [W]számítógépet, //31és //11elindította a programot. (5)

A **szó** szintje az intonáció kialakításának második fontos része. A rendszerben azt a szabályt írtuk elő, hogy az adott mondat minden szavára jelölni kell, hogy milyen lesz a szó belső dallamszerkezete, ami vonatkozhat az egész szóra, illetve vonatkozhat a szó valamely szótagjára, szótagjaira. Kijelentő mondatokban négyféle jelölést kaphatnak a szavak: fókusz, hangsúlyos, neutrális (nincs rajta hangsúly) és negatív hangsúlyos. Jelölésük a szövegben az adott szó (szótag) előtt a következő: [F:] [W:] [N:] [-:]. Ezek közül a teljes szóra csak a neutrális jelzés és a negatív hangsúly vonatkozik, a többi szótag szinten valósítjuk meg. Negatív hangsúly esetén az egész szóban valósítjuk meg a dallamszerkezetet, függetlenül a szó hosszától. Negatív hangsúlyt kaphatnak a névelők, határozók stb. Neutrális jelzés esetén a szóban semmilyen dallamváltozást nem hajtunk végre, csak az alapidallam van jelen (az 1. táblázatból), amit magasabb szinten előírtunk.

**Szótagszinten** valósítjuk meg a hangsúlyozás különböző szintjeit (fókusz, hangsúly), a kérdés–intonáció jellegzetes dallamsúcsait, a kérés, a felszólítás, a parancs jellegzetes dallam meneteit. A szótag szintjére előírt dallamváltozásokat az esetek többségében két szótagban valósítjuk meg (dallamsúcs: dallam emelkedés és csökkenés), ritkábban egyben (például kérdésekben: az alapfrekvencia felugrása és csökkenése ugyanabban a szótagban). Az (1) példamondat a szó szintű jelzésekkel ellátva a következőképpen alakul:

/21 [W:]Holnap //21 [W:]megírom /22 [-:]a [W:]levelet /14 [-:]a [W:]külföldi [N:]partnernek.

A 3. ábrán bemutatjuk ennek a mondatnak a sematikus dallamszerkezetét.



3.ábra

A dallammenet és a ráépített szó szintű alapfrekvencia változások az (1) mintamondatban

Az ábrán a függőleges vonalak a szóhatárokat jelképezik. Az F0 értékek 100Hz-es kezdeti alaphangra vonatkoztatva vannak ábrázolva. A vastag vonalak az 1. táblázat szerinti alapidallamokat (4 külön dalmegeység egymás után kapcsolva), a vékony vonalak a szó, illetve szótag szintű alapfrekvencia-változásokat jelölik.

A **hang** szintű beállításokkal fejeződik be az alapfrekvencia-változás végleges formájának kialakítása. Itt a hang függvényében néhány Hz-es változtatásokat hajtunk végre. Ezek a változtatások tulajdonképpen a mikrointonáció megvalósítását jelentik (Olaszy 1989).



### **Az intenzitás-szerkezet beállítása**

Az intenzitás-szerkezet beállításához három szintről, a hang, a szó és a mondat szintjéről használunk fel információt.

A **hangszintű** beállításnál a magánhangzók hangzóssági tulajdonságainak és a hang helyzetének alapján végzünk hangerő növelést, illetve -csökkentést. A természetes beszéd során ezt a hang képzésének pillanatában automatikusan elvégezzük, itt azonban rögzített elemtárral dolgozunk, amelyben a hangok adott hangintenzitással vannak tárolva. A korrekciók eredménye, hogy az intenzitás szempontjából is kiegyenlített hangzást kapunk, nem lesznek túl erős és túl halk hangok a hangsorban. A hang szintű feldolgozáshoz tartozik az is, hogy olyan esetekben, amikor a hangsúlyozást nem dallamcsúccsal kell megvalósítani, az adott magánhangzó intenzitását növeljük meg.

A szó szintjén a hangsúlyozott szavakban csökkenő intenzitás-struktúrát valósítunk meg. A mondat szintjén pedig kialakítjuk a mondatra jellemző hangintenzitás szerkezetet. Ez a mondat típusától függően változik.

### **A szintetikus beszéd minőségének értékelése**

A fentiekben leírt elembázissal és szabályrendszerrel létrehozott szintetikus beszéd minőségét egyszerű percepciós tesztekkel és társadalmi szintű véleménykéreléssel ellenőriztük. A percepciós tesztekben 240 egy- és kétszótagú szót, valamint 48 mondatot (3-6 szó mondatonként) hallgattatunk meg 6 tesztalannyal (4 férfi és 2 nő, életkoruk 30-45 év közötti). Mindkét esetben az volt a feladatuk, hogy írják le amit hallottak. Az eredmények kiértékelésénél csak azt a szót tekintettük elfogadhatónak, amelyben nem volt hanghiba. Ezzel a kritériummal 83%-os eredményt kaptunk. Abban az esetben, amikor zárhang tévesztését is elfogadtuk jó azonosításnak (például *Pál* helyett *tál*-t értettek), akkor 93%-os szóazonosítási szintet kaptunk. Mondatok esetében a helyes azonosítás elérte a 98%-ot.

A társadalmi szintű értékelést a Westel Mobiltelefon Társaság végezte (a szintetizátor a Westel "Mailmondó" nevű elektronikuslevél-felolvasó szolgáltatásában működik 1999 decembere óta

(Németh et al. 2000)). A kérdőíves felmérés során az előfizetőknek 5 fokozatú skálában kellett értékelni az elektronikus levélfelolvasó beszédminőségét. Az mérés átlageredménye 4,5 volt.

### Irodalom

Olaszy, G. – Gordos, G. – Németh G. : The Multivox multilingual text-to-speech converter. In: Bailly, G. – Benoit, C. – Sawallis, T.R. (Eds.): Talking Machines: Theories, Models, and Designs. Elsevier. Amsterdam 1992, 385–411.

Olaszy, G. – Németh, G. – Olaszi, P. – Kiss, G. – Zainkó, Cs. – Gordos, G.: Profivox – a Hungarian TTS System for Telecommunications Applications. International Journal of Speech Technology. Volume 3/3-4. Kluwer Academic Publishers. 2000.

Olaszy, G.: The prosody structure of dialogue components in Hungarian. International Journal of Speech Technology. Volume 3, number 3-4. Kluwer Academic Publishers. 2000.

Magdics Klára: A magyar beszédhangok időtartama. Nyelvtudományi Közlemények 68. 1966, 125-139.

Koutny, I. – Olaszy, G. – Olaszi, P.: Prosody prediction from text in hungarian and its realization in TTS conversion. International Journal of Speech Technology. Volume 3/3-4. Kluwer Academic Publishers. 2000.

Németh, G. – Zainkó, Cs. – Fekete, L. – Olaszy, G. – Endrédi, G. – Olaszi, P. – Kiss, G. – Kiss, P.: The design, implementation and operation of a Hungarian E-mail reader. International Journal of Speech Technology. Volume 3/3-4. Kluwer Academic Publishers. 2000.

Olaszy Gábor: Beszédatadabázisok tervezése gépi beszédelőállításához. Beszédkutatás '99. Szerk.: Gósy Mária. MTA Nyelvtudományi Intézet. Budapest 1999, 68-89.

Olaszy Gábor: Elektronikus beszédelőállítás. A magyar beszéd akusztikája és formánsszintézise. Műszaki Kiadó. Budapest 1989.

Szende Tamás: A beszéd hangszerelése. MTA Nyelvtudományi Intézet. Budapest 1995.

Varga László: A hanglejtés. In: Strukturális magyar nyelvtan. 2. Fonológia. Szerk.: Kiefer Ferenc. Akadémiai Kiadó. Budapest 1994, 468-549.



# MŰKÖDŐ SZÖVEG A RETORIKÁBAN

L. Aczél Petra

ELTE Tanárképző Főiskolai Kar  
Magyar Nyelvtudományi Tanszék

## A retorika jellege

A retorika divatos; megújult népszerűségét az az egyveleg adja, ami igen kevés nyelvészeti tudományág sajátja: a gyakorlat és elmélet élvezetes alkalmazhatósága, eszköztárrá egyszerűsíthetősége, ugyanakkor a funkcionális tudományosság klasszikus jellege. A retorikai témájú művek – népszerűek vagy jobb esetben korszerűek – a receptet keresik és adják a vágyott sikeres megszólaláshoz és így lassan a görög tekhnéből (mesterség, művészet, gyakorlat) csak a technikát, az alkalmazást hagyják a retorika művelőire.

A retorikát valós igények hívják életre, s bár Arisztotelész filozófiai rendszerében a cselekvés, alkotás, szemlélődés kategóriái közül az alkotáshoz (poiészis) tartozik, tudományos hozzáállásában is megmarad a funkcionális, leíró sajátosság, az a tapasztalati jelleg, amit Crevier a legrészletesebb francia retorika szerzője így írt le: „A rétorok és a nyelvtanok szerzői csupán nevet adtak azoknak a dolgoknak, amire a természet mindnyájunkat megtanít.” (vö. Kibédi Varga 1998, 52). A retorika a társadalom és ember közelítésében végül mégiscsak a gondolatot köti össze a szöveggel, a beszéddel, olyan módon, hogy létrejöhessen a hatás, s a szöveg önmagán túlmutató egységgé váljon. A retorika a nyelv organizmusának olyan működési elvét adja, olyan formátárat biztosít és olyan igényeket támaszt, amelyek eredményeképpen megszülethet a meggyőző, a működő szöveg.

A szó hatalma a gondolon: ez lehetne a retorika szlogenje, s talán illő is korunkhoz, amely a rövidülő megnyilatkozásban nem hagy kiaknázatlanul egyetlen szót, egyetlen hangosítási lehetőséget sem. Célja pedig a meggyőzés, a meggyőzés nyelvi kategóriájában pedig a retorika mutat majd utat.



## **A retorikai szituáció és a beszélő szándéka**

„A beszéd három dologból tevődik össze: a beszédből, amiről beszél és akihez beszél; a beszéd célja az utóbbira irányul, azaz a hallgatóra.” (Arisztotelész 1982; 1358b). Arisztotelész a kommunikációs modell archetípusát adja, azt a beszédhelyzetet írja le, amelyben a beszéd, a szöveg létrejön. A retorikai szituáció egyes értelmezések szerint azon kontextusok összessége, amelyben a szónok beszédeket mond, vagyis viszonyok, események, tárgyak és szereplők szövevénye, más felfogásban azonban leegyszerűsítve azt találjuk, hogy a retorika mint esztétika a formára reflektál, illetve azt alakítja, s mint etika válaszol egy szituációra (vö. Árpás 1997, 93).

Ez utóbbi elképzelés ráirányítja figyelmünket a kommunikációs helyzeten belül érvényesülő és leírható kommunikációs és beszéd-funkciókra. Az antik görög–latin retorika vetette meg annak a funkcionális leírásnak is az alapját, amely a közlésfolyamatban a tájékoztatás, a kifejezés és a felhívás (Jakobson modelljében émotív, konatív és referenciális funkciók) meglétét mutatja meg. Esztétikai, stilisztikai formátára pedig, a poétikai és metanyelvi funkciók létjogosultságát teremti meg.

A szónoki, beszélői szándéktípusok közül első az értelmi szándék, amellyel tanítunk, érvelünk, s erkölcsi tartalmat fogalmazunk meg, ennek kifejezése egyszerű stílusban történik. A második, a mérsékelt érzelmi szándék közepes stílussal a gyönyörködtetést célozza, s kulcsfogalma az éthosz. Ez a beszélőnek szavaiban megnyilvánuló jelleme, a hallgatóság erkölcsi felfogásához viszonyított morális állapota, amely által tetszést vagy nemtetszést vált ki a hallgatóból. A beszédpartner számára az éthosz teremti meg a szavahihetőséget, amely a hozzáértés, az erényesség és a képviselt ügy iránti elkötelezettség kifejezése által jön létre. Az éthosz szerepe a meggyőzésben is lényeges, hiszen a „jó beszéd tudományával”, a retorikával a beszélő a hallgatót az önmagáival azonos lelkiállapotba hozza, s így a tisztességes célokkal a hallgató is azonosulni tud. Az interiorizálás, a bennfoglalás minden olyan nyelvi üzenet célja és eredménye, amely a meggyőzés, vagy az elfogadtatás felé vezet. Azok a reklámok, amelyek szavahihető forrásokat (ismert színészeket, személyiségeket, nevesülő orvosokat, szakértőket) vonultatnak föl szolgáltatásuk népszerűsítésére, éppen ezt

az etikus, a beszélő személyiségéből fakadó, szavahihetőséget eredményező, azonosulást és meggyőzést elindító szónoki szándékot hasznosítják. Harmadikként a páthosz fogalmával dolgozó szenvedélyeket célzó értelmi szándékot találjuk, azt a törekvést, amely felhívja, megindítja a hallgatót fennkölt stílussal. A páthosz a szövegben megnyilvánuló erő, a stílus adta emelkedettség, amely csak a megnyilatkozás bizonyos pontjain fogadható el.

A három szándék megfeleltethető a kommunikációs és jakobsoni funkcióknak egyaránt, így közelebb kerülünk a retorika és az általa létrehozott, jelen esetben működő szövegnek nevezett üzenet közlés-folyamatához.

### **A szöveg részei és a meggyőzés folyamata**

A retorika a sikeres szöveget értelmezi, érti és írja elő, a hatással rendelkezőt, a meggyőzőt. A meggyőzés feltárásában a szöveget tekinteti alapegységnek, s a szöveg építkezésével írja le azt a pszichés folyamatot, amely végén a meggyőzés megváltozása áll. A meggyőzés verbális természetét mutatja meg a klasszikus retorika szövegstruktúrája úgy, hogy ezzel a nyelvről való tudást, a felhasználható nyelvi anyagraktárt is bővíti és kategorizálja.

Korax adja az első beszédfeosztást, bevezetésre, elbeszélésre, befejezésre osztva a szöveget. Tanítványa, Teisziász ehhez a sorhoz illeszti a bizonyítást, azt a beszédrészt, amely az érvek, példák és jelek segítségével beteljesíti a rábeszélést, meggyőzést vagy éppen manipulációt. Tehát minden beszéd legalább négy részből áll: az exordium, a narratio, a confirmatio és a peroratio szövegrészeiből (vö. Kibédi Varga 1998, 51). A római retorika egyik kiemelkedő munkája a *Rhetorica ad Herennium* ezt a beszédrész-sort hét elemre bővíti úgy, hogy az elbeszélés (narratio) mögé fűzi a kitérés (digressio), a téma-megjelölés és felosztás (propositio és partitio) szakaszait, a bizonyítást (argumentatio, confirmatio) követően pedig, a cáfolás (refutatio) szövegrészt illeszti. Eltekintve attól a haszontól, amelyet ez az árnyalt struktúra egy beszéd összeállításakor jelenthet, az egyes beszédrészek szövegfunkcióit kell figyelembe vennünk és vizsgálnunk a meggyőzés folyamatával párhuzamba állítva, hogy a működő szöveget valóban elemezhesük.



A meggyőzés mint eredmény a szociálpszichológia által is elismert hatás. „A meggyőzés tehát a tudat átszervezése valamely nézet, elv elfogadtatása, illetve még meglévő hibás álláspont, beidegződés megváltoztatása; főként racionális és intellektuális momentumokra építő, kommunikálás útján végbemenő pszichikus ráhatás.” (Szabó 1998, 68). A meggyőzés lehetőségét a kommunikációban a beszédpартnerek között létrejött 1. bizalom, 2. összehangoltság, és 3. alkukészség teremti meg. Mindhárom alapját a beszélői éthiszban találjuk.

A meggyező kommunikáció leképezésében olyan állomásokat fedezhetünk fel, amelyeket mind a beszélői szándékokkal, mind a szövegrészek felhasználásával elérhetünk. A meggyőzés hatása felfogható eredményeként is. A hozzáállás változása Hovland és Janis szerint úgy tekintendő, mint bármilyen megfigyelhető változás mögött meghúzódó konceptualizáció. Az attitűd megváltozásából visszavezethető a vélemény, az észlelés, az affektus és a cselekvés változása (vö. Bettinghaus 1977, 171-178). Ezen változásokat indukálják a klasszikus retorika szövegfunkciói a meghatározott szerkezetben.

### **Az észlelés megváltozása**

A bevezetés a szerénységi toposzokkal a jóindulat megragadását, a párhuzamos mondatszerkesztéssel, anaforák vagy epiforák alkalmazásával, kérdések és felkiáltások használatával a figyelem felkeltését, a téma megadásával az informativitást biztosítja. Az elbeszélés, az ügy, a tényállás szereplőit, körülményeit, az eset sajátosságait tárja fel, mintegy jellemzést adva. A bevezetés és az azt követő elbeszélés (narratio) egy adott téma, és egy vállalt ügy kifejtésében az észlelés irányítását, a percepció megváltoztatását célozza. Az ismert a valószínűvel keveri, így a figyelmet átirányíthatja más és más, a beszélő észlelését fókuszáló tartalmi, valóságbeli jegyekre. Például:

1. Szókratész védőbeszéde, bevezető: az észlelés a bírák hazugsága, a beszélő kora és tisztessége felé fordul.
2. „*Ronda és finom*”; csokoládéreklám, amelyhez kép is társul: az elfogadható, azonosítható, észlelhető mellé az új tartalmi jegy már a percepció irányítójaként lép fel.



3. „*Ki ismeri Széphidi Klárát?*”; óriásplakát, amely ezen hirdetési módszer eredményességét volt hivatva bizonyítani, azáltal, hogy az ismeretlen nevet rögzítette.

### **Az affiliáció megváltozása**

A kitérés (digressio) kihagyható ugyan a beszéd építkezéséből, mégsem elhanyagolható az a hatás, amelyet a meggyőzés érdekében elérhetünk vele. A kitérő, elbeszélő szövegrész egy – látszólag a témától elszakadó – történet, saját élmény, anekdota leírása, amely a szemléletességet, személyességet és a hitelességet növelheti a hallgató számára. Ugyanakkor példászerűen később erősítheti az induktív bizonyítást. A történet olvasása közben átélhetővé válik a közölnivaló tartalom, azonosulhatunk a beszélő által jól kiválasztott szereplőkkel. Az affiliációs változások eléréséhez, vagyis ahhoz, hogy a kommunikációs partner érzelmileg közeledjen a beszélő emocionális állapotához, szükség van erre a beszédrészre. A példákban a digresszív részt kurzívan jelöltük.

4. „Az az előadó, aki előre megírja és betanulja beszédét, csak idejét pazarolja és veszélybe sodorja magát. Egész életünkben spontán beszélünk, soha nem állítjuk össze előre a mondatokat. Gondolataink vannak, és ha azok világosak, a szavak olyan természetesen és öntudatlanul jönnek, mint légzéskor a levegő. *Még Winston Churchillnek is jókora fáradságába került, míg mindezt megtanulta. Fiatalkorában, miközben a brit parlamentben beszélt, teljesen leblokkolt. Az égvilágon mindent elfelejtett. Szörnyen zavarban volt és szégyellte magát. Újra és újra nekirugaszkodott az utolsó mondatnak, de a folytatás sehogy sem akart az eszébe jutni. Leült, az arca égett a szégyentől. Attól a naptól kezdve Winston Churchill soha meg sem próbálta betanulni a beszédeit.*” (Carnegie 1999, 44)

5. Pázmány Péter így digresszált egyik nagypénteki beszédében: „*Én is elfáradtam a szólásban, és talán sokan megúnatkoztak a hosszú hallgatásban, mert ez az emberi gyarlóságnak fogyatkozása, hogy csak előszámlálni és meghallani sem győzi, amiket szenvedni akart az isten Fia miérettünk.*”

6. Reklámszöveg a következő jókívánssággal és kiegészítéssel: „A Moulinex kellemes húsvéti ünnepeket kíván Önnek és családjának. *(Ez utóbbi nem része az akciónak, ez csak egy jókívánsság.)*”

### A vélemény megváltozása

A bizonyítás és cáfolás a racionalizáció útján, az érvek logikai felépítésével a hallgatót verbalizált hozzáállásainak és értékeinek megváltoztatásáig juttathatja el.

Arisztotelész a bizonyítást olyan összetett folyamatnak tekinti, amely az éthosz, a páthosz és a logika forrásaiból táplálkozik (vö. Adamik 1993, 68). Ez utóbbi fogalom nyelvi kategóriája nagymértékben az arisztotelészi retorika erénye és érdeme, amennyiben a retorikát a dialektikával párba állítva gondolkodásbeli–logikai alapegységét az enthümémában határozta meg. A dialektika a szillogizmusokkal él, olyan állítássorokkal, melyek egyik premisszáját a másikkal megfeleltetve, a harmadikat következményként is levonhatjuk (például *Szókratész ember. Az ember halandó. Tehát Szókratész halandó*). A formális logika ezt modus ponendo ponensként, állító módként, ennek negatív megfordítását modus tollensként ismeri. Mindazonáltal a retorikai szöveg alapvető tulajdonsága a kommunikativitás, az a sajátság, hogy adott beszédhelyzetben készül, s annak jegyeit magán viseli. Így már nemcsak a beszélő szándéka az információ közlése, mint ahogyan nem pusztán a szónok feladata, hanem a szöveg sajátos hatása is a meggyőzés. Az enthüméma a szillogizmus retorikai változata. A görög kifejezés *en thümé* azt jelenti, az elmében már meglévő, tehát a hallgató tudatában felismeréssé érő következtetés (vö. Barthes 1997, 26). Az enthüméma nem mondja ki a harmadik premisszát, önállóságot hagy a hallgatónak, miközben mindennél erősebben kontrollálja és irányítja a gondolkodását. A kommunikációs partner saját gondolataként engedi verbalizálni az eredeti szövegbe tudatosan kódolt tartalmat. A vélemény megváltoztatásában az okoskodás kihagyásos, enthümematikus formáját választja. Ez nyújt lehetőséget a manipulációra, arra, hogy a beszélő elkerülhesse az implikált tartalmakkal kapcsolatos felelősségvállalást.

A cáfolás, a bizonyításhoz fűzve, az enthüméma egy sajátos átformálását, a retorikai kérdést hasznosítja, amelynek elnevezései már a



klasszikus retorikai alakzattárban sokszínűséget mutatnak attól függően, hogy milyen céllal és beszélői kifejező szándékkal építjük bele a választ. A retorikai kérdés választalansága („szuggesztivitása”, vö. Birkenbihl 1998, 181-182) olyan pragmatikai tulajdonsággal bír, olyan performatív jelleget nyer, ami a hallgató irányított válasz-realizációjában összpontosul. Az interrogáció (retorikai kérdés) a cáfolásban a meggyőzés közvetlen eszközévé válik. Arisztotelész így ír szükségességéről: „Ami a kérdezést illeti, alkalmazása akkor célszerű, amikor az ellenfél már kifejtette ellenkező álláspontját és egyetlen kérdéssel ki lehet mutatni annak lehetetlenségét. (...) amikor két dolog közül az egyik világos, továbbá nyilvánvaló, hogy az ellenfél megengedi a másikat, ha megkérdezzük. A kérdező azonban, miután a kérdésre adott válasszal az egyik premissza birtokába jut, ne kérdezzen rá a másakra, mely nyilvánvaló, hanem vonja le a következtetést. (...) Továbbá, ha azt akarja kimutatni, hogy az ellenfél ellentmond önmagának, vagy állítása paradox. Negyedszer, amikor az ellenfél csak szofisztikus választ tud adni (...)” (1982, 1419a).

#### 7. Kossuth Lajos: Kiáltvány a néphez

*„Még a nyomorult féreg sem tűri, hogy lábbal tapodják. A magyar nép nem leszen, nem lehet gyávább a féregnél.”*

#### 8. Újságcikk: „Mit mondjunk ujját szopó gyermekünknek?

*Amikor valamiről tudjuk, hogy rossz, és mégis csináljuk, az azt jelenti, hogy nincs elég erőnk legyőzni önmagunkat. Az igazán okos emberek nem okoznak bajt saját maguknak. És én tudom, hogy te már nagyon okos vagy.”* (Hölgyvilág, 2000. április)

#### 9. Óriásplakát: „A szoftverhamisítást a törvény 2-5 évig terjedő szabadságvesztéssel büntetheti. Önt ellenőrizték már?”

### **A cselekvés megváltozása**

A befejezés (peroratio) a cselekvésre szólítás imperatívuszát hordozza. A páthosz, az emelkedett stílusjegyek segítségével, az ismétlés által tartalmi rögzítéssel újtára bocsátja a ható, működő szöveget, amely majd fajtája szerint a múlt (törvénytörési beszéd), a jelen (bemutató beszéd) vagy a jövő (tanácskozó beszéd) eseményeiben,



vagy azok megítélésében készíti a hallgatót attitűdváltásra (vö. Plett 1988, 135-136).

10. Arisztotelész, Lüsziasz beszédét idézve (Rétorika): „*Szóltam, meghallgattatok, értitek, ítéljetek.*”

11. Móricz Zsigmond Ady gyászbeszédében: „*A költő, aki fölöttünk fekszik dupla koporsóban: közöttünk van, közöttünk él tovább, bennünk munkál, a mi életünk boldog jövőjén.*

*Emlékezzetek Ady Endrére, emeljétek fel az ő igéit, és hajtsátok végre az ő életének törvényét!*”

A fenti változásfolyamat, az észlelés, az affiliáció, a vélemény, cselekvés, majd attitűd változásai, melyek leírásai a Yale Studies in Communication című kiadványsorozatban jelentek meg, mára túl általánosan és szabályszerűen hatnak a kommunikációelmélet szempontjából (vö. Buda 1994, 38-39). Ugyanakkor a retorika értelmezésében, a sikeres szöveg genezisében a szerkesztés valós, tudatos folyamatát is leírhatják.

### **A retorikai szöveg hangosításáról**

A retorikai szövegnek, ha valóban hangosításra készülő előadás, az előszó, a szövegfonetika eszközszerkeztését is kódolnia kell. Az ilyen szövegekben „erőre kell kapatnunk az önálló gondolkodásra, problémátásra nevelő mozzanatokot” (Bakos 1983, 12). Az átélés folyamatában, a megértést szem előtt tartva rá tudjuk venni a hallgatóságot arra, hogy a felolvasóval vagy az előadóval együtt gondolkodjék. A retorika Bakos szerint a deklamálás, sőt a gesztikuláció megjelenítést is szabályozhatja, úgy, hogy a tartalmi gazdagságot megtartva a jól hangosítható, tisztán tagolható mondatszerkeztést a kolónok (tagmondatok) kiegyenlítésével érjük el. A periódusokban (körmondatokban) kerüljük a monoton, azonos alárendeléseket, figyelünk a mellérendelések funkciójára, s arra, hogy ez ne vezethessen tartalmi redundanciáig. A szóhasználatban a szerkesztési egységek azonos pontjain alkalmazunk ismétlést, fokozást. Szimmetrikus vagy párhuzamos szerkesztéseket alkalmazunk az ismétlésben, a gondolatalakzatok érthetővé tételében. Ezekkel az alakzatokkal a hangosításban a ritmus és tem-

pó színesítését is megcélozhatjuk. Szüneteket kódolunk a szövegbe az ellipszis, zeugma, vagy éppen az előbb említett logikai enthümea révén. Intonációs és hangerőbeli váltásokat indukálunk a modalitás változatosságával.

### **Összefoglalás**

A szöveg a hallgatóra tett várható hatás, meggyőzési folyamat szövegbe tervezésétől válik a retorikai tekhné elméletévé és gyakorlatává. A szándékbeli, szerkesztési, szövegfunkciós és hangosítási szempontok figyelembe vételével a megnyilatkozás, a kommunikációs üzenet a nyelven túli viszonyok megváltozásában, az érzelmek, az értelem, az elhatározás, a tudat átszervezésében vehet részt. A retorikai szöveg működése az a hatás, amelyet a meggyőzés során, annak eredményeképpen érünk el. A retorikai tudás tehát a divatosság és klasszikusság végleteiben az egyéni, emberi megszólalást alakítja az igényes gondolkodás tükrévé.

### **Irodalom**

Adamik Tamás: A retorika történetének tanulságai a mai iskola számára. MNyTK 196. 1993, 63-69.

Arisztotelész: Rétorika. Gondolat. Budapest 1982.

Árpás Károly: Magyar nyelvtan VI. (Rétorika). Szeged 1997.

Bakos József: Felolvasásra és előadásra szánt szövegek alkotása, értelmezése és hangosítása. In: Tanulmányok a mai magyar nyelv szövegtana köréből. Szerk.: Rácz Endre – Szathmári István. Nemzeti Tankönyvkiadó. Budapest 1994, 5-15.

Barthes, R.: A régi retorika. In: Az irodalom elméletei III. Szerk.: Thomka Beáta. Jelenkor. Pécs 1997, 69-175.

Bettinghaus, E.: A meggyőző kommunikáció. In: Kommunikáció I. Szerk.: Horányi Özséb. Közgazdasági és Jogi Könyvkiadó. Budapest 1977, 167-191.

Birkenbihl, Vera F.: Kommunikációs gyakorlatok. Trivium. Budapest 1998.

Buda Béla: A közvetlen emberi kommunikáció szabályszerűségei. Animula. Budapest 1994.

Carnegie, D.: Sikerkalauz 3. Gladiátor Kiadó. Budapest 1999.

Kibédi Varga Áron: Szavak, világok. Jelenkor. Pécs 1998, 48-61.

Szabó István: Bevezetés a szociálpszichológiába. Nemzeti Tankönyvkiadó. Budapest 1998.

Plett, H.: Retorika és stilisztika. In: Tanulmányok az irodalomtudomány köréből. Szerk.: Kanyó Zoltán – Síklaki István. Tankönyvkiadó. Budapest 1988, 131-167.



# A MAGYAR KÍSÉRLETI BESZÉDKUTATÁS TÖRTÉNETE állandó kiállítás

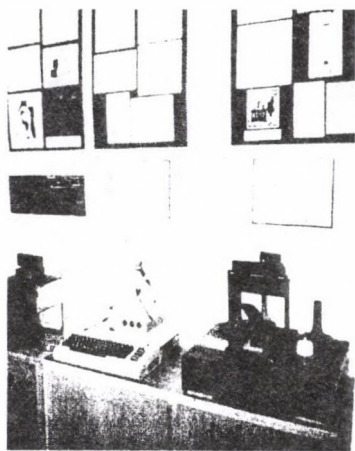
„Milyen más alapokra lehet a  
jövőt építeni, mint a múltra?”

*Szent-Györgyi Albert*

Az MTA Nyelvtudományi Intézetének Fonetikai Laboratóriuma 2000-ben felvette a „Kempelen Farkas Beszédkutató Laboratórium” nevet. Kempelen Farkas magyar nemes, polihisztor, az első kísérleti és alkalmazott fonetikus. Úgy véljük, hogy eddigi munkásságunk, eredményeink alapján jogunk van a nagy előd nevének felvételére, s egyben azt is jelezni kívánjuk, hogy – őhöz hasonlóan – az elméleti és az alkalmazott beszédkutatás elkötelezett hívei vagyunk.

\*

A magyar beszédkutatás haladó hagyományait felvállalva, a Kempelen Farkas Beszédkutató Laboratórium munkatársai (névsorban: Gósy Mária, Kelemen Imréné, Menyhárt Krisztina, Olasz György és Nikléczy Péter) a millennium évében létrehozták azt az egyedülálló kiállítást, amely bemutatja az elmúlt több, mint kétszáz év kísérleti eszközeit, berendezéseit, s ezzel azt az utat, amelyet a hazai beszédkutatás napjainkig bejárt. A kiállított tárgyak, műszerek jó része működtethető, így a kiállítás látogatóinak nemcsak a leírások a-



lapján kell elképzelniük, hogy miként dolgoztak elődeink, hanem tapasztalhatják is azt.

Kempelen Farkas a 18. század végén beszélőgépet épített, amely emberi hangképzés nélkül képes volt szavak és rövid mondatok ejtésére. Ezzel megnyitotta az ember–gép kommunikáció lehetőségét, és kijelölte az utat a beszéd sokrétű vizsgálatához. Munkája sokáig követők és folytatás nélkül maradt. Közel egy évszázadot kellett arra várni, hogy a kísérleti fonetika, az eszközös beszéd kutatás meginduljon. A magyarországi kezdetek csaknem egybeestek a nemzetközivel. A kutatás több tudományterületen indult meg, a fonetikán kívül a gégesztetben, a gyógypedagógiában és a fizikában. A 19. század végén, a 20. század elején a fonetikusok rendszerint saját eszközeikkel dolgoztak, ezek jó részének leírását ma már csak tanulmányaik őrzik. Balassa József például a palatográfia és a labiográfia ösét használta kutatásaiban, Gáti József pedig már 1912-ben rezgésképeket közölt a magyar magánhangzókrol. A kísérleti beszéd kutatás első eredményei „tudományos eufóriát” idéztek elő, sokan úgy gondolták, hogy az eszközökkel számos gyakorlati probléma azonnal megoldható, például a beszédhibák javítása. Noha ezek az elvárások nem teljesültek közvetlenül, a tudományág – nemritkán a kutatók hihetetlen fantáziájának következtében – rendkívüli fejlődésnek indult.

A beszéd rögzíthetősége, a mechanikus és elektromos eszközök, a röntgenográfia, a 'látható beszéd', végül pedig a számítógépes lehetőségek az emberi hangképzést egyre jobban megismerhetővé tették. Mindezek alkalmazási eredményeiként megvalósult a számítógépes beszéd szintézis, a mesterséges beszéden alapuló hallásszűrés, a beszédpercepció folyamatok diagnosztikája, a beszéden alapuló beszélőfelismerés kritériumrendszere, számos gyakorlati megoldás a határ tudományok területén (foniátria, audiológia, gerontológia stb.), és lehetőség nyílt a spontán beszéd sokrétű vizsgálatának megkezdésére.

Meggyőződésünk, hogy ez a kiállítás, az elődök munkájának és eredményeinek ismerete hozzásegít bennünket a múlt és a jelen megfelelő értékeléséhez.

*Kempelen Farkas Beszéd kutatató Laboratórium*

**A kiállítás előzetes időpont-egyeztetés után tekinthető meg. Tel.: 321-4830, 172-es mellék. Cím: 1068 Budapest, Benczúr u. 33. IV. emelet.**





