



# FONETIKAI OLVASÓKÖNYV

Egyetemi e-jegyzet

Szerkesztette

BÓNA JUDIT

ELTE FONETIKAI TANSZÉK

Budapest, 2016

Szerkesztőbizottság

GÓSY MÁRIA

IMRE ANGÉLA

MARKÓ ALEXANDRA

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

Technikai szerkesztés: MARKÓ ALEXANDRA és TAKÁCS KAROLINA

Borítóterv: BEKE ANDRÁS

© A szerkesztő és a szerzők

ISBN 978-963-284-668-2

DOI 10.18425/FONOLV.2016

## Tartalom

BÓNA JUDIT: <i>Előszó</i> .....	5
BALÁZS BOGLÁRKA – BÓNA JUDIT: <i>Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban</i> .....	7
GÓSY MÁRIA: <i>Beszédhangok viselkedése a spontán beszédben</i> .....	19
DEME ANDREA – GRÁCZI TEKLA ETELKA – MARKÓ ALEXANDRA: <i>Különleges beszédképzési módok: éneklés, suttogás, gége nélküli beszéd, hasbeszélés</i> .....	33
AUSZMANN ANITA: <i>A semleges magánhangzóról, azaz a sváról</i> .....	49
GRÁCZI TEKLA ETELKA: <i>A zöngelkedési időről</i> .....	61
NEUBERGER TILDA: <i>Nonverbális hangjelenségek a beszédben</i> .....	75
GYARMATHY DOROTTYA: <i>A zaj hatása a beszédre</i> .....	91
HORVÁTH VIKTÓRIA: <i>A hallássérülés hatása a beszédre</i> .....	105
KOHÁRI ANNA: <i>A beszéd ritmusa</i> .....	119
NÉMETH MÁTYÁS FERENC: <i>Milyen tulajdonságokra következtethetünk a beszédből?</i> .....	131
NÉMETH SZILVIA: <i>Az alkohol hatása a beszédre</i> .....	145
BÓNA JUDIT: <i>Temporális sajátosságok a beszédben</i> .....	159
KREPSZ VALÉRIA: <i>Fonetikai hasonlóságok és különbözőségek a beszéd típusokban</i> .....	175
TÓTH ANDREA – SZABÓ ÁGNES: <i>Szociofonetika</i> .....	189
VÁRADI VIOLA: <i>Milyen egységekre tagolható a beszéd?</i> .....	201

MARKÓ ALEXANDRA – BEKE ANDRÁS:	
<i>Beszél(get)ünk a számítógéppel? A beszéd mesterséges előállítása,     számítógépes beszéd- és beszélőfelismerés</i> .....	217
VARJASI GERGELY:	
<i>Beszédatadabázisok</i> .....	233
VAKULA TÍMEA:	
<i>A fonetika története röviden</i> .....	245
A magánhangzók és a mássalhangzók jelölésére szolgáló IPA-jelek.....	259

## Előszó

A fonetika rohamosan fejlődő tudományterület, amely eredményeinek alkalmazása egyre szélesebb körű a mindennapokban is. A témában írt kézikönyvek és tankönyvek jól használhatók az alapismeretek oktatásához, de egyrészt terjedelmi kötöttségeik miatt nem tartalmazznak minden specifikusabb témakört, másrészt a megjelenésük óta újabb és újabb fontos tudományos eredmények születtek és születnek napjainkban is. Szükséges volt tehát egy olyan kötet létrehozása, amely egyrészt kiegészítésként szolgál az egyetemi fonetika kurzusok anyagához, másrészt a laikus érdeklődőkkel is megismerteti a fonetikai kutatások legújabb eredményeit. Az olvasókönyvet elsősorban az elsőéves Fonetika kurzushoz készítettük, de a benne szereplő tanulmányokat úgy írtuk meg, hogy előképzettség nélkül is érthetők legyenek. Bízunk benne, hogy emiatt a nem fonetikusnak készülő hallgatók és más érdeklődők számára is érdekes és jól érthető összefoglaló tanulmányokat tartalmaz.

A tanulmányok a tudományosság követelményeinek eleget téve, közérthető módon fogalmazzák meg az adott témakörrel kapcsolatos ismereteket. A szakszavakat minden tanulmányban megmagyaráztuk az első előfordulásukkor. A tanulmányok elején a témával kapcsolatos legfontosabb kérdések olvashatók, amelyek a mindennapi életben is felmerülő problémák szemszögéből közelítik meg a bemutatott tudományos kérdéseket. A tanulmányok végén ellenőrző kérdések állnak, ezzel megkönnyítve a hallgatók felkészülését a számonkérésre.

A kötetben szereplő tanulmányok a fonetika szinte minden területét érintik. Tárgyalják a beszéd élettani alapjait, a beszédszervek változásait az életkor függvényében, az alternatív beszédképzési módokat, a hallás zavarait; bemutatják a koartikuláció különböző formáit, szólnak a semleges magánhangzóról és a zöngeskedési időről, a szupraszegmentális elemekről, alkalmazott fonetikai kérdésekről, a szegmentálás problémájáról, a fonetika és a beszédtechnológia kapcsolatáról és a fonetika történetéről is.

Bízunk benne, hogy az olvasókönyvet haszonnal forgatják majd az egyetemi kurzusokon, és több hallgató kap kedvet ahhoz, hogy még jobban elmélyüljön a beszéd tudományában.

*A szerkesztő*



## **Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban**

Ha egy ismeretlennel telefonon beszélünk, általában meg tudjuk állapítani, hogy körülbelül milyen idős lehet, vagy legalábbis azt, hogy gyermek, felnőtt, avagy már idősebb személy-e. Amikor személyesen beszélgetünk valakivel, ösztönösen is alkalmazkodunk hozzá, így a korához is: más hanghordozással, tempóval szólítunk meg egy kicsi gyermeket, mint egy fiatal felnőttet vagy idős személyt. Az utóbbi életkori csoportba tartozó beszélgetőpartnerek esetében – különösen akkor, ha nagyon idősek – hangosabban, tagoltabban, lassabban beszélünk.

Vajon milyen biológiai változások állnak annak a hátterében, hogy másképp hangzik a különböző életkorú személyek beszéde? Milyen fonetikai jellemzői vannak a beszédnek az egyes életkorokban? Hogyan „öregszik” a hallószerv? Mikor és milyen változások következnek be a beszédfeldolgozásban?

### **1. A hangképzés élettanáról**

A beszéd megkülönbözteti az embert minden fejlett élőlénytől, hiszen a gondolatainkat szavakba önteni és így kommunikálni csak mi, emberek vagyunk képesek. „A beszéd gondolataink megformálásának, továbbításának, valamint a gondolatok feldolgozásának eszköze” (Gósy 2004: 11). Körfolyamat, amely beszédprodukcióból és beszédfeldolgozásból áll; és egyben egy akusztikai jel, hangrezgés (Gósy 2004).

A kis amplitúdójú rezgéseket a köznapi életben hangnak nevezzük. A hang fogalma fizikai, élettani és lélektani. A fizikai meghatározás szerint a hang valamely rugalmas közeg állapotának a saját egyensúlyi helyzete körüli ingadozása, amely egy rugalmas hordozóközegben tovaterjed. A hangnak ezt a megnyilvánulását hangjelenségnek nevezzük. Az élettani meghatározás szerint a hang az az érzet, amelyet a nyomásingadozás kelt a hallószervben. Az élettani hangmegnyilvánulást hangérzetnek nevezzük. Lélektanilag a hang csak akkor válik számunkra feldolgozhatóvá, ha valamelyik érzékszervünkön keresztül felfogjuk azt.

A hangképzés bonyolult feladatának előfeltétele a megfelelő idegrendszeri irányítás, tehát az, hogy az idegrendszer – több más feladata között – a hangképző apparátus izmait igen pontosan koordinálja.

Az emberi hang keletkezésének alapja:

1. a kilégzés okozta levegőáramlás;
2. a hangszalagok fonációs (a hangadáshoz szükséges) beállítása, tónusa és maga a hangszalagrezgés;
3. a toldalékcső rezonátori tevékenysége.

Az emberi hangot a hangszalagrezgés kelti. A hang minőségét alkati, hormonális és pszichés tényezők befolyásolják. Az emberi hang magasságát a rezgésszám határozza meg, amelyet egyrészt a hangszalagok hossza és feszültsége, másrészt a tüdőből kiáramló levegő áramlási sebessége szab meg. Az emberi hang erőssége függ a hangszalagok rezgési amplitúdójától, amelyet mindenkor a tüdőből kiáramló levegő ereje és a hangszalagok zárási foka, tehát a szubglottikus nyomás (a hangszalag alatti térségben a levegő nyomásának) mértéke szab meg. Az emberi hang színezetét a toldalékcső és a test rezonátor üregei alakítják ki.

## **2. A beszédszervek és az emberi hang változása az életkor függvényében**

A beszédképző szerveink az életünk folyamán mind méretükben, mind a működésüket tekintve (pl. a mozgató izmok feszessége, a porcok elmeszesedése) változnak, ami hatással van az általuk létrehozott beszédhang minőségére is. Gyermekkorától időskorig változik a tüdőkapacitás, a gége és a hangszalagok nagysága, működése, a toldalékcső mérete és rugalmassága, benne a nyelv nagysága és mozgékonyasága is. Hangképző szervünk, a gége továbbá egész életünkben a hormonháztartás befolyása alatt áll (Balázs 1993). Az alábbiakban megvizsgáljuk, hogyan változnak az életkor előrehaladtával a beszédszervek, és ezeknek milyen beszédbeli következményeik vannak. (Megjegyzendő, hogy az a nagyon bonyolult folyamat, amely a beszédlégzésben, a hangszalagrezgésben és a toldalékcső működésében hangadás alatt fennáll, egyrészt koordinált központi idegrendszeri, másrészt perifériás idegrendszeri és reflexes tevékenység következménye. Az életkor hatással van az agy méreteire és felépítésére is, ennek tárgyalása azonban nem tartozik a jelen tanulmány tárgyába.)

A tüdőkapacitás összefüggésben áll a testmérettel, az életkorral és a nemmel; illetve a cserélt levegő mennyiségét a légzéstechnika is befolyásolja. A tüdőkapacitás tehát újszülött kortól kezdve a test növekedésével együtt nő. Felnőtt korra a nők és a férfiak az eltérő testméreteik miatt különböző tüdőkapacitásúak. Általában egy felnőtt ember tüdejének levegőkapacitása 4,5–5,5 liter, ez a férfiaknál 5–5,5, esetleg 6 liter is lehet (Gósy 2004). A két nemet a tüdőkapacitás mellett eltérő légzéstechnika is jellemzi: a nők többsége mellkasi légzést, a férfiak nagy része pedig inkább hasi légzést használ belégzéskor (Gósy 2004).

A légzőrendszer öregedése következtében időskorban a tüdőkapacitás csök-

ken, emellett a szív munkájának romlása és a renyhébben dolgozó hangszalagzáró izomzat miatt a hangerő is csökken, és a hangtartás megrövidül, ami gyakoribb levegővétellel jár együtt (Levitzky 1984; Huber 2008). Bóna (2015) beszéd közben a fiatal felnőtteknél (21–26 évesek) percenként átlagosan nyolcszor adott hangos lélegzést, míg az időseknél szignifikánsan gyakrabban, percenként átlagosan tizenkétszer. (Néma légzés esetén a felnőttek nyugalmi helyzetben 16–19-szer vesznek levegőt; Gósy 2004.) Bóna (2015) eredményei azt is megerősítették, hogy az idősek belégzései a fiatalokénál rövidebb időtartamúak, ami a már fentebb említett csökkent tüdőkapacitásra vezethető vissza.

A megszületett gyermekek gégeje tekintet nélkül arra, hogy fiúk vagy lányok, egyforma nagyságú (Balázs 1993; Welch–Howard 2002; Vorperian–Kent 2007). Gyermekkorban – mintegy tíz éves korig – nemtől függetlenül a hangszalagok hossza kb. 1 cm. A kisgyermek hangmagassága és hangterjedelme lényegében egyforma, körülbelül egy oktávnyi terjedelmű. Ez a „gyermekhang”.

A hormonális érés következtében a pubertásban (hazánkban 14 éves kor körül) a másodlagos nemi jellegek kialakulásával egy időben, a hangban is létrejön a nemek kettéválása, elkülönül egymástól a női és férfi hang. A mutálás során a gége alkotóelemei nem harmonikusan nőnek, ezért átmenetileg mintegy fél évig hangképző szervi ügyetlenség figyelhető meg. Mutálás nemcsak a férfiaknál van, hanem a nőknél is, azonban a hang változása náluk kisebb fokú. A férfiaknál a gége a születés kori méret ötszörösét éri el, a nőknél a gégeporcok körülbelül a kétszeresükre nőnek. A felnőtt nők hangszalagjai mintegy 1,5–2 cm hosszúak, a felnőtt férfiaké 1,9–2,9 cm közöttiek, és a nők hangszalagjainál vastagabbak. A gége mellső és oldalsó falát alkotó pajzsporc megvastagodhat, kívülről is jól láthatóvá válhat főként a férfiaknál (ez az ádámcsutka) (Gósy 2004).

A hangszalagok eltérő hosszúságának következtében különböző alaphangmagasság jellemzi a nőket és a férfiakat (Gósy 2004). A férfiak hangja általában egy oktávval mélyebb, mint a nőké (Balázs 1993). Az úgynevezett „felnőtthang” a pubertástól a hatvanadik életévig hallható (Frint–Surján 1982) A két nem hangszínezete, átlagos beszédhangmagassága és a hangok karaktere erősen különbözik egymástól, a hangterjedelem mindkét nemnél két oktávra bővül (Balázs 1993).

A hatvanadik életév után évről évre fokozatosan csökken a két nem hangminőségének távolsága. A hangszalagok rugalmatlanabbá válnak, a gégeizomzat leépül, a porcok meszesedése fokozottabb lesz, a hangképző izmok tónusa csökken a nemi hormonok hypofunkciója (‘alulműködése’) miatt, ami a hang gyengüléséhez, szaggatottságához vezet (Balázs 1993). A férfiak és nők hangfekvése és átlagos beszédhangmagassága közeledik egymáshoz. Az alaphangmagasság a férfiaknál általában 1–1,5 hanggal emelkedik, a nőknél 0,5 hanggal mélyebb lesz, vagy nem változik (Frint 1982; Balázs 1993; Russel et al. 1995; Nishio–Niimi 2008). A hangterjedelem beszűkül, általában ismét a

gyermekhangra jellemző egy oktáv lesz (ez alól vannak kivételek, például a profi énekeseknél kisebb mértékű a hangterjedelem szűkülése). Ez a beszűkülés a férfiaknál a mélyebb hangokat érinti (a hangfekvéstől függően az F, G, A, H hangokat), a nőknél a hangterjedelem felsőbb részére esik (a c", d", e" hangokra) (Balázs 1993). A zöngé minősége is veszít a korábbi tisztaságából, zöreijesebb lesz (Bóna 2009). Az ily módon megváltozott beszédhangot „öreghang”-nak nevezzük (Balázs 1993).

Az életkor tehát nagymértékben befolyásolja az alaphangmagasságot és a hangterjedelmet. Az 1. táblázatban néhány adatot közlünk különböző életkorú magyar anyanyelvű beszélők átlagos alaphangmagasságáról, hogy bemutassuk, milyen mértékű az életkor hatása a zöngéképzésre. (A bemutatott adatok keresztmetszeti vizsgálatokból származnak.) Megjegyzendő, hogy a más kultúrákhoz tartozó emberek átlagos alaphangmagassága eltérő lehet. Ennek különböző, az alkati különbségek mellett egyéb, például kulturális tényezők állnak a háttérben (Gósy 2004).

Fontos megemlítenünk a dohányzás ártalmait, amelyek az évek során egyre nagyobb változást okoznak a hangminőségben. A dohányzás miatt az évek alatt kialakuló kisebb-nagyobb megvastagodás – amely még nem tekinthető betegségnek, de a hangszalagok rezgő tömegét növeli – a hang mélyüléséhez vezet (Balázs 1993).

1. táblázat: Különböző életkorú magyar anyanyelvű beszélők átlagos alaphangmagassága

Életkor	Átlagos alaphangmagasság (Hz)	
	Férfiak	Nők
5 évesek (Tóth 2014)	261	262
6–7 évesek (Deme 2012)	288	272
7 évesek (Beke–Horváth 2015)	241	236
8 évesek (Beke–Horváth 2015)	232	235
9 évesek (Beke–Horváth 2015)	224	239
10 évesek (Tóth 2014)	255	246
13 évesek (Auszmann–Neuberger 2014)	146	198
21–60 éves felnőttek (Markó 2015)	114	192
70–85 évesek (Bóna 2013a)	119	151
90–105 évesek (2 férfi, 1 nő; Bóna 2013a)	147	166

A toldalékcsőben is számos változás megy végbe az életkor előrehaladtával. Születéskor a toldalékcső hossza kb. 6-8 cm; a felnőtt nőknél mintegy 14,1 cm, de a 15 cm-t is elérheti, míg a férfiaknál 16,9 cm, vagy akár 18 cm is lehet (Gósy 2004; Vorperian–Kent 2007). A hangszalagoktól az összezárt fogakig tartó zárt toldalékcső térfogata a felnőtt nőknél átlagosan kb. 130 cm<sup>3</sup>, a férfiaknál 170 cm<sup>3</sup>.

A garatüreg hossza a nőknél 6,3 cm, a férfiaknál kb. 8,9 cm; a szájüreg a nőknél kb. 7,8 cm, a férfiaknál 8,1 cm hosszú. Különbség van a nemek között a nyelv térfogatában is: a nőknél 90 cm<sup>3</sup> a becsült értéke, a férfiaknál 110 cm<sup>3</sup> (Gósy 2004). Időskorban nehezítetté válik a nyelvmozgás, ami az artikuláció pontatlanabbá válásával jár együtt; míg a lány szájpad mozgásának renyhülése orrhangzós színezetet okozhat a beszédben (Balázs 1993).

A fogazat részleges vagy teljes elvesztése is hallható változásokat okozhat a beszédben (Götz 2013). Természetesen ez bármely életkorban előfordulhat, mégis két olyan időszak van az emberi életben, amikor nagyobb valószínűséggel fordulhat elő a fogazattal kapcsolatos beszédzavar: fogváltás alatt és után, illetve az idős korban. Az általában 6–7 éves korban kezdődő fogváltás során a tejfogak kiesése és a maradandó fogak kibújása közötti időben a fogazat részleges hiánya átmenetileg megváltoztathatja a beszéd artikulációs (és akusztikai) sajátosságait. A maradandó fogazat kialakulása után sokan fogszabályzót kapnak, amely szintén hatással van a labiális, labiodentális vagy az alveoláris, posztalveoláris hangok képzésére. Az esetek döntő többségében a fogszabályozó készülék által okozott artikulációs zavar átmeneti, mindössze egy–három hétig tart (Götz 2013). Idős korban a fogak részleges vagy teljes elvesztése okozhat artikulációs zavarokat. Ezek korrigálhatók fogpótlással, ám a különböző fogprotézisek megszokása is időigényes, és átmenetileg beszédzavart okoz (Götz 2012).

### 3. Akusztikai fonetikai változások a beszédben

Mindennek természetesen következményei vannak a beszéd akusztikai szerkezetében is. Már évtizedekkel ezelőtt, a komplex hangelemző szoftverek megjelenése előtt, hangszínkép segítségével is megállapítható volt az, hogy idős korban változások következnek be a beszéd akusztikai szerkezetében. Megállapították például, hogy a magánhangzók formánsterületén zörejelemek láthatók. Minél idősebb az ember, annál jellemzőbb az elől képzett magánhangzók, az *e* és az *i* harmonikus elemeinek felbomlása és a zörejelemek megjelenése (Balázs 1993). Az újabb kutatások igazolták, hogy az életkor függvényében változások mutatatók ki például a beszéd időviszonyaiban, avagy a magánhangzók formánsértékeiben is (például Deme 2012; Bóna 2013a; Auszmann 2015). A gyermekek ejtése az egyre több beszédrutin megszerzésével, az artikulációs szervek ügyesedésével fokozatosan kezd hasonlítani a felnőtt nyelvi kiejtésre (például Deme 2012; Auszmann 2015); míg 60 éves kor felett az elfáradó beszédszervek és a vezérlés okozzák a magánhangzók zártabb ejtését, a lassuló artikulációs tempót stb. (Bóna 2013a).

A 2. táblázatban összefoglaltuk az egyes életkori csoportok beszédének főbb fonetikai jellemzőit. Megjegyzendő azonban, hogy a három fő életkori csoport

mellett szokás szűkebb életkori határok között is vizsgálni a beszéd változását. Más-más beszédjellemzői vannak ugyanis az anyanyelvét éppen elsajátító gyermeknek (sőt a csecsemőkori hangadást, sírást is vizsgálták, vizsgálják, vö. Benyó et al. 2002); az óvodásoknak; az írni-olvasni tanuló kisiskolásoknak, a kamaszoknak; avagy a fiatal-időseknek és a nagyon időseknek (például Laczkó 2009; Deme 2012; Bóna 2012; Neuberger 2014; Beke–Horváth 2015).

2. táblázat: A gyermekek, a felnőttek és az idősek beszédének főbb fonetikai jellemzői

	<b>Gyermekkor</b>	<b>Felnőttkor</b>	<b>Időskor</b>
<b>Zöngképzés</b>	a hangterjedelem körülbelül egy oktáv: c'-től c"-ig; nincs különbség a lányok és a fiúk között	a hangterjedelem két oktáv; elkülönül egymástól a férfihang és a női hang; kialakulnak a hangosztályok	a hangterjedelem ismét egy oktáv; a női és férfihang differenciáltsága egyre kevésbé kifejezett; a nők természetes beszédhangfekvése mélyül, vagy nem változik; a férfiaké magasodik
<b>Beszédsebesség</b>	fokozatosan nő, köszönhetően az artikulációs mozgások ügyesedésének, a szókincs bővülésének, a beszédben való egyre nagyobb gyakorlottságnak	az átlagos köznyelvi artikulációs tempó 12,5–14 hang/s; a beszéidő kb. 20–30%-a szünet	az artikulációs tempó csökken; gyakoribb a lélegzetvétel miatti szünettartás
<b>Artikulációs pontosság</b>	egyre pontosabbá válik	az adott nyelvre jellemző artikuláció (kivéve beszédhiba vagy nem anyanyelvi beszélő esetén)	pontatlanná váló artikuláció

#### 4. A hallás szerepe a beszédben

Az állatvilág szinte minden lényének jellemző tulajdonsága, hogy a külvilágban uralkodó viszonyokról információkat képes felvenni, ezeket specifikus szöveteiben fel tudja dolgozni, a különböző jellegű hatásokat egymással kapcsolatba hozni és megfelelő, az egyén létfenntartása szempontjából célszerű intézkedéseket – például menekülés, támadás vagy a táplálék felkeresése –

kidolgozni képes. E biológiai feladatra specifikusan differenciált szervrendszer szolgál: az idegrendszer és az érzékszervek.

A beszédfeldolgozásnak több szintje van: a hallás, az észlelés (a hangok, hangkapcsolatok azonosítása), a megértés és az értelmezés (Gósy 2005). Az ép hallás elengedhetetlen feltétele a felsőbb szintek megfelelő működésének, ugyanakkor nem elégséges feltétele annak, hogy pontosan működjön az észlelés, a megértés és az értelmezés. (Bár ez utóbbi két feldolgozási szint nem része a fonetika tárgyának, alább röviden megemlítyük ezek életkori változásait is.)

A hallást, valamint a fül felépítését és működését az élet „mindennapos csodájának” is nevezhetjük. Hallásnál a hanghullámok a levegőben longitudinális hullámokként sűrűsödve és ritkulva terjednek. A hallócsontok a hullámokat transzverzálissá alakítják. A mozgást a belfül alakítja át kódolt bioárammá. Az itt lévő szörsejtek ingerületét a hallóidegrostok vezetik el.

A hallás környezetünk akusztikai ingereit dolgozza fel. Ennek megfelelően azt a folyamatot nevezzük hallásnak, amelyben hallószervünk a hangjelenségeket érzékeli, felerősíti, irányukat meghatározza és továbbítja a megfelelő agyi központba, ahol a végső feldolgozás megtörténik (Gósy 2004).

#### **4.1. A hallószerv**

A hallószerv két fő részből, a centrális és a perifériás részből áll. A perifériás rész maga a fül, amelyet további három fő részre oszthatunk, ezek a következők: külső, középső- és belfül. Ez az akusztikai ingerek felvételére alkalmas készülék a halántékcsontról a sziklacsontra helyezkedik el.

A külső hallójárat szerepe az, hogy a hangrezgéseket elvezesse a rezgést végző dobhártyáig. A dobhártya határolja a dobüreget, amelyben a hangtranszformációt végző hallócsontok vannak. A dobüreg belfül falán lévő ovális ablak a külső és belfül határa, a kengyeltalp ide juttatja el a légrezgéseket, amelyek ezután folyadékrezgésekké alakulnak, és a csiga membránján hozzák létre a hallható hangokat örvényáramok keletkeztetésével. A hallópálya a csigától a primér hallási kéregmezőig terjed.

#### **4.2. A hallható hallástartomány**

Az audiológiában hangerőt és hangmagasságot mérünk. A hangerőt, azaz a hangintenzitást decibelben adjuk meg. A decibel viszonyszám. Az 1 dB nagyon kicsi érték; a 30 dB a halk emberi beszéd; az 50–60 dB hangos emberi beszéd; a 80 dB hangos zaj, tartós fennállása halláskárosító; a 120 dB a diszkó és a légalapács zaja; a 140 dB a repülőgép motor zaja közlelről. A hangmagasságot a másodpercenkénti rezgésszám adja meg, Hz-ben mérjük. A hallástartomány fiatal korban 16-tól 20 000 Hz-ig tart. Az ép emberi fül 1000 Hz-en mért

hallásküszöbe:  $0,00002 \text{ Pa} = 0,002 \text{ din/négyzetcentiméter}$ .

A hangot akkor halljuk meg, ha erőssége 1,5–2,5 dB-lel meghaladja a környezeti zaj szintjét. Nagyobb intenzitású vagy hosszabb ideig tartó hang hatására a fül „kifárad”. Ha a terhelő hang hosszabb ideig tart, akkor a hangot fokozatosan halkabbnak halljuk. A hang kellemessége vagy kellemetlensége a megszokástól függ.

## 5. Életkori sajátosságok a beszédfeldolgozásban

Az ember hallásállapota az élet folyamán változik. Már a magzat is hall, és reagál az erős hangingerekre az anyaméhben. Születéskor az újszülöttek hallása még nem tökéletes, a hallásküszöb 75–80 dB, szemben a 0 dB-es tökéletes hallással. A születés utáni első két hónapban a csecsemő a hangos zajoktól megijed. A harmadik hónaptól a hang forrását keresi a szemével, mintegy hat hónaposan a fejét a hangforrás irányába fordítja (Fent 2007). Már az újszülöttek is képesek megkülönböztetni az emberi hangot a másfajta hangoktól, és felismerik az emberi beszéd számos sajátosságát is. A gögicsélés, amelynek az egyik oka és fenntartója az akusztikai élmény (vö. Gósy 2005), a 3. és a 6. hét között kezdődik; féléves korra a csecsemő már sokat és hosszan gögicsél. Ebben az életkorban már kimutatható a gögicsélésben bizonyos anyanyelvi sajátosságok jelentkezése, amely a nyelvi háttérű beszédészlelés működésére enged következtetni (Gósy 2005). A következő hónapokban már reagál az ismerős zajokra; tíz hónapos korában megért bizonyos ismerős tárgyakat jelölő szavakat (Fent 2007). A gyermekek észlelése és beszédmegértése fokozatosan finomodik, a kezdeti (egyéves kor körüli) globális beszédmegértés (amikor a megértést elsősorban a beszédhelyzet, az emberi gesztusok és mimika biztosítják) fokozatosan alakul át a beszédhangokat pontosan megkülönböztetni tudó, felnőttekre is jellemző beszédfeldolgozásra (ez a hierarchikus beszédfeldolgozás hároméves kortól jelen van) (Gósy 2005). Ez egy relatíve hosszú folyamat, a hosszú-rövid hangok megkülönböztetése például csak hétéves korra válik stabilá, míg bizonyos pragmatikai sajátosságok (pl. az irónia) észlelése és megértése még későbbi életkorokhoz köthető (Gósy 2005).

Időskorban, 65 éves kor után a hallás romlik, főleg a magas hangok területén észlelhető halláscsökkenés, ez a *presbyacsis* (Frint–Surján 1982; Gósy 2004). Időskorban emellett a beszédészlelés és a beszédmegértés is romlik. Ezek hátterében a halláscsökkenés mellett az emlékezeti működések és a figyelmi teljesítmény romlása is fellelhető. Magyar anyanyelvű fiatalok és idősek (68 év feletti-ek) mondatisméltéseinek vizsgálatakor azt találták, hogy az idősek szignifikánsan kevesebb mondatot tudtak helyesen visszamondani, mint a fiatalok, és a reakcióidejük is szignifikánsan nőtt (Neuberger 2010). Egy szövegméltékezeti

vizsgálatban pedig, amelyben a résztvevőknek hallás után kellett visszamondaniuk az elhangzottakat, azt találták, hogy nemcsak a valódi idősek, hanem már az idősödők (60–74 évesek; az idézett vizsgálatban átlagosan 67 évesek) szöveg- emlékezeti teljesítménye is szignifikánsan csökkent a fiatalokéhoz képest (Bóna 2013b).

Életünk során folyamatosan szerzünk információkat. Kiskorunkban hallással sajátítjuk el ismereteink egy részét, így magát a beszédet, az elvont fogalmak megértését, saját beszédünk kialakítását és a társainkhoz fűződő viszonyainkat is a beszéd útján alakítjuk ki. Az első éveinkben az anyanyelvet szinte játszva sajátítjuk el. Az idősödő korig a kapcsolattartás elsősorban a hallással, a beszéd megértésével történik. Ezzel általában nincs is probléma, a fiatalok jól megbirkóznak a zajos környezetben való gyors és nem egészen tiszta beszéddel is. A nem jól halló gyermek vagy felnőtt azonban nem reagál megfelelően a megszólításokra, nem tud megfelelően részt venni a társalgásban, így az életminősége is romlik. Nagyothallásról akkor beszélünk, ha valakinek a hallása az életkorában elvárt teljesítménynél rosszabb. Nagyothallás esetén már fiatal korban is nehézséget okoz a zajban való beszédmegértés.

A gyermekkori hallásproblémák kiszűrése nem egyszerű feladat, de vannak jelek, amelyekre a szülőnek is fel kell figyelnie. Halláscsökkenést jelezhet, ha az első három hónapban a csecsemő nem reagál a zajokra, az ágy felé közeledő léptekre; nem veszi észre az anya hangját, mielőtt megpillantaná őt. Figyelmeztető jel, ha 4–6 hónapos korában nem fordítja fejét a hang irányába, nem reagál a számára ismerős hangokra: ajtónyikorgás, telefoncsengés, nevének szólítása; nem igyekszik utánozni a mások által előidézett hangokat. További jelzések: 7–12 hónapos korban nem érti meg azokat a szavakat, amelyek jelentését már ismernie kellene: személyek, ételek, játékok; nem engedelmeskedik, ha rászólnak; 1–2 éves kora körül nem gagyog; nem igyekszik primitív módon a környezetében hallott hangokat, szavakat ismételni vagy utánozni; 2–3 éves kora körül nem tapasztalható értelmi képességének fejlődése, nem beszél tagoltan és érthetően, nem tesz eleget, ha összetett feladatokat kellene elvégeznie, esetleg még gügyög vagy érthetetlen szavakat mond. Az átlagosan egészséges leánygyermek 2 éves korára tanul meg beszélni; a fiúgyermek 2,5 éves korára, amelyhez elengedhetetlen az ép hallás. Ahogyan korábban írtuk, az ép hallás azonban nem elégséges feltétele annak, hogy a beszédfeldolgozás hibátlanul működjön. Elmaradás vagy zavar keletkezhet a beszédészlelés és a beszédmegértés szintjén az életkornak megfelelő hallás esetén is (Gósy szerk. 2007).

A hallássérülés típusairól és a nagyothallás beszédre gyakorolt hatásáról ugyanezen kötetben Horváth Viktória (2016) ír részletesen.

## 6. Kitekintés

A tanulmányban körbejártuk, hogy hogyan változik a beszédképzés és beszédfeldolgozás biológiai és fonetikai, pszicholingvisztikai szempontból. Számos vizsgálat szükséges azonban még ahhoz, hogy pontosan megismerjük a beszéd változásának életkori jellemzőit. A gyermekek beszédének és beszédfeldolgozásának mind pontosabb megismerése elengedhetetlen ahhoz, hogy szükség esetén megfelelő fejlesztési terveket dolgozhassunk ki. Az ép, életkornak megfelelő beszédprodukciónak és beszédfeldolgozásnak fontos feltétele az írás-olvasás megtanulása, így az iskolába lépésnek is. Annak érdekében, hogy minél alaposabban megismerhessük a beszéd sajátosságait az óvodásoknál és az iskolásoknál (kiscsoportos óvodás kortól érettségig), 2013-ban megindult az ELTE Fonetikai Tanszékén egy nagyszabású beszédatadatbázis, a GABI (Gyermeknyelvi beszédatadatbázis és információtár) létrehozása. Emellett több más specifikusabb gyermekbeszéd-adatbázis is készült az elmúlt pár évben a magyar anyanyelvű gyermekek beszédfejlődésének vizsgálatához. Ezek azért különösen fontosak, mert korábban – megfelelő eszközök és adatbázisok hiányában – főként szülői jegyzetek alapján, avagy kisszámú gyermek vizsgálata révén volt lehetséges a magyar gyermekek beszédfejlődésének leírása.

A kutatók felismerték annak a jelentőségét is, hogy a felnőtt beszéd jellemzőinek vizsgálata nem korlátozódhat pusztán a fiatal felnőttek csoportjára. Egyre több olyan vizsgálat születik, amely a felnőttkoron belül további kategóriákat különít el, és külön elemzi az idősek beszédét is. Ez azért fontos, mert az öregedés hatásainak pontosabb ismerete segít abban, hogy érzékenyebbé tegyük a társadalmat az időskori kommunikációs problémák iránt, és megelőzhessük az idősek életminőségének (a beszédhez kapcsolódó) romlását.

## Irodalom

- Auszmann Anita 2015. A magánhangzók akusztikai szerkezete 9 és 11 éves iskolások spontán beszédében. *Beszédkutatás* 2015. 164–175.
- Auszmann, Anita – Neuberger, Tilda 2014. Age- and gender-related differences in formant structure during the stabilization process of vowels. In: *Proceedings of the Olomouc Linguistics Colloquium 2014*. 663–676.
- Balázs Boglárka – Gósy Mária 1988. Környezetünk hangjelenségeinek hatása a beszédre. *Fül-orr-gégegyógyászat* 34. 145–150.
- Balázs Boglárka 1993. Az időskori hangképzés jellemzői. *Beszédkutatás '93*. 156–165.
- Beke András – Horváth Viktória 2015. Kisiskolások alaphangmagasságának variabilitása. *Beszédkutatás* 2015. 133–147.

- Benyó Zoltán – Farkas Zsolt – Illényi András – Katona Gábor – Várallyai György 2002. Csecsemők hangelemzése különös tekintettel a hallásképességükre. *Beszédkutatás* 2002. 100–117.
- Bóna Judit 2009. Az idős életkor tükröződése a magánhangzók ejtésében. *Beszédkutatás* 2009. 76–87.
- Bóna Judit 2012. A spontán beszéd sajátosságai idősödő, idős és matuzsálemi korban. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngelkedési időig*. ELTE BTK – MTA Nyelvtudományi Intézet, Budapest, 100–115.
- Bóna Judit 2013a. *A spontán beszéd sajátosságai az időskorban*. Beszéd – Kutatás – Alkalmazás 2. ELTE Eötvös Kiadó, Budapest.
- Bóna Judit 2013b. Narrative recall in the elderly: Content, fluency and speech errors in the narrative speech of young, young-old and old-old speakers. *Acta Linguistica Hungarica* 60/2. 123–142.
- Bóna Judit 2015. Nonverbális hangjelenségek fiatalok és idősek spontán beszédében. *Beszédkutatás* 2015. 106–119.
- Deme Andrea 2012. Óvodások magánhangzóinak akusztikai jellemzői. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngelkedési időig*. ELTE BTK – MTA Nyelvtudományi Intézet, Budapest, 77–99.
- Fent Zoltán 2007. A hallószerv, a hallás folyamata, zavarok. In Gósy Mária (szerk.): *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 44–57.
- Frint Tibor – Surján László 1982. *A hangképzés és zavarai. Beszédzavarok*. Medicina Könyvkiadó, Budapest.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária szerk. 2007. *Beszédészlelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest.
- Götz Gergely 2012. A fogazat és a fogmedernyúlvány eltérésének hatása a hangképzésre. In Navracsecs Judit – Szabó Dániel (szerk.): *A mentális folyamatok a nyelvi feldolgozásban. Pszicholingvisztikai tanulmányok III*. Tinta Könyvkiadó, Budapest, 208–217.
- Götz Gergely 2013. A fogszabályozó készülékek hatása a hangképzésre. *Beszédkutatás* 2013. 271–277.
- Horváth Viktória 2016. A hallássérülés hatása a beszédre. In Bóna Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest, 105–118. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu)
- Huber, Jessica E. 2008. Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory physiology & neurobiology* 164/3. 323–330.
- Laczkó Mária 2009. Tizenévesek beszédének fonetikai és stilisztikai elemzése. *Anyanyelv-pedagógia* 2009/1. <http://www.anyp.hu/cikkek.php?id=151> (A letöltés ideje: 2015. május 21.)
- Levitzky, Michael G. 1984. Effects of aging on the respiratory system. *Physiologist* 27/2. 102–107.
- Markó Alexandra 2015. *A spontán beszéd prozódiai szerkezete. Időzítés és beszéddallam*. Nyelvtudományi Értekezések 166. sz. Akadémiai Kiadó, Budapest.
- Neuberger Tilda 2010. Mondatészlelési sajátosságok fiatal és idős korban. In Gecső Tamás – Sárdi Csilla (szerk.): *Új módszerek az alkalmazott nyelvészeti kutatásban*. Tinta Könyvkiadó, Budapest, 220–225.

- Neuberger Tilda 2014. *A spontán beszéd sajátosságai gyermekkorban*. Beszéd – Kutatás – Alkalmazás 4. ELTE Eötvös Kiadó, Budapest.
- Nishio, Masaki – Niimi, Seiji 2008. Changes in speaking fundamental frequency characteristics with aging. *Folia Phoniatica et Logopaedica* 60. 120–127.
- Russel, Alison – Penny, Lynda – Pemberton, Cecilia 1995. Speaking fundamental frequency changes over time in women: A longitudinal study. *Journal of Speech and Hearing Research* 38. 101–109.
- Tóth Andrea 2014. Gyermekek nemének és életkorának meghatározása a beszédük alapján. *Beszéd kutatás 2014*. 98–111.
- Vorperian, Houri K. – Kent, Ray D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50. 1510–1545.
- Welch, Graham F. – Howard, David M. 2002. Gendered voice in the cathedral choir. *Psychology of Music* 30. 102–120.

### **Kérdések, feladatok**

1. Hogyan változnak a beszédszerveink (tüdő, gége, toldalékcső) az életkor előrehaladtával?
2. Milyen fonetikai jellemzőkben tér el az idősök beszédprodukciója a fiatalokétól?
3. Melyek a legfontosabb figyelmeztető jelek, amelyek a gyermekkori halláscsökkenést jelzik?

## Beszédhangok viselkedése a spontán beszédben

Mindenki tapasztalhatta már, hogy ha olyan idegen nyelvű beszédet hallgat, amit egyáltalán nem ért, akkor az a benyomása, mintha nem is lennének határok a beszédfolyamban, legfeljebb akkor, ha a beszélő szünetet tart. Lehetetlen megállapítani, hogy hol kezdődik, vagy hol végződik egy szó, meddig halljuk ugyanazt a beszédhangot, mikor már egy másikat. Mi okozhatja ezt a benyomást? Miért nem fordul ez elő (vagy csak nagyon ritkán), ha a közlések az anyanyelvünkön hangzanak el?

### 1. Bevezetés

A fenti kérdésekre az a legkézenfekvőbb válasz, hogy azért, mert anyanyelvünk szavait, szókapcsolatait ismerjük, így nem nehéz őket „kihallani” a szövegből. Sőt, a szavakat le tudjuk bontani szótagokra, illetve beszédhangokra is. A válasz helyes, azonban az továbbra is kérdés, hogy mi eredményezi a beszéd folyamatosságát, illetve hogy miként kapcsolódnak egymáshoz a beszédhangok és a szavak.

Induljunk ki abból, hogy a beszédet összehasonlítjuk az írással. Íráskor – akár kézírásra, akár gépelt szövegre gondolunk – a szavakat jól láthatóan elkülönítjük egymástól. Ezáltal nagymértékben megkönnyítjük a szavak felismerését olvasáskor. Nézzük meg, hogy mennyivel nehezebb lenne az olvasás, még a gyakorlott olvasónak is, a szóközök nélkül:

*Elbuktákakötvényeknagyresztmertnemfigyeltekamegfelelőzáradékpontosmegértésére.*

Írásunkban a betűk követik egymást, és a betűsorozatok alkotják az adott szót. A nyomtatott betűk között nincs összeköttetés, a kézírásban általában összekötjük a szomszédos betűket. A beszédben azonban nincsenek szóközök, nem tartunk szünetet minden egyes szó ejtése után. A szavak kiejtése folyamatos, noha egyes prozódiai eszközök (pl. a hangsúly, a dallam, illetve maga a szünet) segíthetik a hallgatót a szavak határainak felismerésében. Az különösen megnehezíti az értelmezést, ha többféle módon is tagolható a közlés, azaz más-más helyen van a szavak eleje és a vége. Például: *tegnapótaittál* (’tegnap óta itt áll’, illetve ’tegnap óta ittál’) vagy *ahatalmasokévekótaatengeren* (’a hatalmasok évek óta a tengeren’, avagy ’a hatalmasok évek óta a tengeren’, sőt ’a

hatalmas ok évek óta a tengeren’).

Vajon mi történik a beszéd során? Hogyan következnek egymás után a beszédhangok, hogyan alakulnak ki a kiejtés során a beszédhangsorozatok, a szókapcsolatok? Van-e valamiféle összekötő szakasz a beszédhangok között, avagy csak gyors egymástutánban képezzük őket, amikor kiejtjük a szavakat? Mindig ugyanúgy ejtjük az egyes beszédhangokat, függetlenül attól, hogy milyen hangkörnyezetben fordulnak elő?

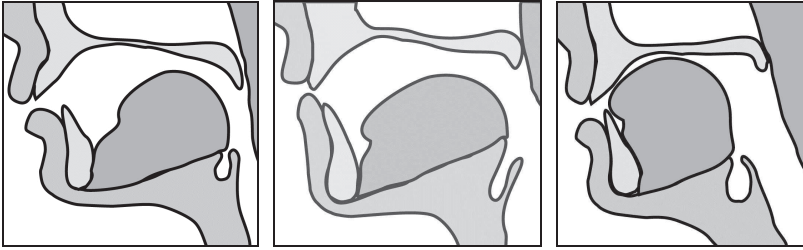
## 2. Beszédhangok és beszédhangok sorozata

A szavakat alkotó, egymást követő beszédhangok nem függetlenek egymástól. Ez azt jelenti, hogy a szomszédos beszédhangok hatással vannak egymás képzésére. Ezt a jelenséget koartikulációnak („együttes hangképzésnek”) nevezzük. A magánhangzókat és a mássalhangzókat másként ejtjük önmagukban (izoláltan), azaz hangkörnyezet nélkül és másképpen hangsorban, vagyis ha megelőzi, avagy követi őket egy másik beszédhang. Az életben számos esetben megfigyelhető, hogy két állapot között szükségszerűen átmenetek vannak (pl. a gyermekkor és a felnőttkor között, a nyár és a tél között, avagy az ülés és a felállás között). Az emberi mozgások folyamata tehát két állapot között bizonyos átmeneteket tartalmaz (egyelőre eltekintve attól, hogy milyen mozgásokról és milyen átmenetéről van is szó), ezeket az 1. ábrán a szürke színárnyalataival szemléltetjük.



1. ábra. Egy-egy állapot és a köztük lévő átmenetek szemléltetése színárnyalatokkal

A beszédhangok, hangkapcsolatok, szavak kiejtése mozgások sorozatának az eredménye. A kiejtés a beszédképző szerveink folyamatos mozgása. Ezeket a mozgásokat az ajkak, a nyelv, a légyszájpadlás hajtják végre a beszéléskor. A folyamatos beszéd során az egyes beszédhangok képzését felfoghatjuk egy-egy állapotnak, ez a rájuk jellemző képzési vagy **artikulációs konfiguráció** (célkonfiguráció). Más szakkifejezéssel ez a tiszta fázis. Ez azt jelenti, hogy az artikulációs szervek az adott beszédhangnak megfelelő képzési együttest mutatják (2. ábra). Az ábra az *o*, a *v* és az *i* beszédhangokra jellemző artikulációs helyzetet szemlélteti, amikor ezeket a magánhangzókat és a mássalhangzót önmagukban, hangkörnyezet nélkül ejtjük ki.

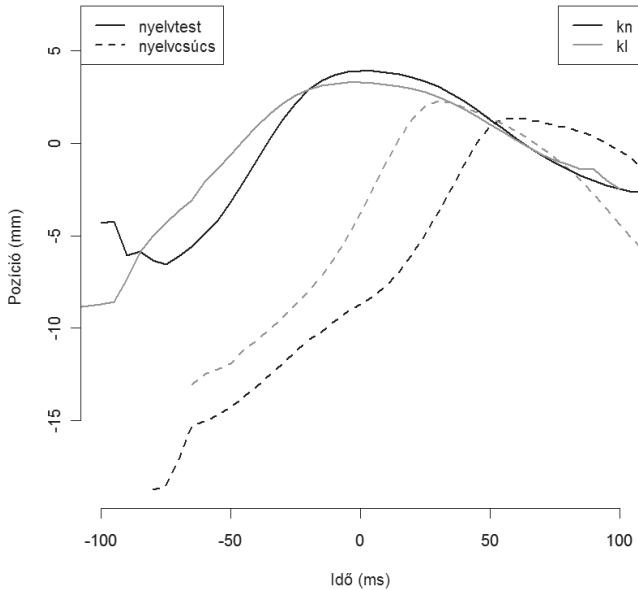


2. ábra. Az *o*, a *v* és az *i* beszédhangok jellemző, „ideális” képzési állapota (artikulációs konfigurációja) hangkörnyezet nélkül kiejtve (röntgenfelvételekről készült sémák alapján)

Látható, hogy a nyelv a szájüreg különböző részében van az egyes beszédhangok esetében, hasonlóképpen eltérő a nyelv és a fogmeder helyzete, valamint a két ajak távolsága. Ha azonban nem önmagukban ejtjük ezeket a beszédhangokat, hanem például az *ovi* szót mondjuk ki, akkor a szomszédos beszédhangok között szükségszerűen átmeneteket fogunk ejteni. Ezeket **artikulációs átmeneteknek** nevezzük. Ezekben az átmeneti hangszakaszokban módosul, változik a részt vevő hangképző szervek helyzete. Az átmenetek létrejöttének az az oka, hogy a beszédszerveknek időre van szükségük ahhoz, hogy az egyik beszédhangra jellemző ejtési állapotból eljussanak egy következő hang ejtési állapotába. Hangképző szerveink relatíve rugalmatlanok, nem képesek „átugrani” egyik helyzetből a másikba. A beszédszerveknek az egyik helyzetből a másik helyzetbe történő kerülése – a beszédszervek relatív tehetetlensége következtében – nem hirtelen következik be, hanem folyamatosan, rövidebb-hosszabb idő alatt megy végbe.

Ha a *v* és az *i* hang közötti átmenetre gondolunk, akkor elmondható, hogy a nyelv előre mozdul a szájüregben és közelít a fogmederhez, az ajkak távolodnak egymástól, módosul az állkapocs nyitásszöge is. Tekintettel arra, hogy egyik beszédhang sem orrhangú, ezért az ínyvitorla (uvula) helyzete változatlan marad.

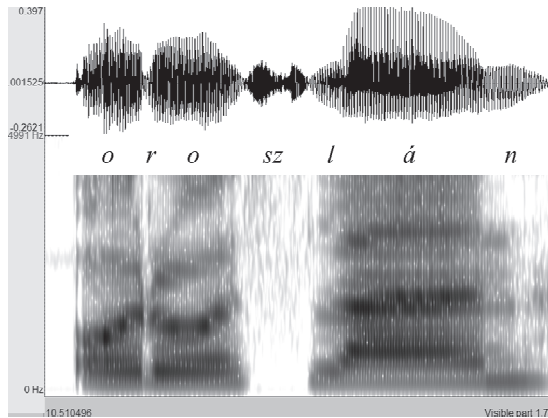
A nyelvtest és a nyelvcsúcs mozgását szemlélteti az idő függvényében a *kn* és a *kl* mássalhangzó-kapcsolatok ejtésekor a 3. ábra. Gondoljunk például a *hakni* és a *cékla* szavakra, amelyekben ezek előfordulnak. Ejtsük ki őket, és próbáljuk megfigyelni a nyelvünk mozgását. Az ábra rajza tükrözi ezeket a mozgássorokat (amelyeket egy speciális szoftverrendszer, az EMU-R felhasználásával rajzoltak meg): a folyamatos vonal a nyelvtest, a szaggatott pedig a nyelvcsúcs mozgását tükrözi az ábrán a *k* mássalhangzó ejtésétől a *n*, illetve a *l* mássalhangzó ejtéséig. A vízszintes tengelyen az idő jelenik meg, a függőleges tengelyen pedig a nyelv helyzete.



3. ábra. A nyelvtest és a nyelvcsúcs mozgásának szemléltetése a *kl* és a *kn* mássalhangzó-kapcsolatok ejtésekor az idő függvényében (Harrington 2010: 139)

### 3. Mik az artikulációs átmenetek jellemzői?

Láttuk, hogy a kiejtés folyamatosságához nélkülözhetetlenek az artikulációs átmenetek. Az artikulációs átmenetek a beszédhangok tiszta ejtési idejéhez képest relatíve rövidek. A kiejtés következtében létrejövő rezgés vizuálisan (képileg) megjeleníthető, ez az akusztikai lenyomat vagy az ún. 'látható beszéd'. A beszéd akusztikai lenyomatait a hangszínképeken (spektrogramokon) láthatjuk, vö. 4. ábra. Az ábra az *oroszlán* szó női ejtéséről készült akusztikai lenyomatot mutatja. Ezek a hangszínképek teszik lehetővé, hogy a kutatók elemezhesék a beszédet, beleértve az artikulációs átmenetek akusztikai következményeit. Ezen a módon megállapították, hogy például ahhoz a helyzetváltoztatáshoz, hogy a nyelv a szájüreg hátsó részéből az elülső részbe kerüljön – például a *gítár* szóban a *g* mássalhangzóra jellemző helyzetből az *i* magánhangzóra jellemző helyzetbe –, mintegy 100 ms-ra van szüksége. Ez az időtartam egy átlagos tempóban ejtett hosszú magyar magánhangzó időtartamának felel meg. Magyar beszédben az artikulációs átmenetek átlagos időtartamai 19 ms és 68 ms közöttiek.



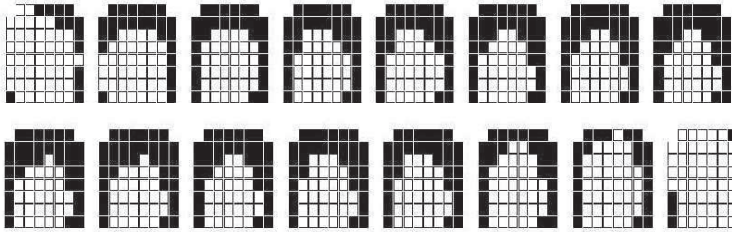
4. ábra. Az oroszlán szó női ejtés alapján készült rezgésképe (felül) és hangszínképe (alul). A hangszínkép függőleges tengelye a frekvenciát, a vízszintes tengely az időt, a feketedés mértéke az összetevők erősségét szemlélteti

A hangátmenet jellemzői attól függenek, hogy a szomszédos beszédhangok képzése milyen mértékben tér el egymástól. Az eltérés mértéke szerint alkalmazkodnak egymáshoz a beszédhangok kisebb vagy nagyobb mértékben. Említettük, hogy a magánhangzók (és a magánhangzószerű mássalhangzók, pl. *j*, *l*) ejtésakor van egy ún. tiszta fázis, ami az adott magánhangzóra jellemző állapotot, az artikulációs konfigurációt tükrözi. Az átmenet megelőzi és követi ezt a tiszta fázist, ha a beszédhangot megelőzi és követi egy másik beszédhang, például a *géz* szóban az *é* magánhangzó előtt és után ejtünk egy mássalhangzót.

A hangképző szervek relatív rugalmatlansága mellett még egy másik ok is hozzájárulhat ahhoz, hogy a szomszédos beszédhangok különböző mértékben hatást gyakoroljanak egymás ejtésére. Ez a kiejtés gazdaságosságára törekvés. Elhagyhatunk például egy-egy hangképzési jellemzőt egy adott hangkapcsolatban a könnyebb ejthetőség érdekében. A *hapci* szóban például a *p* mássalhangzót képezhetjük úgy, hogy a zárpfelpattanás elmarad, és a *p* és a *c* zárszakaszai gyakorlatilag egybeolvadnak. Az *átnyúl* szó ejtésekor a *t* mássalhangzót hátrább képezzük a szokásosnál; ez azt jelenti, hogy a nyelv és a szájpadlás egy pontján történik a zárképzés, nem a nyelv és a fogmeder között, ami a *t* hangra egyébként jellemző. Figyeljük ezt meg például az *étel* szó kiejtésekor. Az a tény, hogy az *átnyúl* szóban hátrább képezzük a *t* mássalhangzót, azt fogja eredményezni, hogy a nyelvnek kisebb mozgást kell végeznie ahhoz, hogy könnyen kiejtsük a *t* után következő *ny* mássalhangzót, amelynek képzésekor a nyelv a szájpadlással érintkezik.

Megfigyelhető az is, hogy a hangkörnyezet hatására nem mindig teljesen ugyanazt a mássalhangzót ejtjük ki. Ez azt jelenti, hogy a szomszédos beszédhangok hatására például a *k* kiejtése módosulhat. Jól érzékelhető ez a *kutya* és a

*kicsi* szavakban. Ha ezt a két szót szókapcsolatként mondjuk ki, úgy, hogy *kicsi kutya*, akkor még jobban megfigyelhető, hogy a *kutya* szóban a *k* képzésekor a nyelvhat hátrább érintkezik a szájpaddal, a *kicsi* szó ejtésekor pedig kissé előrébb. Korszerű kísérleti technikával (elektropalatográfia) vizsgálható a nyelv és a szájpaddás érintkezésének helye és mértéke egy adott beszédhang ejtésekor. Az 5. ábra mutat egy ilyen képsort a braziliai portugál *d* mássalhangzó esetében.



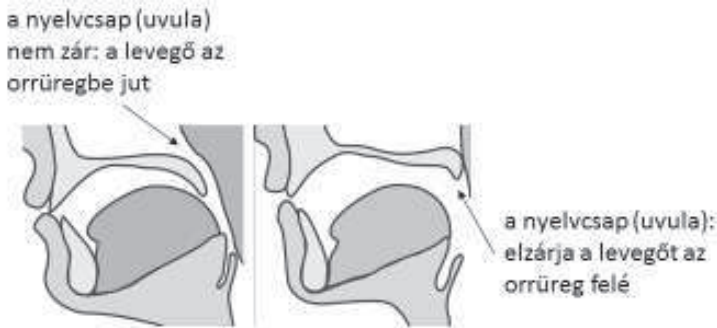
5. ábra. A braziliai portugál nyelv *d* mássalhangzójának ejtéséről készült elektropalatogramok az ejtés kezdetétől a végéig (Jesus–Reis 2012: 257). A nyelv és a szájpaddás érintkezési területeit a feketedések jelzik

A magyarban három nazális (orrüregi képzésű) mássalhangzó van (ezeknek a *m*, a *n* és a *ny* betűk felelnek meg). Folyamatos beszédben azonban két további nazális mássalhangzót is ejtünk a szomszédos beszédhangok hatására. Az egyiket akkor ejtjük, amikor a *m* hangot a szóban *v* vagy *f* követi (pl. *kémvád*, *romfal*). Ekkor nem az eredeti, két ajakkal képzett orrhangú mássalhangzó jön létre, hanem a nazális mássalhangzót kissé hátrább képezzük, a zár a felső fogsor és az alsó ajak között realizálódik (6. ábra). Az ábra bal oldali képe szemlélteti a *v* mássalhangzó hatására módosult *m* hangot. A jobb oldali kép a *v* mássalhangzó ejtésének felel meg. Az eredeti *m* mássalhangzó megőrizte nazális jellegét, a zár képzési helye azonban jól láthatóan nem a két ajak között van, hanem a felső fogsor és az alsó ajak között jön létre.

A másik nazális mássalhangzó, a *n* az *öt* követő *g* vagy *k* hang hatására szintén módosul, a képzési helye hátrább tolódik a szájüregben (pl. a *lángos* vagy az *inka* szavakban). Kis gyakorlással ezek a módosult képzési helyű nazális mássalhangzók önmagukban is kiejthetők.

A nazális mássalhangzók ugyanakkor maguk is hatást gyakorolnak a környezetükben lévő beszédhangokra, ez különösen jól megfigyelhető azokban az esetekben, amikor magánhangzó előtt vagy után ejtjük a nazális mássalhangzót (Beke–Horváth 2009). A magánhangzók kisebb-nagyobb mértékben válnak orrhangúvá, például: *hamar*, *sunyít*, *menet*, illetve *nyenyec*, *mongol*, *manó*. Nyilvánvalóan erőteljesebb lesz a magánhangzó nazális színezete, ha előtte és utána is orrhangú mássalhangzót ejtünk (pl. *mama*), mint ha csak egyik oldalon szom-

szédos a nazálissal (pl. *majális* vagy *karom*).



6. ábra. A *hamvas* szóban ejtett nazális (baloldalon) és az önmagában ejtett *v* mássalhangzó képzési konfigurációja (jobb oldalon) röntgensémák alapján

#### 4. Beszédhangok nyelvspecifikus egymásra hatásai

Évtizedekkel ezelőtt felmerült az igény, hogy megmagyarázzák és rendszerezék a beszédhangok egymásra hatását a folyamatos beszédben. Válaszokat igyekeztek kapni arra, hogy az egymásra hatás miért és miként jön létre, valamint hogy a lezajló folyamatokat modellezzék, és a működésmechanizmusokat értelmezzék. A vonatkozó kutatások számos általános, nyelvtől független és számos nyelvspecifikus, azaz egy-egy adott nyelvre vagy nyelvváltozatra jellemző jelenséget írtak le a beszédhangok viselkedéséről a folyamatos beszédben. Ezeket fonológiai koartikulációknak nevezzük.

Az egymásra hatás megvalósításában nagyon fontos szerepük van a nyelvspecifikus sajátosságoknak. Minden nyelv hangzó változata tartalmaz olyan jelenségeket, amelyek okai nem a beszédszervek működéséhez köthetők, mint például az artikulációs hangátmenetek, hanem azért jönnek létre, mert érvényesülnek az adott nyelvre jellemző kiejtési szabályok. Ez utóbbi esetekben tehát nem az artikulációs szervek rugalmatlansága idézi elő a módosulásokat, hanem az adott nyelvnek vannak olyan saját szabályai, amelyek előírják ezeket az ejtésmódokat (pl. Siptár 2002). Ezek a jelenségek lehetnek hasonulások, kiesések, rövidülések, avagy a hiátus feloldásai (l. bővebben: Gósy 2004).

##### 4.1. Hasonulás

Ha a szomszédos beszédhangok a zöngéesség/zöngétlenség vagy a képzési hely tekintetében változnak egy adott hangkapcsolatban, akkor különböző típusú hasonulások, illetve sorozatos hasonulások következnek be (pl. a *háztól* ejtése

*hásztól*, a *hasból* ejtése *hazsból*, az *egészség* ejtése *egésség*, a *színpad* ejtése *szímpad*, a *gyöngy* ejtése *gyönygy*, a *húnyjál* ejtése *húnnnyál*, a *költség* ejtése *kölcség*).

Ezek az egymásra hatások akkor is létrejönnek, ha két szó folyamatos ejtése következtében kerülnek a beszédhangok egymás mellé. Ekkor az egyik beszédhang az első szó utolsó, a másik beszédhang a második szó első beszédhangja. Például *még tovább*, ejtése: *mék tovább* vagy a *két beteg* ejtése: a *kéd beteg* vagy a *képen gyár van* ejtése: *képenygyár*.

A magyar köznyelvi ejtésben a nazális mássalhangzók és a *v* réshang nem hasonít, vagyis a megelőző zöngétlen mássalhangzó nem válik zöngéssé, például a *rakni* vagy a *hatvan* ejtésekor. Például a *két vár* vagy a *kötvény* szó esetében az ejtésben a *v* nem hat a *t* hangra abban az értelemben, hogy megváltoztatná annak zöngétlen voltát. Van azonban olyan nyelvjárás, ahol ilyenkor is bekövetkezik a koartikulációs hatás, ekkor az ejtés *kéd vár* és *ködvény* lesz. Érdekesség, hogy a gyermek anyanyelv-elsajátítása során egy adott szakaszban túláltalánosítja ezt a koartikulációt, és alkalmazza mind a nazálisok, mind a *v* esetében annak ellenére, hogy a köznyelvi minta ezt nem erősíti meg, például: *zokni* ejtése *zogni*, *lekvár* ejtése *legvár*. Viszonylag gyorsan felismeri azonban (bár nem tudatosan) a szabályt, amely a nazálisokat és a *v*-t a köznyelvben kizárja ebből a koartikulációs jelenségből, és onnantól már a köznyelvi mintának megfelelően ejti az adott szavakat.

## 4.2. Hiátus

A beszédhangok egymásra hatásának egyik különleges esete a hiátus. Mi a hiátus? Ha két magánhangzó egymás mellé kerül egy szóban vagy szavak határán – például *fiatal*, *bácsié*, *leég*, *ráút*, *almaital*, illetve *maci édes*, *arra egyedül*, *földi élet* – akkor a köztük lévő ürt hiátusnak nevezzük. Hiátus csak akkor jön létre, ha a szomszédos magánhangzók nem egy kettőshangzó (diftongus) alkotó elemei. A *tiéd* szóban például hiátus található, mert a szomszédos magánhangzók egyike a szó első, másika a második szótagban foglal helyet (*tiéd*). Ellenpélda lehet a *sziép* nyelvjárásban ejtett szó, amelyben a mássalhangzók között egy kettőshangzó található (félkövérrel jelölve). Ez tehát nem hiátus. (Megjegyezzük, hogy több nyelvben nincs olyan helyzet, hogy két magánhangzó kerüljön egymás mellé sem szóban, sem szavak határán.)

A nyelvek nagy része kevésbé vagy egyáltalán nem tűri el a hiátust, azaz a nyelv belső szabálya az, hogy igyekszik a szomszédos magánhangzók ejtését valamilyen módon feloldani. A cél az, hogy ne kelljen két magánhangzót kiejteni egymás után. Akad olyan nyelv, amelyik teljesen kizárja annak a lehetőségét, hogy két magánhangzó egymás mellett forduljon elő, ilyen például a török vagy a kikereve bantu nyelv. Az alapkérdés az, hogy vajon hogyan valósítjuk meg az

ejtésben a szomszédos magánhangzókat, azaz miként oldjuk fel a hiátus jelenségét. Nem csupán egyetlen elkerülési lehetőség van, a feloldási módzatok nyelvenként változók, sőt egyetlen nyelven belül is többféle megoldás, többféle feloldási stratégia létezhet.

A szomszédos magánhangzók ejtésének elkerülési módzatait – a már ismert és elemzett nyelvek alapján – hat kategóriába sorolják (a példákat l. Gósy 2014-ben).

(i) A magánhangzó-kapcsolat feloldása kettőshangzóval, azaz a két magánhangzó egy szótagot alkot (pl. a *ngiti*, Zairéban beszélt, közép-szudáni nyelvben). Ekkor tehát azáltal egyszerűsödik az ejtés, hogy a beszélőnek nem kell két magánhangzót teljes értékűen egymás után kiejtenie, hanem egyetlen beszédhangnyi időtartamban ejti az egymásba alakuló két magánhangzót.

(ii) A két magánhangzó közé egy mássalhangzó ékelődhet, ezt a jelenséget epentézisnek nevezzük, például az angolban alkalmazzák a hiátus feloldására.

(iii) Egyszerű megoldás, amikor az egymás mellé került két magánhangzó egyikét nem ejtik ki, más szóval az egyik magánhangzó kiesik (pl. az *etszako*, nigériai nyelvben).

(iv) Magánhangzók vegyüléséről beszélünk, amikor egy harmadik magánhangzó lesz a két eredetiből, de ez a létrejött magánhangzó tartalmazza az eredeti magánhangzók egyes artikulációs sajátosságait (pl. a *kimatumbi bantu* nyelvben).

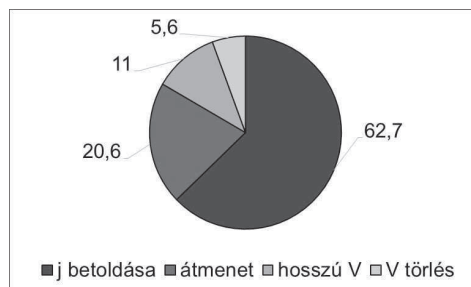
(v) Egyfajta hasonulás történik, amikor az egyik magánhangzó teljes mértékben beleolvad a másikba, és egyidejűleg ennek a magánhangzónak az időtartama megnövekszik (pl. a törökben).

Mi a helyzet a szomszédos magánhangzók ejtésével a magyarban? Hogyan ejtjük a *fiú*, *kiáll*, *beenged* vagy *neon* szavak szomszédos magánhangzóit? A hiátus feloldására elméletileg három lehetőség adódik: (i) a magánhangzók egyikének törlése, (ii) egy hiátustöltő hang közbeékelése és (iii) a két magánhangzó egyikének átalakítása ún. félmagánhangzóvá (Siptár 2002). A két magánhangzó között kötelezően megjelenik a *j*-szerű mássalhangzó akkor, ha a részt vevő magánhangzók egyike az *i* vagy az *í* (pl. *síel* vagy *mai*). Nem következik be azonban ez a fajta hiátustöltés akkor, ha mindkét magánhangzó ajakkerekítéses, a szájüreg hátulsó és alsóbb részében realizálódik (pl. *a* és *o* kerül egymás mellé az ejtésben).

A kérdés az, hogy vajon az elméleti feltételezést igazolják-e a konkrét beszédelemzések? Egy kutatásban a hiátus feloldását vizsgálták a spontán beszédben, és ennek során egyrészt meghatározták a feloldási módzatokat, valamint megmérték a létrejött magánhangzó-kapcsolatok időtartamát. Négy módzatot találtak a hiátus feloldására a magyarban szemben az elméletileg feltételezett három lehetőséggel. Ezek a következők:

- (i) hiátustöltés, vagyis a *j* mássalhangzó beillesztése a két magánhangzó közé (pl. *fiú*, ejtése *fijú*, további példák: *zenei*, *fiatal*, *anyagiak*, *akció*, *lelki-állapot*),
- (ii) az átmenet a két magánhangzó között: ekkor az egyik magánhangzó ejtéséből folyamatosan történik az átmenet a másik magánhangzóba (pl. *beugrik*, *leakaszt*, *operaénekes*, *Budaörs*, *alapvetően*, *kottaolvasó*),
- (iii) a magánhangzók összeolvadása: két azonos szomszédos magánhangzót egyetlen hosszú magánhangzóként ejtünk (pl. a *leesik* ejtése *lesik* hosszú *e*-vel, *rááll*, *kiír*, *kávédés*),
- (iv) az egyik magánhangzó kiesése (pl. *szexuális*, ejtése *szekszális* vagy *ebbe akkor*, ejtése *ebbakkor*).

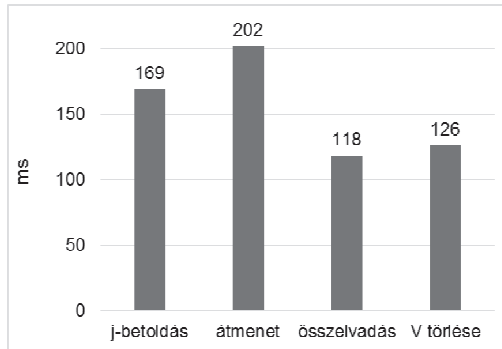
Az egyes jelenségek előfordulása egyrészt a nyelv szókészletétől függ, vagyis attól, hogy hány szó tartalmaz magánhangzó-kapcsolatokat (*mai*, *kakaó*, *idióta*, *fecskéi*), illetve ezek milyen gyakran fordulnak elő a kutatásban vizsgált spontán beszédanyagban (néhány, a mindennapokban valószínűleg ritkábban használatos szó: *mióma*, *pióca*, *kupéi*, *teveitató*, *autentikus*, *tuareg*, *gumiöv*, *duett*, *repertoár*). Az előfordulás gyakorisága másrészt attól függ, hogy az egymás után ejtett szavak esetében adódik-e magánhangzó-kapcsolat, például *a mozi érdekes*, *kese-rű étel*, *felkavaró álom*, *a megye után*, *külföldre akarok*. A vizsgálatban kapott előfordulási arányokról a 7. ábra tájékoztat. A *j* hiátustöltő megjelenése a leggyakoribb, ez azzal magyarázható, hogy a magánhangzó-kapcsolatokban nagy arányban fordul elő az *i*, *í*, és ekkor minden esetben megjelenik a produkcióban az ejtéskönnyítő mássalhangzó.



7. ábra. A hiátus feloldási módozatainak előfordulási arányai spontán beszédben (%)  
(V = magánhangzó)

Érdekes, hogy miként alakulnak az egyes hiátusfeloldási módozatokban az időtartamok. A *j* betoldása esetén a teljes hangkapcsolat ejtésének időtartama átlagosan 169 ms. Az átmenet megvalósulása esetén a teljes hangkapcsolat időtartamának átlaga 202 ms. A magánhangzók összeolvadásakor az ejtés átlagos

időtartama 118 ms. Végül a magánhangzó törlődése esetén a megnyúlt magánhangzók átlagos időtartama 126 ms. Az átlagos időtartamokat a 8. ábra összegzi.



8. ábra. A hiátus feloldásának következményei a magánhangzók időtartamára

Megállapíthatjuk, hogy a szomszédos magánhangzók kiejtése a spontán beszédben jóval összetettebb, mint azt korábban elméletileg feltételezték. A magánhangzó-kapcsolatok időtartamai egyértelműen függenek a szomszédos magánhangzók minőségétől, pontosabban attól, hogy miként történik a hiátus megszüntetése, feloldása. A feloldási módzatok tendenciaszerűen egyfajta ejtési gazdaságosságot mutatnak. Az átmenetek megjelenése az időtartamot tekintve a legkevésbé gazdaságos ejtés. A két magánhangzó kapcsolatának legrövidebb ejtése összeolvadáskor és törlődéskor következik be; azt mondhatjuk, hogy ezek egyben a hiátus legegyszerűbb feloldási módzatai. A hiátustöltő *j* beillesztése nem tekinthető gazdaságosnak abban az értelemben, hogy megnöveli a hangkapcsolat időtartamát, ezáltal növekszik a teljes szó vagy szókapcsolat időtartama is. Ugyanakkor a kiejtési könnyebbség tekintetében ez a legmegfelelőbb artikulációs megoldás.

### 4.3. Nem szabályszerű egymásra hatások

Vannak olyan jelenségek a beszédben, amelyek hagyományos értelemben nem koartikulációs jelenségek, mégis érdemes megemlíteni ezeket mint sajátos egymásra hatásokat.

Különböző okok idézik elő a megjelenésüket, rendszerint az, hogy a beszélő törekszik az ejtési egyszerűsítésre. Ok lehet a lazább kiejtés, ami sokszor jellemző a spontán beszédre, és gyakran tapasztalható gyors beszéd esetén. Ezeknek a jelenségeknek egy része az adott beszélő egyéni ejtései, és egy vagy néhány szót érintenek, más részük az általában gyakrabban előforduló szavakhoz kapcsolódik. Formailag hasonlóak a szabályos koartikulációkhoz, azonban nem a nyelv szabályos működésének következményei. Lehetnek rövidülések, kiesések, hang-

helyettesítések, nemegyszer azonban kifejezetten szócsonkítások. Például: a *-ban/-ben* határozórag ejtése *-ba/-be* formában, egyes szóvégi mássalhangzók eltűnése, például *mer* a *mert*, *ho* a *hogy* helyett, hanghelyettesítések, amelyek egyszerűbb kiejtést tesznek lehetővé, mint *ejébként* az *egyébként* vagy *lejet* a *lehet* helyett. A jellegzetes szócsonkítások általában ugyanazoknál a szavaknál fordulnak elő különböző beszélők ejtésében is, például: *tát* a *tehát*, *tonképpen* a *tulajdonképpen* vagy *szal* a *szóval* helyett.

## 5. Összegző megállapítások a beszédhangok egymásra hatásáról

**5.1.** A folyamatos beszédben a szavak kezdetét és a végét nem különítjük el állandóan szünetekkel, azaz nincsenek közöttük az írás szóközeinek megfelelő csendes szakaszok. Két beszédhang nem csak a szótőben (pl. a *d* és a *r* a *drága* szóban), de a szótő és más morfémák határán (pl. a *d* és a *r* a *földre* szóban), sőt szavak határán is egymás mellé kerülhetnek az ejtés során (pl. a *hold ritkán* közlésrészletben).

**5.2.** A beszédben a szomszédos beszédhangok között átmeneti szakaszokat ejtünk, amelyek biztosítják az adott beszédhangra jellemző képzési sajátosságok kialakítását. Ezek a szakaszok az artikulációs hangátmenetek.

**5.3.** A szomszédos beszédhangok hatással vannak egymás ejtésére, ami azt jelenti, hogy kisebb-nagyobb mértékben módosíthatják az önmagában ejtett beszédhangra jellemző képzési sajátosságokat. Ezek a koartikulációs jelenségek.

**5.4.** A koartikulációs jelenségek két fő típusra különíthetők el aszerint, hogy mi idézi elő az egymásra hatást, illetve hogy az miként valósul meg. Az egyik típusban nem következik be olyan beszédhang ejtése, amely része a nyelv hangállományának, míg a másik típusban ez bekövetkezik. Például a *nocsak* szó első mássalhangzója kismértékben különbözik a *nicsak* szóban ejtett nazálistól, de maga a nazális *n* mássalhangzó azért nem változik meg. Az utánuk következő magánhangzó hatására kismértékben eltolódik a nazális mássalhangzó és a fogmeder közötti zár helye. Ezt nevezik fonetikai koartikulációnak. Ha azonban a *konty* szó ejtésére gondolunk, ami *konyty*, akkor látható, hogy az eredeti *n* mássalhangzó a *ty* hatására megváltozott, *ny* lett, ami olyan beszédhang, amely tagja a magyar mássalhangzó-rendszernek. Ezt nevezik fonológiai koartikulációnak. Idetartoznak a hasonulások, a szabályos hangkiesések, a hiátus.

**5.5.** A szomszédos beszédhangok egymásra hatása lehet előre ható (progresszív) és lehet hátra ható (regresszív). Az előre ható egymásra hatáskor a másodikként kiejtett hang illeszkedik a megelőzőhöz, míg a hátra ható egymásra hatáskor az elsőként kiejtett hang közeledik a másodikként képzett hang ejtési sajátosságaihoz. Az előre ható koartikulációra példa az *anyja* szó *annya* ejtése, amikor a *ny* hatására a *j* mássalhangzó megváltozik; a hátra hatóra példa a

*légtornász* szó *léktornász* ejtése, itt a zöngétlen *t* hatására vált zöngétlenné (*k*) a zöngés *g* mássalhangzó.

**5.6.** A beszédhangok egymásra hatását fiziológiai és nyelvi tényezők idézik elő.

**5.7.** Nem szabályos koartikulációs jelenségek adódhatnak egyéni ejtési sajátosságokból. Az ilyen egyéni ejtési sajátosságok terjedhetnek (pl. a *szóval szal* ejtése).

A koartikulációs jelenségek leírása még nem teljes (sem a magyarban, sem más nyelvekben). Mind a kísérletes megközelítés, mind a nagy mennyiségű spontán beszéd fonetikai vizsgálata közelebb visz a jelenségek mindenre kiterjedő, rendszerszerű leírásához. Ezeknek a jelenségeknek az ismerete a gyakorlatban több területen is meghatározó fontosságú. A magyar mint idegen nyelv tanításában nélkülözhetetlen, hogy a tanár felhívja erre a figyelmet, adott esetben szembeállítsa a tanuló anyanyelvével. Fontossá válhat a patológiás (atipikus) beszéd minősítésében, és ezáltal jelentős hozzájárulás a pontosabb diagnózis felállításához. A beszédhangok egymásra hatásai alapján a mesterséges beszéd-előállításban szabályrendszer alakítható ki a mind jobb hangzás létrehozása érdekében. A mesterséges (gépi) beszédfelismerés ugyancsak nagy hasznát veszi a jól meghatározott beszédhang egymásra hatásoknak a spontán beszéd automatikus felismertetésében, mivel ez csökkenti a felismerés hibaarányát.

## Irodalom

- Beke András – Horváth Viktória 2009. A nazális koartikuláció variabilitása a spontán beszédben. *Beszédkutatás* 2009. 28–45.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris, Budapest.
- Gósy Mária 2014. A palatális közelítőhang kétféle funkcióban. *Beszédkutatás* 2014. 17–40.
- Harrington, Jonathan 2010. *Phonetic analysis of speech corpora*. Blackwell Publishing House, Malden–Oxford.
- Jesus, Marisa S. – Reis, Cesar 2012. Phonetic description of alveolar phones using electro-palatography. *Journal da Sociedade Brasileira de Fonoaudiologia* 24. 255–261.
- Siptár Péter 2002. Hiátus. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 85–98.

## Kérdések, feladatok

1. Hogyan magyarázható a koartikuláció a spontán beszédben?
2. Mit jelentenek a beszédhangok nyelvspecifikus egymásra hatásai a beszédben?
3. Milyen módon oldható fel a hiátus a különböző nyelvekben?



## **Különleges beszédképzési módok: éneklés, suttogás, gége nélküli beszéd, hasbeszélés**

### **1. A különleges beszédképzési módok fonetikai vizsgálatának motivációja**

Az emberi beszédképző szervek számos, a hétköznapiak tekinthető beszédétől eltérő beszédképzési vagy hangképzési módra, így például suttogásra, nyelöcső-beszédre vagy akár több különböző technikájú éneklésre is (pl. operaéneklés, jódli, népi ének „torokéneklés”, vagy más néven *overtone singing* 'felhangéneklés' stb.) képesek – ezek vizsgálatának a fonetika nagy figyelmet szentel. A ki-tüntetett figyelem sok esetben két okra vezethető vissza. Az egyik ok, hogy a mindennapitól eltérő beszédkeltés a beszédképző szervek alapvető működéseiről is sokat elárul: a különleges produkciós módok sokszor a hangadás olyan aspektusait teszik vizsgálhatóvá, amelyeket esetleg a mindennapi beszéd vizsgálata során nem, vagy csak korlátozott mértékben áll módunkban feltárni. A másik ok pedig, hogy az ilyen különleges beszédképzési módokban létrehozott beszédhangok, szavak vagy megnyilatkozások feldolgozásának vizsgálatával az emberi beszédpercepció és beszédmegértés is behatóbban megismerhető. Mivel a különleges beszédképzési módokban ejtett beszédhangok és prozódia a hétköznapi beszédétől akár nagyobb eltérést is mutathatnak, mint a mindennapi beszédben egyébként tapasztalható variabilitás eredményeként létrejött változatok, ezek a produkciós módok mintegy természetes módon feszegetik az egyes percepciók kategóriák, például a beszédhangok kategóriáinak határait. Ilyen módon például a suttogás, a nyelöcsőbeszéd vagy az éneklés mint produkciós mód vizsgálata képes lehet a beszédvizsgálatoknál jobban megvilágítani, hogy milyen információk lehetnek szükségesek vagy elegendők egy-egy nyelvi jel vagy jellészlet, például egy adott beszédhang felismeréséhez.

### **2. Az énekhang fonetikai vizsgálata**

Többek közt az operaéneklés fonetikai érdeklődésű vizsgálata is visszavezethető a fent említett két okra. Az éneklés kutatásának központjában a magánhangzók állnak, mivel – akusztikus sajátosságaikból fakadóan – e hangzók a legalkalmasabbak az énekléshez nélkülözhetetlen alaphangélmény kialakítására (ti. „csak a magánhangzókra lehet énekelni”).

A kutatókat az operaéneklés produkciójára vonatkozóan elsősorban a követke-

ző kérdések foglalkoztatják: Hogyan képesek az énekesek olyan nagy hangerejű hangot produkálni, amely erősítés (mikrofon és hangszórók) nélkül is jól hallható akár nagyzenekari kíséret mellett is? Hogyan képesek a szoprán énekesek olyan magas hangmagasságok elérésére, ami a hétköznapi, énektechnikailag képzetlen beszélők számára elérhetetlennek tűnik? Hogyan lehetséges az, hogy az énekesek az általuk használt igen nagy (több oktávnyi) hangterjedelemben is képesek hasonló hangszínezettel énekelni – miközben a hétköznapi beszélők sok esetben már a beszédhangnak az éneklésnél sokkal szűkebb és mélyebb alulfrekvencia-tartományában is tapasztalják, hogy hangjuk hangszínezetet vált, „elcsuklik” akár már kisebb alaphangemelésnél is?

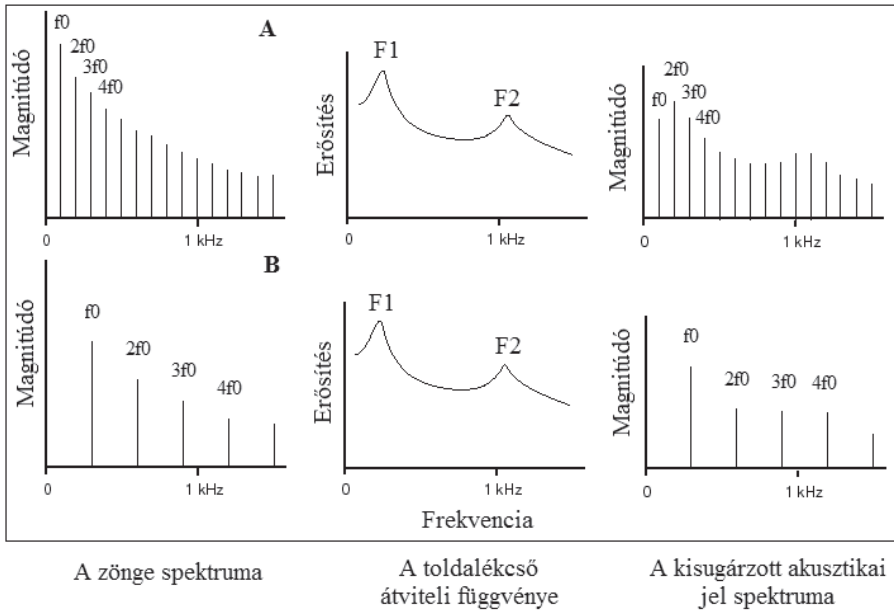
Az énekelt magánhangzókkal kapcsolatos percepciók vizsgálatok legfőbb kérdése, hogy miként befolyásolja az alulfrekvencia az énekelt magánhangzók azonosítását, ugyanis – mint látni fogjuk – az éneklésben az alulfrekvencia emelésével jelentősen megváltozhat a magánhangzók ejtése.

## 2.1. Az alulfrekvencia hatása a magánhangzók akusztikai szerkezetére

A zöngé a természetben előforduló hangok legnagyobb részéhez hasonlóan komplex hang, azaz több összetevőből áll: az alulfrekvenciából vagy alaphangból ( $f_0$ ), valamint annak egész számú többszöröseiből, a felhangokból (egy lehetséges jelölés szerint  $2f_0$ ,  $3f_0$  stb.; Vicsi 2010). Az alulfrekvencia és a felhangok összefüggése miatt akkor, amikor az  $f_0$  értéke megváltozik, a zöngé és a zöngés beszédhangok, például a magánhangzók akusztikai szerkezete is globálisan átalakul, ugyanis az  $f_0$  emelésének hatására az egyes felharmonikusok távolsága is növekedni fog. Ezt szemlélteti szematikusan az 1. ábra. Az ábra mindkét sorában az első grafikonon a zöngé alulfrekvenciája ( $f_0$ ) és felhangjai ( $2f_0$ ,  $3f_0$  stb.) látszanak. Az  $x$  tengely a frekvenciát reprezentálja. Jól látszik, hogy amikor alacsony az  $f_0$  (felső, A ábrásor), a felhangok szorosan követik egymást (hiszen az alacsony  $f_0$  egész számú többszörösei közelebb esnek egymáshoz). Amikor azonban magas az  $f_0$  (alsó, B ábrásor), annak egész számú többszörösei távolabb vannak egymástól. Egy konkrét példán: ha szoprán énekes beszédben 200 Hz-es alaphangot hoz létre, ennek a felhangjai 400, 600, 800 Hz-esek (a távolságuk kicsi). Ha az éneklés magas hangfékvésben történik, pl. 800 Hz-en, akkor ennek felhangjai 1600, 2400, 3200 stb. Hz-esek lesznek (vagyis a távolságuk jóval nagyobb).

Hogyan hat mindez a magánhangzók akusztikai tulajdonságaira? Ennek a kérdésnek a megválaszolásához ismernünk kell a magánhangzók akusztikai természetének kialakulását, illetve az ezt meghatározó tényezőket.

Ismeretes, hogy a beszédben az egyes magánhangzók akusztikailag jól elkülöníthetők (definiálhatók) a két legalacsonyabb frekvenciaértékű formánusok ( $F_1$  és  $F_2$ ) frekvenciaértéke segítségével (Peterson–Barney 1952). A formánusok értéke



1. ábra. Az alapprofundencia hatása a magánhangzók spektrális szerkezetére. Első oszlop: a gégeben létrejött zöngespektruma, második oszlop: a szűrőként funkcionáló toldalékcső rezonanciái (azaz a formánsok), harmadik oszlop: a kisugárzott akusztikai jel (azaz a magánhangzó) spektruma alacsony (A sor) és magas (B sor) alapprofundencián (Epps et al. 1997: 1113 alapján)

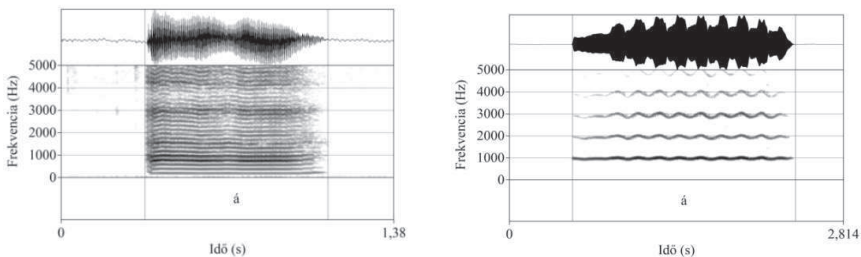
attól függ, hogy az egyes magánhangzók ejtésekor milyen a toldalékcső belső alakja (milyen helyzetben vannak az egyes képzőszervek – ez az artikulációs konfiguráció). A formánsok a toldalékcső ún. sajátrezonanciái (Fant 1960), amelyek az artikulációs konfigurációtól függnnek. Az első két formáns ( $F_1$  és  $F_2$ ) kitüntetett szerepének egyik legfőbb oka, hogy azok frekvenciaértéke szoros összefüggést mutat az egyes (a magánhangzók képzésében fontos) artikulációs szervek, a nyelv és az állkapocs mozgásával. Az  $F_1$  értéke elsősorban a nyelvállás fokával vagy az állkapocsnyitás szögével függ össze, mégpedig fordított arányosságban (a magasabb nyelvállás alacsonyabb  $F_1$ -értéket eredményez). Az  $F_2$ -re elsősorban a nyelv vízszintes irányú pozíciója (az elöl képzett hangzók  $F_2$ -értéke magasabb), másodsorban pedig az ajakműködés (a kerekített hangzók  $F_2$ -értéke alacsonyabb, mint a réses hangzóké) hat (Stevens 1998). Felnőtt női beszélők esetében a felső nyelvállású magyar magánhangzók ( $i, ü, u$ )  $F_1$ -értéke a legalacsonyabb frekvenciájú, ez átlagosan 100 és 400 Hz közötti értéket vesz fel (vö. pl. Gósy 2004).

A toldalékcsőnek ezek a sajátrezonanciái ( $F_1$  és  $F_2$ , az 1. ábrán a középső grafikonok szemléltetik) felelősek tehát a magánhangzó hallható minőségéért:

ezeket dolgozza fel a hallásunk, és ezek alapján tudjuk azonosítani a magánhangzókat. Ahhoz, hogy az  $F_1$  és az  $F_2$  formánsok hallhatók legyenek, az kell, hogy ezek frekvenciaértékének a közelében legyen(ek) a zöngé spektrumában olyan frekvenciaértékű összetevő(k) (felhangok), amely(ek) e formánsok tartományába esik (esnek).

Azokban az esetekben, amikor magas alapfrekvencián a felhangok távolabb kerülnek egymástól a magánhangzó spektrumában, gyakran előállnak olyan helyzetek, amelyek során egyetlen felhang sem esik közel egy adott formáns középértékének frekvenciahelyéhez, sőt akár a formáns teljes sáv szélességébe sem (ezt látjuk a 2. ábra B tablóján az  $F_2$  formáns esetében, mely éppen a  $3f_0$  és  $4f_0$  közé esik). Mivel ebben az esetben az adott formánst nem gerjesztik felharmonikusok, a formáns nem (vagy nem egyértelműen) jelenik meg a kisugárzott hang spektrumában (spektrális csúcsként), így tehát nem is lesz detektálható (és az emberi fül számára érzékelhető).

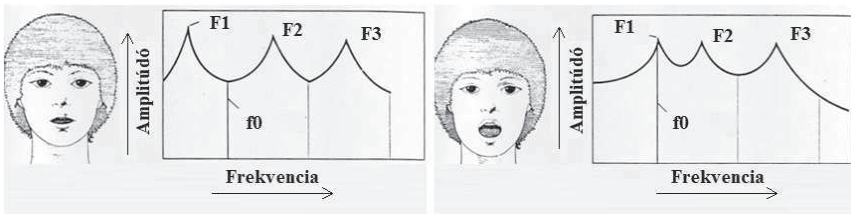
Egy valós beszédmintából származó példát mutat be a 2. ábra. Itt az *á* magánhangzó beszédbeli megvalósítását (bal oldal), valamint a magas alapfrekvenciájú éneklésben rögzített realizációját (jobb oldal) látjuk egy női énekes kitarított ejtésében (állandó alapfrekvencián). A hangszíneképeken (alsó panelek) látható, hogy míg a beszédben vízszintes sötét csíkként jól kivehetők a magánhangzó formánsai, addig az 1 kHz közeli alapfrekvencián már csak az egyes felharmonikusok rajzolódhatnak ki. Így bár magas alapfrekvencián a felharmonikusok sötétedésük mértéke alapján sejtetni engedik, hogy beleesnek-e valamely formáns sáv szélességébe, a formánsok sáv szélessége és középértéke a hagyományos eszközökkel már nem mérhető megbízhatóan. A jelenség természetesen a beszédpercepcióra nézve is következményekkel jár: a magánhangzók azonosításának szempontjából hagyományosan fontosnak tartott formánsok ugyanis nem, vagy csak korlátozottan érhetők el az emberi beszédpercepció számára a magas alapfrekvenciájú beszéd- és énekhangban.



2. ábra. Beszélve (bal oldal,  $f_0 \sim 200$  Hz) és magas alapfrekvencián énekelve (jobb oldal,  $f_0 = 988$  Hz) ejtett *á* hullámformája (felső panelek) és keskenysávú hangszíneképe (alsó panelek).

## 2.2 Az énekhangképzés és a magánhangzóejtés ellentmondása

Az éneklés során az egyes hangosztályok, de főként a szoprán hangosztály énekesei nem ritkán olyan magas alaphangon ( $f_0$ ) énekelnek, amelynek a frekvenciaértéke magasabb, mint az egyes magánhangzók beszédben jellemző  $F_1$  (vagy ritkábban  $F_2$ ) formánsának értéke. Ezekben az esetekben, ha az énekes változatlanul (a beszédben megszokott) artikulációval ejtené a magánhangzókat, annak igen jelentős akusztikai következményei lennének. Az  $f_0$  komponens erősítésmentesen (kis intenzitással) jelenne meg a spektrumban, azaz a hang hirtelen lehalkulna, és mind a magánhangzó hangszíne, mind pedig az énekhang sajátos hangszínezete is feltűnően megváltozna azokhoz a helyzetekhez képest, amikor az  $F_1$  frekvenciaértéke magasabb az  $f_0$ -énál (3. ábra, bal oldal).



3. ábra. A magas alaphfrekvenciájú magánhangzók ejtése változatlan (a beszédejtésben jellemző) artikulációval (bal oldal), és az operaéneklésben (módosított artikulációval, jobb oldal) (Sundberg 1977: 91 alapján)

Az operaéneklésben a fent említett változások azonban nemkívánatosak. A képzett énekeseknél megfigyelhető, hogy hangjuk igen nagy hangerejű, szép és homogén hangszínezetű, függetlenül az énekelt zenei hang magasságától (alaphfrekvenciájától). Egyúttal viszont azt is gyakran tapasztalhatjuk, hogy az énekesek által ejtett szövegek érthetősége (elsősorban a magánhangzók eltérő minősége miatt) a magas hangokon rosszabb.

## 2.3. Hogyan módosítja a hangzók ejtését az énekes magas alaphfrekvenciákon?

Magas alaphfrekvencián való éneklés esetén tehát előfordulhat, hogy az  $f_0$  frekvenciaértéke magasabb az  $F_1$ -nek a beszédben jellemző értékénél ( $f_0 > F_1$ ). A szakirodalomban mára elterjedt feltételezés szerint az ilyen helyzetekben fellépő nemkívánatos módosulásokat az énekesek úgy kerülik el, hogy az  $F_1$  értékét a megemelt  $f_0$  értékére (vagy valamivel a fölé) hangolják (jelölése a továbbiakban  $F_1 : f_0$ , lásd 3. ábra, jobb oldal). Ez azt jelenti, hogy az énekesek az  $F_1$  rezonancia frekvenciájának kialakításáért felelős artikulációs szervek beállítá-

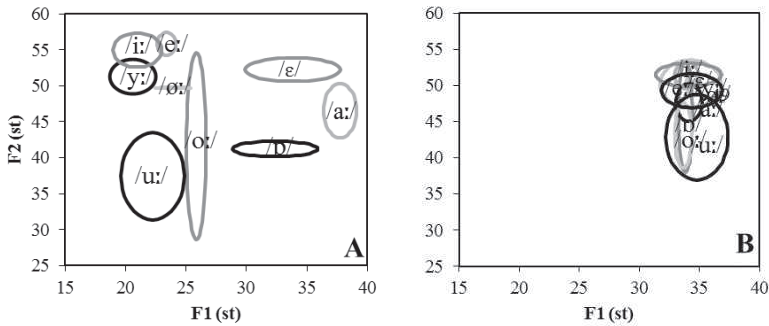
sát módosítják: növelik az állkapocsnyitás szögét, illetve csökkentik a nyelvállás fokát (vö. Sundberg 1987). Ebből az következik, hogy a két cél elérése ellentmondásba kerül: a homogén hangszínű, csengő hangú éneklés és az észlelés számára azonosítható minőségű magánhangzóejtés a zártabb magánhangzók esetében. Ez a konfliktus kivétel nélkül az éneklés javára dől el; a magas alaphangfrekvenciájú éneklésben (az  $f_0 > F_1$  helyzetekben) ugyanis az énekes (az alaphangfrekvencia emelésével egyre) nyíltabban ejti a magánhangzókat. E hangolási stratégia alkalmazásával az énekesek képesek minden további vokális erőfeszítés (például az egységnyi idő alatt a tüdőből kiáramoltatott levegő mennyiségének a növelése, vagyis kiabálás) nélkül is akár 20–30 dB hangnyomásszint-növekedést, azaz hangerő-növekedést elérni (Sundberg 1987).

Az énekesek formánshangolási stratégiáit vagy az artikuláció alapján, közvetett úton, vagy közvetve, akusztikai adatokra támaszkodva vizsgálják. A kutatók artikulációs eredményeiből nagyjából egységesen az a következtetés vonható le, hogy míg a felső és középső nyelvállású hangok esetében az énekesek az alaphangfrekvencia emelésével egyre növelték az állkapocs nyitásának szögét már közel 300–400 Hz alaphangfrekvenciától kezdve (ahol is az  $f_0$  elérte a magánhangzó  $F_1$  értékét), a legalsó nyelvállású *á*-szerű hangzókat akár még 600–700 Hz (a zenei  $f''$  hang) környéki alaphangfrekvenciákon is változatlanul (a beszélt ejtéshez hasonlóan) ejtették.

A teljes magyar magánhangzó-rendszert vizsgáló kísérlet szerint a hangolás megjelenése nagyon erősen a magánhangzó beszédben jellemző nyíltságától, azaz első formánsának ( $F_1$ ) beszédben jellemző értékétől függött: a felső nyelvállású hangzók esetében már 350 Hz alaphangfrekvencia környékén megjelent, míg az alsó és legalsó nyelvállásúaknál csak 1 kHz környékén (Deme 2014a). Emellett a vizsgálatok azt is kimutatták, hogy az  $F_1 : f_0$  hangolás velejárájaként az  $F_2$  értéke is megváltozott, mégpedig úgy, hogy a velárisok és a palatálisok  $F_2$ -értéke az  $f_0$  emelésével konvergálni kezdett, tehát az alaphangfrekvencia emelésével az elől- és hátul képzett magánhangzók egyre jobban hasonlítottak egymásra.

Összefoglalóan elmondható, hogy az  $F_1$  és  $F_2$  formánsokkal meghatározható akusztikai magánhangzótér (vö. 4. ábra) az alaphangfrekvencia emelésével fokozatosan lezsugorodik úgy, hogy az *á* pozíciója felé mozdul, mígnem – megközelítőleg 1 kHz alaphangfrekvencián – bármely ejteni szándékozott magánhangzó helyett már csak egyetlen, uniform akusztikai szerkezetű magánhangzót képes ejteni az énekes (még akkor is, ha szándéka szerint különböző magánhangzókat produkál).

A fentiekkel összefüggésben a percepció vizsgálatok legfőbb kérdése, hogy miként azonosíthatók/különíthetők el egymástól az alaphangfrekvencia emelésével egyre jobban uniformizálódó (és egyre ritkább spektrumú) magánhangzók az ejtési szándék szerint.



4. ábra. Az F<sub>1</sub> és F<sub>2</sub> formánsokkal meghatározható akusztikai magánhangzótér (A) a beszédben (f<sub>0</sub> ~ 200 Hz) és (B) magas alapfrekvenciájú éneklésben (f<sub>0</sub> = 698 Hz) három énekes ejtésében. (Az ellipszisek középpontja: az F<sub>1</sub> és F<sub>2</sub> átlaga; sugara: kétszeres szórás.) (Deme 2014a)

## 2.4 Hogyan azonosítjuk az énekelt magánhangzókat?

Mivel az énekelt magánhangzók ejtése az alapfrekvencia emelésével egyre nyiltabb (magas alapfrekvenciákon pedig egyre inkább az *á* képzéséhez hasonló) lesz, magától értetődőnek tűnik az a feltételezés, hogy ezeket egyre nyiltabb (magasabb alapfrekvenciákon pedig *á*-szerű) magánhangzókként is azonosítják a hallgatók (vö. pl. Scotto di Carlo–Germain 1985). Ezt a feltételezést azonban sem a beszédhangok percepciójának szakirodalmá, sem pedig az énekelt magánhangzók egyes újabb kísérletes vizsgálatai nem támasztják alá.

A beszédtudományokban közismert, hogy a magánhangzók percepciójában fontos szerepe van az F<sub>1</sub> és az F<sub>2</sub> formánsnak – pszichoakusztikai kísérletek tanúsága szerint azonban ezeknek nem az abszolút frekvenciaértéke, hanem érzeti viszonya számít, azaz a formánsfrekvenciák távolsága a fül akusztikus felbontásának figyelembevételével. A kísérletek szerint ugyanis a magánhangzók zártságának észleletét a beszédben az F<sub>1</sub> és az f<sub>0</sub> érzeti távolsága (Traunmüller 1981), az elől és a hátul képzettség érzetét pedig az F<sub>2</sub> és az F<sub>3</sub> érzeti távolsága (Carlson et al. 1970) határozza meg. Ráadásul magasabb alapfrekvenciákon ez az összefüggésrendszer még tovább bonyolódik, aminek következtében bár 370 Hz alapfrekvencia alatt az F<sub>1</sub> és f<sub>0</sub> csökkenő távolsága egyre zártabb észleletet okoz, 370 Hz fölött leginkább csak felső és legalsó nyelvállású hangzókat azonosíthatunk a beszédben (Traunmüller 1981). Mivel az észlelt nyíltság szempontjából az f<sub>0</sub> és az F<sub>1</sub> távolsága számít kritikusnak (de ez az alapfrekvencia-tartománytól függően változik), valamint tudjuk, hogy az éneklésben az F<sub>1</sub> : f<sub>0</sub> hangolás lép fel, valójában egyáltalán nem indokolt az a feltételezés, hogy a magas alapfrekvencián énekelt magánhangzókat „eredeti” minőségüknél nyiltabbakként ismernek fel a hallgatók.

Az éneklés percepcióját vizsgáló kísérletek eredményei szerint az énekelt magánhangzóknak az ejtési szándéknak (azaz a szövegben leírt magánhangzóknak) megfelelő azonosítása fokozatosan csökken az alaphangfrekvencia emelésével. A vizsgált legmagasabb alaphangfrekvenciákon (kb. 1 kHz) a hallgatók a legnyíltabb *á*-t és *a*-t azonosítják bármely ejteni szándékozott magánhangzó helyett (80–90%-ban); míg az *á*-t és *a*-t nagy arányban azonosítják az ejtési szándék szerinti az egész szoprán alaphangfrekvencia-tartományon (l. pl. Scotto di Carlo–Germain 1985; Deme 2014b, 2015). 1 kHz alaphangfrekvencia alatt azonban az egyetlen jól kirajzolódó tendencia mindösszesen az, hogy az  $f_0$  emelésével egyre bizonytalanabb a megváltozott akusztikai szerkezetű magánhangzók észlelete, azaz egyre kisebb számban azonosítják azokat a hallgatók az ejtési szándék szerint (már 300–400 Hz alaphangfrekvencián 60–70%, 700 Hz körül pedig mindössze kb. 15–20%-ra csökken az azonosítás), és egyre nagyobb a válaszok szóródása a válaszlehetőségek mentén (Deme 2014b, 2015).

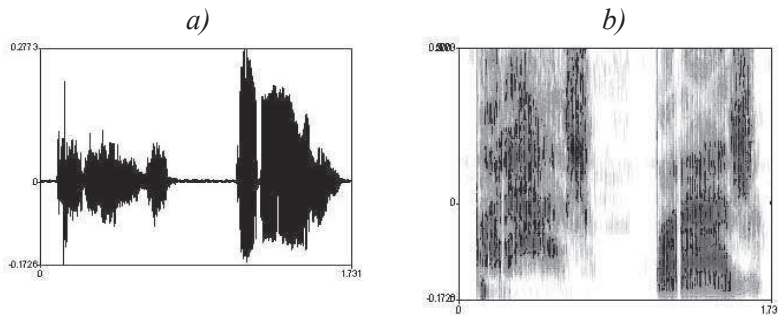
A magas alaphangfrekvenciájú szoprán éneklés során tehát gyakorlatilag nem ejtethetők és nem is ismerhetők fel a kotta szövegében feltüntetett magánhangzók. Feltehető, hogy ezek kizárólag azért ismerhetők fel mégis az ejtési szándék szerint, mert azokat az énekesek értelmes szavakban, mondatokban ejtik, ez pedig egészen egyszerűen leszűkíti a felismerésben szóba jöhető magánhangzókat, tehát nagyobb eséllyel „típelünk helyesen” az azonosítás során (Deme 2014b, 2015).

### 3. A suttogás

A suttogás a beszédképzés természetes módja. Általában olyan helyzetekben alkalmazzuk, amikor a társadalmi konvenciók tiltják a hangos beszédet, vagy ki akarjuk zárni az információközlésből a tágabb hallgatóságot. Ez a beszédmód az artikuláció tekintetében a hangszalagok működésében, az akusztikai szerkezet tekintetében pedig több paraméterben is eltér a normál fonációjú (zöngéképzésű) beszédjeltől, ami a beszédfeldolgozásra is hatással lehet. Az elsődleges artikulációs különbség a hangszalagok működésében található: míg normál ejtés esetén a szűk légzőállás (zöngétlen mássalhangzók létrehozásakor), a *h*-állás (a *h* mássalhangzó képzésekor) és a zöngéállás (magánhangzók és zöngés mássalhangzók létrehozásakor) felelős a hangképzésért, addig suttogáskor a suttogóállás. A jelen téma szempontjából a zöngé- és a suttogóállás különbségei a legérdekesebbek. A zöngé képzésekor a hangszalagok egymásnak feszülnek, és a tüdőből kiáramló levegő összegyűlik a szubglottális (gége alatti) területen. Amikor a nyomás elér egy bizonyos szintet, fölnyitja a hangszalagokat (a kannaporcok továbbra is egymásnak feszülnek), majd a nyomáscsökkenés hatására a hangszalagok újra összezárulnak, és a levegő újból elkezd a szubglottális területen feltorlódni.

A zöngéképzéskor ez a körfolyamat folyamatosan ismétlődik (mioelasztikus aerodinámiás elmélet, vö. Gósy 2004). Az így létrejövő kváziperiodikus rezgés szolgál a beszédképzés alapjául.

Suttogáskor a hangszalagok teljes hosszukban zárva vannak; a kannaporcok azonban nyitottak, és a tüdőből kiáramló levegő a kannaporcok közötti résen távozik. A kannaporcokon sűrűlő levegő miatt suttogáskor a hangképzés alapja aperiodikus rezgés. Normál ejtésű beszéd létrehozásakor a glottális területnek sokkal nagyobb részét használjuk föl a hangképzéshez (kb. 60–95%-át), mint suttogásnál (kb. 25%-át). Jellemző még a suttogásra egyfajta „préselés”, melyet a lefelé és befelé nyomódó hangszalagok eredményeznek. Ezen hangképzésbeli különbségek eredményeképpen a következő főbb vonások jellemzik a suttogott beszéd akusztikai szerkezetét a normál ejtésűvel szemben (5. ábra). A beszéd akusztikai formájára a zöngé hiánya, a beszéd zöreijessége (aperiodicitása), valamint a kisebb hangerő a jellemző. 500 Hz alatt a spektrum nagyon alacsony intenzitású, 500 és 2000 Hz között pedig kevésbé változatos intenzitás- és frekvenciaszerkezetű. A beszédhangokra jellemző akusztikai szerkezet eltérhet a normál fonációjú beszédben tapasztalható értékektől: a magánhangzók esetében például az első formáns gyengül, illetve a formánsértékek (főként az elsőé) megemelkednek (vö. Gósy 2002; Morris–Clements 2002). A hangsor intenzitásvizsgálatai is megváltoznak: míg a zöngétlen mássalhangzók intenzitását alig befolyásolja a fonáció módja, a magánhangzókét jelentősen csökkenti.



5. ábra. A parázs szó a) hullámformája és b) hangszínképe ugyanazon beszélő suttogott (bal oldalon) és normál (jobb oldalon) ejtésében

Magyar beszélők ejtésében felvett logatomok (jelentés nélküli „szavak”), szavak és mondatok vizsgálatában azt találták, hogy a mássalhangzók tévesztése 24–37% között változott annak függvényében, hogy logatomban (értelmes) szóban vagy mondatban hallották a vizsgált mássalhangzót, emellett a suttogott ejtésnek megfelelően a zöngés mássalhangzók azonosítása volt nehezebb a hallgatók számára. Értelmes szavakban kevesebb tévesztés fordult elő, mint a

másik két feladatban. Két magánhangzó közötti helyzetben több helyes észleletet kaptak, mint szókezdő vagy szóvégi helyzetben. A mondatok úgy voltak összeállítva, hogy az utolsó szó mássalhangzóinak helyes azonosítása alapján volt értelmes vagy értelmetlen a mondat. A hallgatók feladata az volt, hogy ítéljék meg, a hallott mondat tartalmilag értelmes vagy sem. Az eredmények alátámasztják, hogy az észlelést a beszédmegértési szint felülírhatja. A mássalhangzó zöngességétől függetlenül gyakoribb volt az értelmetlen mondatok értelmesként való azonosítása, mint az értelmes mondatok értelmetlenként való megítélése (Grácz 2005). Ezzel szemben a magánhangzók esetében az azonosítás jobb volt (a tévesztések aránya logatomokban 17,56%, szavakban 1,54% (Grácz 2006)). Jellemzően a nyelvvállásfokot tévesztették a hallgatók. Ennek oka az énekelte magánhangzóknál leírt észlelési mintázat: azaz a vokális zártságát az  $F_1 : f_0$  viszony alapján azonosítjuk, azonban ez az akusztikai kulcs suttogásnál a hangszalagrezgés hiányában nem elérhető az észlelés számára. A jelentés nélküli hangsorok esetében lényegében csak a felső, míg a szavakban a középső nyelvvállásfokú magánhangzóknál is fordult elő téves azonosítás.

A hangszalagrezgés hiányában suttogásban hiányzik az az alaphang is, amely alapján a beszéddallamot észleljük. Ugyanakkor esetenként-egyénekenként eltérő módon hordozzák ezek a jegyek a kérdő funkciót: formánsemelkedéssel, a formáns magasabb fekvésével, de olyan is tapasztalható, hogy nincs ilyen különbség a kijelentés és a kérdés között, stb. (Fónagy 1968). Egy kutatásban (Markó 2008) a kontextusukból kiemelt suttogott kijelentések és kérdések modalitását kellett a hallgatóknak megállapítaniuk. A suttogott kijelentések és kérdések között nem volt különbség a felismerési arány átlagában: az előbbi 71%, az utóbbi 73% volt. Ez azt jelenti, hogy a kísérleti személyek ugyanolyan arányban észlelték kérdésnek az állítást, mint fordítva. Bár a véletlen találatnál nagyobb mértékű az azonosítás, az adatok jól mutatják az észlelés bizonytalanságát is. Nyilvánvalóan a mindennapi közlési helyzetekben az azonosítás nagyobb mértékű, hiszen segíti a kontextus és/vagy a szituáció.

#### **4. Beszéd a gége eltávolítását követően: nyelőcsőbeszéd, elektromos gége és hangprotézis**

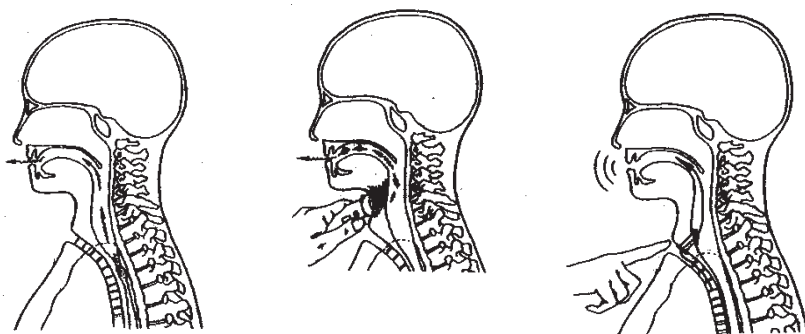
Magyarországon évente több száz esetben végeznek teljes gégeeltávolítást (laryngectomia), általában valamilyen rákos megbetegedés következményeképpen. Ezt követően – hangszalagok hiányában – a beszélő nem tud a megszokott módon zöngét létrehozni, hanem valamilyen alternatív beszédképzési módot kell elsajátítania. Ezek közül a két leggyakoribb a nyelőcsőbeszéd (latin eredetű kifejezéssel *özofágusbeszéd*) és a külső gerjesztéssel, azaz elektromos gégével létrehozott beszéd. A műtéti technika fejlődésének köszönhetően egyre elterjedtebb az ún.

hangprotézis, más néven Provox segítségével történő zöngé-előállítás is.

A nyelőcsőbeszéd megtanulása a betegek számára annak a technikának a begyakorlását jelenti, hogy a gyomorba nyelt levegőt a nyelőcsőn át kiengedve és a nyelőcső felső szakaszán lévő nyálkahártya-redőzetet mintegy póthangszalagként használva hozzák létre a zöngét helyettesítő hangot. A betegek egy része olyan jó minőségű hangot tud így előállítani, hogy a hallgatók úgy érzékelik, mintha csak egy kissé rekedt lenne a beszélő.

A külső gerjesztés használata esetén a beszélő a nyaki lágyrészre illeszti az elektromos gégét, ez a készülék hoz létre egy állandó (100 Hz körüli) frekvenciájú alaprezgést, miközben a toldalékcső beszédszervei a megfelelő artikulációs mozgásokat kivitelezik.

A hangprotézis (Provox) segítségével történő zöngé-előállítás esetén a nyelőcső és a légcső közé illesztett szelep teszi lehetővé a levegő átáramlását (és gátolja meg az ételnek a légcsőbe jutását). A hangadáshoz szükséges levegő a tüdőből, a légcső felől, a szelepen keresztül jut a nyelőcsőbe, ahol a garatfűző izomzatból és a nyálkahártyaredőkből kialakult póthangszalag rezgései veszik át a kivett gége hangképző szerepét (6. ábra; vö. Kiefer–Répássy 1997).



6. ábra. A nyelőcsőhangképzés (balra), az elektromos gége (középen) és a hangprotézis (jobbra) működési elve (Forrás: Kiefer–Répássy: 1997)

A vizsgálatok azt mutatják, hogy ezen hangképzési módok esetében a beszédjel akusztikai szerkezete – függetlenül az adott nyelvtől – jelentősen megváltozik, így a percepcióra gyakorolt hatásai is eltérnek a normál fonációjú beszédétől (vö. Christensen–Weinberg 1976, Robbins et al. 1984, Balázs–Gósy–Szabó 1996, Gósy 2002).

Egy kutatássorozatban (Markó–Grácsi 2007) elektromos gégével és nyelőcsővel létrehozott beszédet vizsgáltak különféle percepciós tesztekben. Ezek az ún. attitűdvizsgálatok azt az általános kérdést hivatottak megválaszolni, hogy a gé-

geeltávolításon átesett betegeknek a társadalomba való visszailleszkedését milyen kommunikációs tényezők befolyásol(hat)ják, és ezek mennyiben függenek az alkalmazott beszédképzési technikától. A főbb eredmények szerint az elektromos gégét használók elégedettebbek a beszédmódjukkal és a kommunikációs hatékonyságukkal, mint a nyelőcsőbeszédet alkalmazók; ugyanakkor a(z egészséges) hallgatók értékelése szerint mind a beszéd minősége, mind pedig az érthetősége szempontjából a nyelőcsőbeszéd bizonyult jobbnak. Azok a beszélők, akik a nyelőcsőbeszédet használják, azért választották az elektromos gégével szemben ezt a beszédmódot, mert az elektromos gégével gerjesztett hangot túlzottan gépinek találták. Ez felveti azt a magyarázati lehetőséget is, hogy a mindennapjaikban csak az elektromos gégét használók – legalábbis részben – azért döntöttek emellett, mert nem tudták vagy nem akarták megtanulni a nyelőcsőbeszédet (ami komoly erőfeszítést igényel). Ugyanakkor az elektromos gégével beszélők közül sokan a nyelőcsőbeszédéről mondták azt, hogy mind a létrehozásának a módját, mind a hangzását természetellenesnek találják. Felmerül persze, hogy ez a vélemény valamiféle kompenzálás következményeként alakult-e ki bennük. (Megjegyezzük, hogy a gégeeltávolítás után az orvosok mindenkinek felírják az elektromos gégét, a beteg választásától függ, hogy használja-e, illetőleg meddig használja az eszközt, valamint, hogy vált-e nyelőcsőbeszédre.) A nyelőcsőbeszédet használók a beszéd minősége és érthetősége tekintetében nagyobb mértékben térnek el egymástól, mint az elektromos gégét alkalmazók. Ez valószínűleg azzal függ össze, hogy a nyelőcsőbeszéd minőségét meghatározza az elsajátítás és az alkalmazás sikeressége, míg az elektromos gége használata nem igényel különös képességet, körülbelül azonos szintű/minőségű beszédet tudnak vele előállítani a betegek, nincsenek nagy egyéni különbségek. Az ezekhez a beszédmódokhoz való hozzászokás a potenciális hallgatók részéről a mindennapi életben fontos tényező. A hozzászokás ugyanakkor a kísérleti helyzetben csak a beszédminőség megítélését érintette. Az érthetőség mértéke a nyelőcsőbeszéd (akárcsak a normál zöngéképzés) tekintetében nem függ az alkalmazkodástól, ugyanakkor az elektromos gége érthetősége második hallásra gyengébbnek mutatkozott, ami a hallgatói figyelem nagyfokú igénybevételére utal. Ezek szerint a nyelőcsőbeszéd a hallgató szempontjából közelebb áll a mindennapokban megszokott normál képzésű beszéd akusztikumához, hiszen ennek érthetőségét nem befolyásolta a megszokás, vagyis nem nehezítette meg jelentősen a feldolgozást, hogy a hallgatók közül a legtöbben még nem hallottak ilyen beszédet. Egybevág ezzel, hogy a nyelőcsőbeszédet alkalmazók nagy része szerint beszédpartnereik azt gondolják, hogy rekedtek, de a hangzás alapján nem feltételeznek nagymértékben eltérő beszédképzést.

Ezekben a beszédmódokban – a normál zöngeműködés és alaphérfrekvencia-moduláció hiánya miatt – különösen lényeges problémakör az (lexikálisan nem jelölt) eldöntendő kérdés létrehozása elektromos gégével és nyelőcsőbeszédvel,

illetőleg hogy a hallgatók ezeket a közléseket kérdésként azonosítják-e (vö. Markó–Grácz 2007 és 2008). A magyar eldöntendő kérdő mondatot a szakirodalom alapján egyértelműen a szakaszvégi dallam különbözteti meg a kijelentőtől (vö. pl. Fónagy–Magdics 1967). A hangtan elmélete a dallamot hagyományosan (és értelemszerűen) artikulációsan a hangszalagok rezgésének következményeként, akusztikailag pedig „az alapprofrecuencia ( $f_0$ ) folyamatos és célzott változtatásának (modulálásának)” (Gósy 2004: 187) eredményeként határozza meg. A hazánkban legelterjedtebb Servox típusú elektromos gégek állandó alaphangot gerjesztenek, és automatikusan sem a hanglejtés, sem az intenzitás modulációjára nem adnak módot. Ugyanakkor megtanulható az a módszer, amellyel – ha nem is természeteshez közeli moduláció, de – frekvenciaváltás idézhető elő. Ezen az elektromos gégen ugyanis két gomb van egymás felett, és ezek megnyomásával két különböző frekvenciájú alaphangot lehet előállítani. Ily módon a gombok közti „lépkedéssel” hangnemet lehet váltani, és kvázi lokális frekvenciacsúcsot lehet „szerkeszteni” a mondanivalóba. Az újabb típusú elektromos gégek (pl. TruTone, Blom-Singer) alkalmasabbak a dallammodulálásra, ezek azonban hazánkban még nem terjedtek el.

A nyelvcsőbeszéd és a hangprotézis (Provox) technikájával megoldható a hang modulálása – ha nem is olyan mértékben és olyan értékhatárok között, mint a normál zöngé esetében (vö. Sáfrán–Subosits 1980). Egy percepciós vizsgálatban (Markó 2008) a nyelvcsőbeszéd esetében közel 15%-os eltérés volt tapasztalható az állítások és a kérdések felismerésének sikeressége között: az előbbi modalitást átlagosan 99%-ban, az utóbbit 84%-ban azonosították a hallgatók. (Megjegyezzük, hogy a suttogott kérdés esetében rosszabb volt a találati arány.) A mindennapi beszédhelyzetekben nyilvánvalóan a szituáció és a kontextus is segítheti e közlések kérdésként való azonosítását. Ugyanakkor a nyelvcsőbeszédet alkalmazó interjúalanyok is arról számoltak be, hogy bár általában nem okoz gondot számukra a kérdezés, előfordul, hogy nekik is meg kell erősíteniük ezt a szándékukat verbálisan.

## 5. A hasbeszélés

A hasbeszélés (idegen szóval *ventriloquia*) természetesen valójában nem hasból beszélés. Olyan beszédképzésről van szó, amely közben az állkapocs és az ajkak mozgásai minimálisra redukálódnak, és a hangzók minőségének előállítása a beszédképző szervek módosított pozícióival történik. A beszéd így érthető marad, de több tekintetben eltér a normál képzési módtól. Erre a képzési módra jellemző, hogy a rekeszizom – akárcsak belégzéskor – mindvégig le van süllyesztve. Ezáltal a levegő sokáig bent tartható, és felhasználható beszédképzésre. A gégeizmok feszítettebbek, mint normál zöngéképzéskor, emiatt a hasbe-

szélésre általában magasabb alaphang jellemző. A gégefedő porc részlegesen zárja a gégeenyílást, és a garat is szűkebb, mint a normál beszédprodukciónban. Ezek is azt a célt szolgálják, hogy a levegő lassabban távozzon, így egy levegővel hosszabb beszédegységeket lehessen létrehozni (Kassai 1998). Bár a beszédképző szervek nagyfokú alkalmazkodóképességének révén az egyes beszédhangok még az állkapocs- és ajakmozgás redukciója mellett is viszonylag jól megvalósíthatók és elkülöníthetők egymástól a produkcióban, egyes, főleg az ajkakkal képzett beszédhangok (pl. *m*) ejtése ellehetetlenül, így ezeket a hasbeszélő helyettesíteni kényszerül (pl. *m* → *n*). Ugyanakkor a „hasbeszéd” még így is általában teljesen érthető, annak köszönhetően, hogy a mindennapi közlések megértésekor nem pusztán az akusztikai információkra (azok dekódolására) hagyatkozunk. A mindennapi közlésekben ugyanis a beszédhangokat mindig szavakba, mondatokba, tágabb kontextusba ágyazva halljuk, így bizonyos hangsorokat nagyobb valószínűségűnek súlyozunk másoknál, azaz „azt halljuk, amit hallani szeretnénk”. Emiatt ha egy hasbeszélő *Simon* helyett *Sinont* ejt, valószínűleg nem fogjuk észrevenni a csalást. Ehhez az is hozzájárul, hogy a „művész” kezében lévő bábura figyelünk.

### Irodalom

- Balázs Boglárka – Gósy Mária – Szabó Iván 1996. A gége nélküli beszéd fonetikai sajátosságai. *Beszédkutatás* 1996. 58–71.
- Bresch, Erik – Narayanan, Shrikanth 2010. Real-time magnetic resonance imaging investigation of resonance tuning in soprano singing. *Journal of the Acoustical Society of America* 128/5. EL335–EL341.
- Carlson, Rolf – Granström, Björn – Fant, Gunnar 1970. Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 2–3. 19–35.
- Chistovich, Ludmilla A. – Lublinskaya, Valentina V. 1979 The ‘center of gravity’ effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research* 1/3. 185–195.
- Christensen, John M. – Weinberg, Bernd 1976. Vowel duration characteristics of esophageal speech. *Journal of Speech and Hearing Research* 19. 678–689.
- Deme, Andrea 2014a. Formant strategies of professional female singers at high fundamental frequencies. In Fuchs, Susanne – Grice, Martine – Hermes, Anne – Lancia, Leonardo – Mücke, Doris (eds.): *Proceedings of the 10th International Seminar on Speech Production Cologne (ISSP)*. Köln, 90–93.
- Deme, Andrea 2014b. Intelligibility of sung vowels: the effect of consonantal context and the onset of voicing. *Journal of Voice* 28/4. 523.e19–523.e25.
- Deme Andrea 2015. Speech perception at its best: Extracting linguistic information from phonetically underspecified input. In: *Proceedings of the 18th International Congress of Phonetic Sciences*. University of Glasgow, Glasgow, UK. <http://www.icphs2015.info/pdfs/Papers/ICPHS0168.pdf>

- Epps, J. – Smith, John R. – Wolfe, Joe 1997. A novel instrument to measure acoustic resonances of the vocal tract during speech. *Measurement Science and Technology* 8. 1112–1121.
- Fant, Gunnar 1960. *Acoustic theory of speech production*. Mouton & Co., The Hague.
- Fónagy Iván 1968. Suttogott dallam? *Magyar Nyelvőr* 92. 253–262.
- Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama*. Akadémiai Kiadó, Budapest.
- Gósy Mária 2002. Beszédképzés zöngé nélkül. *Beszéd kutatás 2002*. 18–37.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Joliveau, Elodie – Smith, John – Wolfe, Joe 2004. Vocal tract resonances in singing: The soprano voice. *Journal of the Acoustical Society of America* 116/4. 2434–2439.
- Kassai Ilona 1998. *Fonetika*. Nemzeti Tankönyvkiadó, Budapest.
- Kiefer Gábor – Répássy Gábor 1997. A hangrehabilitáció akusztikai-fonetikai eredményei teljes gégeeltávolítás és Provox hangprotézis implantációja után. *Beszéd kutatás 1997*. 180–187.
- Markó Alexandra 2008. A magyar eldöntendő kérdő mondat megvalósulásai különböző beszédmódokban. *Beszéd kutatás 2008*. 77–92.
- Markó Alexandra – Grácz Tekla Etelka 2007. Gégeeltávolításon átesett betegek beszédének hallgatói megítélése. *Alkalmazott Nyelvtudomány* VII/1–2. 39–55.
- Markó Alexandra – Grácz Tekla Etelka 2008. Az eldöntendő kérdő modalitás megvalósulása a nyelöcsőbeszédben. *Alkalmazott Nyelvtudomány* VIII/1–2. 147–158.
- Morris, Robert W. – Clements, Mark A. 2002. Reconstruction of speech from whispers. *Medical Engineering and Physics* 24. 515–520.
- Peterson, Gordon E. – Barney, Harold L. 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24/2. 175–184.
- Robbins, Joanne – Fisher, Hilda B. – Blom, Eric C. – Singer, Mark I. 1984. A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders* 49. 202–210.
- Sáfrán Antal – Subosits István 1980. A nyelöcsőbeszéd és a hangprotézises beszédproduktó dallamformái. In: *A Bárczi Gusztáv Gyógypedagógiai Tanárképző Főiskola Évkönyve IX*. BGGyTF, Budapest, 337–346.
- Scotto di Carlo, Nicole – Germain, Aline 1985. A perceptual study of the influence of pitch on the intelligibility of sung vowels. *Phonetica* 42/2. 188–97.
- Sundberg, Johan 1975. Formant technique in a professional female singer. *Acta Acustica united with Acustica* 32/2. 89–96.
- Sundberg, Johan 1977. The acoustics of the singing voice. *Scientific American* 234/3. 82–91.
- Sundberg, Johan 1987. *The science of the singing voice*. Northern Illinois University Press, DeKalb Illinois.
- Stevens, Kenneth N. 1998. *Acoustic phonetics*. The MIT Press, Cambridge, Massachusetts–London.
- Traunmüller, Hartmut 1981. Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America* 69/5. 1465–1475.
- Vicsi Klára 2010. A beszéd fizikai jellemzése. In Németh Géza – Olasz Gábor (szerk.): *A magyar beszéd. Beszéd kutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest, 38–56.

### Kérdések, feladatok

1. Ítélje meg, igaz-e a következő állítás (válaszát indokolja!): Az éneklés során a magasabb alaphérfrekvenciákon a szopránok nyíltabban ejtik a magánhangzókat, amiket a hallgatók is nyíltabb magánhangzókként azonosítanak.
2. Milyen beszédhelyzetekben jellemző a suttogott beszéd?
3. Keressen rá videómegosztó portálokön a *laryngectomy*, *provox*, *electrolarynx*, *esophageal speech* kifejezésekre, és figyelje meg, milyen módon hoznak létre beszédet a teljes gége-eltávolításon átesett páciensek. Milyen ezeknek a beszédmódoknak a hangzása?
4. Keressen rá a *ventriloquism* kifejezésre videómegosztó portálokön, és figyelje meg a hasbeszélés sajátosságait!

## A semleges magánhangzóról, azaz a sváról

Biztos mindannyian megfigyeltük már magunkon és beszédpartnereinken, hogy minél fesztelenebb egy társalgás, annál „lazábban” beszélünk, vagy épp ellenkezőleg, formális szituációban törekszünk a jól artikulált beszédre. Hogyan hat vajon ez az egyes hangok képzésére? Gondolkodott már azon, hogy ténylegesen melyik hang volt az, amelyiket gyermekként először kimondott? Vajon van olyan magánhangzó a hangrendszerünkben, amelyik több funkcióval rendelkezik, mint a többi? Mindezen kérdésekre választ kaphatunk, ha közelebbről megvizsgáljuk a semleges magánhangzót, azaz a svát.

### 1. A magánhangzók képzése és akusztikai jellemzői

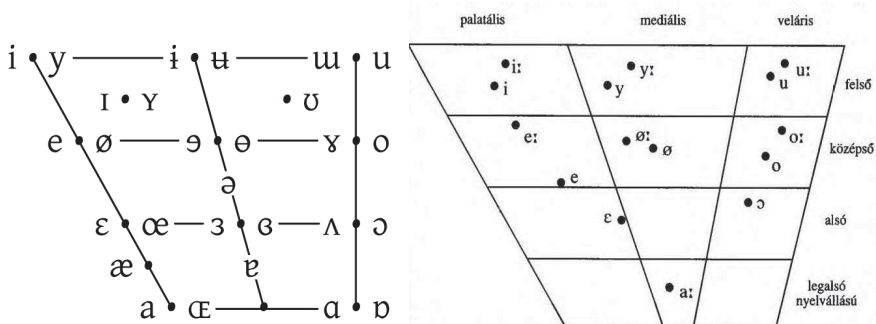
Az egyes beszédhangokat a képzési jegyeik határozzák meg, amelyek megadják, hogy képzésükkor a beszédszervek milyen pozícióban vannak, vagy milyen mozgást végeznek (= artikulációs gesztusok). Ezeknek az összessége az artikulációs konfiguráció. A magánhangzók képzésére a teljes szájüreg a rendelkezésünkre áll, és a relatív rugalmassága nagymértékben érvényesülhet az artikuláció folyamán (Gósy 2006). Vannak azonban esetek, amikor nem sikerül a megfelelő artikulációs konfigurációt, azaz célkonfigurációt megvalósítani. Beszédünk során az egyes magánhangzók kisebb-nagyobb mértékben módosulhatnak, redukálódhatnak (Gránczi–Horváth 2010). Ilyen, a célkonfigurációtól eltérő fonéma realizáció lehet a semlegesedés, vagyis az, amikor egy adott nyelv magánhangzórendszerében bizonyos magánhangzók, amelyek képzésére egyébként nem a centrális ejtés jellemző, mégis középpontivá válnak. Lindblom (1990) elméletében a semlegesedést a beszélői szándék felől értelmezi. A beszédprodukciónál két fő szempontot feltételez a beszélőnél: az érthetőségre törekvést, valamint az energiaráfordítás minimalizálását.

A tüdőből kiáramló levegő rezgésbe hozza a hangszalagokat, ha azok a megfelelő állásban vannak. Így jön létre a zöngé, amely a magánhangzók és zöngés mássalhangzók képzésének alapja. A garat-, száj- és orrüreg rezonátorként működik, ami a zöngé frekvenciáiból bizonyos felhangokat/felharmonikusokat felerősít. Ezeket nevezzük formánsoknak, amik megjelennek a magánhangzó akusztikumában. Az egyes magánhangzókat az első két formáns ( $F_1$ ,  $F_2$ ) egyértelműen meghatározza (Gósy 2004). Az első formáns ( $F_1$ ) értéke az állkapocs nyitásszögével, a nyelvállás fokával van összefüggésben, míg a második

formáns ( $F_2$ ) értéke a nyelv vízszintes mozgásával és az ajakműködéssel. Számos, különböző szempontú vizsgálat irányult már a magánhangzók formánsszerkezetére: különböző nyelvekben (pl. amerikai angol, svéd, német) végeztek például kutatásokat arra nézve, hogy különböző beszédstílusokban hogyan valósul meg a magánhangzó formánsszerkezete. Az eredmények alapján feltételezhető, hogy a lezser, társalgási beszédstílusban a magánhangzók első két formánsértékei nagyobb átfedést mutatnak, mint pl. a tiszta, érthető beszédre való törekvés esetében.

## 2. Mi a svá?

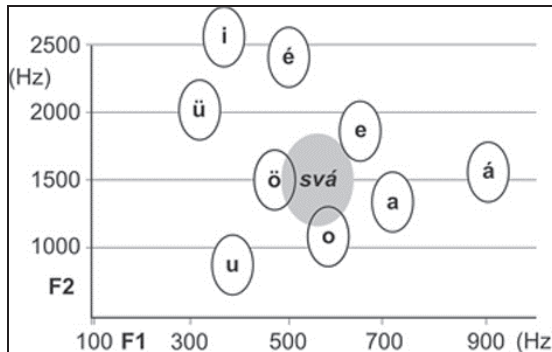
Annak érdekében, hogy le tudják írni a magánhangzókat, Daniel Jones 1922-ben megalkotta a kardinális magánhangzók rendszerét (1. ábra). Ebben a rendszerben eredetileg nyolc magánhangzó volt található. Ezek nem egy adott nyelv magánhangzóit, hanem a szájüregi lehetőségek és az észlelés összefüggésének eredményeképpen meghatározott magánhangzó-lehetőségek, tulajdonképpen viszonyítási pontok (Gósy 2004). Később Jones létrehozta a másodlagos kardinális magánhangzók rendszerét a pontosabb, részletezőbb leírás érdekében, amely további 22 típust tartalmazott.



1. ábra. Balra a Jones-féle IPA trapéz alakú magánhangzórendszer látható, amely körülbelül leképezi a nyelv mozgását a szájüregben; jobbra ennek a Bolla Kálmán által a magyar nyelvre készített átirata (Bolla 1995: 211)

Többen kritikával illették ezt a rendszert legfőképp amiatt, hogy összekeverednek benne az artikulációs és az észlelési kritériumok. Mindezek ellenére a mai napig ez a legelterjedtebb és legjobban használható rendszer egy nyelv magánhangzóinak leírására. Ennek a magánhangzó-rendszernek a közepén található a semleges magánhangzó, a svá (2. ábra), ugyanis ejtésekor a nyelv a szájüregben lényegében középen van a nyugalmi helyzethez nagyon közel, a

vokális csatornára a semleges konfiguráció a jellemző, az ajkak nem kerekítettek (Gósy 2006).



2. ábra. A svá helye a magánhangzó-rendszerben (Gósy 2007: 53)

A semleges magánhangzót többféleképpen nevezik a szakirodalomban: homályos, meghatározhatatlan, redukált, centrális, mormolt, illetve hangsúlytalan szótagi magánhangzó. Ugyanakkor a leggyakoribb megnevezése a svá, jelölése: [ə]. A szó héber eredetű, jelentése 'semmisség'. Hangzásában a magyar *ö* magánhangzóhoz áll a legközelebb. Gyakran határjelző hangváltozatként, átmeneti hangként, hangátmenetként jelenik meg. Így felvetődik a kérdés, hogy melyik hanghoz kell sorolni a svát, illetve okvetlenül más hanghoz tartozónak kell-e tekinteni.

A svá megjelenése a nyelvek többségében a hangsúly hiányával összefüggő jelenség, mivel a nem hangsúlyos magánhangzók helyett áll a hangsorban, tehát egy fonéma redukált változata. Van azonban egy másik változat is, ez pedig a hangsor különböző pontján egymást követő mássalhangzók között vagy mögött keletkező átmeneti hang, amelyről a fonetikai kézikönyvek általában nem vesznek tudomást, s ha mégis, akkor mint „helytelen” említik.

## 2.1. A svá első megjelenései a szakirodalomban

A megfordított *e* betű [ə] régóta megtalálható a fonetikában a maival megegyező fonetikai szimbólumértékkel. Elsőként Johann Andreas Schmeller használta 1821-ben. A svá megnevezés azonban csak később, 1895-ben az angolban Peter Giles egy filológiáról szóló tankönyvében jelent meg először.

A magyar hangtan történetében hárman foglalkoztak vele. Kyss Sándor *vocalis elementaris*-nak nevezte, és úgy gondolta, hogy orrhangú magánhangzó, amit jellemzően zárt szájjal ejtenek és csak négy mássalhangzóval alkothat szótagot, ezek a [l r m n]. Matúsik András azonban megkérdőjelezi mindazokat a

kijelentéseket, amit Kyss Sándor tett a semleges magánhangzóról: nem tekinti orrhangúnak, és a zárt ejtést sem tartja jellemzőnek. Mindkettlen megfigyelték ugyanakkor, hogy a betegek nyögésében, a sírásban, a nevetésben, valamint a morgásban is ez a hang hallható (Vértes 1980: 119). Kerekes Ferentz (1836) már a felpattanó zárhangok ejtésével hozza kapcsolatba a semleges magánhangzó megjelenését, amelyet a héber „scheva” hanghoz hasonlónak talál.

## 2.2. A svá artikulációs és akusztikai jellemzői

A svá jellemzően középen képzett, középső nyelvállású és ajakréses magánhangzó, amely általában rövidebb időtartamú, mint a nyelvileg rövid magánhangzók. Spontán beszédben leggyakrabban a nem pontos artikulációnak és bizonyos fonetikai helyzetekben a fiziológiai kényszernek köszönhetően fordul elő (Gósy 2004).

A semleges magánhangzó artikulációs konfigurációja látszólag egyértelműen meghatározott, mégis összevetve más magánhangzókkal azt találjuk, hogy sokkal nagyobb a megvalósításának a variabilitása. A nyelv vízszintes és függőleges helyzetének kismértékű, de különféle változása (pl. kissé magasabban, hátrább vagy előrébb képezve a szájüregben) is eredményezheti a semleges magánhangzó észleletét. A világ nyelveiben különbözőképp jellemezhetők a svák. Az artikulációs sokféleségnek köszönhetően a nemzetközi kutatások a svá első és második formánsaira jelentős mértékben eltérő adatokat közöltek (Gósy 2006).

Számos kutatás foglalkozott annak vizsgálatával, hogy beszéd során az egyes magánhangzók milyen mértékben redukálódnak és ez mely paraméterben érhető tetten. Koopmans-van Beinum (1994) munkatársaival végzett mérései azt mutatták, hogy az első formáns redukció során erőteljesebben változik, mint a második formáns, ugyanakkor más kutatások ennek az ellenkezőjét találták: a semlegesedést a második formáns értékeiben lehet tapasztalni.

A semlegesedés tekintetében vizsgálták a nemek között eltéréseket is. Az eredmények azt mutatták, hogy a férfiak beszédében jellemzőbb a célalulmúlás (Henton 1995), azaz az artikulációs gesztusok nem érik el a megfelelő konfigurációt; a nők magánhangzótere pedig tágabb a férfiakénál. A nemzetközi szakirodalomban a magánhangzóter a magánhangzók  $F_1$  és  $F_2$  koordinátái mentén ábrázolt vokális térként jelenik meg (vö. Vorperian 2007, IPA Handbook 1999).

Beke (2009) összevetette a magyar nyelv esetében a veláris magánhangzó alap- és svás realizációit, és kutatásában szignifikáns eltérést adatolt az időtartam és a második formáns értékeiben a két realizáció között. Az első formáns esetében nem volt statisztikailag jelentős eltérés, azonban tendenciájában magasabb volt a semlegesedett ejtésben, mint az alaprealizációban. Eltérő volt a semlegesedés annak függvényében, hogy az adott szótag hangsúlyos vagy hangsúlytalan volt: a hangsúlytalan szótagok magánhangzói gyakrabban semlegesedtek.

Beke és Grácsi (2010) 19 beszélő esetében vizsgálták a magyar magánhangzók spontánbeszédbeli semlegesedését. Az eredmények azt mutatták, hogy az egyes magánhangzók képzési jegyeik alapján változó arányban semlegesednek a spontán beszédben. Az időtartam tekintetében a magánhangzók alaprealizációi között nagymértékű különbségeket adatoltak, de az alap- és semleges realizációk összevetésében csak kevés esetben volt szignifikáns a különbség. A második formáns minden esetben, azonban az első formáns csak a kardinális magánhangzók akusztikai szerkezetében volt eltérő az alap- és semleges realizációk között. A beszélők nemének tekintetében nagymértékű különbségeket találtak. Eltérések voltak a férfiak és a nők ejtésében a semlegesedés aránya mellett az alap- és a semleges realizációk, valamint azok eltérésének mintázata is.

Az, hogy a korábbi kutatások nagymértékben különböző akusztikai adatokat közöltek a semleges magánhangzóra, kérdéseket vet fel:

Hogyan határozható meg a semleges artikuláció?

Nyelvspecifikus-e a semleges magánhangzó képzése? Az adott nyelv magánhangzórendszere, illetőleg a nyelvet beszélők észlelése határozza-e meg a semleges minőségűnek tekintett állapotot?

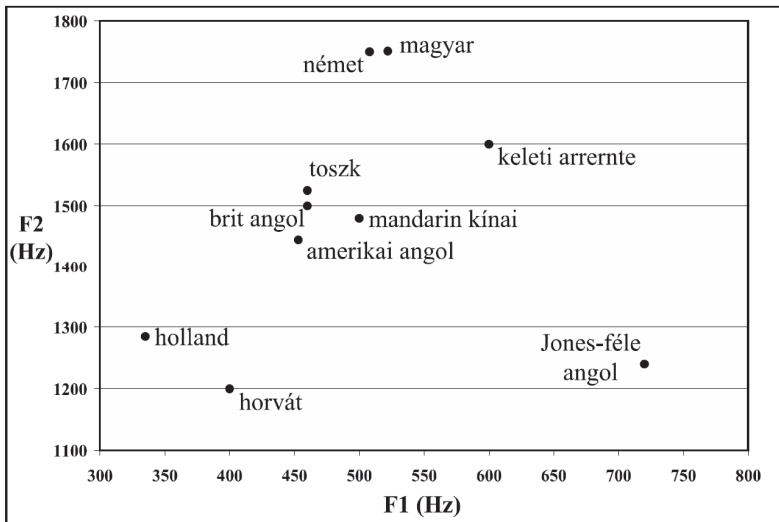
### **2.3. A svá a különböző nyelvekben**

A svá a világ számos nyelvében megtalálható, de ejtése nagymértékben különbözik (3. ábra). A svá artikulációs konfigurációja elméletileg semleges, azonban az aktuális frekvenciaszerkezete nagy eltéréseket mutat nemcsak az egyes nyelvek között, hanem sok esetben egy adott nyelven belül is (Gósy 2006).

A különféle nyelvek magánhangzórendszerében azt látjuk, hogy a semleges magánhangzó:

- lehet fonéma, vagyis a szavak jelentéseit különítheti el (pl.: az amharik, a sindi vagy a szlovén nyelvben);
- megjelenhet egyszerű változatként, mellékvariánsként (pl.: a horvátban, a katalánban vagy az írben);
- létrejöhet valamiféle szabály eredményeként (pl.: a hollandban);
- lehet egyidejűleg fonéma és variáns (pl.: a németben);
- jelezheti másik magánhangzó hiányát (pl.: a héberben);
- helyettesíthet más magánhangzókat a gyors beszédtempó és elnagyolt artikuláció eredményeképp (pl.: az angolban 22,9%-ban, a németben 30%-ban, a franciában 7,6%-ban, a magyarban akár 30–40%-ban).

Akusztikai vizsgálatok azt mutatják, hogy a svá ejtése nem annyira stabil, mint a többi magánhangzóé, mindig az adott nyelv magánhangzórendszerében értelmezve tekinthető a svá centrálisnak, így semlegessége inkább a percepcióban fogható meg. Ez azt jelenti, hogy más nyelvek semleges magánhangzóit pél-



3. ábra. A svá formánsértékei a különböző nyelvekben (Gósy 2006: 10)

dául egy magyar anyanyelvű nem feltétlenül fogja svának hallani az elhangzásukkor (Gósy 2006, 2007). Az albániai toszk nyelvben azt találták, hogy a svát egyre inkább a szájüreg hátsó részében képezik a beszélők, ezzel szemben a holland mérések azt mutatták, hogy nyelvükben a svá relatíve stabil helyzetű a többi magánhangzó viszonylatában. A magyar nyelvben azt találták, hogy a svá létezik a semlegesnek tekinthető ejtésben, ugyanakkor tendenciaszerűen egyre gyakrabban előfordul, hogy a beszélők előrébb képzik.

## 2.4. A svá funkciói a magyar nyelvben

A magyar nyelvben a svá és az *ő* magánhangzó akusztikai szerkezete döntően különbözik. Ahogy korábban említettük a svá középen képzett, középső nyelvállású és ajakréses magánhangzó, az *ő* magánhangzó pedig elől képzett, középső nyelvállású és ajakkerekítéses. Még a gögicsélt svák és a gögicsélt *ő* hangok is szignifikáns különbséget mutatnak (Gósy 2006). Ez azt jelenti, hogy a svá nemcsak nyelvenként lehet különböző, de egy nyelven belül is több nyelvi funkcióval rendelkezik, a beszélők feltételezhetően nem tudatosan, de bizonyos fokig eltérően ejtik a semleges magánhangzót attól függően, hogy milyen nyelvi funkcióban használják. A magyarban összesen nyolc különböző funkcióban fordulhat elő (a gögicsélésen túl), és ezek a beszédprodukciónak különböző szintjeihez kapcsolódnak.

**Gögicsélés:** Olyan „nem mássalhangzó”, amely gyakorlatilag bármely más magánhangzó-minőség kiváltására képes. E funkciójú svá lényegében úgy jön

létre, hogy a csecsemő zöngét képez, a nyelve azonban semleges helyzetben, mintegy középen van a szájüregben, s az ajkai kismértékben nyílnak szét. A felnőtt környezet pedig a hangélményt az *ö*-vel azonosítja.

Számos esetben előfordulhat, hogy a beszélő különböző (érzelmi, értelmi vagy egyéb) okokból **hangsúlyozni** kíván valamit a beszéde során. Erős nyomatók/feszítettebb ejtés: ha feszítettebb izmokkal képezzük a hangokat, nagyobbak a hangképző szervek kitérései, ezáltal szinte minden mássalhangzó után fellép a svá mint a hangképző szervek tehetetlenségének, a lassúbb elernyedésnek a következménye.

Gyakran a **gondos hangképzés szükségszerű velejárója**: amikor a két hang közül az első képzésének a végén a hangképző szervek nem a legrövidebb úton térnek át a második hangra jellemző helyzetbe, akkor a megteendő út jelentősen meghosszabbodik, amely úton átmeneti hang keletkezik. Ez a siketeknél gyakran hallható jelenség.

Szintén megjelenik svá a **szóhatár, illetve a morfémahatár jelzése** céljából: bizonyos szegmentális homonímia kialakulása esetén megakadályozza az igazodások létrejöttét – ehhez az kell, hogy az egymáshoz igazodó hangokat egy vokális elem elválassza például: *halálok* de *halál ok*. Ez tudatosnak tekinthető.

Megjelenik **koartikulációs következményként**; ekkor a beszélő szintén szándéka ellenére (fiziológiai kényszerből) ejti bizonyos hangkörnyezetben a semleges magánhangzót. Három jellegzetes kontextusfüggő előfordulása van: a) a felpattanó zárhang és a nazális kapcsolatában (pl. *gnú* = *g[ə]nú* vagy *teknő* = *tek[ə]nő*), b) bizonyos mássalhangzók után abszolút szóvégi helyzetben (pl. *adag[ə]* vagy *véd[ə]*) (4. ábra), valamint c) a pergőhang képzésében. Annál erősebb ez kényszer, minél eltérőbb hangok követik egymást.

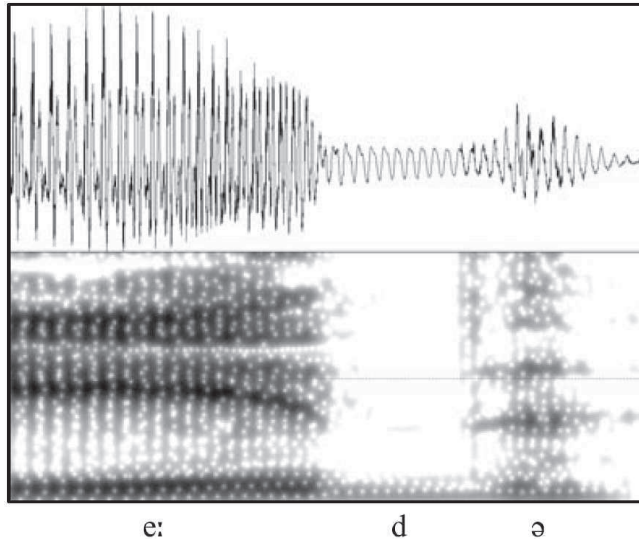
Előfordulhat **helyettesítő** funkcióban is, ilyenkor a kívánt magánhangzó-mi-nőség helyett a beszélő a svát ejti. Illetve hármas és négyes mássalhangzó-kapcsolatok esetében, ha az összes fonémát meg akarjuk valósítani a beszédben, akkor a kieséstől fenyegetett hangot meg kell támogatni egy hangzósabb vokális elemmel.

Gyakran találunk svát **szóvégi helyzetben**, amely az artikulációs tervezéssel van kapcsolatban. A mássalhangzók lecsengése eredményeként a szó végén előforduló egyes mássalhangzók (ha nem követi őket újabb szó) csak úgy ejthetők ki, hogy hallható lesz a svá.

**Pergőhang része**: a hangszalagok már rezegnek az *r* mássalhangzó képzése előtt és a nyelv a semleges magánhangzó pozíciójában van. Az *r* mássalhangzónk szó eleji helyzetben egy rövid svát tartalmaz, azonban ezt észre sem vesszük.

A spontán beszéd egyik jellemzője, hogy a beszélő a tervezési bizonytalanság, illetve az átmeneti szótalálási nehézségek következtében megakad, ami nem-egyszer hangadással jár. Ezt nevezzük **hezitálásnak**, ismertebb nevén „özés-

nek”. Az a törekvés hozza létre, hogy beszéd közben a beszéd folyamatossága ne szakadjon meg hosszú időre (Kassai 1978; Gósy 2006).



4. ábra. A semleges magánhangzó a *(v)éd* szóban mondatvégi helyzetben (Gósy 2006: 12)

## 2.5. A svá formánsszerkezete a különböző funkciókban

Egyre inkább a fonetikai kutatások középpontjába került az a kérdés is, hogy vajon a funkcióknak megfelelően van-e különbség a semleges magánhangzók képzésében, és ez akusztikailag kimutatható-e. A svák időtartamának elemzései azt igazolták, hogy szignifikáns a különbség a hezitáláskor előforduló és a más nyelvi funkcióban létrejövő svák között (5. ábra). A magánhangzók minőségét meghatározó akusztikai összetevők, a formánsok alakulása is egyértelműen elkülönül a funkció függvényében.

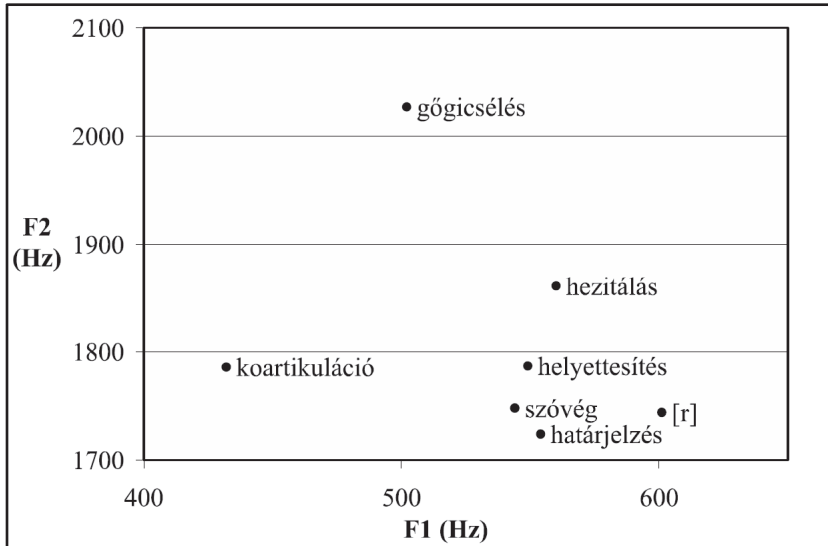
## 3. A hezitálás

A magyar beszédben hezitálásként legnagyobb arányban a semleges magánhangzó különféle időtartamú ejtése jelenik meg.

A svá különböző előfordulásai hezitálásként:

- a svá és a nazális kombinációja (előfordulása 10% alatti);
- a svát egy [h]-szerű zörejelem követi (előfordulás kb. 2,1%);
- a hezitálás *ömh* hangsorként realizálódik;

- a svát két [h]-szerű zörejelem fogja közre;
- az alsó nyelvállású, palatális, ajakréses /ε/ realizálódott a svával, illetve egy [h]-szerű zörejelemmel kombinálva.



5. ábra. A svá formánsértékei a funkció függvényében (Gósy 2006: 20)

### 3.1. A hezitációs funkciójú svá alakulása felnőtteknél a két nem esetében

Horváth Viktória (2014) kutatásának eredményei szerint mind a férfiak, mind a nők a semleges magánhangzót preferálják a néma szünet kitöltésére. A férfiaknál némileg gyakoribb ez a forma, de a nők változatosabb realizációkkal valósítják meg a hezitálást. A férfiaknál négyféle hezitálásra talált példát (*ö, m, öm, öh*), ezzel szemben a nők ennél kétszer több fajtát produkáltak, csak náluk fordultak elő az *ömh, höh, eö* formák. Összevetve az egyes hezitálások időtartamát Horváth Viktória azt találta, hogy a hangkapcsolatból álló hezitációs jelenség időtartama hosszabb, mint az egyes hangokból álló kitöltött szüneteké.

A hezitálás különféle funkciókat tölthet be a beszédben: jelezheti a bizonytalanságot, a hibát vagy a beszédszándékot. Beszédszándék esetén a hezitálás nem a beszédtervezés valamely szintjén bekövetkező problémát jelzi, sokkal inkább azt, hogy a beszélő beszélni kíván, nem akarja átadni a szót beszédpartnerének – egyfajta tudatosság jellemezheti ezt a funkciót a sokkal kevésbé tudatos bizonytalansággal szemben. Választ kerestek arra, hogy vajon a semleges magánhangzó időtartama függ-e attól, hogy milyen funkcióban jelenik meg. Horváth Viktória azt találta, hogy a hezitálás funkciójú svá időtartama bizonyos mértékben függ a beszédbeli funkciójától. A bizonytalansági megakadasként

megjelenő semleges magánhangzó átlagosan a legrövidebb, míg a beszélők beszédszándékát kifejező semleges magánhangzó időtartama a leghosszabb. Köztes helyet foglal el a hiba környezetében megjelenő semleges magánhangzó időtartama. A hiba környezetében megjelenő svá időtartama nem különbözik szignifikáns mértékben a másik két funkcióban megjelenő svá időtartamoktól, ugyanakkor a bizonytalansági megakadásként megjelenő és a beszédszándékot kifejező svá időtartama közötti különbség statisztikailag szignifikáns, ami a funkcióval magyarázható.

Horváth Viktória (2014) megvizsgálta a hezitálás funkcióban megjelenő semleges magánhangzók formánsszerkezetét is, és kutatásában azt találta, hogy sem a nők, sem a férfiak ejtésében nincs jelentős különbség az  $F_1$ -értékekben attól függően, hogy a svá milyen szerepet tölt be a közlésben. Azonban a semleges magánhangzó funkciófüggő realizációja a férfiak ejtésében az  $F_2$  értékében kimutatható: a beszédszándékot kifejező funkcióban ejtett svá második formánsa szignifikánsan alacsonyabb értéken realizálódik, mint amikor bizonytalanságot jelez a beszédben. A férfiak a nyelv vízszintes mozgását tekintve hátrébb képzik a szájüregben a beszédszándékot jelző semleges magánhangzót. A nők esetében a svá  $F_2$ -értékében nem találtak szignifikáns különbséget a funkció tekintetében, ugyanakkor az adatok azt mutatják, hogy a hibajelzés funkcióban a legalacsonyabb az  $F_2$ , vagyis ekkor képződik a svá leghátrébb a szájüregben, valamint a semleges magánhangzó  $F_2$ -értéke beszédszándék funkcióban a legmagasabb. Vagyis a nők ebben a funkcióban képzik legelőrébb a semleges magánhangzót.

A kutatások választ kerestek arra a kérdésre is, hogy a javított és a javítatlan téves kivitelezések környezetében megjelenő hezitációk időtartama között van-e különbség. Az eredmények azt mutatták, hogy a javított hibákat kísérő svák időtartamátlagos hosszabb, míg a nem javítottaké valamivel rövidebb. Ebből arra lehet következtetni, hogy ha a beszélő – tudatosan vagy sem – úgy dönt, hogy kijavítja a téves kivitelezést, akkor hosszabb időtartamú (kitöltött) szünetre van szüksége a korrekciós folyamatok elvégzéséhez (Horváth 2014).

### 3.2. A hezitálás gyermekkorban

Kutatások azt bizonyították, hogy a hezitálások gyakoriságának szempontjából az életkor meghatározó tényező. Minden életkorban semleges magánhangzóként realizálódik a leggyakrabban a hezitálás, de aránya 8–9 éves korban körülbelül 60%, míg felnőttkorban körülbelül 80%. Gyermekkorban még nem dominál annyira a svá használata (ekkor gyakoribbak az *ö*, *m* típusok); majd a gyermekek egyre inkább elsajátítják a felnőtt mintát, hiszen ez a forma bizonyul a legmegfelelőbb időnyerési stratégiának, amikor a beszédtervezés folyamata nehézségekbe ütközik (Horváth 2014).

Az egyéni különbségeket vizsgálva ugyanakkor óriási különbségeket adatl-

tak egy-egy életkorban; például a legtöbbet hezitáló óvodáskorú gyermek egy perc alatt akár 8 alkalommal is produkálhat kitöltött szünetet, de sok gyermek beszédében egyáltalán nem szerepel ez a jelenség. Óvodáskorú gyermekek spontán beszédét összevetve azt találták, hogy a lányok nemcsak valamivel kevesebbet hezitálnak, mint a fiúk, de a kitöltött szünetek időtartama is rövidebb beszédükben.

A beszédben betöltött funkció szerint a gyermekek mindhárom funkció esetén a semleges magánhangzót ejtik a legnagyobb arányban. A svá dominanciája leginkább a hibajelzés funkcióban figyelhető meg, a bizonytalanság jelzésére körülbelül 80%-ban használták a svát a gyermekek. Míg a beszédészándék funkcióban különféle hezitációs jelenségek jelentek meg, ezek közül ugyan a svát használták a legnagyobb arányban, de dominanciája nem olyan mértékű, mint a másik két funkcióban.

A svá átlagos időtartamában a gyermekeknél nincs különbség attól függően, hogy a bizonytalanságára utal, avagy hiba környezetében jelenik meg. Ugyanakkor a beszédészándék kifejezésére a gyermekek hosszabban ejtik a semleges magánhangzót, amellyel azt üzenik beszédpartnereik számára, hogy közölni kívánnak valamit, de még nem biztosak a folytatás mikéntjében, ezzel mintegy időt akarnak nyerni a gondolkodáshoz.

A semleges magánhangzó átlagos időtartamában nincs különbség a felnőttek és a gyerekek között. A felnőttek beszédében a bizonytalanságból adódó funkcióban megjelenő hezitálások 79,5%-a svá, a gyermekeknél szinte ezzel az aránnyal megegyező: 79,4%. A felnőtteknél a hibákat kísérő kitöltött szünetek 81,8%-a svá, a gyermekeknél ez 91,7%. Beszédészándék kifejezésére a felnőtteknél nagyobb a svá dominanciája (69,1%), mint a gyermekeknél (44,4%). Az elemzések azt mutatták, hogy a gyermekek mindhárom funkcióban rövidebben ejtik a svát, mint a felnőttek.

#### **4. A svá kutatásának fontossága**

Mint láthattuk, a semleges magánhangzó a beszédben számos funkcióban megjelenő, mégis szinte észrevétlen maradó „jelenség”. Vizsgálata során így választ kaphatunk mind az artikulációt, mind a beszédtervezés magasabb szintjét érintő kérdésekre. A semleges magánhangzók automatikus felismerésével kapcsolatos vizsgálat (Beke és Szaszák 2009) eredményei továbbá hozzájárulnak a magánhangzók realizációinak megismeréséhez, és felhasználhatók a beszédészintézis, illetve a beszéd felismerés területén.

## Irodalom

- Beke András 2009. A veláris magánhangzók stabilitása a spontán beszédben. In: Gecső, Tamás – Csilla Sárdi (szerk.): *A kommunikáció nyelvészeti aspektusai*. Kodolányi János Főiskola – Tinta Könyvkiadó, Székesfehérvár–Budapest, 27–31.
- Beke András – Szaszák György 2009. A svávariációk automatikus felismerése magyar nyelvű spontán beszédben. *Beszédkutatás 2009*. 148–170.
- Beke András – Grácsi Tekla Etelka 2010. A magánhangzók semlegesedése a spontán beszédben. In: Navracsics Judit (szerk.): *Nyelv, beszéd, írás*. Pszicholingvisztikai tanulmányok I. 57–64.
- Bolla Kálmán 1995. *Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei*. Nemzeti Tankönyvkiadó, Budapest.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. A semleges magánhangzó nyelvi funkciói. *Beszédkutatás 2006*. 8–23.
- Gósy Mária 2007. A semleges magánhangzó: a funkció, a kiejtés és az akusztikum összefüggései. *Magyar Tudomány* 167/1. 50–54.
- Grácsi Tekla Etelka – Horváth Viktória 2010. A magánhangzók realizációja spontán beszédben. *Beszédkutatás 2010*. 5–16.
- Henton, Caroline 1995. Cross-language variation in the vowels of female and male speakers. In: *Proceedings of the XIIIth International Congress of Phonetic Sciences (ICPhS 95)*. 420–423.
- Horváth Viktória 2009. *Funkció és kivitelezés a megakadásjelenségekben*. Doktori értekezés. ELTE.
- Horváth Viktória 2014. *Hezitációs jelenségek a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- International Phonetic Association 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University, Cambridge.
- Kassai Ilona 1978. Svá-jelenségek a magyar beszédben. *Magyar Fonetikai Füzetek* 1. 92–96.
- Koopmans-van Beinum, Florian J. 1994. What's in a Schwa? *Phonetica* 51/1–3. 68–79.
- Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modelling*. Kluwer, Dordrecht, 403–439.
- Vértés O. András 1980. *A magyar leíró hangtan története az újgrammatikusokig*. Akadémiai Kiadó, Budapest.
- Vorperian, H. K., – Kent, R. D. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50/6. 1510–1545.

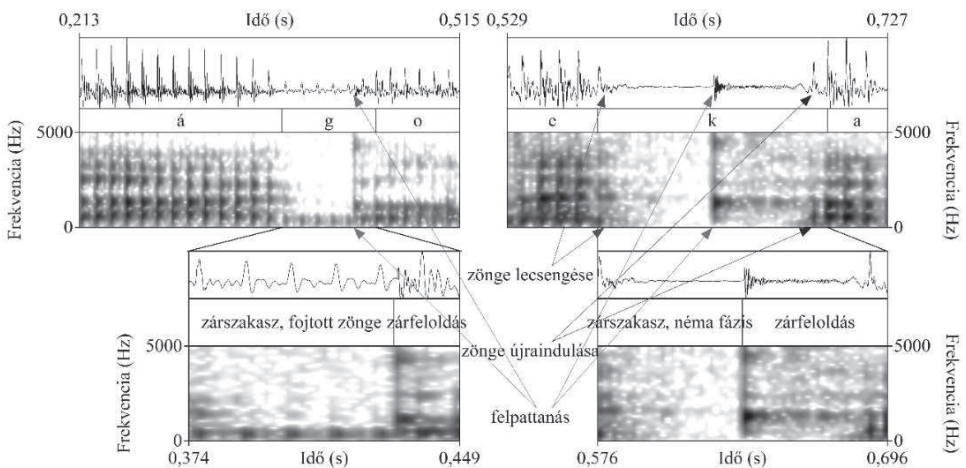
## Kérdések, feladatok

1. Mi határozza meg a magánhangzó formánsszerkezetét?
2. Mi jellemzi a semleges magánhangzó képzését?
3. Milyen funkciókban jelenhet meg a svá?

## A zöngkezdesi időről

### 1. A felpattanó zárhangok szerkezete

Az obstruensek (zörejhangek) képzése során a szájüregi akadály létrehozása, fennállása és feloldása, valamint a hangszalagok működésének egymáshoz képesti időzítése sokszínű mintázatot mutat az egyes nyelvekben (vö. pl. Ladefoged–Maddieson 1996). Így például a magyarban és az angolban is zöngés és zöngétlen felpattanó zárhangokat (explozívákat), szoktunk megkülönböztetni, mégis gyakorta okoz problémát nyelvtanulás során például, hogy míg a magyarban a *b* hangzó ejtésekor folyamatosan rezegnek a hangszalagok, a *p* esetében pedig nem képzünk zöngét, addig az angol szókezdő *b* konzonánst a magyar *p*-hez hasonlóan ejtik, a *p* képzése pedig úgynevezett hehezettel valósul meg. Ahhoz, hogy ezt az eltérést részletesen elemezhesük, vegyük át először a felpattanó zárhangok képzését, felépítését (1. ábra).



1. ábra. Az explozívák szerkezete

[Balra az *ágo*, jobbra az *ek a* hangsor az *A minap önmagát kiáltotta ki a legnagyobb énekesnek a világon.* mondat ejtéséből. Férfi bemondó a BEA adatbázisból (Gósy et al. 2012), 84 Hz átlagos alaphangmagasság.]

Az explozívák képzése során a szájüregben (esetleg a glottiszban a gégezárhang esetében) zárat alkot két beszédképző szerv. A zár fennállása során a tüdőből ki-

áramló levegő a zöngétlen hangok esetében a fűvő állásban álló hangszalagok között távozik, míg a zöngés beszédhangok esetében megrezegteti a hangszalagokat, és így létrehozza a fojtott zöngét. Az 1. ábrán a láthatjuk a [g] hang esetében a zárszakasz alatti fojtott zöngét: Az oszcillogramon a kis intenzitású rezgésekként, a hangszínképen pedig az alsó frekvenciákon a periodikus sötétedésekként jelenik meg. Ezzel szemben a [k] hang esetében a zárszakasz alatt nem jelenik meg egyik sem, ez a néma fázis. Mind a zöngés, mind a zöngétlen explozíva esetében a zár felnyitásakor felpattan a zár, ez és a követő magánhangzó megjelenéséig tartó szakasz a zárfeloldás. A zöngétlen felpattanó zárhang képzéséhez a zárszakasz folyamán leáll a megelőző beszédhang zöngéje (amennyiben volt), majd a zárfeloldás után újraindul (megindul) a zöngéképzés (ha a követő hangzó zöngés).

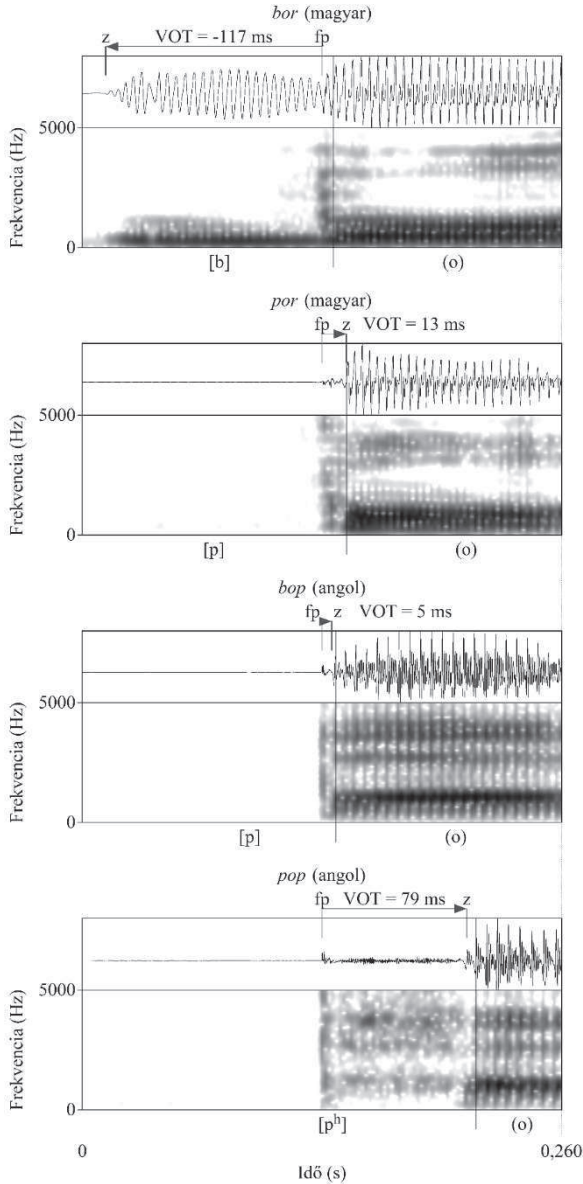
## 2. A zöngkezdesi idő

A magyar és angol explozívak közötti említett eltérést abban az időzítési mozzanatban ragadhatjuk meg, hogy a szájüregi zárfeloldás és a zöngéképzés egymáshoz képest mikor indul meg. **Azt az időtartamot, amely a zár feloldásának kezdetétől a zöng megindulásáig telik el, zöngkezdesi időnek hívjuk (VOT = voice onset time; 2. ábra).** Abban az esetben, amikor a zár feloldása után indul meg a zöng, pozitív, amikor előtte, akkor negatív értéket kapunk. Így tehát a magyarban a zöngés explozívak és affrikáták jellemzően negatív zöngkezdesi idővel rendelkeznek. A hehezet vagy aspiráció a tüdőből kiáramló levegő sűrűsödése a fűvő állásból zöngállásba záródó hangszalagok között. Ez a hehezet a 2. ábrán az angol /p/-realizáció felpattanása és a magánhangzó között látható *h*-szerű zörejelem. Azokat az explozívakat, amelyeknél ez megjelenik, hehezetes/aspirált mássalhangzónak hívjuk.

A zöngesség mellett több jeggyel is próbálták korábban megragadni a zöngességi szembenállást. Elterjedt (volt) például az artikulációs erő alapján a fortisz (zöngétlen) és lénisz (zöngés) megnevezés, de ezek nem voltak egyértelműen megfelelőek (vö. pl. Lisker–Abramson 1964; Ladefoged–Maddieson 1996).

## 3. Nyelvek közötti eltérések

A magyarhoz hasonlóan például a spanyolban vagy az oroszban is a zöngés felpattanók teljes képzési ideje alatt jellemző a hangszalagrezgés, amely a zöngétlenek esetében a zár feloldása után rövid idővel (~0–30 ms) indul meg. Az angolhoz hasonlóan a németben és például a dánban azonban abszolút szókezdő helyzetben (szünetet követő pozíció) a (fonológiai) zöngés felpattanóknál a zár időtartamában, annak vége felé legfeljebb rövid ideig jelenik meg zöng; de

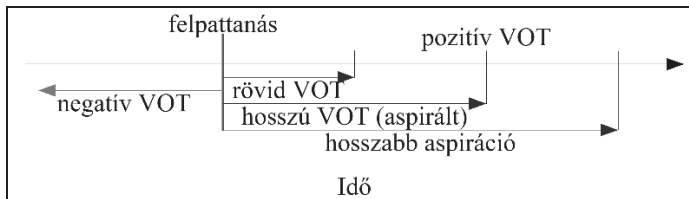


2. ábra. A magyar és az angol zöngés és zöngétlen felpattanó zárhangok VOT-típusai (Ladefoged 2005: 136 alapján).

A felső két ábra a magyar *bor* és a *por*, az alsó kettő az angol *bop* és a *pop* szó szókezdő explozíváit és a követő magánhangzó egy részét mutatja. A „z” feliratos függőleges vonal a zöngé kezdetét, az „fp” feliratos függőleges vonal a felpattanást, a harmadik függőleges vonal pedig a zárhang és a követő magánhangzó határát jelöli. A nyíl a zöngelkedési idő tartamát mutatja.

általában jellemzőbb, hogy a magyar zöngétlenhez hasonlóan csak a zár feloldása után kezdődik a zöngékepzés. A zöngétlen zárhangok képzése során pedig a zöngé mindig csak a zár feloldása után hosszabb idő (40–100 ms) elteltével (pl. Lisker–Abramson 1964) jelenik meg. Az egyéb fonetikai helyzetek, pl. az intervokális (két magánhangzó közötti) vagy a szó(tag)végi realizációk további variabilitást, eltérést mutathatnak, eltérően viselkednek egyes nyelvek között. A /b d g/ az angolban néhányszor, kevés beszélő ejtésében rövid előzöngével is megvalósulhat. Két magánhangzó között (intervokális helyzetben) azonban az angolban és a németben a nem aspirált felpattanó zárhangok általában (részben) zöngés zárszakasszal is megvalósulhatnak, míg az izlandiban a szókezdő pozícióhoz hasonlóan zöngétlenek (Beckman et al. 2013). Szóvégi helyzetben ugyancsak eltérő mintázatokat kapunk. Vannak olyan nyelvek, amelyekben legalább fonetikailag nincsen szóvégi explozív (pl. bobo, tiwi), vagy csak valamilyen korlátozással fordulhat elő (pl. a tagalogban elmarad a zár feloldása), esetleg csak zöngétlen változat jelenhet meg (pl. német, lengyel, finn; Keating et al. 1983), stb.

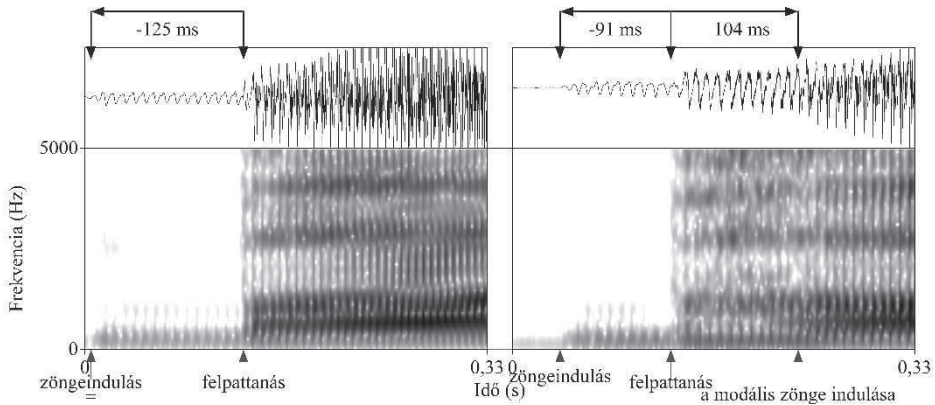
Vannak olyan nyelvek, amelyekben előzöngével rendelkező és zöngétlen aspirált felpattanó zárhangok alkotnak oppozíciót (pl. svéd: Helgason–Ringén 2008), előaspirációt (a zárszakasz kezdetén megjelenő hehezetet) használnak (Helgason 2002), illetve amelyek mindhárom időzítést felhasználják, így hármas szembenállást alkotva a zöngés, a zöngétlen és az aspirált felpattanó zárhangok között (pl. thai, örmény). Az előzöngé megléte nem jelenti azonban, hogy a teljes zárszakasz alatt rezegnek a hangszalagok, csak annyit, hogy az a zár feloldása előtt megindul. Olyan nyelvek is előfordulnak, amelyek három zöngétlen kategóriát különböztetnek meg (3. ábra), ilyen például a koreai nem aspirált, egy rövidebb és egy hosszabb hehezettel képzett explozívájának szembenállása (Lisker–Abramson 1964).



3. ábra. A VOT-értékek időzítése

Más nyelvekben (pl. hindi, marati) további fonológiai kategóriák is megjelennek, akár négyes vagy többes **oppozíciót** (szembenállást) is alkotva. Ezekben a nyelvekben a hangszalagrezgés variabilitásának lehetőségeit használja ki a nyelvi rendszer (pl. **levegős** vagy **irreguláris zöngeminőség**; részletesen: Ladefoged–Maddieson 1996). E nyelvekben maga a zöngékezdési idő természetesen

nem elégséges a fonémacsoportok elkülönítésére. A 4. ábra a hindi /b/ és /b<sup>h</sup>/ (zöngés aspirált) fonémák tipikus megvalósulásait mutatja be. A jobb oldali hangszínekben a magánhangzó formánsai zörejesek. Ilyenkor a hangszalagok csak annyira nyitottak, hogy a kiáramló levegő nem csak sűrűdik közöttük, hanem meg is rezegteti azokat.



4. ábra<sup>1</sup>. A hindi /b/ (bal) és /b<sup>h</sup>/ (jobb) fonémák oszcillogramja és spektrogramja a követő magánhangzó egy részével  
(A függőleges nyilak a zárfeloldása a zöngé és a modális zöngé kezdetét jelölik.)

A szakirodalom a nyelveket az abszolút szókezdő (szünetet követő) explozívák fonetikai megvalósulása alapján szokta rendszerezni. Az 1. táblázat ezt a rendszerezést mutatja be néhány nyelv alapján. Fontos hangsúlyozni, hogy a VOT-hosszúság az adott nyelven belül a beszédhangok csoportjai közti különbséget hivatott jelezni. Természetesen a nyelvek közötti eltérést is szemlélteti, de nagyobb mértékű átfedések jelenhetnek meg a nyelvek egyes kategóriáit összevetve egyes realizációk mért fizikai időtartamában, mint egy nyelven belül, ahol az egyes kategóriák fonológiai oppozíciót képviselnek.

<sup>1</sup> Az ábra Peter Ladefoged *Vowels and consonants* (2005) című könyvének online segédanyagában (<http://www.phonetics.ucla.edu/vowels/chapter12/hindi.html>) elérhető hanganyagból készült.

1. táblázat. Az egyes nyelvek VOT szerinti kategóriabesorolása  
(A pozitív VOT-k hosszúsága az adott nyelven belül relatívan értendő!)

VOT	előzöngge, negatív VOT	rövid pozitív VOT	hosszabb pozitív VOT	még hosszabb pozitív VOT
explozívák kategóriái	zöngés	zöngétlen	zöngétlen aspirált	zöngétlen hosszabb aspirációval
magyar	X	X		
angol, német		X	X	
svéd	X		X	
koreai		X	X	X
hindi	X + zöngés aspirált!	X	X	
dán	X (rövid előzöngge vagy zöngétlen)		X	

#### 4. A zöngképzési idő nyelven belüli variabilitása

Természetesen az adott nyelv zöngés-zöngétlen fonémáira jellemző zöngkezési idők nem adhatóak meg egy-egy konkrét értékkel. A zöngkezési idő függ egyrészt az adott mássalhangzó képzési jegyeitől, másrészt a fonetikai kontextustól, de eltérhet a közlés egyéb jellemzői és a beszélő sajátosságai mentén is. Ezek megértéséhez röviden kitérünk ennek az artikulációs-akusztikai, aerodinamikai okaira.

A zöngképzés során a hangszalagok zöngéállásban állnak, így a tüdőből érkező levegő a hangszalagok alatt feltorlódik, azaz a nyomás megemelkedik. Ha ez a nyomásszint eléri a hangszalagok felnyitásához szükséges szintet, azok felnyílnak, a levegő kiáramlik, majd a csökkent nyomás miatt visszazáródó hangszalagok alatt ismét felgyűlhet a levegő (Tonndorf 1925, idézi Gósy 2004). A szájüregi akadály miatt mind a zöngés, mind a zöngétlen explozíva esetében feltorlódik a szájüregben a hangszalagoktól érkező levegő. (A zöngés explozívak esetében tehát kétszer torlódik fel a levegő, mielőtt kiáramolhatna a szájüregből: először a hangszalagok alatt, hogy létrehozza a zöngét, majd másodszer a szájüregben a zár következtében, ami a felpattanást követő turbulens zörejt hozza létre. A zöngétlen explozívak esetében a levegő csak a szájüregi akadály mögött torlódik fel.) A nyomás emelkedését passzív és aktív artikulációs gesztusok lassíthatják (előbbi pl. a szájüreg tágulása, amelyet ugyancsak a felgyülemlő levegő okoz, utóbbi pl. a gége süllyesztése, amit a beszélő végez; összefoglalóan vö. Maddieson 1996). Ahhoz azonban, hogy a tüdőből kiáramló levegő felnyithassa a hangszalagokat, a felettük lévő nyomás sem haladhat meg egy adott szintet (amelynek alacsonyabbnak kell lennie a szubglottális nyomásnál; Baer 1971,

idézi Solé 2011). (A zöngéképzés le is állhat, de erre a jelen tanulmányban nem térünk ki, vö. pl. Grácz 2012.)

A mássalhangzó képzési helye meghatározza, hogy mekkora a térfogat az akadály és a hangszalagok között. Minél hátrább képzett a beszédhang, annál kisebb ez a térfogat, ezért annál nagyobb nyomás tud felgyülni ugyanannyi idő alatt. Ezért a zöngés hangok esetében a zárszakasz, azaz az előzöngé időtartama rövidebb lesz a hátsóbb képzési helyeken, a zöngétlen (akár aspirált, akár nem) beszédhangok esetében pedig annál több levegőnek kell az akadály felnyílása után távozni, ami hosszabb VOT-t eredményez (pl. Cho–Ladefoged 1999; magyarra: Gósy 2000a, b; Gósy–Ringen 2009; Grácz 2012).

A hangszalagok zöngéhez történő beállítása zöngétlen képzésből vagy légvéltelből nyilvánvalóan más artikulációs feltétel, mintha már zöngés hangzó után másik zöngés hangzó következne. Ehhez hasonlóan, ha a követő hangzó zöngés, nem pedig zöngétlen, vagy szünet, az a zöngé megtartását (hosszabb fenntartását) is segíti (Westbury–Keating 1986). Emellett vegyünk magánhangzó által követett pozíciót. Ha a magánhangzó nyílt, a szájüreg térfogata nagyobb a mássalhangzó után (és a nyitódás is gyorsabb), így a felgyülemlt levegő hamarabb távozhat. Azaz a zöngé is előbb meg tud indulni a követő hangzóhoz. Magyarra először Gósy Mária mutatta be a magánhangzó képzési jegyeinek a VOT-re gyakorolt hatását zöngétlen felpattanó zárhangokra (Gósy 2000a, b). A magánhangzónak nem csak a nyelvvállásfoka, hanem a további képzési jegyei is befolyásolták az eredményeket. A nyelv vízszintes mozgása és az ajakműködés pedig a mássalhangzó képzési helyének függvényében van rá hatással (hasonló artikulációs és nyomásviszonyokbeli okok miatt; további irodalom: Gósy 2001; Grácz 2011a, b; Grácz 2013).

Eltérő zöngelkedési-idő-mintázatokat mutatnak az egyes beszédmódok: az olvasott szöveg és mondatok (Lisker–Abramson 1964), illetve az olvasás (szavak vagy szöveg) és spontán beszéd (Gósy 2000a, b; de vö. Bóna 2011, 2012; Neuberger 2014). Spontán beszédben jellemzően nagyobb az értékek szóródása, kisebb mértékű a követő magánhangzó hatása, de a főbb tendenciák azonosak a két beszédmódban. Az artikulációs tempó (elsősorban pl. a hangzóidőtartamok révén) is befolyásolja a VOT-t. Gyorsabb beszédben rövidebb a zöngelkedési idő (összefoglalóan: Yao 2009).

## 5. A zöngéképzési idő és speciális beszélői csoportok

A zöngéképzési idő megvalósulása – mint minden fonetikai realizáció – kisebb-nagyobb mértékben eltérhet az egyes beszélők között. Kimutathatók eltérések például az életkor és a nem függvényében. Az anyanyelv-elsajátítás során a motoros készségek és a fonémaazonosítás fejlődésével, idős korban annak gyen-

gülésével, illetve a beszédképző szervek állapota (idős korra pl. az izomzat renyhülése, szövetek állapota, tüdőkapacitás változása) hatással van beszédünkre, így a zöngelkedési időre is (magyarra Bóna 2011, 2012, 2013; Bóna–Auszmann 2014; S. Tar 2015). A nők és a férfiak (Gósy–Ringen 2009; Bóna 2011, 2012), illetve a lányok és fiúk ejtése zöngelkedési idejének mintázata is eltér (Bóna–Auszmann 2014; S. Tar 2015).

Néhány betegség a beszédképzést is érinti. Az agy, a beszédképző szervek sérülése, motoros problémák vagy a tervezés, programozás és kivitelezés sérülése miatt (afázia, dyslalia, dysarthria vagy apraxia) a beszéd egyes jellemzői eltérnek az ép beszélők ejtésétől. Ezek bekövetkezhetnek például stroke, agyvérzés, Parkinson-kór, szklerózis következtében. Ezek az eltérések hozzájárulnak a beszéd olyan aspektusainak a megismeréséhez (az artikuláció tervezésétől a kivitelezésén át akár a beszédfeldolgozásig), amelyeket normál beszédképzés esetében nem figyelhetünk meg. Emellett az orvosok számára a vizsgálatok kiinduló pontjával szolgálhat, vagy a diagnózis felállításához nyújthat segítséget, ha a kutatók következetesen igazolható eltérést tapasztalnak a beszédben, amely vagy az adott betegségre, vagy betegségcsoportra jellemző, és a nem patológikus mintázatoktól egyértelműen elkülöníthető. Maga az eltérés lehet következetes, azaz néhány mérés alapján a további esetek jósolhatóak – ez segítheti az orvosi munkát kellően feltárt mintázatok esetében –, avagy rendszertelen (az előbbi jellemzi például a dysarthriát, az utóbbi az apraxiát). Az egyes okozó betegségek (tehát pl. csak Parkinson-kóros dysarthriás betegek), vagy általában a bekövetkezett nyelvi, beszédbeli betegség (tehát az alapbetegségtől függetlenül például afáziás vagy dysarthriás betegek) alapján is végeznek kutatásokat, hogy feltárják a jellemző nyelvi tünetegyüttest. A VOT sokat kutatott jellemző ezek között, mivel a laryngális és a szupraglottális artikuláció összehangolását igényli. Az eredmények általában alátámasztják az időzítés problémáját. Kent és Rosenbek (1983) angol, Itoh és kollégái (1979) pedig japán anyanyelvű beszélők esetében bizonyították, hogy apraxia esetében a VOT eltér a tipikus, ép beszélők ejtésében kapott értékektől. Mindkét nyelvben a zöngés fonémák (angolban ez általában zöngétlen nem aspirált beszédhang, japánban pedig előzöngés) megvalósulásai tértek el nagyobb mértékben a tipikustól. Az angol apraxiás beszélőknél hosszú (akár 250 ms-os) zöngés zárszakaszt is mértek, a japánoknál pedig pozitív VOT-tel realizálták a vizsgált beteg beszélők ezeket a fonémákat. Az eltérések azonban változatosak, nem minden apraxiás beszélőre jellemzőek, és az érintettek ejtésében sem elég következetesek. Hasonlóan nem diagnosztikus értékű VOT-eredményeket találtak dysarthria esetében is (Fischer–Goberman 2010: Parkinson-kóros betegek vizsgálata). A betegek és a kontrollcsoport eredményei nem tértek el szignifikánsan. Ebben a kutatásban a szerzők artikulációs kompenzációt feltételeznek az eredményeik mögött.

A beszédtervezést, -kivitelezést, artikulációt érintő betegségeken túl a hallás és a beszédfeldolgozás károsodása is befolyásolja a beszédet. Például a nagyot-hallás vagy a siketség esetében nincs akusztikai visszacsatolás, illetve a korai, az anyanyelv-elsajátítás alatti vagy előtti halláskárosodás esetében a hallott minta hiányában eltérő a beszédtanulás. További eltérést okozhat a hallássérült személyek esetében, hogy ritkábban beszélnek, mint az ép hallók, és így eltér(het) a légzéstechnikájuk is. Magyar nyelvre siket beszélők beszédében elemezték a /p t k/ mássalhangzókat (Ausmann 2014). A realizációk belső időszerkezete eltért az ép halló adatközlők értékeitől: a zárszakasz hosszabb, a zöngelkedési idő rövidebb volt a siketek felolvasásában.

Elsődleges nyelvi zavarral diagnosztizált gyermekek ejtésében 6 éves korban is változó a fojtott zöngé megjelenése, de egyéb (másodlagos) akusztikai kulcsok<sup>2</sup> megjeleníthetik az észlelet számára az oppozíciót (S. Tar 2013).

## 6. A zöngéképzési idő és percepció

A zöngelkedési idő nem csak artikulációs és akusztikai, hanem percepció-s szempontból is fontos fonetikai jellemző. Vegyük az angolhoz vagy a koreaihoz hasonló nyelveket, ahol a felpattanó zárhangok fonetikailag zöngétlenek, azaz két vagy több pozitív VOT között kell észlelés alapján a fonémát azonosítani. Ha figyelembe vesszük, hogy a képzési helyek alapján is eltér az adott explozívára jellemző VOT, akkor felmerül, hogy a képzési hely észleletében is szerepet játszhat. Néhány kutató más fonetikai jellemzőket (pl. a követő magánhangzó első formánsának hangátmeneti értékét, Pind 1999) tart fontosabb akusztikai kulcsnak. Ugyanakkor számos kísérlet igazolta a zöngelkedési idő szerepét (magyarban a képzési hely azonosítására pl. Gósy 2000a). Ezt alátámasztják állatokkal (pl. csincsillák: Kuhl–Miller 1978) végzett fonémaazonosítási vizsgálatok is, ahol az állatok a(z angol) tipikus VOT-értékeket hallották, majd közben-ső értékkel rendelkező beszédhangokra kellett a betanult reakciók valamelyikét adni. A „fonémadöntésük” küszöbértéke az ember észleletével egyezett. Természetesen az észleletnek tipikus beszédkörül-mények között nem csak egy akusztikai kulcsra kell hagyatkoznia, hanem a mássalhangzó egyéb jellemzői (másodlagos akusztikai kulcsok), a hangkörnyezet a koartikuláció révén (Ohala 1983), a kontextus (az elhangzottak értelmezése) is segít (összefoglalásul lásd Gósy 2005). Ezek ugyanakkor meg is téveszthetik a fonémaazonosítást (spiránsokra vö. Stevens et al. 1992).

---

<sup>2</sup> Magyar nyelvben felnőttek beszédében a megelőző magánhangzó időtartama nem minden esetben, de a mássalhangzó zárszakaszának és teljes időtartama eltért a zöngésségi párok között (Grácsi 2011a, b).

Meg kell azonban említeni, hogy nem csak az artikuláció, hanem a percepció esetében is vannak speciális csoportok. Ezek közül néhányat említünk meg. Nittrouer és Lowenstein (2008) tanulmányában számos olyan vizsgálatot foglal össze, amelyek a zöngésség azonosításában eltérést mutattak a gyermekek életkori csoportjai között, illetve a felnőttektől is eltérő mintázatot mutattak. Általában jellemzőnek találták (nem csak VOT tekintetében), hogy a gyermekek észlelete a formánsmenetekre alapul más jegyekkel szemben. Diszlexiás gyermekek zöngékezdesiidő-észlelését vetették össze korban, illetve olvasási szintben illesztett kontrollcsoport percepciójával. Az eredmények mindkét kontrollcsoporttól eltérő észleleti mintázatot adtak, azaz ezeknél a gyermekeknél a beszédfeldolgozás is atipikus (Bogliotti et al. 2008).

## 7. A zöngékezdesi idő és a nyelvtanulás

Ugyancsak speciális beszélői csoportot képeznek azok a személyek, akik több nyelvet beszélnek. Ha a két nyelv explozíváira eltérő VOT-kategóriák jellemzőek – mint pl. a magyar és az angol esetében –, akkor az adott nyelvet nem anyanyelvként beszélő személy számára mind a beszédészlelés, mind a beszédképzés esetében meg kell tanulnia a célnyelv mintázatait. Meg kell tanulnia az anyanyelvi percepciós és artikulációs bázisukban nem szereplő beszédhangokat (pl. magyar anyanyelv esetében az angol aspirált zöngétlen felpattanó zárhangjai), illetve az azokban szereplő, de a célnyelvben eltérő fonémához tartozó alapvariánsaként megjelenő használatát (pl. a magyarok számára az angol rövid VOT-val rendelkező explozívák). Azok a többnyelvűek, akik párhuzamosan sajátították el a két nyelvet, vagy elég fiatal korukban, általában az adott nyelvek egy-nyelvű beszélőivel azonosnak tekinthető zöngékezdesi időt produkálnak (és általában egyéb akcentusjelenségük is ritkán van). Ezzel szemben a második nyelv késői tanulása során jellemzőbb, hogy az idegen nyelvi beszédprodukciónban a VOT vagy anyanyelvi, vagy a két nyelv közötti értékekkel valósul meg, illetve a kettő közötti skálán mozog – olykor a célnyelvi értéket eltalálva – (pl. Flege 1980, 1991; Thornburgh–Ryalls 1998). A jelenség háttere, hogy 6-7 éves kor között lezárul az első biológiai sorompó (Gósy 2005), így ez után már kisebb sikerrel jár, avagy több, speciális oktatás alapján tud a beszélő idegen nyelvi fonetikai-fonológiai mintázatot elsajátítani mind a percepcióban, mind az artikulációban.

## Irodalom

- Auszmann Anita 2014. Zöngétlen explozívák időszerkezete siket beszélők szövegfelolvasásában. In Váradai Tamás (szerk.): *Alknyelvdok8. Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA Nyelvtudományi Intézet, Budapest, 1–12.
- Baer, Thomas 1975. *Investigation of phonation using excised larynxes*. Doktori értekezés. Massachusetts, MIT.
- Beckman, Jill – Michael Jessen – Catherine Ringen 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49. 259–284.
- Bogliotti, Caroline – Serniclaes, Willy – Messaoud-Galusi, Souhila – Sprenger-Charolles, Liliane 2008. Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. *Journal of Experimental Child Psychology* 101. 137–155.
- Bóna Judit 2011. A [p, t, k] mássalhangzók zöngelkedési ideje idők és fiatalok spontán beszédében. *Beszédkutatás* 2011. 61–72.
- Bóna Judit 2012. A zöngétlen explozívák zöngelkedési ideje idők és fiatalok folyamatos beszédében. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai kiadó, Budapest, 121–137.
- Bóna Judit 2013. *A spontán beszéd sajátosságai időskorban*. Beszéd – Kutatás – Alkalmazás 2. ELTE Eötvös Kiadó, Budapest.
- Bóna Judit – Auszmann Anita 2014. Voice onset time in language acquisition. In *Proceedings of 10th International Seminar on Speech Production*. 41–44.
- Cho, Taheong – Ladefoged, Peter 1999. Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27. 207–229.
- Fischer, Emily – Goberman, Alexander M. 2010. Voice onset time in Parkinson disease. *Journal of Communication Disorders* 43. 21–34.
- Flege, James Emil 1980. Phonetic approximation in second language acquisition. *Language Learning* 30/1. 117–134.
- Flege, James Emil 1991. Age of learning effects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of Acoustic Society of America* 89/1. 395–411.
- Gósy Mária 2000a. A [p t k] mássalhangzók zöngelkedési ideje. *Magyar Nyelvőr* 124. 195–204.
- Gósy Mária 2000b. The VOT of the Hungarian voiceless plosives in careful and spontaneous speech. *Govor* 17. 15–28.
- Gósy Mária 2001. The voice onset time of the Hungarian voiceless plosives in words and in spontaneous speech. *International Journal of Speech Technology* 4. 75–85.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris, Budapest.
- Gósy Mária – Ringen, Catherine O. 2009. Everything you always wanted to know about VOT in Hungarian. Előadás. *International Conference on the Structure of Hungarian*. Budapest, 2009. szept. 1. Elérhető: [http://icsh9.unideb.hu/pph/handout/Ringen\\_Gosy\\_handout.pdf](http://icsh9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf)

- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczki Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 9–24.
- Gráczki Tekla Etelka 2011a. Explozívák a zöngésségi oppozíció függvényében. In Várad Tamás (szerk.): *V. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. MTA Nyelvtudományi Intézet, Budapest, 51–66.
- Gráczki Tekla Etelka 2011b. Intervokális explozívák a zöngésségi oppozíció függvényében. *Beszédkutatás* 2011. 32–46.
- Gráczki Tekla Etelka 2012. *Zörejangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében*. PhD-disszertáció. ELTE BTK, Budapest.
- Gráczki Tekla Etelka 2013. Az explozívák és affrikáták zöngésségének időviszonyai. *Beszédkutatás* 2013. 94–120.
- Helgason, Pétur 2002. *Preaspiration in the Nordic languages. Synchronic and diachronic aspects*. PhD-disszertáció. Stockholm University.
- Helgason, Petur – Ringen, Chaterine O. 2008. Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36. 607–628.
- Itoh, Motonobu – Sasanuma, Sumiko – Tatsumi, Itaru F. – Kobayashi, Yo 1979. Voice onset time characteristics of apraxia of speech. *Annual Bulletin of RILP* 13. 123–132.
- Keating, Patricia – Linker, Wendy – Huffman, Marie 1983. Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics* 11. 277–290.
- Kent, R. D. – Rosenbeck, John C. 1983. Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research* 26. 231–249.
- Kuhl, Patricia K. – Miller, James D. 1978. Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America* 63. 905–917.
- Ladefoged, Peter – Maddieson, Ian 1996. *The sounds of the worlds languages*. Blackwell Publishers, Oxford, Massachusetts.
- Ladefoged, Peter 2005. *Vowels and consonants*. Blackwell Publishers, Oxford.
- Lisker, Leigh – Abramson, Arthur 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20. 384–422.
- Maddieson, Ian 1996. Phonetic universals. In Hardcastle, William J. – Laver, John (eds.): *The handbook of phonetic sciences*. Blackwell Publishers, Oxford, 619–639.
- Nittrouer, Susan – Lowenstein, Joanna H. 2008. Spectral structure across the syllable specifies final-stop voicing for adults and children alike. *Journal of the Acoustical Society of America* 123. 377–385.
- Neuberger, Tilda 2014. Zöngétlen explozívák időviszonyai a beszéd típus és az egyéni sajátosságok tükrében. *Beszédkutatás* 2014. 40–53.
- Neuberger, Tilda 2015. Zöngétlen zárhangok időszerkezete a fonológiai hosszúság függvényében. *Beszédkutatás* 2015. 5–20.
- Ohala, John J. 1983. The origin of the sound patterns in vocal tract constraints. In MacNeillage, Peter F. (ed.): *The production of speech*. Springer Verlag, New York, 189–216.
- Pind, J. 1999. The role of F<sub>1</sub> in the perception of voice onset time and voice offset time. *Journal of the Acoustical Society of America* 106/1. 434–437.
- Solé, Maria Josép 2011. Articulatory adjustments in initial voiced stops in Spanish, French and English. In: *Proceedings of the 17<sup>th</sup> International Congress of Phonetic Sciences*. Vol. 3. 1878–1881.

- Stevens, Kenneth N. – Blumstein, Sheila E. – Glicksman, Laura – Burton, Martha – Kurowski, Kathleen 1992. Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of Acoustic Society of America* 91/5. 2979–3000.
- S. Tar Éva 2013. A zöngességi kontraszt elsajátíttóságának mintázata atipikus nyelvfejlődés esetén. *Beszédkutatás* 2013. 194–209.
- S. Tar Éva 2015. Szóeleji zárhangok zöngelkedési ideje: Beszédproduktions adatok az óvodás korosztályra vonatkozólag. *Beszédkutatás* 2015. 148–163.
- Thornburgh, Dianne F. – Ryalls, John H. 1998. Voice onset time in Spanish-English bilinguals: Early versus late learners of english. *Journal of Communication Disorders* 31/3. 215–229.
- Tonndorf, W. 1925. Die Mechanik bei der Stimmlippenschwingung und beim Schnarchen. *Zeitschrift für Hals-, Nasen- und Ohrenheilkunde* 12. 159–168.
- Westbury, John R. – Keating, Patricia 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22. 145–166.
- Yao, Yao 2009. Understanding VOT variation in spontaneous speech. In: *Proceedings of 18th International Congress of Linguists (CIL XVIII)*. 29–43.

### **Kérdések, feladatok:**

1. Sorolja fel, mely nyelvi és nem nyelvi tényezők befolyásolják a zöngelkedési időt.
2. Miért érdemes a zöngelkedési időt alkalmazni az explozívák zöngességével kapcsolatban?
3. Milyen helyzetekben nincs zöngelkedési ideje az explozíváknak?



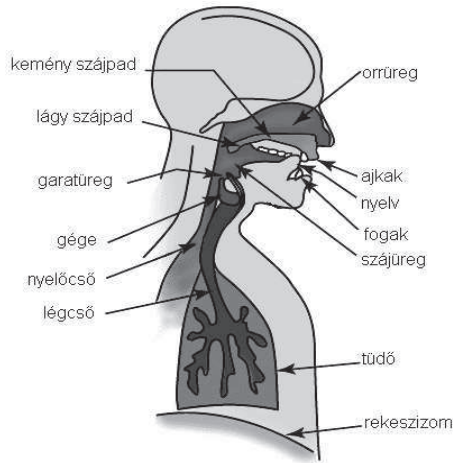
## **Nonverbális hangjelenségek a beszédben**

A mindennapi kommunikáció leggyakrabban a verbális és a nonverbális csatornán keresztül együttesen zajlik. A verbális kommunikáció alapja a nyelv és a beszéd, amely az ismeretek megszerzésének és továbbításának, gondolataink kifejezésének az eszköze. A beszéd azonban nem az egyetlen információközvetítő csatorna, ugyanis szorosan együtt járnak vele a nem nyelvi, nonverbális jelek, amelyek lehetnek a látható és a hallható csatornán keresztül érkező jelenségek.

A tanulmány arról ad áttekintést, hogy miként lehet csoportosítani a nonverbális kommunikáció elemeit, különös tekintettel a hallható-hangzó csatorna jelenségeire. Választ keresünk arra, hogy milyen információkat szállítanak, milyen funkciókat tölthetnek be a vizsgált hangjelenségek. Felvázoljuk továbbá, hogy mire irányulnak a témában folytatott nyelvészeti, fonetikai kutatások. Néhány nemzetközi és hazai elemzés ismertetése során körvonalazni kívánjuk, hogy miért is szükséges, valamint milyen haszonnal járhat a beszéd kísérőjelenségeinek vizsgálata.

### **1. Verbális és nonverbális kommunikáció**

A **verbális csatornán** keresztül közvetített beszéd az emberi kommunikáció legtermészetesebb módja. A beszéd fajspecifikus, biológiailag meghatározott jelenség. A humán hangképző apparátus felépítése adott, vagyis minden embernél azonos beszédszervek állnak rendelkezésre a beszéd létrehozásához (antropofonikus elmélet, l. pl. Gósy 2004; 1. ábra). A hangképzésben részt vevő szervek elsődlegesen nem a verbális kommunikáció céljait szolgálják, hanem az életben maradáshoz szükséges oxigénfelvételt, a táplálkozást stb. A beszéd – mint a nyelvhasználat hangzó formája – tehát másodlagos funkció. A beszédképző mechanizmus sokféle hang létrehozására alkalmas, az egyes nyelvek eltérő hangzását a beszédszervek különbözőképpen való működtetése, a különböző artikulációs gesztusok adják. A hangképzési lehetőségek tehát univerzálisak, de az egyes nyelvek specifikusan válogatnak a lehetőségekből: a tüdő, a hangszalagok, az üregrendszerek, a nyelv és az ajkak működése, illetve mozgása szerint. Az így produkált hangok, hangkapcsolatok szavakká és nagyobb egységekké rendeződnek, a szóbeli üzenetek hordozzák a mondanivaló tartalmi jegyeit, a nyelvi jelentést.



*1. ábra.* A hangképzésben részt vevő szervek

A nyelvi kommunikáció mellett fontos szerepet kap a **nonverbális csatorna**. A befogadás módja szerint megkülönböztetjük a látható, vizuális jeleket és az auditíve feldolgozható, vokális elemeket. A nem verbális vizuális kommunikációnak, a kinezikának több összetevője is van, mint például a mimika, a gesztusok vagy a testtartás. Az arc sokféle érzelm kifejezésére képes, valamint informál a társas interakció menetéről is. A hangképző szervek nem verbális használata is lehetséges, amellyel olyan nonverbális vokális jelenségeket hozunk létre, mint a nevetések, torokköszörülések, sóhajok stb. A beszédjelben a nyelvi információn túl számos tényező is kódolva van a beszélőről, amelyet a hallgató a beszédpercepció során (nem tudatosan) dekódol. Ezek lehetnek fizikális (pl. nem, életkor, testalkat, egészségi állapot), szociális (pl. regionális hovatartozás, szociális helyzet, iskolázottság, foglalkozás), valamint pszichológiai jellemzők (pl. személyiségjegyek, érzelmi állapot, hangulat) (Laver 1994: 14).

A szándékosság szempontja szerint két csoportba sorolhatjuk a jelenségeket: önkéntelen, nem tervezett, spontán jelzések (szimptomák), valamint szándékosan létrehozott jelek, amelyekhez szükséges tudás birtokába tanulás útján kerülnek a beszélők. A nem nyelvi jelek egy része tehát szándéktalan velejáruja a beszélők; érzelmi vagy egészségi állapotunkat, hangulatunkat, érzékelésünket mindig valamilyen gesztus vagy mimika kíséri, ilyen például a mosoly, a tüszentés vagy a csodálkozáskor tágra nyílt szem. Másik részük szándékosan létrehozott, nyelvileg is értelmezhető közlés, ebbe a csoportba tartozik többek között a tagadó fejrázás, a vállrándítás vagy a rosszallást kifejező köhintés. Utóbbiak konvencionális jellegét mutatja, hogy a különböző kultúrákhoz tartozó emberek különbözőképpen értelmezhetik őket. Ezek a jelek egy nyelven keresztül működnek, ebből fakadóan előfordulhat, hogy bizonyos afáziás betegeknél a

beszédkészség elvesztésével párhuzamosan leépülnek a másodlagos jelek is (Kenesei szerk. 2000).

A nonverbális jelzések információkat közvetítenek a feladótól a címzethez, és hozzájárulhatnak ahhoz, hogy a beszélgetőpartnerek jobban megértsék egymást. A természetes kommunikáció során a hallgatók egyidejűleg dekódozzák a különböző modalitásokból érkező (vizuális és auditív) információkat, a multimodális percepció során a feldolgozási műveletek nem csupán a hallottakra, hanem a látottakra is kiterjednek. A társalgásban egyazon funkciót képesek betölteni különböző modalitású jelenségek, például egy új téma felvetését jelezhetjük egy nyelvi elemmel, vagy érzékeltethetjük a prozódiaival, esetleg egy gesztussal (Rodrigues 2008). A nyelvi és nem nyelvi kommunikációs jelek általában (a sikeres, együttműködésen alapuló társalgásban) kiegészítik, támogatják egymást, és bár a nyelvi jelek vizsgálata könnyebben kivitelezhető, a verbális és nonverbális viselkedés együttes elemzése többet árul el a társalgás mikéntjéről. Noha a nem vokális anyag elemzése a beszéd tanulmányozásán kívül esik, fontos része a kommunikációs helyzetnek. A nyelvi tartalomtól független hangjelenségeknek sok hasonló és sok különböző tulajdonságuk van, amelyek az artikuláció módjában, az így létrehozott akusztikus szerkezetben, a szegmentális és szupraszegmentális aspektusban nyilvánulnak meg, s hatnak a percepcióra. Érdeemes a kvantitatív eredményeket kvalitatív elemzéssel (kontextuselemzéssel) kiegészíteni, hogy pontosabb képet kapjunk a jelenségek sajátosságairól: mind az előfordulási gyakoriságukról, mind a lehetséges funkcióikról.

## 2. Tipológia

A verbális és nonverbális kommunikáció elemeinek nincs egységes rendszerezése sem a nemzetközi, sem a hazai szakirodalomban. A kategóriarendszerek felállítását mindig az adott kutatási terület és cél határozta meg, ebből kifolyólag eltérő terminusok használatosak a különböző osztályozásokban. Gyakran használatosak a *nem nyelvi*, a *nyelven kívüli*, a *paralingvális*, az *extralingvális*, a *prozódiai*, a *szupraszegmentális* kifejezések. Ezek az osztályok egyrészt átfedéseket mutatnak a különböző kutatók munkáiban, másrészt néhány beszédtevényt több mint egy kategóriába is besorolható (vö. Schötz 2002). A következőkben a nonverbális kommunikáció elemeinek néhány lehetséges osztályozását mutatjuk be.

George L. **Trager** (1961) a szóbeli közlések tartalmán kívüli jelenségeket a paranyelv terminus alá sorolta. Tipológiájának leírásakor három kategóriát állított fel:

a) voice set: a beszélők körüli beszédhelyzetet, kontextust jelenti, magában foglalja a beszélők életkorát, nemét, aktuális hangulatát stb.,

b) voice qualities: az egyéni hangszínezet, a hangerő, a tempó és az artikulációs sajátosságok összessége,

c) vocalization: az érzelmek kifejezése (például nevetés), az üzenetátadás stílusa (például kiabálás) és a visszacsatolások (például hűmmögések) csoportja.

**Roach és munkatársai** (1998) különbséget tettek a paranyelvi és a nem nyelvi kommunikáció között. Előbbihez olyan tényezőket soroltak, amelyeket a beszélő szándékosan hoz létre, míg az utóbbi csoport tényezői nem hozhatók létre szándékosan (pl. fiziológiai sajátosságok, reflexek).

**Traunmüller** (2000) a beszéd akusztikumában kódolt információkat négy csoportba sorolta.

a) Nyelvi információ, amely konvencionálisan, szociálisan meghatározott. Az üzenet, a dialektus, a szociolektus, a beszédstílus sorolható ide, s a beszédhangok, a szavak, a prozódiai mintázatok közvetítik.

b) Organikus információ, amely a beszélők közötti anatómiai különbségekből adódik. Az életkorról, a nemről, a patológiás jelenségekről adott információk tartoznak ide, amelyeket többek között a gége mérete vagy a toldalékcső hossza befolyásol.

c) Expresszív információk, amelyek a beszélőn belüli variációt mutatják. Ezek az érzelmek, az attitűdök, valamint a környezethez való alkalmazkodás jelzői, amelyeket kifejezhetjük a fonáció típusával, a hangterjedelemmel, a hangerő vagy a beszédtempó változtatásával stb.

d) Perspektivikus, vagyis fizikai, térbeli információk, mint például a hely, az irányultság vagy a csatorna. Ide tartoznak a beszélőtől a hallgatóig terjedő út jellemzői, az átviteli csatorna zaja, a beszélő hangereje, különböző akusztikai hatások (háttérzaj, visszhang).

A magyar szakirodalomban **Vicsi és munkatársai** (2011) a nem verbális hangjelenségek két nagy csoportját különböztették meg:

a) A nyelvi tartalommal együtt megjelenő érzelmi, hangulati tartalom, amely a szupraszegmentumok által jut kifejezésre. Ilyen például a különböző érzelmek (szomorúság, vidámság stb.), illetve az egyetértés és egyet nem értés prozódiai eszközökkel való érzékeltetése.

b) A nyelvi tartalomtól elhatárolt hangesemények, amelyek vagy kifejeznek valamilyen jelentést, vagy nem rendelkeznek jelentéssel. Előbbiek szándékosan létrehozott, közlésértékű hangjelenségek, mint például a sóhajtás, a nevetés vagy a hűmmögések. A hűmmögések konvencionális motivált jelek, amelyek a kommunikációnak gyakori elemei; önálló jelentéssel rendelkeznek, amelyet a szegmentális és szupraszegmentális szerkezet együttesen hordoz. A hallgatók bizonyos akusztikai kulcsok alapján dekódolják őket, s ezek segítségével több típust képesek elkülöníteni, mint például az 'igen', a 'nem' jelentésű vagy a kérdő funkciójú hűmmögést (Markó 2005a; 2005b; 2006). Ide sorolhatjuk még a kitöltött szüneteket (vagy hezitálást), amelynek leggyakrabban semleges magánhang-

zóként, sváként realizálódnak, és számos funkciójuk lehet a beszédben (Horváth 2014). Időt biztosítanak a tervezési folyamatokhoz, s egyúttal reflektálnak arra, hogy a hallgató a közlésre szánt gondolatokhoz válogat a rendelkezésre álló lexémák, szerkezetek között. Jelzik a beszélő bizonytalanságát, vagy megjelenhetnek hiba kísérelésekként. A társalgásban beszédzándék jelzésére szolgálhatnak, illetve utalhatnak arra, hogy a beszélő át kívánja adni a szót, vagy épp ellenkezőleg, magánál akarja tartani a szót. A jelentéssel nem rendelkező hangjelenségek közé sorolhatjuk a különféle fiziológiai reflexeket, a szándéktalan testhangokat, mint például a köhögés, a lélegzés vagy az ásítás. A nem verbális hangesemények között említhetjük meg azokat a hangokat is, amelyek nem vokális eredetűek, mint például a taps – jóllehet ezek nem a beszédhez tartoznak, de akusztikailag jelen lehetnek a felvételeken (Vicsi et al. 2011).

A **társalgáselemzésben** háttércsatorna-jelzéseknek nevezett jelenségek olyan vokális és nem vokális jelzéseket takarnak, melyek lehetnek verbálisak vagy nem verbálisak. Közös jellemzőjük, hogy a beszélő fordulója alatt hangzanak el, és a hallgató figyelmét, társalgásban való részvételét jelzik a beszélő számára (Hámori 2006). A hümmögések és a nevetések fontos háttércsatorna-jelzéseknek számítanak.

### 3. Kutatások

A beszéd kutatás hagyományosan a verbális csatorna működésének megismerésére és leírására fókuszál, a spontán beszéd vizsgálata azonban magával vonta a nem verbális kommunikáció leírásának és megértésének igényét. A nonverbális csatornán keresztül érkező információk kutatása beszédtechnológiai indíttatásból az utóbbi évtizedekben indult meg, főként a beszédben kifejezésre kerülő érzelmek detektálásával. Az ilyen jellegű kutatásokhoz nagy mennyiségű spontán hanganyag gyűjtésére és feldolgozására van szükség, erre kiválóan alkalmasak a különféle adatbázisok. Az egyes beszédadatbázisok eltérő módon kezelik a nem verbális hangjelenségek megjelenését, az átiratok a fejlesztési és kutatási céltól függően bizonyos esetekben mellőzik jelölésüket, másokban valamilyen módon jelölik őket. A természetes gépi beszéd előállításához, illetve a spontán beszédet felismerő rendszerek megvalósításához elengedhetetlen a nem verbális információkat hordozó hangjelenségek annotálása, amelyre több lehetőség is adódik, de nincs egyetlen általánosan elfogadott, sztenderdizált módszer a jelölésre. A beszédben megjelenő érzelmek vizsgálatára (főképp az automatikus érzelmfelismerő és osztályozó rendszerek számára) specifikus adatbázisok is készültek. Ezek között említhetjük meg például a dán DES adatbázist (Danish Emotional Speech Database, Engberg et al. 1997), a német Emotional Speech adatbázist

(Burkhardt et al. 2005), vagy a brit angol SAVEE-t (Surrey Audio-Visual Expressed Emotion, Haq–Jackson 2010).

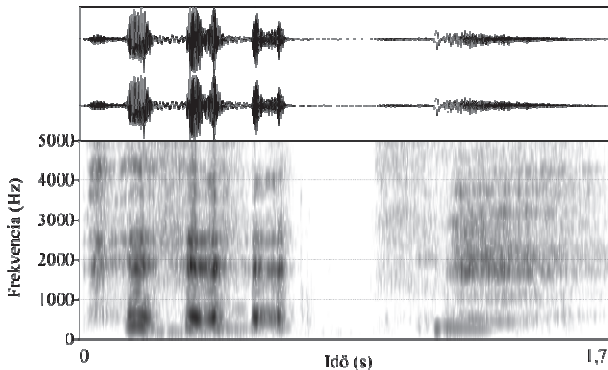
A kutatások kiindulási alapja az volt, hogy a beszélők hangulata és érzelmeik hallható paralingvisztikai jegyek alapján felismerhetők. A szubjektív tesztek mellett a gépi beszéd- és beszélőfelismerésben általában különféle automatikus osztályozási technikák (GMM – Gauss-keverék modell, SVM – szupport vektor gép, MLP – többrétegű perceptron hálózatok) alkalmazásával vizsgálják az érzelmekre jellemző spektrális és prozódiai paramétereket (pl. Seppänen et al. 2003; Navas et al. 2007; Troung–van Leeuwen 2007). Az érzelmet tükröző jellemzők közé sorolják az alaphang-frekvencia, a hangintenzitás különböző értékeit (keretenkénti, maximum, minimum, átlagértékei stb.), amelyek az intonációról és a hangerőről hordoznak információt, valamint a szótagok, beszédhangok akusztikai időtartamértékeit, amelyek meghatározzák a beszéd tempóját (Yildirim et al. 2004; Vicsi et al. 2011). A percepció tesztek eredményei alapján azt láthatjuk, hogy az alapérzelmeik a beszédben igen pontosan felismerhetők, például a finn beszélt nyelvben átlagosan 77%-osan (Seppänen et al. 2003), a németben 74%-osan, a (brit és amerikai) angolban 68%-osan, a magyarban 60%-osan (Tóth et al. 2008), de az automatikus osztályozó eljárások eredményei is hasonlóan biztatók (60–80%) (Scherer et al. 2001; Whiteside 1998).

### 3.1. Nevetések

A különböző nem verbális hangesemények közül főként a nevetések szolgáltatnak információt a beszélőről. Ezen hangjelenségek változatos funkcióval bírnak, jellemzően az öröm fajspecifikus jelzői, de megjelenhetnek agresszív viselkedés részeként („a képébe röhög valakinek”), kora gyermekkorban (az anya-gyermek interakció során) pedig a kötődés kialakításának eszközeként (Rothgänger et al. 1998). A nevetés a szociális interakció része, már négy hónapos korban megjelenik. Evolúciós aspektusát tekintve nevetésszerű hangadások az állatvilágban is előfordulnak; a különböző fajok esetében a jó közérzetet, a pozitív érzelmet, az elfogadást, illetve befogadást jelzik, vagyis azt, hogy a hang produkálója nem tervez támadni (Bryant–Aktipis 2014). A kutatók számára azért lehet fontos a nevetések különböző szempontú vizsgálata, mert általa többet tudhatunk meg az emberi viselkedésről, a szociális interakció szerveződéséről.

A nevetések akusztikai jellemzőit számos tanulmány elemezte a nemzetközi szakirodalomban (Bickley–Hunnicut 1992; Rothgänger et al. 1998; Bachorowski et al. 2001; Trouvain 2003 stb.), a magyar nyelvre vonatkozóan azonban alig akad ilyen jellegű munka (Neuberger 2012; Neuberger–Beke 2013; Bóna 2015). A kutatások többsége elkülönít nevetéstípusokat (voiced song-like, unvoiced grunt-like, unvoiced snort-like, mixed stb.), amelyek fő megkülönböztető jegye, hogy a nevetés zöngével vagy a nélkül valósult meg. A vizsgálatokból megtudhat-

juk, hogy a nevetések akusztikai jellemzői ( $f_0$ , formánsok, amplitúdó, zöngeménység) a beszédhez hasonlatosak, azok hehezetes CV /hV/ „szótagok” sorozataként realizálódnak, bár a beszédhez képest hosszabb zöngétlen résszel valósulnak meg. A 2. ábrán egy tipikus nevetés rezgésképe és hangszínképe látható a Praat programban. A szöveges részekről való elkülönítésükben (a nevetések detektálásában) tehát nagy szerepet játszik a zöngétlen-zöngés rész arány.



2. ábra. Egy tipikus nevetés rezgésképe (felül) és hangszínképe (alul): a bal oldalon /hV/-szerű szótagok, a jobb oldalon kilégzés

Két amerikai angol beszélő (egy nő és egy férfi) nevetéseit vizsgálva kimutatták, hogy egy szótagnyi nevetés átlagos időtartama 204, illetve 224 ms, és a nevetések átlagosan 6,7, illetve 1,2 szótagból állnak (Bickley–Hunnicut 1992). A további eredmények szerint a nevetés produkációjában másodpercenként átlagosan 4,7 szótag adatható, ami nagy hasonlóságot mutat az (angol, francia és svéd) olvasott mondatok másodpercenkénti szótagszámával.

Német és olasz anyanyelvű beszélők esetében azt találták, hogy a nevetések átlagos időtartama 798 ms nőknél és 601 ms férfiaknál, valamint hogy az alaphangmagasság átlagosan 472 Hz nőknél és 424 Hz férfiaknál, tehát a beszéd és a nevetések megkülönböztetésében az alaphangfrekvencia értékének is jelentős szerepe van (Rothgänger et al. 1998).

Bachorowski és munkatársai (2001) összehasonlítottak szakirodalmi adatokat a nevetések átlagos alaphangfrekvencia-értékeire vonatkozólag, amelyek nők esetében 160 és 502 Hz-es, férfiak esetében 126 és 424 Hz közötti értékkel jelentek meg a különböző kutatásokban.

A magyar nevetéseket a BEA adatbázis (Gósy et al. 2012) hanganyagaiban vizsgálták, és azt találták, hogy az átlagos időtartamuk 911 ms (átlagos eltérés: 605 ms), átlagos  $f_0$ -értékük 207 Hz (átl. elt.: 49 Hz) férfiaknál, 247 Hz (átl. elt.: 40 Hz) nőknél. A nevetések számos akusztikai paraméterben (jitter, shimmer, jel-

zaj viszony,  $f_0$ -átlag) szignifikánsan különbséget mutattak a beszédsegmentumokhoz (jelen esetben szavakhoz) képest, ami megkönnyítheti az elkülönítésüket a szöveges részekről.

A nevetések egy része azonban nem önállóan, hanem beszéddel egyidejűleg jelenik meg a társalgásokban (Provine 1993; Nwokah et al. 1993), de gyakoriak az együttnevetések is, vagyis amikor két beszélő szimultán nevetése hangzik el. Kennedy és Ellis (2004) tanulmányukban az átfedő nevetéseket is elemezték, és automatikus osztályozó algoritmussal igyekeztek detektálni azokat. A rendszerükkel 87%-os helyes eredményt értek el. Két másik tanulmányban ennél is magasabb helyes osztályozási eredmény mutatkozott, Truong és van Leeuwen (2005, 2007) 87,6%-ost, míg Knox és Mirghafori (2007) 90% fölötti teljesítményt ért el a nevetések detektálásában. Hasonlóan, egy magyar vizsgálatban is, 90% fölötti eredményt mutatott a nevetések és a beszédsegmentumok osztályozása (Neuberger–Beke 2013).

A BEA adatbázis hanganyagaiban vizsgált nevetések egyik funkciójaként megemlíthetjük a nyelvi tartalomhoz kapcsolódó érzelm kifejezést, a beszélők nevetéssel fejezték ki a saját vagy a beszélgetőpartner közlésére adott reakciójukat, így gyakorta fordultak elő a saját beszéddel egyidejűleg, illetve a társalgó partner közlései alatt. Ebben az esetben tehát a nyelvi tartalomhoz való visszacsatolás jelölői; ennél fogva a szemantikának meghatározó szerepe van. A beszélő saját közléséhez kapcsolódó nevetésre hozhatjuk az (1) példát, illetve az adatközlő (A) nevetéssel reagál a társalgó partner (TP) mondanivalójára a (2) példában (Neuberger 2012):

- (1) A: Én jelenleg másodéves egyetemista vagyok a BTK-n alkalmazott nyelvész, illetve nem, elsőéves [nevetés].
- (2) TP: Mondta, hogy hát igen, kanyarodunk, akkor dőlj be a kanyarba. De-hogy dőltem, én pont ellenkező irányba dőltem.  
A: [nevetés]

Ahogy más nonverbális elemek, a nevetések esetében is felvetődik a kérdés, hogy megvalósításuk milyen mértékben kultúra-, illetőleg nyelvfüggő. Egy tanulmányban olasz és német anyanyelvű beszélők nevetéseit elemezve nem találtak szignifikáns különbséget nyelvek szerint a jelenségek akusztikai fonetikai jellemzőiben (Rothgänger et al. 1998), egy másik kutatás azonban feltételezi a kultúraspecifikus nevetések létezését abból kiindulva, hogy a japán anyanyelvűek nevetése bizonyos fokig eltér attól, amit a nyugati kultúrákban produkálnak (Campbell et al. 2005).

### 3.2. Hümmögések

A hümmögésekkel kapcsolatban is felmerült annak lehetősége, hogy funkcióit és formáit tekintve a nyelvek között nem feltétlenül van megegyezés, de a perceptációs tesztek ezt nem támasztották alá (Markó 2005a). Artikulációs jellemzőit tekintve a hümmögés az orrüregen keresztül távozó zöngéből, esetleg ehhez társuló [h]-szerű zörejből áll (Markó 2005a: 91). Markó (2005a) öt hümmögéstípus prozódiai sajátosságait elemezte: az 'igen', a 'nem' jelentésű, a figyelemjelző, a semleges és az emocionális kérdő típusúét. Eredményei alapján elmondható, hogy az elemzett három alaptípus (igen, nem, kérdés) tagoltságukban és a dallamenet jellemzőiben eltér egymástól. Ezek a hangjelenségek önálló jelentéssel rendelkeznek, amely a szegmentális és szuprasegmentális szerkezet együtteséhez kapcsolódik, s e kettő nem választható szét.

Egy másik kutatás (Markó et al. 2014) a diskurzusban betöltött szerepük alapján vizsgálta a különböző hümmögések gyakoriságát, időtartamát és dallamenetét: figyelemjelző [(3) példa], megértést jelző [(4) példa] és egyetértést jelző hümmögések [(5) példa] esetében.

- (3) TP: én ugye még a régi rendszerben érettségiztem  
A: ühüm
- (4) TP: mesélj egy kicsit arról hogy milyen szakos vagy, illetve hogy mivel akarsz majd későbbiekben foglalkozni  
A: ühüm öö hát én ugye magyar szakra járok
- (5) TP: az olvasáshoz hozzátartozik a színház is nyilván és az is ilyen ellenkezéseket szokott kiváltani a diákokból  
A: aha igen persze ühüm

Húszy beszélő több mint 15 órányi anyagában azt találták, hogy a társalgásban produkált hümmögések leggyakrabban a hallgatók figyelmét jelzik a beszélő számára (visszacsatolásként), de különbségek mutatkoztak a nők és a férfiak között (előbbieket a figyelemjelző funkciójú hümmögéseket produkálták legnagyobb számban, utóbbiak pedig a megértést jelzőket). A kutatás fő hipotézise, miszerint a különböző funkciójú hümmögések prozódiai mintázatukban eltérnek, részben igazolódott: a típusok az akusztikai paraméterek közül az időparaméter mentén szignifikánsan különböztek.

### 3.3. Egyéb nonverbális hangjelenségek

Vicsi és munkatársai (2011) nagy mennyiségű spontán hanganyag feldolgozásával létrehozta egy olyan korpuszt, amely nyelvi tartalomtól független hangjelenségeket tartalmaz. A Hanggesztustárnak elnevezett gyűjtemény a csoportok-

ba sorolt hangeseményeket az akusztikai jellegzetességeikkel, valamint egy-egy minta hangképével együtt tárolja. A folyamatosan bővülő tárban megtalálhatók olyan hanggesztusok, mint például a nevetés, az igenlő hümmögés, a *hát*, az *izé*, a *na*, továbbá a kitöltött szünetek, a nem vokális eredetű hangok és a testhangok (sóhaj, szipogás stb.).

Két magyar tanulmány elemezte a BEA adatbázis felvételeiben, hogy bizonyos nonverbális vokális elemek (a ki-/belégzés, a köhögés, a nyelvcsettintés, a hallható nyelés, a nevetés) milyen gyakoriságban, időtartamban, helyzetben és milyen lehetséges funkciókkal fordulnak elő a spontán beszédben (Neuberger 2012; Bóna 2015).

Az egyik vizsgálat (Neuberger 2012) ötórás korpuszában a hangesemények közül a ki-, illetve belégzések valósultak meg leggyakrabban, a köhögés vagy torokköszörülés pedig a legritkábban. Összehasonlításként megemlíthetjük, hogy a Verbmobil korpusz mintegy 34 órás anyagát elemezve a nonverbális vokális elemek megjelenése hasonló arányokat mutatott, mint a jelen korpuszban: a legnagyobb számban ki- és belégzések, legritkábban torokköszörülések voltak adatolhatók (Prylipko et al. 2012). Ez összefüggésben állhat azzal, hogy az egyes típusok mennyire természetes velejárói a kommunikációnak: míg a levegő biztosítása alapvető fontosságú a beszélés során, a nyelvcsettintés az artikulációhoz, a nevetés a verbális tartalomhoz, a szemantikához kapcsolódik, addig a torokköszörülések, köhögések atipikus működést jelezhetnek, a beszélő (rossz) egészségi állapotáról tanúskodhatnak.

Az egyes típusok eltérő időtartamban realizálódtak, ami az artikulációs megvalósításukból fakad, valamint más-más pozíciókban fordulnak elő leggyakrabban a beszédszakaszok és szerkezetek szerint, ami összefügghet a funkciójukkal. Az egyes típusok különböző funkciókat tölthetnek be, és egy jelenség több funkciót is képes betölteni. A nyelvcsettintések többségénél a szövegekben való helyzetükből láthattuk, hogy a közlés indítását vagy a szóátvételt jelzik [(6) példa], más esetekben a mentális lexikonban való válogatásra utalnak [(7) példa]:

(6) TP: Fölfedezted három hónap alatt a várost vagy?

A: [nyelvcsettintés] hát igen igen, dolgoztam is.

(7) A: van a Budapesten a KICS a [nyelvcsettintés] Központi Irányító Csoport.

A ki- és belégzések általában a közlések tagolására, illetve az artikulációhoz szükséges levegő biztosítására szolgálnak. Ezt jelzi az a tény is, hogy sokszor fordulnak elő szerkezethatáron, valamint kötőszó előtt vagy után, mint például az (8) példában:

(8) az így elhíresült [belégzés] és nagyon sokan jönnek.

Ahogy a néma szünetek, az észlelhető levegővétellel kitöltött szünetek is időnyerési lehetőséget is jelentenek a beszédtervezéshez, gyakran együtt járnak más időnyerő stratégia alkalmazásával (pl. hezitálással), mint a (9) példában:

(9) hogy én a [belégzés] ööö blogokkal szeretnék foglalkozni.

Más esetekben a kilégzések, illetve a levegő sóhajszerű kifújása az üzenet tartalmához fűződő viszonyulást fejezi ki, mint a (10) példa mutatja:

(10) TP: faramuci dolgok vannak azért így a rendőr berkeken belül is  
A: [kilégzés]

Bóna (2015) tanulmánya az életkor függvényében elemzi a nonverbális hangeseményeket. Tíz fiatal (21–26 éves) és tíz idős (70–81 éves) beszélő mintegy kétórás hanganyagában azt találta, hogy az idősek beszédében gyakrabban fordultak elő a vizsgált jelenségek, az egyes típusok gyakorisága azonban eltérő volt a két életkorban. Az időseknél szignifikánsan gyakoribb volt a hangos lélegzés, mint a fiataloknál, és tendenciaszerűen gyakrabban fordult elő a köhögés és a nyelvcsettintés. Ez olyan fiziológiai, biológiai változásokra vezethető vissza, mint a szárazabb nyálkahártyák vagy a beszédszervek renyhébb működése. A fiataloknál szignifikánsan gyakoribb volt a nevetések előfordulása, ami a beszédpartnerek azonos életkorára, és így a felszabadultabb társalgásra vezethető vissza. A hangos nyelés előfordulása is tendenciaszerűen gyakoribb volt a fiataloknál; az egyes esetek elemzése pedig azt mutatta, hogy ezeket a beszédtervezési bizonytalanságok feloldására használták az adatközlők. A testhangok időtartama az idősebbeknél bizonyult rövidebbnek, ami a kisebb tüdőkapacitással magyarázható.

#### 4. A nonverbális hangjelenségek kutatásának haszna

Mindennapi tapasztalatunk és a rögzített hanganyagok tanúsága szerint a nem verbális hangjelenségek relatíve gyakori kísérőjelenségei a spontán beszédnek. Irodai környezetben rögzített felvételekben például a közlések 10%-a tartalmaz valamilyen nonverbális hangot, illetve zajt. Ezek forrása lehet maga a beszélő (torokköszörülés, lélegzetvétel stb.), vagy lehet környezeti zaj (telefoncsörgés, ajtócsapódás stb.) (Ward 1989). A nem beszéd jellegű hangokat az automatikus beszéd felismerés során általában zavaró tényezőként tartják számon, amelyek félrevehetik a szavak felismerését; ezért alapvető fontosságú elkülöníteni őket a szövegtartalomtól. Kellő mennyiségű előfordulás gyűjtése és akusztikai fonetikai elemzése után lehetőség nyílt az adott hangesemény akusztikai modelljének felépítésére (Vicsi et al. 2011). Más helyzetekben, például a beszéd szintézis esetén azonban pozitív szerepet játszanak, hiszen hozzájárulnak a beszéd termé-

szetes hangzásának kialakításához (Butzberger et al. 1992; Ruinskiy–Lavner 2007). Továbbá azt is kimutatták, hogy a beszélőfelismerésben is jótékony hatással lehetnek, főként telefonos beszédmintában (Janicki 2012). Emellett a nonverbális vokális jelek tanulmányozása szerepet játszhat a beszélők megítélésében (Mohammadi et al. 2010), valamint a kriminalisztikai szakterületen. A beszéd nyelvi tartalmán kívül információt adnak a beszélő személyéről, érzelmi állapotáról, attitűdjéről, így támogatást adhat az ember-gép kommunikáció számára is. Számos nem verbális hangesemény háttérchatorna-jelzésként jelenik meg, szabályozza a beszélgetés szerkezetét, előrejelzi a beszélésváltást, melyek a társalgáselemzés vizsgálati tárgyai. Mivel bizonyos hangesemények szerepet játszanak a társalgás szerveződésében, ezért a konverzációelemzésbe is integrálhatók a kutatási eredmények. A nonverbális hangjelenségek a társas interakció szerves részei, így elemzésük fontos adalékot nyújt az emberi kommunikáció pontosabb megértéséhez.

### Irodalom

- Bachorowski, Jo-Anne – Smoski, Moria J. – Owren, Michael J. 2001. The acoustic features of human laughter. *Journal of the Acoustical Society of America* 110/3. 1581–1597.
- Bickley, Corine – Hunnicutt, Sheri 1992. Acoustic analysis of laughter. In: *Proceedings of the International Conference on Spoken Language Processing*. Banff, Canada, 927–930.
- Bóna Judit 2015. Nonverbális hangjelenségek fiatalok és idősek spontán beszédében. *Beszéd-kutatás 2015*. 106–119.
- Bryant, Gregory A. – Aktipis, C. Athena 2014. The animal nature of spontaneous human laughter. *Evolution and Human Behavior* 35/4. 327–335.
- Burkhardt, F. – Paeschke, A. – Rolfes, M. – Sendlmeier, W. F. – Weiss, B. 2005. A database of German emotional speech. In: *Interspeech 2005*. Vol. 5. 1517–1520.
- Butzberger, John – Murveit, Hy – Shriberg, Elizabeth – Price, Patti 1992. Spontaneous speech effects in large vocabulary speech recognition applications. In: *Proceedings of the 1992 DARPA Speech and Natural Language Workshop*. Morgan Kaufmann, New York, 339–343.
- Campbell, Nick – Kashioka, H. – Ohara, R. 2005. No laughing matter. In: *Proceedings of the Interspeech 2005*. Lisbon, Portugal, 465–468.
- Engberg, I. S. – Hansen, A. V. – Andersen, O. – Dalsgaard, P. 1997. Design, recording and verification of a Danish Emotional Speech Database. In: *Proceedings of the Eurospeech '97*, Rhodes, Greece. [http://www.mirlab.org/conference\\_papers/International\\_Conference/Eurospeech%201997/pdf/wmc/A0385.pdf](http://www.mirlab.org/conference_papers/International_Conference/Eurospeech%201997/pdf/wmc/A0385.pdf) (A letöltés ideje: 2015. március 7.)
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Grácsi Tekla Etelka – Beke András – Neuberger Tilda – Nikléczy Péter 2012. BEA: Beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Budapest, Akadémiai Kiadó, 9–24.
- Hámori Ágnes 2006. Dominancia és barátságosság a Kulcsár-kihallgatáson. *Magyar Nyelvőr* 130. 129–165.

- Haq, S. – Jackson, P. J. B. 2010. Multimodal Emotion Recognition. In: Wang, W. (ed.): *Machine Audition: Principles, Algorithms and Systems*. IGI Global Press, 398–423.
- Horváth Viktória 2014. *Hezitációs jelenségek a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- Janicki, Artur 2012. On the impact of non-speech sounds on speaker recognition. In Sojka, Petr – Horák, Ales – Kopeček, Ivan – Pala, Karel (eds.): *Text, Speech and Dialogue. Proceedings of the 15th International Conference*. Brno, Czech Republic, Springer-Verlag, Berlin, Heidelberg, 566–572.
- Kenesei István (szerk.) 2000. *A nyelv és a nyelvek*. Corvina Kiadó, Budapest.
- Kennedy, Lyndon S. – Ellis, Daniel P. W. 2004. Laughter detection in meetings. In: *Proceedings of the NIST Meeting Recognition Workshop at the IEEE Conference on Acoustics, Speech and Signal Processing*. Montreal, Canada, 118–121.
- Knox, Mary Tai – Mirghafori, Nikki 2007. Automatic laughter detection using neural networks. In: *Proceedings of Interspeech 2007*. 2973–2976.
- Laver, John 1994. *Principles of phonetics*. Cambridge University Press, Cambridge.
- Markó Alexandra 2005a. „Szavak nélkül”. Nonverbális vokális közlések fonetikai elemzése. *Magyar Nyelvőr* 129. 88–104.
- Markó Alexandra 2005b. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. Doktori disszertáció, ELTE, Budapest.
- Markó Alexandra 2006. Nonverbális vokális jelek a társalgásban. *Beszédkutató* 2006. 57–68.
- Markó, Alexandra – Gósy, Mária – Neuberger, Tilda 2014. Prosody patterns of feedback expressions in Hungarian spontaneous speech. In: *Social and Linguistic Speech Prosody: Proceedings of the 7th International Conference on Speech Prosody*. Dublin, 482–486.
- Mohammadi, Gelareh – Vinciarelli, Alessandro – Mortillaro, Marcello 2010. The voice of personality: Mapping nonverbal vocal behavior into trait attributions. In: *Proceedings of the 2nd international workshop on Social signal processing*. ACM. 17–20.
- Navas, Eva – Hernáez, Inmaculada – Luengo, Iker – Sainz, Iñaki – Saratxaga, Ibon – Sanchez, Jon 2007. Meaningful parameters in emotion characterisation. In Esposito, Anna – Faundez-Zanuy, Marcos – Keller, Eric – Marinaro, Maria (eds.): *Verbal and nonverbal communication behaviours*. Springer-Verlag, Berlin–Heidelberg, 74–84.
- Neuberger Tilda 2012. Nonverbális hangjelenségek a spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 215–235.
- Neuberger, Tilda – Beke, András 2013. Automatic laughter detection in spontaneous speech using GMM-SVM method. In Habernal, I. – Matousek, V. (eds.): *Text, Speech, and Dialogue: TSD 2013*. Springer Verlag, Berlin–Heidelberg, 113–120.
- Nwokah, Eva E. – Davies, Patricia – Islam, Asad – Hsu, Hui-Chin – Fogel, Alan 1993. Vocal affect in three-year-olds: a quantitative acoustic analysis of child laughter. *Journal of the Acoustical Society of America* 94. 3076–3090.
- Provine, Robert R. 1993. Laughter. *American Scientist* 84. 38–45.
- Prylipko, Dmytro – Vlasenko, Bogdan – Stolcke, Andreas – Wendemuth, Andreas 2012. Language modeling of nonverbal vocalizations in spontaneous speech. In Sojka, Petr – Horák, Ales – Kopeček, Ivan – Pala, Karel (eds.): *Text, Speech and Dialogue. Proceedings of the 15th International Conference*. Springer-Verlag, Berlin–Heidelberg, 488–495.

- Roach, P. – Stibbard, R. – Osborne, J. – Arnfield, S. – Setter, J. 1998. Transcription of prosodic and paralinguistic features of emotional speech. *Journal of the International Phonetic Association* 28. 83–94.
- Rodrigues, Isabel Galhano 2008. Verbal and nonverbal modalities in face-to-face interaction: How they function as conversational signals. *Estudos Linguísticos/Linguistic Studies* 2. 211–227.
- Rothgänger, Hartmut – Hauser, Gertrud – Cappellini, Aldo Carlo – Guidotti, Assunta 1998. Analysis of laughter and speech sounds in Italian and German students. *Naturwissenschaften* 85. 394–402.
- Ruinskiy, Dima – Lavner, Yizhar 2007. An effective algorithm for automatic detection and exact demarcation of breath sounds in speech and song signals. In: *IEEE Transactions on Audio, Speech, and Language Processing* 15/3. 838–850.
- Scherer, Klaus R. – Banse, Rainer – Wallbott, Harald G. 2001. Emotion inferences from vocal expression correlate across language and cultures. *Journal of Cross-Cultural Psychology* 32. 76–92.
- Schötz, Susanne 2002. Linguistic and paralinguistic phonetic variation in speaker recognition and text-to-speech synthesis. In: *Speech technology* [http://www.speech.kth.se/~rolf/gslt\\_papers/SusanneSchotz.pdf](http://www.speech.kth.se/~rolf/gslt_papers/SusanneSchotz.pdf) (A letöltés ideje: 2015. április 22.)
- Seppänen, Tapio – Väyrynen, Eero – Toivanen, Juhani 2003. Prosody-based classification of emotions in spoken Finnish. In: *Proceedings of 8th European Conference on Speech Communication and Technology, Eurospeech*. Geneva, Switzerland, 717–720.
- Tóth, Szabolcs Levente – Sztahó, Dávid – Vicsi, Klára 2008. Speech emotion perception by human and machine. In: *Proceedings of COST Action 2102 International Conference. Revised papers in verbal and nonverbal features of human-human and human-machine interaction*. Springer, 213–224.
- Trager, George Leonard 1961. The typology of paralanguage. *Anthropological Linguistics* 3/1. 17–21.
- Traunmüller, Hartmut. 2000. Evidence for demodulation in speech perception. Contribution to *ICSLP 2000*. <http://www.ling.su.se/staff/hartmut/demod.pdf> (A letöltés ideje: 2015. 04. 22.)
- Troung, Khiet P. – van Leeuwen, David A. 2007. Automatic discrimination between laughter and speech. *Speech Communication* 49. 144–158.
- Trouvain, Jürgen 2003. Segmenting phonetic units in laughter. In: *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*. Barcelona, Spain, 2793–2796.
- Vicsi Klára – Sztahó Dávid – Kiss Gábor 2011. Nem verbális hangjelenségek spontán társalgásban. *Beszédkutatás* 2011. 134–147.
- Ward, Wayne 1989. Modelling non-verbal sounds for speech recognition. In: *Proceedings of the workshop on Speech and Natural Language*. Association for Computational Linguistics, 47–50.
- Whiteside, Sandra P. 1998. Simulated emotions: An acoustic study of voice and perturbation measures. In: *Proceedings of the International Conference on Spoken Language Processing*. Sidney, Australia, 699–703.
- Yildirim, Serdar – Bulut, Murtaza – Lee, Chul Min – Kazemzadeh, Abe – Busso, Carlos – Deng, Zhigang – Lee, Sungbok – Narayanan, Shrikanth 2004. An acoustic study of emotions expressed in speech. In: *Proceedings of the International Conference on Spoken Language Processing*. Jeju, Korea, 2193–2196.

### **Kérdések, feladatok**

1. Hogyan lehet csoportosítani a nonverbális kommunikáció elemeit?
2. Milyen funkciókat tölthetnek be a nonverbális hangjelenségek?
3. Ismertessen néhány kutatást, amely a nonverbális hangjelenségekre irányul!



## A zaj hatása a beszédre

### 1. Bevezetés

A verbális kommunikáció sikerét vagy hibátlan működését számos tényező befolyásolja, melyek közül talán a legmeghatározóbb a megfelelő akusztikai környezet. Az érthetőség döntően függ a beszéd hangerejétől, a beszélő és a hallgató közti távolságtól, a környezeti zajoktól, illetve ezek összhangjától (Tarnóczy 1984). Mivel a világban minden, ami él és mozog, hangot kelt (Pap 2002), az emberek közötti kommunikáció szinte mindig valamiféle hangzó közegben megy végbe. A különféle hanghatások, zajok, amelyekkel mindennapjainkban találkozhatunk (pl. járművek, épületek zaja, természeti jelenségek, állatok hangja, szórakoztató elektronikai eszközök hangja stb.), nagymértékben befolyásolják közléseink sikerességét.

Hangnak – fizikai értelemben – a rugalmas közegben (levegő, folyadék, szilárd anyagok) fellépő mechanikai rezgéseket és hullámokat nevezzük. Ez a rezgés az emberi hallószervbe jutva az agyban hangérzetet kelt. A hang erőssége a nyomásingadozástól, a nyomáshullámok nagyságától függ (egyenes arányban). Az emberi füllel érzékelhető legkisebb hangnyomást **hallásküszöbnek**, míg a már fájdalmat okozó hangnyomást **fájdalomküszöbnek** nevezzük. A két küszöbérték közötti tartomány 12 nagyságrendnyi, ami azt jelenti, hogy a leghangosabb, még éppen elviselhető hang  $10^{12}$ -szer erősebb a leghalkabb hangnál. A hang fizikai erőssége és a keltett érzet nagysága között exponenciális kapcsolat van, tehát egy adott hangnál a fizikailag tízszer erősebbet egy egységnyivel érezzük hangosabbnak (a százszor erősebbet kettő, az ezerszer erősebbet három egységnyivel). Emiatt használnak a hangintenzitás mérésére logaritmikus skálát, amelynek mértékegysége a decibel (dB). A hallásküszöb tehát 0 dB, az ennél százszor erősebb hang 20 dB, a százezerszer erősebb 50 dB. A hang terjedési sebessége a közvetítő közeg függvényében változik; levegőben 330-340 m/s, míg folyadékokban és szilárd anyagokban sokkal nagyobb (vízben kb. 1400 m/s, acélban 5000 m/s). Légüres térben ezek a mechanikus rezgések nem terjednek.

Az emberi fül (ép hallás esetén) a 20 Hz és 20 000 Hz közötti frekvenciatartományban képes érzékelni környezetünk hangjelenségeit, ezt nevezzük **hallható hangnak**. Az ezen kívül eső hangokat ugyan nem halljuk, mégis hatással lehetnek szervezetünkre. A 16 Hz alatti, nagyon alacsony frekvenciájú hangok az **infrahangok**, melyek magas intenzitás (nagy hangerő) esetén dezorientálják, megrémítik, megzavarják az egyént; nyugtalanságot, rossz közérzetet, szédülést,

egyensúlyvesztést eredményeznek. Gátolják továbbá az érzékszervek funkcióit; fejfájást, émelygést, hányingert és bélgörcsöt idézhetnek elő. Szélsőséges esetekben a magas intenzitású infrahangnak kitett személyek lélegzétvétele akadozóvá válhat, félelemérzetük támadhat; egyeseknek epilepsziás rohamaik lehetnek. Nagy valószínűséggel ezek az alacsony frekvenciájú hangok állnak a szellemlátások és a kísértettörténetek hátterében. Brit kutatók több „szellemjárta” helyszín esetében is igazolták az infrahangok jelenlétét. Az infrahangok figyelése az élet számos területén nagy jelentőséggel bír; segítségükkel előre jelezhetők a földrengések, felkutathatóak az olaj- és földgázlelőhelyek, tanulmányozható a szív működése (ballisztó-kardiológia és szeizmo-kardiológia), illetve észlelhetővé válnak az esetleges nukleáris robbantások is. Mivel bármely közegben kevéssé csillapodnak, nagy távolságokat, akár több száz kilométert is képesek áthidalni, terjedésüket nagyon bonyolult meggátolni, így akár szilárd építmények falain is könnyedén áthatolnak (ennek segítségével kommunikálnak pl. az elefántok és a bálnák).

A 20 000 Hz fölötti rezgések, az **ultrahangok**, amelyek magas hangteljesítményszintnél fülzúgáshoz, hányingerhez, szédüléshez, illetve fejfájáshoz vezethetnek. Az ultrahangot feloszthatjuk aktív és passzív tartományokra; az előbbi főként az anyag tulajdonságainak megváltoztatására, míg utóbbi az anyagtulajdonságok vizsgálatára használható. A passzív ultrahang segítségével például kimutathatóak a fémekben, keménysajtokban lévő törések, repedések; felkutathatóak a barlangok, kőzetüregek, de akár a halrajok is. Elemezhetőek általa a talajrétegek, így olajipari alkalmazásokra is használható; de az orvosi képkalkító eljárások is ezt alkalmazzák. Az aktív ultrahang nagy jelentőséggel bír a biotechnológiában, az élelmiszeriparban (pl. tisztítás, csírátlanítás, húspácolás, ill. tézsták kelesztésének gyorsítása, húsok, csontok vágása, rizs és más szemes termények mosása stb.); kémiai alkalmazási területei közé tartozik például a kristályosodási fok, valamint az anyagi minőség meghatározása, illetőleg az ultrahangos zsírtalanítás. A felsoroltakon túl az aktív ultrahang terápiás célokra is felhasználható, úgy mint fogkö-eltávolítás, vesekőzúzás, kozmetikai célok (mélymasszázs, hajbeültetés), rákos sejtek elpusztítása, ultrahangsebészet, génmanipuláció; de a háztartási felhasználási lehetőségei is szerteágazóak (zsíros, odaégett szennyeződések eltávolítása, ékszeres tisztítása, textilfehérítés).

A hétköznapiakban rengeteg hallható hang vesz minket körül, melyek egy részét kellemesnek, másik részét kellemetlennek értékeljük. A **zaj**, a fizikai definíciója szerint végtelen sok rezgést tartalmazó, szabálytalan szerkezetű hang; köznap értelemben véve pedig az adott körülmények között zavaró hangot jelent (Pap 2002). Azt gondolhatnánk, hogy a zaj az ipari, technológiai fejlődés és a modern korok velejárája, azonban különféle feljegyzésekből tudjuk, hogy már az ókori emberek is szenvedtek tőle. Marcus Valerius Martialis költő az I. században írt arról, hogy a városnak ugyan megvannak az előnyei, de egyben idegesítő és fárasztó is, mert túl sok az ember és túl sok a zaj. Seneca, az ókori filozófus

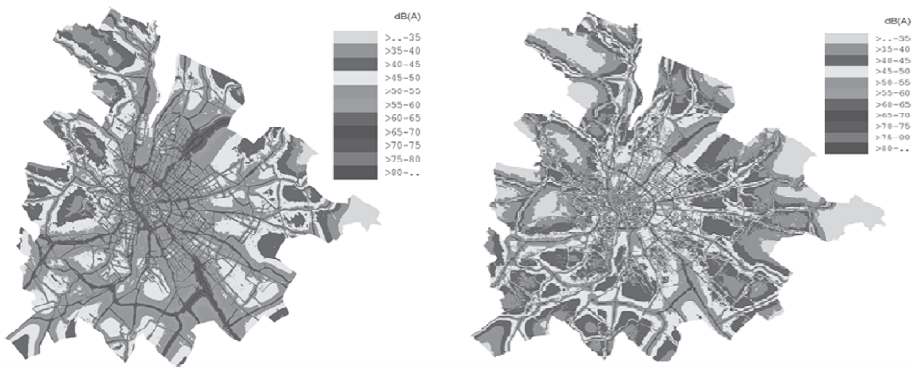
pedig, aki egy fürdő felett lakott, egyik írásában következőképp panaszkodott a felhallatszó zajok elviselhetetlenségéről: „Akik ott súlyzókkal edzenek, azok nyögnek, zihálnak. Ha valaki éppen masszíroztatja magát, a masször tenyerének csattogása visszhangzik különböző hangokon, aszerint, hogy nyitott tenyérrel, vagy begörbítettel paskolja éppen kliensét. Van, aki hangosan fecseg, veszekszik, vagy nagy lendülettel veti bele magát a vízbe. A szőrtelenítő kiabálva ajánlja szolgáltatásait és csak akkor hallgat, amikor valakinek a szőreit tépi, akkor viszont az ordít, akit kezelésbe vesz”. Julius Caesar a városok zajosságának csökkentésére dekrétumot adott ki, amely előírta, hogy vásár- és ünnepnapokon csak gyalog szabad közlekedni az utcákon; Claudius császár pedig kitiltott Róma városából mindenféle kerekeken guruló járművet. A középkorban a csendháborítókat büntetésképpen zúgó harang alá ültették.

A zaj mérésére elsőként 1959-ben Kryter tett kísérletet. Meghatározta a zajossági skálát, mértékegységül pedig az 1 noy-t választotta, ami 900 Hz és 1100 Hz között 40 dB-es intenzitásértéket jelent (Tarnóczy 1984). Noha a minket minden nap körülvevő zajok (közlekedés, gyermekzsivaj, rádió, tévé, a lakókörnyezet zajai stb.) már-már annyira megszokottá, természetessé váltak számunkra, hogy csak ritkán érezzük elviselhetetlennek őket, még ha nem is figyelünk fel rájuk, károsítanak és fárasztanak. Az emberi szervezetre gyakorolt hatások feloszthatók **hallószervre**, illetve **nem hallószervre** (pszichikai és vegetatív) **gyakorolt hatásokra** (Tulipánt 2004). A tartós zajterhelés elsődleges következménye a halláskárosodás, ami a hallásküszöb emelkedését vonja maga után. Kevésbé ismertek azonban a pszichikai, illetve vegetatív hatások. A zaj zavarja, és lassítja a gondolkodási folyamatokat, aminek következtében az egyén ingerültté, illetve fáradttá válhat. A szellemi feladatmegoldó készség kismértékű romlása már 50 dB zajszintnél (ami a halk beszéd, illetve a villanyborotva hangerejével egyezik meg) tapasztalható, a 65 dB feletti zajok pedig már magatartásbeli változásokat és stressztüneteket eredményezhetnek. A magasabb zajszint a szervezetből vegetatív hatásokat is kivált. Ennek következtében felléphetnek anyagcserezavarok, megváltozik a keringési rendszer működése, nő a pulzusszám és a vérnyomás, a pupilla pedig kitágul (vö. 1. táblázat). Mindezek mellett előfordul, hogy megnő az agyfolyadék nyomása, ami fejfájáshoz vezet (Tulipánt 2004).

1. táblázat. A zaj emberi szervezetre gyakorolt hatása a hangosság függvényében

30 dB zajszint	pszichés hatás
65 dB zajszint	vegetatív hatás
90 dB zajszint	hallószervi hatás
120 dB zajszint	fájdalomküszöb
120–130 dB zajszint	maradandó halláskárosodás
160 dB zajszint	dobhártyarepedés
175 dB zajszint	halálos

A WHO kezdeményezésére 2002–2003-ban nyolc európai országban lefolytatott LARES-tanulmány a lakókörnyezetben tapasztalható zajszennyezés és az orvosilag diagnosztizált megbetegedések összefüggését vizsgálta. Az eredmények alapján megállapították, hogy a krónikus zajterhelés kapcsolatba hozható olyan betegségek kialakulásával, mint a kardiovaszkuláris rendellenességek, keringési zavarok, légzőszervi panaszok (pl. bronchitis), mozgásszervi panaszok (pl. ízületi gyulladás), allergia, migrén, illetve depresszió (Niemann et al. 2005). Kimondható tehát, hogy a zaj a modern civilizációs betegségek egyik rizikófaktor, amely egyrészt veszélyezteti egészségünket, másrészt pedig rontja az egyes települések akusztikai minőségét. Az EU lakosságának mintegy 40%-a napközben 55 dB-t is meghaladó közúti zajterhelésnek van kitéve, 20%-a pedig 65 dB-nél is magasabb zajszintben él. Éjjel a lakosság 30%-a kénytelen 55 dB-nél is nagyobb zajt elviselni, amely már bizonyítottan zavarja a pihentető alvást, így egészségkárosító hatása van (Brendt–Bibók 2003). (A Budapestre vonatkozó adatokat lásd az 1. ábrán).



1. ábra. Budapest stratégiai zajtérképe, bal oldal: nappal, jobb oldal: éjjel (forrás: Brendt–Muntág 2007)

A közúti közlekedési zaj hosszútávon **halláskárosodáshoz** vezethet; ilyenkor megemelkedik a hallásküszöb (romlik az egyén hallása), és a kialakult nagyothallás (halláscsökkenés) akadályozza az egyén mindennapi életben való tájékozódását. **Romolhat a beszédértés**, beszűkülnek az ember kommunikációs lehetőségei, megváltozik a magatartása, és az sem kizárt, hogy ennek hatására zavarok keletkeznek a személyiségében. Orvosi szempontból a környezeti zajokkal kapcsolatos talán a legjelentősebb probléma az **alvás zavarása**. Az éjszakai nyugodt pihenés ugyanis a munkaképesség és az egészség megőrzésének, valamint a jó fiziológiai és mentális működéseknek az előfeltétele. Az alvászavar leggyakrabban az elalvás nehézségében, többszöri éjszakai felébredésben, az alvás fázisainak szabályosságában vagy mélységében fellépő változásokban nyilvánul meg. Mindez

olyan fiziológiai hatásokat von maga után, mint a megnövekedett vérnyomás, a szívritmus változása, arhythmia, növekvő pulzus, erek összehúzódása, légzésritmus változása, krónikus fáradtságérzet, idegesség, ingerlékenység. Mindezek következtében az egyén rosszkedvűvé válik, teljesítménye és koncentrációképessége csökken. A zaj hatással van a **mentális egészségre** is; szorongást, emocionális stresszt, idegi panaszokat, hányingert, fejfájást, hangulati ingadozásokat, impotenciát okozhat, illetőleg előidézhethet általános pszichiátriai rendellenességeket, mint a neurózis, pszichózis, hisztéria. Negatívan befolyásolja továbbá a **magatartást** is: tartós zajterhelés hatására az egyén barátságtalanná, kedélytelenné, agresszívvá válik, csökken az együttműködési hajlandósága (Barótfi 2000).

Kutatásokkal igazolták, hogy a fő közlekedési utak mentén lakó gyermekeknek magasabb a pulzusszámuk és a vérnyomásuk; iskolai teljesítményük elmarad az átlagostól (Gatschnegg 2003). A zajos környéken lévő iskolákban csökken az oktatás hatékonysága, a kórházakban hosszabbodhat a gyógyulási idő. A zajos munkahelyeken a teljesítőképeség csökkenését, a tevékenység lelassulását, a koncentráció romlását, a figyelmetlenség és a feszültség növekedését dokumentálták, amely csökkenti a munkaintenzitást, és növeli a balesetveszélyt (Póta 2006).

Ahogy azt korábban említettük, a zajos környezet negatívan befolyásolja a verbális kommunikációt, mind a hallgató, mind pedig a beszélő részéről. A közelmúltban több kutatás foglalkozott a beszéd érthetőségével zajos környezetben, melyekből kiderült, hogy a gyermekek beszédmegértési teljesítményét már az alacsony intenzitású háttérzaj is rontja, míg a felnőttek a zajban elhangzó beszédet jobban képesek észlelni és megfelelően feldolgozni (Hygge 2003; Kloepper et al. 2006).

A zaj hatására azonban nem csupán a beszéd érthetősége, de a beszélő személy beszédprodukciója is romlik. A Lombard-hatást, miszerint háttérzaj hatására a beszéd intenzitása szignifikánsan megnő, elsőként Etienne Lombard francia fül-orr-gégész, audiológus ismertette egy 1911-ben megjelent tanulmányában. Az intenzitásnövekedéshez vezető artikulációs működések automatikusak, a beszélő tehát ösztönösen növeli a hangerejét a zajos környezetben. Ahhoz, hogy a beszéd még érthető maradjon, legalább 15 dB-lel kell hangosabbnak lennie a háttérzajnál. Egy átlagos hangerejű beszélgetés (55–60 dB) esetén tehát a környezeti zajoknak nem szabadna meghaladni a 40–45 dB-t ahhoz, hogy a kommunikáció zavartalan legyen (Pompetzki 2000). Az 2. táblázatból kiderül, hogy ez a kritérium csak kevés esetben teljesülhet. Az esetek többségében a beszélő a környezeti zaj miatt ösztönösen meg kell, hogy emelje a hangerejét, hogy közlése érthető legyen; a beszélőt a produkciós folyamatokban, míg a hallgatót a percepciók folyamatokban akadályozza a zaj.

2. táblázat. Hangforrások hangnyomásszintjei (Tarnóczy 1982 alapján)

Mesterséges hangok	Természetes hangok, zenei hangok	dB
	testzörejek (légzés)	0
	falevélsusogás	10
karóraketyegés (1 méterről)	szellő	20
hangversenytermi alapzaj	suttogás	30
halk rádiószó	vízcsobogás	40
villanyborotva	halk beszéd	50
porszívó (3 méterről)	átlagos beszéd	60
személyautó	hangos beszéd	70
autóforgalom (10 méterről)	tipsorkán	80
nagy gépjármű	walkman (közepes hangerő); viharos szélzúgás	90
légkalapács	diszkó (közepes hangerő)	100
kovácsműhely	Niagara; walkman (maximális hangerő)	110
légcsavaros repülő (10 méterről)	diszkó (maximális hangerő)	120
	tengeri vihar	130
sugárhajtású repülő (1 méterről)	villámcsapás (10 méterről)	140

## 2. A zaj hatása a beszédre

A folyamatos környezeti zaj nem csupán a spontán beszéd folyamatait befolyásolja; a zajhatás következtében megváltozott beszéd feldolgozása is más stratégiákat igényel a hallgató részéről. Nyelvészeti szempontból a zajos környezet hatásaival eddig magyar nyelven kevés kutatás foglalkozott. Balázs Boglárka és Gósy Mária az 1980-as évek végén vizsgálták különböző zajoknak beszédre (produkción) és annak feldolgozására (percepción) gyakorolt hatását (Balázs–Gósy 1988). Kísérletükben 6-féle, szubjektíve kellemesnek és kellemetlennek ítélt zajt (metróállomás, kutyaugatás, elektromos írógép, társalgás, présgép, zene) használtak fel. A zajok mind időbeli lefolyásukban, mind frekvenciaeloszlásukban különböztek egymástól; voltak köztük folyamatosak és kvázi-periodikusak (bizonyos időtartamú megszakításokkal ismétlődtek a hangeffektusok), illetve a beszédfrekvenciákra és a spektrum más tartományaira esők is. A zajokat három decibelértéken rögzítették (55, 70, és 85 dB), minden szinten 25 másodpercnyi időtartamban. A kísérletben részt vevő 30 adatközlő feladata az volt, hogy *A kis herceg* című meseregényből olvassanak fel hangosan pár oldalt. 25 másodpercnyi hangos olvasás után a fejükre fülhallgatót tettek, így kellett további 25 másodpercen át olvasniuk, majd ezután a fülhallgatón keresztül három különböző intenzitásértéken (egyre hangosabban) hallották a zajokat (egy kísérleti személy kétfajta zajt). A kísérlet után az adatközlőknek megértést ellenőrző kérdésekre kellett válaszolniuk, melyek egy része a csendben, más része a zajban olvasott szövegrészekre vonatkozott. Az eredmények igazolták, hogy zaj hatására nő a

beszélők hangereje, hangképzésük préselté, feszítetté, a normál, egészséges képzéstől eltérővé válik; beszédmegértésük jelentős mértékben romlik. A különböző zajok eltérő mértékben bizonyultak zavarónak; legkevésbé a zene és a társalgás, míg a leginkább a prégép és az írógép zavarta a beszélőket.

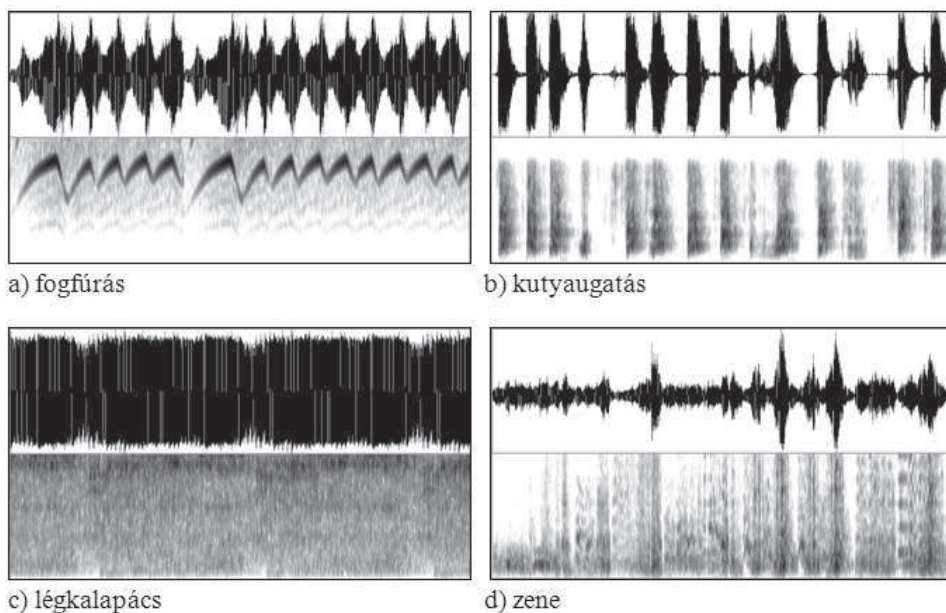
Egy másik magyar kutatás arra kereste a választ, hogy miként befolyásolja a beszédzaj a fiatalok beszédét és szövegértését (Gósy 2008). A főbb kérdések között szerepelt többek között, hogy i) a zaj hatására csupán az intenzitás (hang-erő) változik-e, avagy a beszéd más paraméterei is módosulnak; ii) hatást gyakorol-e a zaj a beszédtervezés rejtett működéseire; iii) igazolható-e, hogy a középiskolások számára a relatíve kis intenzitású beszédzaj (amelyben a mindennapjaikat töltik) nem okoz feldolgozási nehézséget. A kísérlet első része a beszédprodukció vizsgálatára irányult. A 16 (8 férfi, 8 nő) 20 és 30 év közötti adatközlőtől először egy kétperces spontán beszédet rögzítettek, amely közben (anélkül, hogy azt előzetesen a beszélők tudtára hozták volna) egy perc elteltével szabad hangtérben beszédzajt közvetítettek; a továbbiakban ebben a háttérzajban kellett beszélniük. Az átlagos jel/zaj arány 15 dB volt, ami azt jelenti, hogy az adatközlők beszéde átlagosan 15 dB-lel volt hangosabb, mint a háttérzaj. Az eredmények igazolták, hogy a beszéd zajhatásra hangosabbá, és egyben monotonabbá válik, a hangképzés préseltebb, levegősebb lesz. Az adatközlők a zajos beszéd részben gyakrabban és hosszabb szüneteket tartottak. Artikulációs tempójuk gyorsult, tehát az egyes beszédhangok képzésének időtartama rövidebbé vált; beszédtempójuk viszont (amibe a szünetek is beleszámítanak) lassult. A beszélők tehát láthatóan igyekeztek szabadulni a számukra kellemetlen helyzetből, a háttérzaj miatt azonban nehezebbé vált számukra gondolataik nyelvi átalakítása, amit jól jelez a megnövekedett szünettartás, illetve a megakadásjelenségek (nyelvbottlások) gyakoribbá válása. A kutatás második részében a beszédmegértést vizsgálták: az adatközlőknek egy tudományos ismeretterjesztő szöveg meghallgatása után kellett 10 megértést ellenőrző kérdésre válaszolniuk. A kísérletben 88 16 és 18 éves diák vett részt (4 osztályból, osztályonként 22 fő). Kétféle hanganyagot hallgattattak le velük, egy zaj nélkülit és egy beszédzajjal elfedettet (a jel/zaj arány itt is 15 dB volt); mindkét évfolyamból az egyik csoport a zaj nélkülit, a másik a zajos szöveget hallgatta meg. A két évfolyam szövegértése sem a zajos, sem a zaj nélküli szöveg esetében nem mutatott különbséget, de mind a 16, mind a 18 évesek teljesítménye romlott a zaj hatására. Ilyenkor csökken a megértés biztonsága, gyengül az értelmezés pontossága, a beszédfeldolgozás korlátozottá válik.

Szintén a beszédzajt (társalgás) hívták segítségül a kutatók, hogy választ kaphassanak arra, miként működnek az önellenőrzési folyamatok, ha megakadályozzák, hogy a beszélő a saját, hallásalapú észlelési és megértési folyamatainak segítségével ellenőrizze saját beszédét (Gósy–Bóna 2011). A téma fontossága abban áll, hogy hétköznapijainkban a verbális kommunikáció szinte mindig

valamilyen háttérzajban történik; amely gyakorta maga is beszéd. Több kutató is azon a véleményen van, hogy a beszélő számára ez a fajta zaj jelenti a legnagyobb nehézséget produkciós folyamataik működtetése során (pl. Hativa 2000; Vickers 2007; Lu–Cook 2008). A kutatás fő kérdései a következők voltak: i) változnak-e a rejtett ellenőrzési folyamatok, ha a külső ellenőrzést gátolt; ii) mindez milyen hatással lesz a nyelvbtlások megjelenésére, temporális sajátosságaira és javítására iii) a felszíni ellenőrzés hiányában azonos vagy eltérő mintázatot mutat-e a rejtett és a felszíni hibajavítás. A kiinduló feltételezés szerint a beszédzaj előidézte nehezített kommunikációs helyzet a nyelvbtlások nagyobb arányú előfordulásához vezet; illetőleg a felszíni ellenőrzés lehetőségének kizárása miatt jobban fog érvényesülni a rejtett ellenőrzés. A kísérletben 30 fiatal felnőtt vett részt, akiknek két történelmi mondát kellett fülhallgatón keresztül meghallgatniuk; mindkettőt zajmentesen. Az első szöveg meghallgatása után csendes körülmények között kellett elmesélniük a lehető legpontosabban a hallottakat, míg a második szöveg tartalomösszegzése alatt fülhallgatón keresztül folyamatos beszédzajt közvetítettek a fülükbe. A beszédzaj átlagosan 30 dB-lel volt hangosabb az átlagos beszédnél. Az eredmények szerint sem az artikulációs (a hangok képzésének időtartama), sem a beszédtempó nem változott csoportszinten a zaj hatására, egyéni különbségek azonban adódtak; volt akinél csökkentek, volt akinél nőttek a tempóértékek. Szintén nem találtak jelentős különbséget a szünettartásban; sem az előfordulási arányuk, sem a hosszuk nem változott döntő mértékben. Megakadások (nyelvbtlások) a zajos beszéd részben valamivel kevesebbszer fordultak elő, az egyes típusok aránya azonban változott; többféle hiba jellemezte a zajban történt tartalomösszegzéseket. A hibajavítás mértéke a zajhatásra drasztikusan csökkent, azonban a beszélők jellemzően gyorsabban javították tévesztéseiket, amely az ellenőrzési folyamatok gyorsabb működésére utal. Miután a beszélő nem számíthat a felszíni ellenőrzésre, az akusztikus visszajelzésre, ösztönösen fokozottabban ellenőrzi beszédét rejtett módon, növeli figyelemkoncentrációját.

Az előzőektől némiképp különbözik az a két kutatás, amely négy, a nagyvárosi élet szerves részét képező környezeti zaj (kutyaugatás, léghalpapács, fogfúrás, zene) spontán beszédre (Gyarmathy 2008) és az önmonitorozó mechanizmus működésére gyakorolt hatását (Gyarmathy 2012) elemezte. Megnézték, hogy a beszéd mely paramétereiben tapasztalható változás, továbbá hogy a zaj hatással van-e a beszédtervezés rejtett működésére. Választ kerestek arra, hogy az általuk szubjektíve kellemetlennek, illetve kevésbé kellemetlennek ítélt zajok hatására miként módosul a beszélők, artikulációs és beszédtempója, szünettartási stratégiája, a megakadásjelenségeik percenkénti előfordulása, továbbá az egyes típusok gyakorisága. Ezeknek a paramétereknek a megváltozása utalhat a tervezési folyamatok nehezítettségére. A zaj hatására feltehetően a különböző produkciós szinteken is több probléma jelentkezik, amely befolyásolhatja a

korrekciós folyamatok működését is. A kísérlet során az egyes zajokat fülhallgatón keresztül juttatták az adatközlők fülébe, akadályozva az akusztikai visszacsatolást, amivel korlátozták a felszíni önmonitorozás működését. Ennek megfelelően elvárták, hogy az önkorrekció a zajos beszédhelyzetben rosszabbul működik majd. Feltételezték, hogy az egyes zajok eltérő mértékű zavaró hatása megmutatkozik a produkcióban is. A zavarás mértékétől függően nőni fog a megakadásjelenségek száma, illetve a típusok gyakorisága, továbbá a korrekciós folyamatok is kevésbé lesznek sikeresek. A kutatásban négyféle külső hanginger (2. ábra), a fogfűrés, a kutyaugatás, a légkalapács és a klasszikus zene (Maurice Ravel: *Pavanne egy infánsnő halálára* című műve) spontán beszédre, illetőleg az önmonitorozásra gyakorolt hatását vizsgálták 20 fiatal női beszélő részvételével. Adatközlőink a négy zajtípus alapján öt-öt fős csoportokat alkottak. Feladatuk egy 10 képből álló képsor alapján egy történet (minél részletesebb) elmesélése volt. Az instrukció szerint, bármi zavaró tényező is lépjen fel a felvétel során, a történetmesélést folytatniuk kellett. Előzetesen nem közölték velük, hogy zajos környezetet teremtenek. A hatodik képtől fülhallgatón keresztül (ami a kísérlet elejétől a fejükön volt) 60 dB erősségű zajt közvetítettek.



2. ábra. A kutatáshoz felhasznált négy zajtípus rezgés- (a regisztrátumon felül) és hangszínképe (a regisztrátumon alul)

Elsőként a beszédidők alakulását elemezték a csendes és a zajos környezetben. Az adatközlők eltérő mértékben ugyan, de mindegyik csoportban átlagosan hosszabb ideig beszéltek zajban, mint csendben. A legnagyobb mértékű növekedés a kutyaugatás csoportjába tartozó beszélőknél volt adatolható. Ez több okra is visszavezethető. Egyrészt elképzelhető, hogy mivel a kutyaugatás mint háttérzaj bizonyos többletjelentést hordoz (veszélyhelyzet, betörés stb.), jobban elvonja a beszélők figyelmét az éppen aktuális közlésről. Másrészt lehetséges, hogy a vizsgált zajok közül (kellemetlensége miatt) a kutyaugatás hátráltatja leginkább a beszédprodukción (ezt a megakadások nagy száma is alátámasztani látszik). A beszédidő növekedésének további okaként meg kell említenünk, hogy az ebbe a csoportba tartozó összes beszélő beleszótta valahogyan a történetmesélésbe a kutyaugatást, ezzel mintegy kibővítve és meghosszabbítva a cselekményt. Például: *bentről egy kutya, kutya ugatását hallja a hőstünk, szőke hőstünk, de nem zavartatja magát, örülnek a találkozásnak; vagy: a kutya az nem szereti ezt a ezt az embert, de a gazdája se zavartatja magát, és hát bemennek a lakásba, bezárják a kutyát a fürdőszobába*. Sokszor arra is reflektáltak a történetmesélés közben, hogy a kutyaugatás zavarja őket: *hát oda akarja kötni gondolom, vagy föl akarja húzni? nem igaz ez a hülye kutya! szóval igaz, hogy a kutya tovább ugat, de szőke főhőstünk leengedi a kötelet*. Jelentősen hosszabbodott a beszédidő továbbá a légkalapács és a zene hatására, az okok azonban feltehetőleg nem azonosak. Elemezték a megakadásjelenségek előfordulását az egyes csoportokban. Mind a négy háttérzaj hatására – eltérő mértékben ugyan, de – csökkent a megakadásjelenségek gyakorisága. A legnagyobb mértékű csökkenést a fogfúrás esetében adatolták, ezt követte a zene, a légkalapács, majd a kutyaugatás csoportja. A háttérzaj hatására a beszélők vélhetően jobban odafigyelnek a tervezés és a kivitelezés részfolyamataira, hogy a mindenkori cél, a beszéd érthetősége teljesüljön. A legtöbb megakadást percenként a zajos beszédrészben a kutyaugatásnál adatolták, úgy tűnik tehát, hogy ez zavarja leginkább a tervezés és a kivitelezés összehangolását. A hiba típusú megakadásjelenségek közül a zajos beszédrészben többféle típust volt adatolható, ami abból adódik, hogy a beszélő a tervezési folyamat több szintjén is problémával küzd. A kontaminációs hibák megjelenéséből következtethetünk arra, hogy a beszélő a zaj hatására az egymással versengő szerkezetek közül nem képes időben választani, így egyfajta furcsa, kevert alak jelenik meg a felszínen, például: *várja ezt a második fiatalembert egy szakállas, morcona úr marcona úr (morcos × marcona)*. A zaj zavaró hatása a „nyelvem hegyén van” jelenség előfordulásában is megmutatkozik. A beszélő a tervezés során nem képes hiba nélkül eljutni a lemma szintről a lexéma szintre (pl.: *a kalapos aj nem ajtónálló, a aki kalapos barát*).

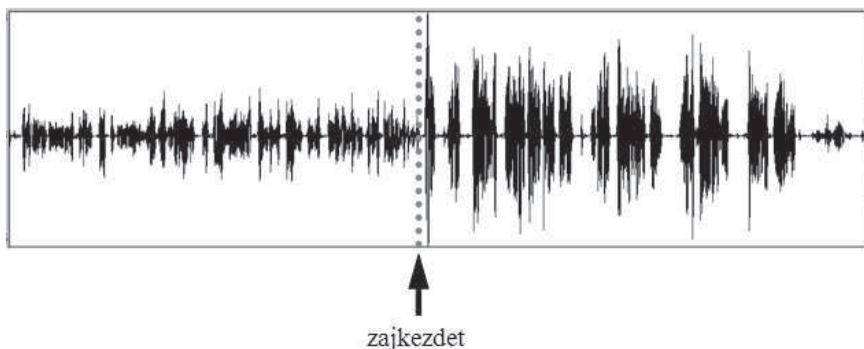
Az egyes zajtípusoknál eltérően alakult a megakadásjelenségek előfordulása. A hiba típusú jelenségek aránya a kutyaugatás és a fogfúrás esetében a zaj hatására gyakorlatilag nem változott, a többi zajnál kismértékű növekedés volt

adatokkal. Jelentősebb változás csupán a megakadástípusok gyakoriságában mutatkozott. A kutyaugatás kivételével minden zaj hatására többféle hiba jellemezte az adatközlők beszédét; a fogfúrás esetén például 18 fajta megakadást produkáltak a kísérleti személyek. A kutatáshoz választott háttérzajok közül ki kell emelni a zenét, amelynek hatása nem azonos a zajéval, a pszichés teljesítő-képességre gyakorolt hatását az eddigi kutatásoknak még nem sikerült egyértelműen tisztázni. Egyénre gyakorolt hatása természetesen függ a zene milyenségétől és hangerejétől (egy diszkóból kihallatszó hangos zene a környező lakosok számára zajként realizálódhat), a vizsgálati eredmények szerint azonban a halk zene a pszichés állapotra kedvezően hat, nem csökkenti, hanem esetenként még növeli is a figyelmet (Tulipánt 2004). Már a Bibliában is megemlíti a zene jótékony, gyógyító hatását: „Az Úr lelke elhagyta Saul, s egy gonosz lélek zaklatta, az Úr küldte. Szolgái így szóltak Saulhoz: »Minden bizonytalansággal Istennek egy gonosz lelke zaklat. Parancsolja hát meg urunk, s szolgái keresnek neki valakit, aki tud hárfázni. Aztán, ha rád tör Istennek a gonosz lelke, megpendíti a húrokat és jobban leszel.« [...] S ahányszor csak rátört Saulra az Isten lelke, Dávid fogta a hárfát és játszott. Ilyenkor Saul megnyugodott, jobban lett, és a gonosz lélek odébbállt” (I. Sám 16.14–16.16, 16.23).

A kísérlet eredményei alapján általánosságban elmondható, hogy a zaj negatív befolyásolja az önjavítást. A kutyaugatás azonban ebből a szempontból elkülönülni látszik a többi háttérzajtól. Egyedül ebben a csoportban fordult elő, hogy a korrekciós folyamatok jobban működtek a zaj hatására.

A szünetek elemzéséből megállapítható, a beszélők zajban kevesebb, de hosszabb szüneteket tartottak. A szünetek számának csökkenése arra utal, hogy az egyén egyrészt igyekszik minél előbb szabadulni a kellemetlen beszédhelyzetből, másrészt pedig még inkább törekszik arra, hogy közlése folyamatos legyen. A zaj zavaró hatása miatt azonban neheztette válik a gondolatok megfogalmazása, nyelvi formába öntése, a tervezés több szintjén is probléma adódhat, ami miatt megnő a szünetek hossza.

A zaj hatása nem csupán a megakadásjelenségek előfordulásában és az önmonitorozó mechanizmus megváltozott működésében érhető tetten; a beszéd akusztikai paraméterei is másként alakulnak a zajos környezetben. A Lombard-hatás, tehát a beszéd intenzitásának növekedése a háttérzaj hatására valamennyi csoport beszélőinél tapasztalható volt (3. ábra). A beszéd átlagos intenzitása a zene hatására növekedett a legkevésbé. A legnagyobb intenzitásnövekedést a fogfúrás esetében adatoltuk. Ezt követte a kutyaugatás és a légalapács csoportja.



3. ábra. A beszéd intenzitásviszonyai csendes (bal oldal) és zajos (jobb oldal) környezetben ugyanazon beszélőnél

Arra vonatkozóan, hogy a zaj hogyan hat a beszédfolyamatra, nem lehet általános és egyértelmű következtetéseket levonni, a Lombard-hatás azonban a felhasznált háttérzaj típusától függetlenül minden esetben érvényesül. A beszéd feldolgozásán túl zajos környezetben a beszéd tervezési folyamatai is nehezítetté válnak, amit a megakadásjelenségek nagyobb arányú előfordulása jelez, illetőleg az, hogy a beszédben többféle hibatípus jelenik meg. Az önmonitorozó mechanizmus az akusztikai visszacsatolás korlátozottsága miatt kevésbé hatékonyan működik; a beszélők hibáikat csak kisebb mértékben képesek korrigálni. A beszélők tehát zajos körülmények között jobban elfáradnak, kevésbé tudják hosszabb időn át az eredeti beszédszándékaiknak megfelelően létrehozni spontán közléseiket.

### Irodalom

- Balázs Boglárka – Gósy Mária 1988. Környezetünk hangjelenségeinek hatása a beszédre. *Fül-Orr-Gégegyógyászat* 34. 145–150.
- Barótfi István 2000. *Környezettechnika*. Mezőgazda Kiadó, Budapest, 135–154.
- Bartha Tibor 2004. Személyek elleni akusztikus fegyverek, mint nem halálos eszköz. *Hadtudomány* 2004/2. 98–109.
- Boersma, Paul – Weenink, David 2009. *Praat: Doing phonetics by computer*. [Computer program] (Version 5.0.20). <http://www.fon.hum.uva.nl/praat/>
- Braun, Angelika – Künzel, Hermann J. – Eysholdt, Ulrich 1992. *Einfluß von Alkohol auf Sprache und Stimme*. Kriminalistik Verlag, Heidelberg.
- Braun, Angelika – Künzel, Hermann J. 2003. The effect of alcohol on speech prosody. In: *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, 2645–2648.
- Brendt Mihály – Bibók Zsolt 2003. *Közlekedjünk környezetkímélően*. Környezetvédelmi és Vízügyi Minisztérium, Budapest, 38–41.

- Brendt Mihály – Muntág András 2007. *Stratégiai zajtérképezés - új eszköz a környezeti zaj elleni védelemben*. VII. Környezettudományi Konferencia. Győr.
- Gatschnegg, Wolfgang 2003. Lärm macht krank. *Verkehr und Umwelt*. Heft 2003/3–4. 38–40.
- Gósy Mária 2008. A zaj hatása a beszédre. *Beszédkutatás 2008*. 5–22.
- Gósy Mária – Bóna Judit 2011. Beszédfolyamatok monitorozása. *Magyar Nyelvőr* 135. 210–223.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukcóra. *Beszédkutatás 2007*. 108–121.
- Gyarmathy Dorottya 2008. Különböző zajok hatása a beszédprodukcóra. *Alkalmazott Nyelvtudomány VIII/1–2*. 135–147.
- Gyarmathy Dorottya 2010. A beszédellenőrzés működése alkoholos állapotban. In Bárdosi Vilmos (szerk.): *Világkép a nyelvben és a nyelvhasználatban*. Tinta Kiadó, Budapest, 125–135.
- Gyarmathy Dorottya 2012. Az agyi monitorozás módosulása zajhatásra. In Navracsics Judit – Szabó Dávid (szerk.): *Mentális folyamatok a nyelvi feldolgozásban*. Pszicholingvisztikai tanulmányok III. 170–181.
- Hativa, Nira 2000. *Teaching for effective learning in higher education*. Kluwer Academic Publishers, Dordrecht.
- Hoffmann Ildikó – Németh Dezső 2006. Nyelvi tünet, klinikai kórkép. *Nyelvtudomány XLIV*. 79–93.
- Hygge, Staffan 2003. Classroom experiments on the effects of different noise sources and sound levels on long-term recall and recognition in children. *Applied Cognitive Psychology* 17. 895–914.
- Kloepfer, Michael – Griefahn, Barbara – Kaniowski, Andrzej M. – Klepper, Gernot – Lingner, Stephan – Steinebach, Gerhard – Weyer, Heinrich W. – Wysk, Peter 2006. *Leben mit Lärm?* Springer, Berlin–Heidelberg.
- Lu, Youyi–Cooke, Martin P. 2008. Speech production modifications produced by competing talkers, babble and stationary noise. *Journal of the Acoustic Society of America* 124. 3261–3275.
- Niemann, Hildegard – Maschke, Christian – Hecht, Karl 2005. Lärmbedingte Belästigung und Erkrankungsrisiko – Ergebnisse des paneuropäischen Lares-Survey. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* 48. Springer Medizin Verlag, Berlin–Heidelberg. 315–328.
- Pap János 2002. *Hang – ember – hang*. Vincze Kiadó, Budapest.
- Pompetzki, Wulf 2000. Lärm und Gesundheit. *Jahresbericht 2000*. Landesumweltamt Nordrhein-Westfalen, Essen, 37–52.
- Póta Györgyné 2006. *Zajosak vagyunk*. Környezetvédelmi és Vízügyi Minisztérium, Budapest, 23–25.
- Szépe Judit 2004. Nyelvbötlések és parafáziák szótagszervezési sajátosságai. In Kukorelli Katalin (szerk.): *Alkalmazott Nyelvészeti Konferencia Füzetek VI. Dunaújvárosi Nemzetközi Alkalmazott Nyelvészet, Nyelvvizsgáztatási és Medicinális Lingvisztikai Konferencia*. Dunaújvárosi Főiskola, Dunaújváros, 193–203.
- Szépe Judit 2006. A közlésfolyamat fonológiai szerveződése: szerkezetisméltó műveletek afáziások beszédében. In Heltai Pál (szerk.): *Nyelvi modernizáció. Szaknyelv, fordítás, terminológia. XVI. Magyar Alkalmazott Nyelvészeti Kongresszus III*. MANYE – Szent István Egyetem Pécs–Gödöllő, 108–114.

Tarnóczy Tamás 1984. *Hangnyomás, hangosság, zajosság*. Akadémiai Kiadó, Budapest.

Tulipánt Gergely 2004. A zajhatások vizsgálata a közlekedés területén. *Közlekedéstudományi Szemle* 54/1. 22–27.

Vickers, Earl 2007. *Coping with speech noise in the modern workplace*. [http://chatterblocker.com/whitepapers/conversational\\_distraction.html](http://chatterblocker.com/whitepapers/conversational_distraction.html). (A letöltés ideje: 2011. 08. 12.)

### **Kérdések, feladatok**

1. Milyen hatást gyakorolhatnak az emberi füllel nem hallható hangok a szervezetünkre?
2. Milyen civilizációs megbetegedésekért tehető felelőssé a krónikus zajterhelés?
3. A kutatások tanúsága szerint milyen hatással van a zaj a beszédre?

## A hallássérülés hatása a beszédre

Az egészséges emberi fül 400 000 hang megkülönböztetésére képes. A környezet hangjait és az emberi beszédet akkor tudjuk tökéletesen feldolgozni, ha hallásunk megfelelően működik. A legtöbb ember számára a jó hallás fontossága csak akkor tudatosodik, ha hallóképessége már lényegesen megromlott. Milyen tünetek esetén kell arra gondolni, hogy a hallás nem tökéletes? Mi okozhatja a halláscsökkenést és milyen típusú zavarokat különböztethetünk meg? A hallásvesztés negatív hatással van a beszéd feldolgozására és produkciójára is. Hogyan segíthetnek az orvosok, a szurdopedagógusok (nagyothallókkal foglalkozó gyógypedagógus) és a különböző beszédfejlesztő szoftverek abban, hogy a nagyothallók kommunikációja minél jobb legyen?

### 1. A hallás és zavarai

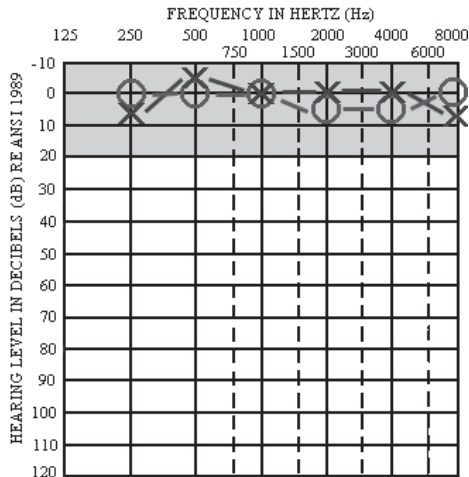
A hallás folyamata során a hallószerv a környezet hangingereit érzékeli, felelősíti és továbbítja az agy hallókérgébe, ahol a végső hangfeldolgozás megtörténik. A hang a levegőben rezgésként terjed. A rezgés elérve a fület, a hallójáraton át a dobhártyán, majd a hallócsontokon keresztül a belfülbe jut. Itt alakul át a mechanikus rezgés elektromos impulzusokká (a Corti-szerv szőrsejtjei által), amelyeket a hallóideg a kérgi hallóközpontba vezet (Fent 2007). A felszálló hallópálya mellett az agykéregből, illetőleg bizonyos agytörzsi magvakból a belfül szőrsejtjei felé a leszálló hallópálya is szállít ingerületet; és ez visszahat a fül működésére.

A hallás változik az életkorral. A hallás már magzati korban elkezd működni. A magzat az anyaméhben töltött utolsó három hónapban már meghallja a különféle „zajokat”, beszédet, zenét; az erős hangingerre reagál. Születéskor az újszülött hallása majdnem olyan teljes, mint egy felnőtté; a hallásküszöbe viszont alacsonyabb, ezért megriad a váratlan és ismeretlen hangoktól. Kamaszkorban hallunk a legjobban, időskorra pedig szinte kivétel nélkül romlik a hallásteljesítmény. A 65 év feletti lakosság egyharmada szenved különböző mértékű halláscsökkenésben. Az időskori hallásvesztés elsősorban a magas frekvenciákat érinti, ezért van az, hogy az idősek nehezen hallják meg például a telefoncsörgést vagy a csengő hangját. Az időskori hallásromlás lassú és fokozatos, de nehezebb a beszélgetésekben való részvételt, különösen háttérzajban.

Hallásprobléma nemcsak időskorban jelentkezhet, hanem gyermekeknél is.

Hallássérülés szempontjából fokozottan veszélyeztetettek azok a gyerekek, akiknél az anya terhessége alatt rubeola, toxoplasmosis, cukorbetegség vagy alkoholizmus állt fenn; de a traumás szülés, a fertőzés, éretlenség is rizikótényezőknek számítanak. Az esetleges rizikótényezők mellett a szülők figyelmét egyéb más tünetek is felhívhatják arra, hogy a gyermeknek hallásproblémája van. Csecsemőkorban a gögicsélés késése, elszegényesedése, hiánya, vagy a kialakulást követő abbamaradása lehetnek az első jelei a hallászavarnak – a gyermek nem hallja saját hangját, az akusztikai visszacsatolás hiánya nem ösztönzi a hangadásra. Ha a gyermek nem fordul a hangok irányába, később pedig nem szereti hallgatni az éneket, mondókákat, meséket; szintén felmerül a hallás problémája. Későbbi tünetek: gyakori visszakérdezés, figyelemzavar, tanulási nehézségek, pontatlan artikuláció. Hallászavar gyanúja esetén fül-orr-gégészeti, illetve audiológiai hallásvizsgálat javasolt.

Az audiológiai hallásvizsgálat során különböző frekvenciájú szinuszhangokat (egyszerű rezgés, sípolásra emlékeztető hang) juttatnak fülhallgatón keresztül felváltva a két fülbe. A vizsgált személynek jeleznie kell, amikor leghalkabban érzékeli a hangot. Így határozható meg a hallásküszöb, ami jellemzi a gyermek vagy a felnőtt hallásállapotát. Azt, hogy az adott magasságú hangot a személy milyen hangerősségen képes meghallani, azaz a különböző frekvenciákon mekkora a hallásküszöbe, a küszöbgörbe mutatja meg. Ép hallás esetén a küszöbgörbe 250 és 8000 Hz közötti tartományban 0 és 10 dB között húzódik (1. ábra).



1. ábra. Ép hallás küszöbgörbéje audiogramon (Forrás: Hain 2015)

(A függőleges vonalak a különböző hangmagasságot, a vízszintesek a hangerősséget jelzik. A vizsgáló az audiogramon jelöli, hogy a vizsgált személy milyen hangerősségen hallotta meg az adott frekvenciájú szinuszhangot. A pontokat összekötve látható a hallási küszöbgörbe.)

A hallássérülésnek különböző fokozatai vannak. A 25 és 40 dB közötti küszöbgörbe enyhe halláscsökkenést jelez (ez azt jelenti, hogy az adott személy nem képes a 25 dB-nél halkabb hangok érzékelésére). Ha a küszöbgörbe 40 és 60 dB között húzódik, akkor közepes hallásvesztésről van szó. A 60 és 90 dB közötti küszöbgörbe súlyos hallásvesztést jelent. Siketség akkor áll fenn, ha a jobban halló fülön mérve a küszöbgörbe 91 dB felett húzódik (azaz az adott személy nem érzékeli a 91 dB-nél halkabb hangokat; Csányi 2005).

A halláscsökkenésnek különféle típusai vannak. A típusok szerint létezik vezetékes halláscsökkenés, ennek oka a hallójárat, a közép- és belfül területén jelentkezik (például fülzsírdugó, zajártalom). A másik fő típus, az idegi eredetű károsodás oka lehet például jóindulatú daganat vagy a központi idegrendszer károsodásai (Kiefer 1998). A harmadik a kevert típusú nagyothallás, amikor vezetékes és idegi típusú halláskárosodás együttesen jelentkezik. A halláscsökkenések további osztályozásának szempontja a fennállási időtartam: egy részük átmeneti, más részük maradandó. Gyermekkorban gyakoribb az átmeneti; rendszerint a megnagyobbodott orrgaratmandula az oka, a fülkürt nem működik tökéletesen, a középfül nem szellőzik megfelelően. Ez az állapot akár közép-súlyos halláscsökkenést is eredményezhet, amely fájdalom hiányában sokáig rejtve is maradhat, mivel kevés a feltűnő tünet (nyitott száj, esetleg a gyakori visszakérdezés). A gyermekkori átmeneti halláscsökkenés másik gyakori oka a krónikus gennyes középfülgyulladás, amelynek fő tünete a fülfolyás. Fülzsírdugó vagy hallójáratgyulladás szintén okozhat átmeneti vezetékes halláscsökkenést (Pytel 1998). A maradandó halláskárosodások nagy része idegi eredetű vagy kevert típusú.

A halláskárosodás a kezdeti ideje alapján lehet veleszületett vagy szerzett zavar. Megkülönböztetjük továbbá a prelingvális vagy posztingvális hallássérülést. Az első azt jelenti, hogy a beteg egy nyelv elsajátítása előtt vált nagyothallóvá vagy siketté, a másik esetben a hallásvesztés a halló társadalomba való szocializálódás után, egy nyelv elsajátításának egyes szakaszait követően, vagy a kialakult anyanyelv meglétét követően történt (Csányi 1990).

A WHO 2002-es felmérése szerint a világon mintegy 250 millió hallássérült (nagyothalló és siket) van, ez a népesség kb. 4%-a. A magyar lakosság mintegy 10%-a él valamilyen mértékű halláskárosodással; ebbe beletartoznak az időskor miatti halláscsökkenést mutató emberek is. 3–4%-uk súlyos fokban nagyothalló, illetve siket (mintegy 60 ezer fő), ők alkotják az ország harmadik legnagyobb nyelvi kisebbségét (Bartha–Hattyár 2002).

## 2. A hallókészülék és a cochleáris implantáció

A részleges halláscsökkenésben szenvedők számára a hallókészülék segítséget nyújthat abban, hogy beszédfeldolgozásuk és ezáltal kommunikációjuk javuljon.

A hallókészülék a beérkező hangrezgéseket felerősíti; mikrofont, erősítőt, hangszórót és elemeket tartalmaz. Képes lehet bizonyos mértékig szűrni a háttérzajokat, zajos környezetben csökkenteni az erősítést, illetve hangosítani, ha szükséges. A hallókészüléket lehet az egyik vagy mindkét fülben viselni. Létezik hallójárat, illetve fülkagylóra illeszthető hallókészülék. A hallójárat készülékek viselése kényelmesebb, de ezek csak enyhe halláscsökkenés esetén használhatóak, mert nagy erősítésre nem képesek (2. ábra). A fül mögötti készülékek előnyei, hogy nem igényelnek annyi karbantartást, cseréjük és kezelésük egyszerű, erősítésük nagyobb jobb hangminőséget biztosítanak (3. ábra).



2. ábra. Hallójárat hallókészülék (Forrás: <http://test14.jvic.hu/hallokeszulekek.html>)



3. ábra. Fülkagylóra illeszthető hallókészülék (Forrás: <http://sejk.oep.hu/?s=segedeszkoz&ttt=310197789&termek=siemens-intuis-sp-dir>)

A digitális hallókészülékek kifinomultabbak, mint az analógok, de nagyon drágák lehetnek, hiszen beültetett számítógépes technológiájuknak köszönhetően, erősítésük a halláscsökkenés profiljának megfelelően egyénileg változtatható.

tó. Léteznek már szemüvegszárba épített készülékek is; illetve olyanok, amelyeket sebészileg ültetnek a fül mögötti csontba.

A cochleáris implantáció egy olyan eljárás, amely a súlyosan nagyothallók és siketek hallásjavítását szolgálja – amennyiben a hallóidegek épen működnek és a gyermek vagy felnőtt az előzetes felmérések alapján alkalmas az implantátum viselésére. Az implantátum vevőtekercsét a fül mögött, a koponyacsontra vájt mélyedésben helyezik el, míg az elektródát a középfülön keresztül a csigába vezetik. Az elektródák különböző hosszúságúak: a rövidek a csiga elején végződnek és a magas frekvenciát érzékelő idegeket ingerlik, míg a hosszabbak a csiga belsejéig mennek és az alacsony frekvenciákért (mélyebb hangokért) felelősek. Az elektróda által keltett elektromos impulzusokat az idegsejtek érzékelik, és az agyba továbbítják, hallásérzetet keltve a kérgi központban (4. ábra).



4. ábra. A cochleáris implantáció

(Forrás: [https://hu.wikipedia.org/wiki/Cochle%C3%A1ris\\_implant%C3%A1tum](https://hu.wikipedia.org/wiki/Cochle%C3%A1ris_implant%C3%A1tum))

A cochleáris implantációval elérhető hallás optimális esetben elérheti a 25–35 dB-es szintet, amely már elegendő lehet a beszéd feldolgozásához. A javulás azonban egyénenként nagyon különböző lehet; és az implantátum erősen torzít a beérkező hangon. A korábban ép halló cochleáris implantáltak *robotszerű, csipogó, fém* szavakkal jellemzik a CI hangvilágát. Korán implantált gyerekeknél ennek ellenére általában az ép hallókra jellemző beszédképesség alakul ki. A rehabilitáció során szoktatják hozzá a beteget a CI-vel való halláshoz. A páciens ugyanis a beszédprocesszor bekapcsolásakor egyből hall, de ezek a hangimpulzusok nagyon távol állnak a normális hallástól, mivel a CI nem képes olyan hangérzetet produkálni, mint a szőrsejtek. A hangingerek idegi feldolgozása azonban olyan rugalmas, hogy a páciens – megfelelő tréning mellett – hamarosan hozzászokik a CI-hez.

### 3. A hallássérülés hatása a beszéd feldolgozására

Az ép hallás alapvető feltétele, hogy a hallgató képes legyen az elhangzott közlések feldolgozására, azonban a jó hallás önmagában még nem biztosítja a sikeres beszédfeldolgozást. A kutatások fő kérdése, hogy a hallgató hogyan képes az akusztikai jelből a nyelvi tartalmak kinyerésére és feldolgozására, például a beszédhangok és a szótagok felismerésére, a szavak és hosszabb közlések azonosítására. Az elhangzó beszéd feldolgozása csecsemőkortól indul fejlődésnek. Az anyanyelv-elsajátítás során fejlődik a percepciós bázis, amely egy adott nyelvre jellemző beszédfeldolgozási működéssorozat. Ez teszi lehetővé, hogy a gyermek képessé váljon az elhangzottak azonosítására.

A beszédfeldolgozási folyamatban a hallásra épül a beszédészlelés, amely a beszédhangok, hangkapcsolatok, hangsorok azonosításáért felelős. Az észlelésre épülő beszédmegértés a szavak, mondatok, szövegek szemantikai és grammatika szerkezetének azonosítását, a tartalmak feldolgozását jelenti. A beszédfeldolgozás legmagasabb szintje az értelmezés/asszociációk szintje, ahol az elhangzottakat az asszociációs kapcsolatok révén a gyermek beépíti a már tárolt ismereteibe (Gósy 2005).

A beszédészlelés három nagy folyamatra osztható: akusztikai, fonetikai és fonológiai észlelés. A feldolgozás első lépcsőjeként a beérkező akusztikai jel elsődleges elemzése az idő, a frekvencia és az intenzitás paraméterek mentén történik, például: milyen jellegű volt a hang (zene/beszéd/ének), gyors volt vagy lassú, halk vagy hangos. Az eredmények egy tába kerülnek, az akusztikai elemzés ebből indul ki: a hang tartalmazott-e zöngét, milyen volt a frekvencia- és intenzitás szerkezet. Az akusztikai elemzés eredményére épül a fonetikai elemzés, amely során a beszédhangok azonosítása történik. A fonológiai szinten soroljuk a beszédhangokat a megfelelő fonémaosztályba. Például az *jégpálya* szóban elhangzó [k] valójában a /g/ fonéma realizációja. Az egymásra épülő akusztikai, fonetikai és fonológiai észlelést számos más működés egészíti ki (ritmusészlelés, a beszédhangok megkülönböztetése stb.).

A hallássérülés természetesen az egész beszédpercepciós folyamatra negatív hatással van. Az akusztikai és fonológiai észlelés működése súlyosan nehezített a nagyothalló gyermekek esetében. Egy kutatás szerint a szövegértésük is jelentősen elmarad ép halló kortársaikétól. A fonetikai észlelés és a mondatértés ugyanakkor megközelíti az életkorban elvárható értéket (Bombolya 2007).

### 4. A hallássérülés hatása a beszédprodukción

A halláscsökkenés nemcsak a beszédfeldolgozás folyamataiban okoz nehézségeket, hanem befolyásolja a beszéd kivitelezését is. A hallássérüléssel született

csecsemő egyrészt nem képes megfelelően feldolgozni környezete hangjait. Az akusztikai visszacsatolás korlátozottsága, illetve hiánya másrészt azt eredményezi, hogy nem indul el időben a beszédprodukció, mert a csecsemő nem hallja saját hangját sem, így nem motivált a hangadásra (a gögicselés hiánya erre utalhat). Ha tehát a gyermek nem jól vagy egyáltalán nem hallja környezetét, késik a hangadás megindulása, emiatt maga az egész beszédfejlődési folyamat nem lesz tipikus (Oller–Eilers 1988; Koopmans-van Beinum et al. 2001).

A hallássérülés hatással van a teljes beszédprodukcióra, így érinti a beszédhangok képzését is. A nagyothallók artikulációjára a beszédhangok hibás ejtése jellemző. Az is gyakran előfordul, hogy a nagyothallók beszédéből egyes beszédhangok/beszédhang-csoportok teljesen hiányoznak, ezeket a beszélők más hangokkal helyettesítik. A sorozatos mássalhangzó-tévesztéseket otogén diszliániának nevezik (Kiefer 1998). Egy kutatásban (Beke 2010) azt vizsgálták, hogy középsúlyosan és súlyosan nagyothalló tanulók magánhangzóinak időtartama (spontán beszédben) hogyan alakul az ép halló kortársaikéhoz viszonyítva. Csúppán néhány magánhangzó esetében (*a*, *o*, *ö*, *ü*) volt nagy különbség a magánhangzók időtartamában a nagyothalló és az ép halló iskolások között (a nagyothallók ezeket hosszabban ejtették). A vártnál kisebb eltérés oka az lehet, hogy a nagyothalló gyermekek számára készült beszédterápia nagy hangsúlyt fektet a beszédhangok megfelelő képzésére a beszéd érthetőségének érdekében. További magyarázat lehet az olvasástanulás során a rövid és hosszú hangok megkülönböztetése, amit az íráskép is megerősít. Siket beszélők szövegfelolvasásai ugyanazt az eredményt mutatták: beszédükben a magánhangzók megvalósulása óriási változatosságot mutat, időtartamuk hosszabb, artikulációjuk pontatlanabb, mint az ép hallású beszélőnél (Auszmann 2013a).

Siket beszélők felolvasásait vizsgálva a kutatási eredmények azt mutatták, hogy a siketek zöngétlen felpattanó zárhangjait hosszabb zárszakasz jellemzi, mint az ép beszélőkét. Ez a különbség egyrészt a lassabb artikulációs és beszédtempóból adódik; de visszavezethető a kevesebb beszédgyakorlatra is, aminek következménye a hibás légzéstechnika és a renyhe artikuláció (Auszmann 2014).

A beszéd prozódiai szerkezete is jellegzetesen eltér az ép beszédre jellemzőtől a nagyothallók esetében. Csecsemőkorban még nem igazoltak nagy eltérést az alaphangmagasságban ( $f_0$ ) ép halló és nagyothalló gyermekek között, de az életkor előrehaladtával nő a két csoport közötti különbség (Iyer–Oller 2008). A beszéddallam megvalósításában óvodás- és iskoláskorban már jellegzetes eltérés adatható a gyermekek között a hallássérülés függvényében (Csányi 1990; Gósy 2008). Kisiskolás korban a nagyothalló fiúk átlagos alaphangmagassága magasabb volt egy kutatás szerint, mint az ép halló kortársaiké; a lányok esetében nem volt különbség. Az alaphangmagasság értékeinek változatossága nagyobb a nagyothalló gyermekek beszédében, mert ők az akusztikus visszacsatolás hiányában, illetve a monotonitás elkerülése miatt sokszor nagyobb és gyors-

sabb ingadozással valószínűsítik meg a beszéddallamot. Az alaphangmagasságot befolyásolja a halláscsökkenés mértéke is. A súlyosan nagyothalló gyermekek átlagos alaphang-értéke alacsonyabb, mint a középsúlyosan nagyothalló társaiké. A különbség azzal magyarázható, hogy a súlyosan nagyothalló gyermekek kevésbé tudják elsajátítani a prozódiai mintázatokat (Beke–Horváth 2011).

A halláscsökkenés a beszédszünetekre és az időzítési viszonyokra is hatással van. A nagyothalló beszélők nehezebben tudják beosztani a beszédhez szükséges levegőmennyiséget, aránytalan a levegő felhasználása (bizonyos szótagok meghangosításához túl sokat használnak, ezért a levegőjük hamar elfogy, a szótagok nem tudnak közlésekké összeállni). Nagyothalló iskolások spontán beszédében a szünetek aránya 38%, ép halló kortársaiknál 20% (Beke–Horváth 2011). Siket felnőttek szövegfelolvasásában 15–27% volt a szünetek aránya (Auszmann 2013b), mondatisméltéleikben pedig 40% (Nickerson et al. 1974). A nagyothalló és a siket beszélők nemcsak gyakrabban tartanak néma szünetet, mint az ép hallók, ezek időtartama is jóval hosszabb (Beke–Horváth 2011; Auszmann 2013b). Ugyanakkor a nagyothallóknál rövidebbek a két szünet közötti közlések időtartamai is. Mindezek miatt beszédük kevésbé tűnik folyamatosnak a hallgatók számára.

A nagyothalló és a siket beszélők beszédtempója és artikulációs tempója is lassabb, mint ép halló kortársaiké, 3–4 hanggal ejtenek kevesebbet náluk másodpercenként (Nickerson et al. 1974; Auszmann 2013b). A hangszínezetet tekintve a siketek beszéde erősen nazalizált (Stevens et al. 1976).

A hosszú időn keresztül, illetve tartósan fennálló beszédzavar nyelvi zavart és általános kommunikációs nehézségeket eredményezhet. A nagyothalló gyermekek közlései döntően tanult sémákra épülnek; szórendjük általában megfelelő, de a grammatikai szerkezetek sokszor nem felelnek meg a szándékolt gondolatnak és a szabályoknak (Horváth 2010). Előfordul a határozott és az általános ragozás keveredése (*szeretem beszélgetni*), a ragok és a vonzatok tévesztése (*utána pedig gyaloggal*), a szám és a személy téves egyeztetése is (*van neki gyerekiük*). Jellemző a szóvégi toldalékmorféma teljes elhagyása is, például: *én elmegyek anya* (anyával) *erdőbe*; *aztán jöttünk vissza Görögország* (Görögországból). A nagyothalló gyermekek közlései döntően tartalmas szavakból állnak, a funkciószavak használata korlátozott. A társalgási interakcióban való kisebb jártasság miatt a szöveg kohézióját nehezen teremtik meg; gondolataik kifejezése sokszor szervezetlen (Wood–Webster 1994).

Egy magyar kutatás (Bombolya 2007) szerint a vizsgált középsúlyosan és súlyosan nagyothalló 6–7 éves gyermekek többségének szókinccse jelentősen elmarad ép halló, iskolába lépés előtt álló kortársaikétól, akikkel egy osztályban kezdik meg általános iskolai tanulmányaikat. Az elmaradás a nagyothalló gyermekek több mint felénél 2–3 évre tehető. A szűk szókinccsel rendelkező gyermekek beszédét továbbá több hangra kiterjedően hibás artikuláció, szótöredékek hasz-

nálata, egyszerű mondatalkotási szint (szómondatok, töredékes szómondatok, sőt rövid, agrammatikus mondatok) jellemzi.

A mindennapi kommunikációban való részvételt természetesen befolyásolja a halláscsökkenés mértéke. Az enyhe fokú hallássérülés még lehetővé teszi a beteg számára a társalgó beszéd megértését, de a feldolgozás már támaszkodik a vizuális információra, a szájról olvasásra. A szókincs beszűkülhet, és felléphetnek beszédhibák is. Közepes fokú hallássérülés esetén a beteg számára már csak a hangos társalgás érhető, még fontosabb a szájról olvasás szerepe, ezért a csoportos beszélgetés erősen nehezített. Súlyos hallássérülés esetén a környezeti zajok egy részének felismerése általában még lehetséges, de a mindennapi kommunikációban való részvétel erősen korlátozott, a nyelv használata torzulhat (Csányi 2005). A nagyothallók és siketek a verbális kommunikáció mellett a jelnyelvet és az ujjábécét használ(hat)ják.

## 5. A jelnyelv és a daktil (ujjábécé)

A jelnyelv a vizuális kódrendszerek egyik legfejlettebb formája, amelyet főként, de nem kizárólag a siketek használnak. A jelnyelvi jelek önálló jelentéssel nem bíró, de jelentésmegkülönböztető szerepű egységekből, ún. kerémákból épülnek fel (vö. Bartha et al. 2006). A 2009. évi CXXV. törvény a magyar jelnyelvről és a magyar jelnyelv használatáról a magyar jelnyelvet a hivatalos nyelv szintjére emelte Magyarországon. A siketek nem ugyanazt a jelnyelvet használják minden országban, de vannak hasonló jelek az egyes jelnyelvekben. A jelnyelvnek is van egy sajátos hangtana, alaktana és mondattana. A magyar jelnyelvre például az alany – többi mondatrész – állítmány szórend jellemző, a fókusz gyakran a mondat végén szerepel. A jelnyelv ugyanolyan tempójú kommunikációt tesz lehetővé, mint a beszéd; de törekedni kell a jel pontos kivitelezésére; fontos a mimika és a testtartás. A jelelők sokszor néma artikulációval kísérik a jelelést a jobb megértés érdekében.

A daktil (más néven ujjábécé) szintén egy kommunikációs eszköz. A jelrendszerben egy kézzel mutatott jel egy hangnak felel meg. Általában akkor használják, ha a kifejezendő fogalomnak nincs szimbolikus jele. Előnye, hogy teljesen pontos, mindent ki lehet vele fejezni, hátránya viszont, hogy a jelkészlet nyelvenként egymástól eltérő. Gyakorlattal igen gyorsan lehet vele jelezni, de ez még mindig lassabb, mint a jelnyelvi gesztusjelek használata. A daktilban ugyanis egy jel egy fonémának felel meg – a jelnyelv egyben gesztusnyelv is, tehát egyetlen mozdulat jelenthet egy mondatot is (*becsukom az ajtót, fenyeget*). A daktil tehát főleg akkor praktikus, ha idegen neveket, külföldi személyneveket, szakszavakat betűnként kell eljelezni.

## 6. A siketek és nagyothallók oktatása

Magyarországon a siketek és nagyothallók tanítása nagy múltra tekint vissza. Cházár András tanulmányai alapján a bécsi mintát követve 1802-ben Vácon megnyitották az első tanintézetet a siketek számára. A bécsi módszer lényege az volt, hogy a siketeket jelelni és beszélni egyaránt tanították. Később bevezették a tisztán orális módszert, tiltották a jelnyelv használatát. Napjainkban is az orális módszert preferálják a siketek oktatása során. A cél a hangzó nyelv hangzó és írott formájának elsajátíttatása; a jelnyelv csak kiegészítő eszköz az oktatásban – ez nehézséget jelenthet azok számára, akik a jelnyelvet sajátították el anyanyelv-ként (Hattyár 2008).

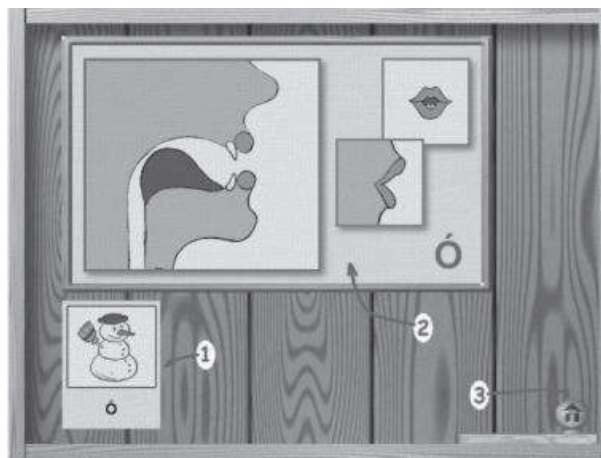
A korai hallásgondozás, az esetlegesen szükséges hallókészülékekkel való ellátás sokat javíthat a beszédfeldolgozás, ezáltal a beszédprodukciónak minőségén is. A nagyothalló (de nem siket) gyermekek egyre nagyobb része (részben az egyre szélesebb körű ellátásnak köszönhetően) integráltan folytatja tanulmányait. Az integráció nyelvi fejlődésre ható előnye egyrészt a környezet pozitív beszédmin-tája, másrészt az, hogy a nagyothalló gyermek társaival és a pedagógusokkal csak a hangzó nyelv használatán keresztül tud kommunikálni, ez ösztönzi arra, hogy gondolatait nyelvi formába öntve fogalmazza meg.

Az integrált oktatás során mindezek miatt fokozott figyelmet kap nemcsak az, hogy a nagyothalló gyermek hogyan képes feldolgozni az elhangzottakat, hanem az is, hogy miként tudja magát hangzó nyelvi eszközökkel kifejezni. A nagyot-halló gyermekek beszédprodukciónak vizsgálata tehát egyre fontosabb nem-csak a beszédtudomány, a szurdopedagógia, hanem a gyakorlati pedagógia szá-mára is.

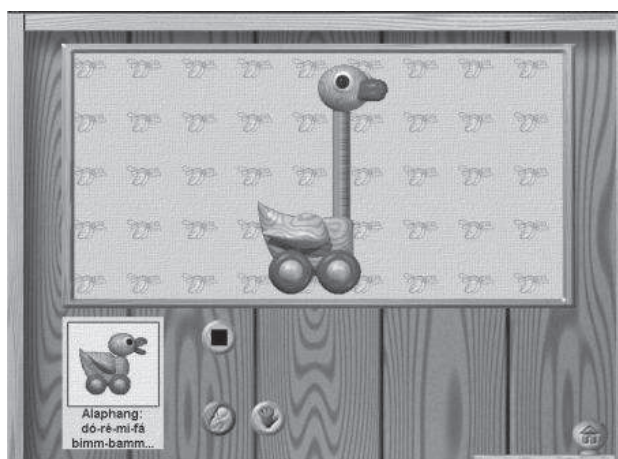
### 6.1. Szoftverek

A hagyományos oktatás mellett a nagyothallók beszédfejlesztését szoftverek is segítik.

A *Varázsdoboz* szoftver a Budapesti Műszaki és Gazdaságtudományi Egyete-men készült (<http://alpha.tmit.bme.hu/speech/sulinet.php>). A program segítséget nyújt ép halló beszédhibás, valamint nagyothalló gyermekek és felnőttek helyes beszédképzésének kialakításában. A beszédhangok helyes képzésének gyakorol-tatása mellett lehetőség van megfelelő hangerő, hangmagasság, ritmus, beszéd-dallam és hangszín kialakítására (5. és 6. ábra). A rendszer visszajelzést is végez: a beszédképzéskor keletkező hang jellemző paramétereit a számítógép képer-nyőjén megjeleníti (hangkép). Ha a hangképzés hibás, akkor a hangkép erősen eltér a helyes ejtést jellemző hangképtől. A képi megjelenítéssel egyidőben a korrekt módon ejtett hangminta: szótag, szó vagy mondat hangszórón keresztül hallható.



5. ábra. A hangképző szervek helyzetének ábrázolása az egyes beszédhangok helyes ejtése esetén (Forrás: <http://alpha.tmit.bme.hu/speech/sulinet.php>)



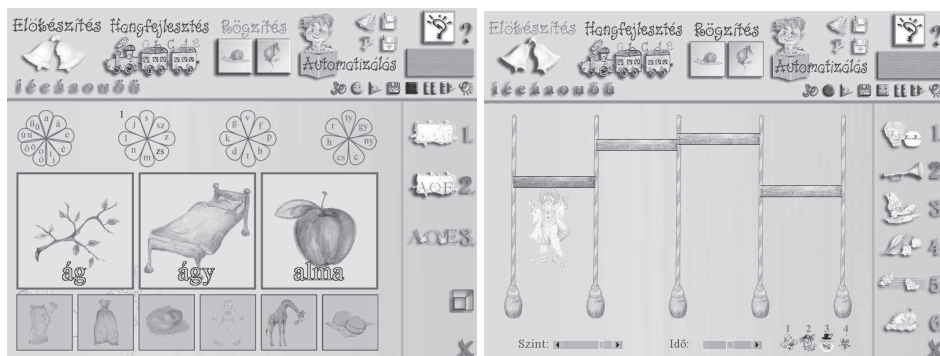
6. ábra. Az alfahangmagasság gyakoroltatása a programban (A kacska nyaka megnyúlik, ha a hangmagasságot növeljük.) (Forrás: <http://alpha.tmit.bme.hu/speech/sulinet.php>)

A programot több iskolában használják sikerrel évek óta. A hatásvizsgálat során kiderült, hogy ép halló beszédhibás és nagyothalló gyermekek fél éves oktatása után 20%-kal érthetőbb beszédet produkáltak azok a gyermekek, akiknek oktatásakor a *Varázsdobozt* mint oktatási segédeszközt rendszeresen használták.

Az Oktatási Minisztérium támogatásával készült a *Beszédmester* szoftver (<http://www.inf.u-szeged.hu/projectdirs/beszedmester/>), amely beszédjavítás-terápiára és olvasásfejlesztésre is alkalmas (Kocsor et al. 2006). A programot a Szegedi Tudományegyetem munkatársai a kaposvári Siketek Iskolájának

szakembereivel együttműködve hozták létre. A hallássérültek beszédképzésének terápiája hagyományosan óriási türelmet, hosszú időt és a szurdopedagógus állandó jelenlétét igényli. A helyes hangképzés és a pontos artikuláció elsajátításához sok ismétlésre és a pedagógus folyamatos korrekciójára van szükség. Ezt az automatizálási folyamatot egyszerűsítheti és felgyorsíthatja a *Beszédmester*; a gyakorlás ugyanis részben önállóan is végezhető. Lehetőség van a hallott beszéd rögzítésére is; gépi beszédfelismerésen alapulva a szoftver kiértékeli a gyermek által megvalósított hang minőségét. A megfelelő használathoz pontos metodikai és didaktikai leírások olvashatóak. A programban található feladatok sorrendje tetszőlegesen választható. A *Beszédjavítás-terápia* részben a zöngéadás, a hang-erő, a ritmus, a hangutánzás, a hangfejlesztés gyakorlására találhatók játékos feladatok (7. ábra).

A *Beszédmester* eredményességét felmérve az eredmények azt mutatták, hogy az öt hónapos terápia alatt látványos fejlődés következett be a magánhangzók helyes ejtésében különböző életkorú siket gyermekeknél. A gyermekek nagyon szívesen gyakorolták a játékos feladatokat, örültek a monitoron megjelenő pozitív visszajelzésnek. A nagyothalló és cochleaimplantált gyerekek esetében is óriási előny, hogy a számítógép által automatikusan elvégzett visszajelzés eredményeként saját kiejtésüket tudják javítani.



7. ábra. A Beszédmester gyakorlófelületei

## Irodalom

- Auszmann Anita 2013a. Magánhangzó-minőségek siketek szövegfelolvasásában. *Alkalmazott Nyelvtudomány* XIII/1–2. 23–39.
- Auszmann Anita 2013b. Siketek szövegfelolvasásának szupraszegmentális jellegzetességei. In Váradai Tamás (szerk.): *Alknyelvdok7: doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA Nyelvtudományi Intézet, Budapest, 17–25.
- Auszmann Anita 2014. Zöngétlen explozívák időszerkezete siket beszélők szövegfelolvasásá-

- ban. In Váradí Tamás (szerk.): *Alknyelvdok8: Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA Nyelvtudományi Intézet, Budapest, 1–12.
- Bartha Csilla – Hattyár Helga 2002. Szegregáció, diszkrimináció vagy társadalmi integráció? – A magyarországi siketek nyelvi jogai. In Kontra Miklós – Hattyár Helga (szerk.): *Magyarok és nyelvtörvények*. Teleki László Alapítvány, Budapest, 73–123.
- Bartha Csilla – Hattyár Helga – Szabó Mária Helga 2006. A magyarországi siketek közössége és a magyarországi jelnyelv. In Kiefer Ferenc (főszerk.): *A magyar nyelv*. Akadémiai Kiadó, Budapest, 852–906.
- Beke András 2010. Magánhangzó-minőségek nagyothalló gyerekek spontán beszédében. *Beszédkutatás 2010*. 223–243.
- Beke András – Horváth Viktória 2011. A spontán beszéd fonetikai mintázata hallássérült gyermekeknél. In Navracsics Judit – Lengyel Zsolt (szerk.): *Lexikai folyamatok egy- és kétnyelvű közegben. Pszicholingvisztikai tanulmányok II*. Tinta Könyvkiadó, Budapest, 27–35.
- Bombolya Mónika 2007. Hallássérült gyermekek beszédfeldolgozási folyamatai. In Gósy Mária (szerk.): *Beszédeszélelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 72–83.
- Csányi Yvonne 1990. *Hallás-beszéd nevelés*. Tankönyvkiadó, Budapest.
- Csányi Yvonne 2005. A hallássérült gyermek. A hallássérült gyermek az óvodában. In Csendes Krisztina (szerk.): *Kisgyermek, nagy problémák*. RAABE Tanácsadó és Kiadó Kft., Budapest, 1–16.
- Fent Zoltán 2007. A hallószerv, a hallás folyamata, zavarok. In Gósy Mária (szerk.): *Beszédeszélelési és beszédmegértési zavarok az anyanyelv-elsajátításban*. Nikol Kkt., Budapest, 44–57.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2008. Nagyothallás, beszédfejlődés és nyelvhasználat. *Fejlesztő Pedagógia 2008/1*. 7–11.
- Hain, Timothy C. 2015. Audiometry. <http://www.dizziness-and-balance.com/testing/hearing/audiogram.html>
- Hattyár Helga 2008. *A magyarországi siketek nyelvajátításának és nyelvhasználatának szociolingvisztikai vizsgálata*. Doktori disszertáció. ELTE BTK, Budapest. [http://doktori.btk.elte.hu/lingv/hattyar/diss\\_nem.pdf](http://doktori.btk.elte.hu/lingv/hattyar/diss_nem.pdf)
- Iyer, Suneeti Nathatni – Oller, Kimbrough D. 2008. Fundamental frequency development in typically developing infants and infants with severe-to-profound hearing loss. *Clinical Linguistics & Phonetics 22/12*. 917–936.
- Kiefer Gábor 1998. Cochleáris implantáció utáni beszédrehabilitációs eredmények. *Beszédkutatás '98*. 204–213.
- Kocsor András – Bácsi János – Mihalovics Jenő 2006. *Beszédmester: számítógépes olvasás-fejlesztés és beszédjavítás-terápia*. <http://epa.oszk.hu/00000/00035/00101/2006-03-mu-Tobbek-Beszedmester.html>
- Koopmans-van Beinum, Florian J. – Clement, C. J. – Dikkenberg-Pot, Ineke 2001. Babbling and the lack of auditory speech perception: A matter of coordination? *Developmental Science 4/1*. 61–70.
- Nickerson, Raymond S. – Stevens, K. – Boothroyd, A. – Rollins, A. 1974. *Some observations on timing in the speech of deaf and hearing speakers*. Bolt, Beranek and Newman Inc., Cambridge.

- Oller, Kimbrough D. – Eilers, Rebecca E. 1988. The role of audition in infant babbling. *Child Development* 59. 441–449.
- Pytel József 1998. A hallás és a beszéd fejlődése. A hallászavarok felismerése és kezelése. In Aszmann Anna (szerk.): *Iskola-egészségügy*. Anonymus, Budapest, 82–88.
- Stevens, K. – Nickerson, R. – Boothroyd, A. – Rollins, A. 1976. Assessment of nasalization in the speech of deaf children. *Journal of Speech and Hearing Research* 19. 393–416.

### **Kérdések, feladatok**

1. Mikor beszélünk nagyothallásról, és milyen fokozatai vannak?
2. Milyen hatással van a halláscsökkenés a beszédprodukcióra?
3. Milyen műszeres eljárások léteznek a halláscsökkenés javítására?

KOHÁRI ANNA

## A beszéd ritmusa

A beszédritmus igen összetett jelenség, amelynek pontos fogalmi meghatározásáról máig nincs egyetértés a szakirodalomban. A különböző elméletekben azonban található egy közös feltételezés, amely szerint a beszédritmust a beszéd bizonyos elemeinek, jelenségeinek szabályos időközönként történő, **rendszerű ismétlődése** hozza létre (Gósy 2004; Fox 2000). A következőkben azt tekintjük át, hogyan próbálták megragadni, megközelíteni a beszédritmus jelenségét, egyáltalán melyek lehetnek azok az ismétlődő elemek a beszédben, amelyek létrehozzák a beszédritmust, és hogyan jellemezhető a hallható beszédből kinyerhető információk alapján.

### 1. Milyen tényezők kelthetik a ritmus érzetét a beszédben?

Az **idő** alapvető szerepet játszik a beszédritmus kialakításában. Hasonlóan ahhoz, ahogy a zenében meghatározhatók ritmusképletek, a beszédben is feltételezhető egy időbeli szerkezet, amely a ritmus alapját képezi. Az egyik legfőbb kérdés, hogy mit tekinthetünk a beszédritmus alapegységének, amelynek változása létrehozza ezt az időbeli mintázatot. Amíg a zenében különböző vagy éppen azonos magasságú hangokhoz társítanak hosszabb vagy rövidebb időtartamot, addig a beszédben több különböző méretű egységet is tekinthetünk a beszédritmus alapjának. A legkisebb egységek a beszédhangok, amelyeknek időtartama egymásután létrehoz egy időszerkezetet. A beszédhangoknak két nagy csoportja van: a magánhangzók és a mássalhangzók. Ezen két csoport váltakozása, egymáshoz viszonyított időtartamaránya szintén kialakít egy sajátos időbeli struktúrát. A beszédhangokból felépülő szótagok vagy az azokból felépített szavak, mondatok stb. is alapját képezhetik a beszédritmusnak (Bailey et al. 1999; Grabe–Low 2002). Ezeken az egységeken érzékelhető, sőt mérhető is az **artikulációs tempó változása**. Az artikulációs tempó meghatározása szerint megadja, hogy a beszélők adott idő alatt hány hangot, szótagot vagy szót stb. ejtenek ki, ha a szüneteket nem vesszük figyelembe (Gósy 2004).

Ezeknek az alapegységeknek (hang, szótag, szó stb.) azonban nem csak az időbeli szerkezete hozhatja létre a beszédritmust, és keltheti a ritmusosság érzetét a hallgatóban. A beszédben léteznek bizonyos lüktetések; **hangsúlyos** és **hangsúlytalan** hangok, illetve szótagok váltakoznak egymás után, amelyek szintén ismétlődő mintázatokat alkothatnak (Honbolygó 2009). A zenében ezeket

a csoportosulásokat egyrészt az intenzitás (hangerő) változtatásával érik el, ugyanakkor a zene ritmusának érzékeléséhez hozzájárul a dallam, azaz a hangmagasság változása is. A zenéhez hasonlóan a beszédben is találtak kapcsolatot a hangsúlyos helyzet és az **intenzitásváltozás** (a hangerő változása), valamint az **alaphangmagasság-változás** (más néven: a hanglejtés) között (Cumming 2011; Tilsen–Johnson 2008). A beszédben lévő hangsúlyviszonyok azonban nem írhatók le olyan könnyen, mint ahogyan lekottázunk egy dallamot. Máig nincs egyetértés például abban, hogy a hangsúlyhoz pontosan milyen jelenségeket, mérhető tényezőket társíthatunk a beszédben, mi a leírásának a megfelelő módja, vannak-e fokozatai stb. (Fox 2000). A hangsúly alapján szerveződő beszédritmussal foglalkozó elméleti terület a metrikus fonológia, amely olyan szabályokat, elveket igyekszik megfogalmazni, amelyek a hangsúlyos és hangsúlytalan szótagok sorrendjét, előfordulásait korlátozzák, valamilyen módon szabályozzák. Minden nyelvre igaznak tekintik például azt az elvet, hogy minden szóban léteznie kell egy főhangsúlynak, amely hangsúlyosabb a többi szótagnál. Ez azt jelenti, hogy nem jöhet létre ún. hangsúlyütközés, azaz egy szónak csak egyetlen főhangsúlya lehet (bővebben l. Honbolygó 2009).

A beszédritmussal nemcsak magát az elhangzó beszédet és annak jellemzőit szokták kapcsolatba hozni, hanem jelentőséget tulajdonítanak a szüneteknek is. A ritmus lényeges elemének tekintik a szünetek elhelyezkedését és hosszát is. Habár a kutatások nem mutatnak egyértelmű összefüggést a szünetek hossza és a beszédritmus érzékelése között (Olaszy 2006), a szünetek megjelenése kimutathatóan az artikulációs tempó csökkenéséhez vezet, továbbá hatással van az alaphangmagasságra és az intenzitásra is. Ez alapján tehát azt mondhatjuk, hogy valamilyen módon szerepet játszik a beszédritmus kialakításában.

Összefoglalva tehát a beszédritmus több akusztikai jellemzővel (artikulációs-tempó-változás, intenzitásváltozás, alaphangmagasság-változás) hozható kapcsolatba. Feltehetően nem pusztán az imént felsorolt tényezők egyikének megjelenése vagy hiánya hozza létre a beszédritmust vagy annak érzetét a hallgatóban, hanem a különböző tulajdonságok együttesen alkotnak egy összetett rendszert (Cumming 2011).

## 2. A különböző nyelveknek hasonló vagy egymástól eltérő a beszédritmusa?

Egy múlt századbeli elmélet szerint a különböző nyelvekben különbözőféleképpen érhető tetten a beszédritmus. Először két nagy csoportot különítettek el egymástól: a **hangsúly-időzítésű** és a **szótag-időzítésű nyelvek** osztályát (Lloyd 1940; Pike 1945). A szótag-időzítésű nyelvek hangzását a puska ropogásához hasonlították, ezekben a nyelvekben ugyanis egyenletes ritmust véltek felfedezni. Ezzel szemben a hangsúly-időzítésű nyelvek hangzásáról úgy tartották, hogy

a Morse-kódra emlékeztet, mivel szabályos, de nem egyenletes lüktetést véltek felfedezni benne. Ebből a percepció (hallható, észlelhető) élményből kiindulva más beszédritmus alapegységeket határoztak meg a különböző osztályokban. A szótag-időzítésű nyelvek esetében az észlelhető egyenletességet a szótagok időtartamához kapcsolták, és feltételezték, hogy a kiejtett szótagok időtartama közel azonos. Ezzel szemben a hangsúly-időzítésű nyelvekről úgy gondolták, hogy a szabályos, de nem egyenletes ritmusérzetet más alapegység időtartambeli azonossága kelti. Feltételezték, hogy a hangsúlytól hangsúlyig terjedő szakaszok időtartama nagyjából megegyezik (Pike 1945; Abercombie 1967). A két különböző osztályban lévő nyelvekben tehát közös, hogy a beszéd egyenlő időtartamú részekre tagolódik, de a beszédritmus alapegysége eltér egymástól.

A későbbiekben a két kategória mellé egy harmadikat is felállítottak, az ún. **moraidőzítésű nyelveket** (Ladefoged 1975). A moraidőzítésű nyelvek esetében az egymást követő morák (azaz a szótag súlyát megadó fonológiai egységek, amelyek lényegében megfeleltethetők a magyar időmértekes verselésben lévő hosszú vagy rövid szótagoknak) időtartamáról feltételezték, hogy közel azonos. Tipikusan hangsúly-időzítésű nyelvnek gondolták például az angolt, a hollandot, szótag-időzítésű nyelvek közé sorolták a franciát, a spanyolt, és tipikusan moraidőzítésű nyelvnek tekintették a japánt. Később több nyelvről megpróbálták megállapítani a percepció élmény, a szótagfelépítési szabályok és az adott nyelvben meglévő hangsúlyozási elvek alapján, hogy mely osztályba tartozhat. Az 1. táblázatban található egy összefoglalás a vizsgált nyelvek besorolásáról (Grabe–Low 2002; Mairano 2007).

1. táblázat. A nyelvek beszédritmus kategóriák szerinti csoportosítása  
(Grabe–Low 2002; Mairano 2007)

Beszédritmustípus	Nyelvek
Hangsúly-időzítésű	angol, arab, cseh, holland, német, portugál, thai
Szótag-időzítésű	francia, görög, olasz, spanyol, tamil
Mora-időzítésű	japán
Kevert	katalán, lengyel

A magyar nyelvet nem tudták egyértelműen besorolni. Feltételezték, hogy alapvetően a szótag-időzítésű nyelvek közé tartozik, de a rövid és hosszú magánhangzók miatt a szótagok hosszát változónak gondolták, így nem sorolták tisztán a szótag-időzítésű nyelvek közé (Varga 2002).

A beszédritmusosztályokról szóló elméletet megpróbálták mérésekkel is igazolni, hogy vajon tényleg közel azonosak-e a szótagok, a hangsúlytól hangsúlyig terjedő egységek vagy a morák időtartamai a különböző nyelvekben. A kutatások viszont cáfolták ezt az elméletet. A szótagok időtartama nagyon nagy változatos-

ságot mutatott, és a többi alapegységet sem találták azonosnak a különböző nyelvekben (Roach 1982). Az alapegységek időtartamának nagy eltérései ugyanakkor nem jelentik azt, hogy a nyelvek nem különülnek el beszédritmusuk alapján. Több kutatás azt mutatta, hogy a kisgyerekek, köztük az újszülöttek is képesek megkülönböztetni az anyanyelvükön elhangzott mondatokat más beszédritmusosztályba sorolt nyelvek mondataitól. A nyelvek beszédritmusuk szerinti csoportosulását támasztja alá többek között az a percepciós kísérlet is, amelyben francia újszülötteket vizsgáltak. Úgy torzították a felvett hanganyagot, hogy csak az alacsony frekvenciákat tartották meg. Ez lényegében azt jelenti, hogy az észlelés szempontjából előtérbe kerültek a beszéd szupraszegmentális eszközei (pl. beszédritmus, beszédhang, tempó, hangerő stb.). A csecsemők reakcióit az alapján számszerűsítették, hogy milyen ütemben szívták a cumisüvegből az ételt. Az eredmények szerint az újszülöttek a felvételeket hallva különbséget tudtak tenni angol (hangsúly-időzítésű nyelv) és japán (moraidőzítésű nyelv) mondatok között, de a holland és az angol mondatok között, azaz két hangsúly-időzítésű nyelv között nem (l. Ramus 1999).

### 3. A beszédritmusosztályok összefüggése mérhető időbeli tényezőkkel

Néhány percepciós kutatás tehát azt mutatta, hogy a nyelvek csoportosíthatók beszédritmusuk szerint, ám a beszédben sokáig nem sikerült olyan tényezőket találni, amelyek alapján elkülöníthetők lettek volna a beszédritmusosztályok egymástól. Ramus és munkatársai (1999) kutatása hozta meg az áttörést. Olyan mérőszámokat (számszerűsíthető jellemzőket) fejlesztettek ki, amelyek kapcsolatot mutattak a korábban felállított beszédritmusosztályokkal.

#### 3.1. A beszédritmus egyik lehetséges alapegysége

Modelljünkben a szótagok, hangsúlytól hangsúlyig terjedő szakaszok, illetve a morák helyett egységes beszédritmus-alapegységet hoztak létre. Minden nyelven az ún. **magánhangzós szakaszok** és **mássalhangzós szakaszok** időtartamát vizsgálták. Mássalhangzós szakaszon az egymást közvetlenül követő mássalhangzók együttesét értjük, amelyeket nem választ szét magánhangzó. Így a hangsorok felépítésétől függően egy mássalhangzós szakasz állhat egy, kettő vagy akár több mássalhangzóból is. A mássalhangzós szakasz időtartamát pedig a benne lévő beszédhangok időtartamának összege adja. Az *ablaka* szóban például a [b] és a [l] együttesen alkot egy mássalhangzós szakaszt, mivel a két mássalhangzó közvetlenül egymás mellett van. E két hang időtartamának összege az adott mássalhangzós szakasz időtartama is egyben. A példában a [k] beszédhang pedig önmagában alkot egy mássalhangzós szakaszt a definíció értel-

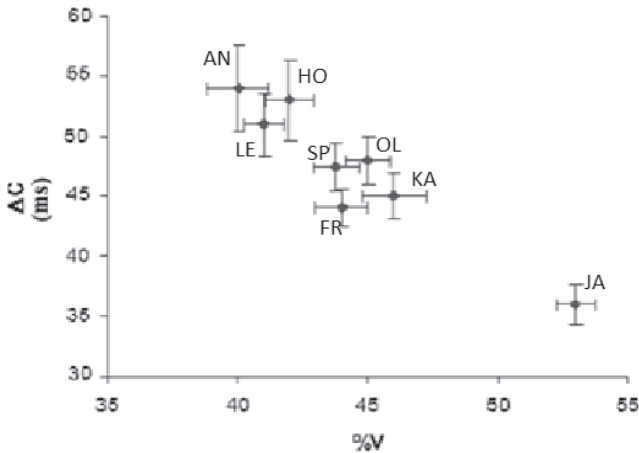
mében. A magánhangzós szakasz hasonló módon definiálható a mássalhangzós szakaszhoz. Az egymást közvetlenül követő egy, két vagy több magánhangzó együttese, amelyet nem választ szét mássalhangzó, alkot egy magánhangzós szakaszt. A magánhangzós és mássalhangzós szakaszok átívelnek a szavakon, sőt a szüneteken is oly módon, hogy a szünetek időtartamát nem veszik figyelembe. Tehát a beszéd teljes folyamatát úgy határozták meg, mint a magánhangzós és mássalhangzós szakaszok váltakozása. Azaz a definíció alapján a magánhangzós szakaszt csak mássalhangzós szakasz követheti, a mássalhangzós szakaszt pedig csak magánhangzós szakasz egészen addig, amíg véget nem ér a beszéd, vagy a kiválasztott beszédrészlet (Ramus et al. 1999). Ez alapján ezt a példamondatot így tagolhatnánk: |A | n|a|gym|a|m|a| sp|e|c|ia||i|t|ás|a| k|é|t|s|é|g|e|t| k|i|z|á|r|ó|a|n| a| b|a|r|a|ckb|e|f|ő|tt|.

### 3.2. A beszédrítmus mérésére fejlesztett eszközök és az ezekkel kapott eredmények

A magánhangzós és mássalhangzós szakaszok időtartamának segítségével olyan mérőszámokat hoztak létre, amelyekkel lehetővé vált az egyes nyelvek beszédrítmus-sajátosságainak mérése. Megmérték egy adott nyelvre az összes mássalhangzós szakasz időtartamát, és meghatározták a szórását (azt mutatja meg, hogy az értékek milyen mértékben térnek el a számtani átlagtól) (a mérőszám jele:  $\Delta C$ ). Ez a mérőszám lényegében azt mutatja meg, hogy a mássalhangzós szakaszok időtartama mennyire változatos egy adott nyelven. Minél szélsőségesebb időtartamértékek jellemeznék egy nyelvet, ez a mérőszám annál nagyobb értéket mutat. Amennyiben az időtartamértékek nagyon hasonlók egymáshoz, ez a mérőszám alacsony értéket vesz fel. Az eredmények szerint a mássalhangzós szakaszok időtartamának szórása ( $\Delta C$ ) szoros összefüggésben van a korábbi elmélet szerint meghatározott beszédrítmusosztályokkal. Magasabb értékeket mutat a klasszikusan hangsúly-időzítésűekhez sorolt nyelvekben (angol, holland), mint a szótag-időzítésűekhez sorolt nyelvekben (francia, olasz, spanyol) (Ramus et al. 1999).

A másik mérőszám, amely összefüggést mutatott a korábban meghatározott beszédrítmusosztályokkal a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának arányára épül. Összeadták az összes magánhangzós szakasz időtartamát, és kiszámolták, hogy a teljes időtartamnak hány százaléka a magánhangzós szakaszok időtartama (a mérőszám jele: %V). Ez a mérőszám azt mutatja meg, hogy mennyire kiegyensúlyozott a magánhangzók és a mássalhangzók időtartamának aránya, tehát ha ez a mérőszám 50%-nál kisebb értéket vesz fel, akkor a magánhangzók időtartamának összege kisebb a mássalhangzók időtartamának összegénél. Azt találták, hogy ez a mérőszám (%V) magasabb értékeket vesz fel a hagyományosan szótag-időzítésűekhez sorolt nyelvekben,

mint a hangsúly-időzítésűekhez sorolt nyelvekben. A két mérőszám ábrázolható egy koordinátarendszerben (%V az  $x$  tengely,  $\Delta C$  az  $y$  tengely az 1. ábrán), amelynek segítségével csoportosíthatók a nyelvek beszédritmusuk szerint. Jól látható, hogy a tipikusan hangsúly-időzítésűnek tartott nyelvek (pl. angol, holland) elkülönülnek a tipikusan szótag-időzítésű nyelvektől (pl. francia, olasz, spanyol), és a japán mint moraidőzítésű nyelv mindkét csoportosulástól távol esik (Ramus et al. 1999).



1. ábra. A nyelvek beszédritmus-mérőszámok alapján történő csoportosítása (AN = angol, HO = holland, OL = olasz, FR = francia SP = spanyol, KA = katalán, JA = japán, LE = lengyel) (Ramus et al. 1999: 9)

A magánhangzós és mássalhangzós szakaszokból számtalan más mérőszámot is létrehoztak, hogy minél jobban feltérképezhessék a beszédritmus és az időzítés kapcsolatát. Az egyik leggyakrabban használt ezek közül Grabe és Low (2002) által felállított mérőszámtípus (PVI), amely figyelembe veszi a magánhangzós és mássalhangzós szakaszok időtartama mellett a szakaszok sorrendjét is. A szakaszok időtartamát ugyanis összeveti a környező szakaszok időtartamával, lényegében az egymás melletti szakaszok különbségeit számolja ki, és matematikai módszerekkel megadja ezeknek a különbségeknek a változatosságát. Ennek a mérőszámnak a jelentőségét egy életből vett példán szemléltetjük a következőkben. Van egy sportoló, akinek 4 teljes kört kell lefutnia az atlétikapályán. Az első kört megteszi 1 perc alatt, a másodikat már csak 2 perc alatt, a harmadikat 3 perc alatt, a negyediket végül csak 4 perc alatt teljesíti. Másnap a futó megpróbál taktikát váltani, és nem kezd olyan gyorsan. Az első kört 2 perc alatt teljesíti, a másodikat 4 perc alatt, a harmadikra nagyon felgyorsul, és 1 perc alatt teszi meg, negyedik körét pedig 3 perc alatt fejezi be. A futó egyik taktiká-

val sem járt jobban, ugyanannyi idővel teljesítette a távot, ugyanúgy volt olyan köre, amelyet 1, 2, 3 és 4 perc alatt teljesített. Ennek ellenére egészen más ritmusmintázatot látunk. Míg az első nap egyenletes mértékben, mindig kicsit változott a tempója, addig második nap nagyon nagymértékű ritmusváltásokkal futott. A PVI mérőszámok pontosan ezeknek az egymást követő ritmusváltásoknak a nagyságát ragadják meg a beszédben. Minél egyenletesebb a tempó, azaz hasonlóbb az egymást követő egységek időtartama a beszédben, annál alacsonyabb ez a mérőszám. Minél nagyobb tempóbeli váltások vannak a beszédben, azaz hirtelen gyorsulások és hirtelen lassulások követik egymást (ahogy például ezt a futó második napján láttuk), annál magasabb ez az érték. Az eredmények azt mutatták, hogy a PVI mérőszámok alapján is elkülönülnek a tipikusan hangsúly-időzítésűnek tartott nyelvek (angol, holland) a klasszikusan szótag-időzítésűekhez sorolt francia és spanyol nyelvtől.

### 3.3. Milyen tényezők befolyásolhatják egy nyelv beszédritmusát és ezáltal a beszédritmus-mérőszámok értékét?

A felállított mérőszámok összefüggésben lehetnek az adott nyelvben található szótagtípusokkal. Egyrészt előfordulhatnak olyan szótagok, amelyekben egy mássalhangzó és egy magánhangzó van (C a mássalhangzó és V a magánhangzó jele, így CV-vel jelölhetjük). Másrészt vannak olyan szavak is, amelyekben mássalhangzó-torlódás figyelhető meg, és a szótagokban kettő, három, négy vagy több mássalhangzó is előfordulhat (CVC, CVCC, CVCCC stb.). Minél többféle szótagtípus fordul elő egy közlésben, vagy egy nyelvben, annál szélsőségesebb értékeket vehetnek fel a mássalhangzós szakaszok időtartamai, így az ezekből számolt mérőszám ( $\Delta$ ) értéke magas lesz. Ugyanakkor minél több mássalhangzó-torlódás van egy közlésben vagy egy nyelvben, annál nagyobb a mássalhangzók időtartamának aránya a magánhangzók időtartamának arányához képest. Ha minél több magánhangzó-torlódás van az adott beszédrészetben, vagy az adott nyelvben megkülönböztetik a hosszú és rövid magánhangzót, és több hosszú magánhangzó fordul elő, akkor annál nagyobb a magánhangzók időtartamának aránya a mássalhangzók időtartamának arányához képest (Ramus et al. 1999; Grabe–Low 2002). Összefoglalva tehát egy mondat **szótagjainak a felépítése** hozzájárul a beszédritmus kialakításához.

A nyelvek szavainak fonológiai felépítettségén (szótagtípusokon, rövid-hosszú magánhangzók meglétéén vagy hiányán) túl a mérőszámok érzékenyek a beszédhangok időtartamát meghatározó minden szabályszerűségeire. Több tényezőről feltételezik, hogy megjelenésük illetve hiányuk okozhatja az eltéréseket a nyelvek beszédritmusa között. Az egyik ilyen tényező a **magánhangzó-redukció**. A magánhangzók időtartama hangsúlytalan helyzetben rövidebb, mint hangsúlyos helyzetben, és olykor formánsszerkezetük sváyszerű (semleges) magánhang-

zóra egyszerűsödik. Ilyen tényező még a **hangsúlyos pozíció** is. Több nyelvben (angol, holland, német, olasz, arab stb.) kimutatták, hogy a magánhangzók hangsúlyos helyzetben hosszabbak, mint hangsúlytalan helyzetben. A nyelvek közti beszédritmusbeli különbséget pedig az okozhatja, hogy ezek a jelenségek kevésbé lelhetőek fel vagy kisebb mértékben meghatározóak a szótag-időzítésű nyelvekben, mint a hangsúly-időzítésű nyelvekben (Barry et al. 2001). Habár itt most csak pár tényezőt emeltünk ki, amelyek a különböző nyelvekben eltérően befolyásolhatják a magánhangzós és mássalhangzós szakaszok időtartamát, nem szabad megfeledkeznünk arról, hogy még számtalan más tényező (pl. a hangsorban elfoglalt hely, az adott szó hosszúsága, a beszédmód stb. l. van Santen 1992; Gósy 2004; Gósy–Beke 2010) is hatást gyakorol rájuk. Ma is kutatás tárgyát képezi, hogy melyek azok a tényezők, amelyek szoros kapcsolatot mutatnak a beszédritmus-mérőszámokkal, és meghatározó szerepet töltenek be a beszédritmus érzetének kialakításában.

### **3.4. A beszédritmus-mérőszámokon alapuló kutatások problémái és kritikája**

Az egyik problematikus kérdés a beszédritmus-mérőszámokkal kapcsolatban, hogy az általuk kapott eredmények alapján besorolható-e a nyelvek beszédritmusosztályokba vagy sem. Mivel a mérőszámok alapján a különböző típusú nyelvek nem különülnek el élesen, és találtak olyan nyelveket, amelyek semmilyen csoportba nem sorolhatók (pl. görög, maláj, román, tamil, walesi nyelv), ezért ma már nem tekintik őket szigorú osztályoknak. Helyette a nyelvek beszédritmusát a főbb kategóriák közti átmenetekként határozzák meg, egyfajta kontinuum mentén képzelik el őket (Ramus et al. 1999; Grabe–Low 2002; Mairano–Romano 2007).

A beszédritmus-mérőszámokat eredetileg csak a különböző egységek időtartamának vizsgálatára használták. A beszédritmust illetve a beszéd rendszerszerű ismétlődésének érzetét azonban nemcsak a szakaszok időtartambeli változatosága okozhatja a befogadóban, hanem számtalan más akusztikai tényező is létrehozhatja. A beszédritmus-mérőszámokat megpróbálták alkalmazni más paraméterekre (pl. alaphangmagasság-változás) is (Cumming 2010), de a beszédritmus és más mérhető, akusztikai tényezők összefüggéseinek átfogó, alapos feltárása mindeztől még nem történt meg.

Összefoglalva elmondhatjuk, hogy a beszédritmus-mérőszámok a beszédritmusnak csak egyetlen aspektusával foglalkoznak; az időzítést mérik, és csak ezt a dimenziót lehet velük vizsgálni. Habár megbízhatóságáról, alkalmazási szabályairól, felhasználási lehetőségeiről ma is éles vita folyik (l. White–Mattys 2007; Arvaniti 2012), tagadhatatlan, hogy ezen beszédritmus-mérőszámok alkalmazása jelenti az egyetlen olyan módszert, amely bármilyen kapcsolatot mutatott a

klasszikus osztályozással és a percepciók tesztek eredményeivel. A beszédritmus-mérőszámok ugyan nem alkalmazhatók szigorú beszédritmusosztályok elkülönítésére, de használatuk iránymutató lehet a nyelvek közti beszédritmusbeli különbségek mérésében, illetve egy adott nyelven belül az időzítésbeli jellemzők feltárásában.

#### 4. A magyar nyelven végzett beszédritmus-kutatások

Magyar nyelven végzett kutatásokban a beszédritmust – a nemzetközi szakirodalomhoz hasonlóan – számtalan tényezővel (hangsúly, dallammenet, intenzitás, szünetezés stb.) próbálták összefüggésbe hozni. Habár történt kísérlet az intenzitás és a beszéddallam akusztikai mérésére a beszédritmussal kapcsolatban (Kecskés 1966), a kutatások középpontjában a beszédritmus időbeli dimenziója állt. A következőkben ezeknek a magyar nyelvre végzett időtartammérések eredményeit mutatjuk be.

Mint korábban írtuk, a magyar nyelvről feltételezték, hogy alapvetően a szótag-időzítésű nyelvek közé tartozik, de a rövid és hosszú magánhangzók megléte miatt nem lehet tisztán ebbe a csoportba sorolni (Varga 2002). Több olyan akusztikai vizsgálatot is elvégeztek, amely során a szótagok időtartamát mérték meg és hasonlították össze magyar nyelvű szövegekben. A mérések nem támasztották alá a szótagok időtartamának egyenlőségét, hasonlóan a nemzetközi eredményekhez (Kecskés 1966; Gósy 2000). Ugyanakkor ezen tanulmányok több összefüggést is feltártak a beszédritmussal kapcsolatban, a következőkben eredményeiket részletesen is ismertetjük.

Kecskés (1966) vizsgálataiban elsősorban a versek ritmusával foglalkozott. Mérései azt mutatták, hogy a nem skandált, normál ejtésű versekben a rövid szótagok és a hosszú szótagok aránya nem közelít az 1 : 2-es arányhoz, valamint az ütemhangsúlyos verselésben nem talált összefüggést a felírt ritmusképletek és a ténylegesen mérhető időtartamok között. Felfigyelt azonban egy olyan összefüggésre, hogy a versekben a rövid és a hosszú magánhangzót tartalmazó nyílt szótagok időtartama átlagosan kisebb a rövid vagy hosszú zárt szótagok időtartamánál. (A kutatásban nem a ténylegesen mérhető időtartamot használta fel, hanem egy olyan arányszámot, amely függetlennek tekinthető a beszédtempótól. A beszédtempó ugyanis nagymértékben befolyásolja a szótagok időtartamát, és egy szöveg vagy vers elmondása során folyamatosan változik. Hogy ténylegesen csak a szótagtípusok hatását vizsgálhassa, és az eredményeket ne befolyásolja a beszédtempó, ezért a szótagokra mért időtartamot elosztotta az adott sorra kapható beszédtempóval.) Tehát az eredmények alapján a mássalhangzó-többlet, illetve mássalhangzó-torlódás jelentősebben növeli egy szótag időtartamát, mint az, hogy a szótag hosszú magánhangzót tartalmaz rövid magánhangzó helyett.

Gósy (2000) kutatásában reppelt szöveget, időmértékes verset és normál ejtésű szöveget is vizsgált. Valamennyi szövegtípus elemzése azt mutatta, hogy az egymást követő szótagok időtartama folyamatosan változik. Műfajtól függetlenül megfigyelhető, hogy a rövidebb és a hosszabb szótagok tendenciaszerűen váltakoznak egymás után, amely egyfajta ritmikus lüktetésként fogható fel. Ez a váltakozás viszont nem ragadható meg valamilyen egyértelmű szabállyal vagy szabályrendszerrel. A vizsgált három műfaj közül a reppelt szövegek több sajátosságot is mutattak. Az eredmények azt mutatták, hogy reppelésben igen ritkák a szünetek, és időtartamuk is meglehetősen rövid a normál ejtésű szövegekhez képest. Ráadásul az artikulációs és beszédtempójuk igen magas a normál ejtésű szövegekhez képest, a szótagok időtartama jóval rövidebb a reppelt szövegekben, amely meghatározó lehet a ritmusélmény szempontjából. Az időmértékes vers soraiban a szótagok időtartama megfeleltethető volt a ritmusképletben előírtaknak, tehát a rövid-hosszú különbségeknek.

A szótagok időtartamának elemzése mellett a magyar nyelv beszédritmusát a beszédritmus-mérőszámok segítségével is vizsgálták már. A felolvasott mondatokra, szövegre, spontán beszédre kapott eredmények a tipikusan szótag-időzítésűekhez sorolt nyelvekéhez közelítenek. A különböző beszélők és a különböző beszédmódok beszédritmusa ugyanakkor nagy eltérést mutat (Kohári 2013). A spontán beszédben a magánhangzós és mássalhangzós szakaszok időtartama szélsőséges értékeket vesz fel, nagyobb variabilitás jellemzi, mint a felolvasásban mért beszédritmusértékeket. A két beszédmód beszédritmusbeli eltéréseinek feltehető oka, hogy a spontán beszédben a beszéd tervezése és az ily módon megfogalmazott szöveg meghangosítása egyidejűleg zajlik, míg felolvasáskor a beszélőnek csak a korábban megtervezett szöveg kiejtésének megvalósítására kell koncentrálnia. A spontán beszédben lévő többletfeladat, illetve annak nehézségei az időviszonyok szélsőséges ingadozásához vezetnek (Gósy 2005).

Más magyar nyelvre végzett tanulmányok a tempó változására fókuszáltak. A beszédritmus időbeli dimenziója ugyanis felfogható az artikulációs tempó folyamatos változásának: a beszélők hol gyorsabban, hol lassabban beszélnek. Azt találták, hogy az intonációs frázisokon belül a szavak tendenciaszerűen lassulnak az egységek vége felé (Váradí–Beke 2013), és a tagmondatok többsége is lassuló tendenciát mutat. A lassulásról megállapítható, hogy nem szigorúan hangról hangra vagy szóról szóra történik, tehát nem arról van szó, hogy az egymást követő hangok, szótagok vagy szavak időtartama mindig egyre nagyobb. Ez a lassulás inkább fokozatosan következik be, azaz a tagmondat vagy intonációs frázis végéhez közeledve egyre nagyobb valószínűséggel találunk hosszabb és hosszabb hangokat, szótagokat, szavakat stb. (Kohári 2014). A lassulások és gyorsulások gyakoriságának, mértékének pontos leírása még nem történt meg, ezen összefüggések feltárása a jövő egyik kutatási feladata lesz.

A beszédritmusról láthattuk, hogy egy igen összetett jelenség, amelynek létre-

jöttéhez több tényező (tempóváltozás, intenzitásváltozás, hangmagasság-változás, a szavak fonológiai felépítettsége stb.) is hozzájárul, sajátosságainak alapos megismeréséig azonban még sok vizsgálatra lesz szükség. Ezen vizsgálatokból szerzett tudás ugyanakkor több gyakorlati területen is hasznosnak tekinthető. A beszéd gépi előállításában például a beszédritmussal kapcsolatos ismereteinket a minél természetesebb, jobb hangzása elérése érdekében alkalmazzák. Nagy igény mutatkozik a beszédritmus további jellemzőinek feltárására, a ritmusváltások helyének és gyorsaságának minél pontosabb megismerésére, amely a szövegfelolvasó programok továbbfejlesztését segítené elő. Az idegen nyelvek minél tökéletesebb elsajátításában is fontos szerepe lehet ezeknek a kutatásoknak, hiszen a beszédritmus alapvető jellemzőinek feltérképezésével könnyebben tudatosítható a tipikus kiejtés a nyelvtanulóban, továbbá az anyanyelv és idegen nyelv közti különbségek célzottabban válnak gyakoroltathatókká.

### Irodalom

- Abercrombie, David 1967. *Elements of general phonetics*. Edinburgh University Press, Edinburgh.
- Arvaniti, Amalia 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40. 351–373.
- Bailey, Todd M.– Plunkett, Kim – Scarpa, Ester 1999. A cross-linguistic study in learning prosodic rhythms: rules, constraints, and similarity. *Language and Speech* 42/1. 1–38.
- Barry, William J. – Andreeva, Bistra – Koreman, Jacques 2009. Do rhythm measures reflect perceived rhythm? *Phonetica* 66. 78–94.
- Cumming, Ruth E. 2010. The interdependence of tonal and durational cues in the perception of rhythmic groups. *Phonetica* 67. 219–242.
- Cumming, Ruth E. 2011. Perceptually informed quantification of speech rhythm in Pairwise Variability Indices. *Phonetica* 68. 256–277.
- Fox, Anthony 2000. *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford University Press, Oxford, USA.
- Gósy Mária 2000. A beszédritmus elemzésének egy lehetséges megközelítése. *Magyar Nyelvőr* 124. 273–287.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.
- Grabe, Esther – Low, Ee Ling 2002. Durational variability in speech and the Rhythm Class Hypothesis. In Gussenhoven, Carlos – Warner, Natasha (eds.): *Papers in Laboratory Phonology* 7. Cambridge University Press, Cambridge.
- Honbolygó Ferenc 2009. *A beszéd prozódiai jellemzőinek észlelése*. Doktori disszertáció. ELTE, Budapest.

- Kohári Anna 2013. Az időzítés és a beszédrítmus néhány összefüggése magyar spontán beszédben. In Várad Tamás (szerk.): *VII. Alkalmazott Nyelvészeti Doktoranduszkonferencia kötete*. MTA Nyelvtudományi Intézet, Budapest, 93–103.
- Kohári Anna 2014. Tempóváltozások a vizsgált szakasz nagyságának függvényében. *Beszédkutatás 2014*. 220–236.
- Kecskés András 1966. A komplex ritmuselemzés elvi kérdései. *Irodalomtörténeti Közlemények* 1–2. 106–140.
- Mairano, Paolo – Romano, Antonio 2007. A. Inter-subject agreement in rhythm evaluation for four languages (English, French, German, Italian). In Trouvain, Jürgen – Barry, William J. (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken, 1149–1152.
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Akadémiai Kiadó, Budapest.
- Pike, Kenneth 1945. *The intonation of American English*. University of Michigan Press, Ann Arbor.
- Ramus, Franck – Nespors, Marina – Mehler, Jacques 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 72. 1–28.
- Roach, Peter 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In Crystal, David (ed.): *Linguistic controversies*. Edward Arnold, London, 73–79.
- van Santen, Jan 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.
- Tilsen, Sam – Johnson, Keith 2008. Low-frequency Fourier analysis of speech rhythm. *Journal of the Acoustical Society of America* 124/2 EL. 34–39.
- Várad Viola – Beke András 2013. Az artikulációs tempó variabilitása felolvasásban. *Beszédkutatás 2013*. 26–41.
- Varga László 2002. *Intonation and stress: Evidence from Hungarian*. Palgrave Macmillan, New York.
- White, Laurence – Mattys, Sven L. 2007. Calibrating rhythm: first language and second language studies. *Journal of Phonetics* 35. 501–522.

### Kérdések, feladatok

1. Milyen mérhető akusztikai tényezőkkel hozható kapcsolatba a beszédrítmus?
2. Mely tényezők befolyásolhatják a magánhangzós és mássalhangzós szakaszok időtartamát és ezáltal a beszédrítmus-mérőszámok értékét?
3. Mely nyelvek beszédrítmusához hasonlítható a magyar nyelv beszédrítmusa?

## Milyen tulajdonságokra következtethetünk a beszédből?

A beszédben megnyilvánuló információtartalom nem csupán a kommunikációs üzenetet és annak jelentését hordozza magában, hanem egyéb információkat is tartalmaz. Ezek olykor a beszélő szándéka ellenére is jelen vannak a közlésben. Ezek a sokszor rejtve maradók, a beszélőről is árulkodó tartalmak nemcsak egy nyelvészeti szakértő vagy épp egy rendőrségi kihallgatást végző számára érdekesek. A mindennapok során számos olyan beszédhelyzettel találkozunk, amelyek során magunk is akarva-akaratlan információkat vagy vélt információkat nyerünk ki a hallott beszédprodukciókból. Emellett magunk is felismerhetünk olyan beszédhelyzeteket, amelyekben igyekszünk annak megfelelően viselkedni, beszélni, amelyről azt gondoljuk, hogy a helyzet megkívánja, vagy ahogy szeretnénk, hogy a beszédünket halló személy bennünket észleljen.

A beszélő beszédéből általában a következő tulajdonságokra szoktunk következtetést levonni:

- a beszélő életkora;
- a beszélő testalkata (magasság, testsúly);
- a beszélő iskolázottsága;
- a beszélő foglalkozása;
- a beszélő vonzó tulajdonságai;
- a beszélő személyisége;
- a beszélő állapotára vonatkozó információk, amelyek lehetnek a pillanatnyi vagy éppen egy tartós állapot mutatói (pillanatnyi: érzelmi állapot, egészségi állapot, alkoholos befolyásoltság; tartós: dohányzás, idült betegségek, balesetek vagy betegség hatására bekövetkezett maradandó elváltozás).

Ezek a következtetések egyes esetekben igazak, más esetekben tévesek. A fentiekben felsorolt lehetőségek legtöbbje tudományos kutatások témájaként is szerepelt már. Az alábbiakban számba vesszük azokat a vizsgálatokat, amelyekben a beszédből kinyerhető nem nyelvi információk megjelenésével, bizonyításával foglalkoztak a kutatók.

### 1. A beszéd hatására kialakuló benyomások létrejötte

A hallás folyamata révén az agyba kerülő információk összetett folyamatokon és ehhez kapcsolódó asszociációkon keresztül kerülnek feldolgozásra. Egyszerűen fogalmazva: a hallott információ nyomán az agyban létrejön a hallottaknak

tulajdonított jelentés, valamint az ehhez kapcsolódó benyomások, asszociációk, amelyek lehetnek természetesen igazak, de tévesek is. Ezek egy részét tudatosan vagy tudattalanul észleljük, más részük csak műszeres vizsgálatok során mutatható ki a beszédből.

A beszédből nyerhető információk a következőképp osztályozhatók (Gocsál 1998: 122–134):

1. „A nyelviileg értelmezhető információk, tehát szavak, mondatok szemantikai értelemben vett jelentése, kiegészítve fonetikai megjelenésükkel (hangsúly, intonáció stb.)”

2. A beszélő személyét jellemző információk, tehát azok a tulajdonságok, amelyek alapján a beszélő felismerhető, pusztán a hangja alapján (anélkül, hogy látnánk), vagy legalábbis megbecsülhető a neme, az életkora, illetve egyéb jellemzői.

3. A beszédhelyzettel kapcsolatos információk, amelyeket nem a beszélő hangja hordoz, de pragmatikai jelentőséggel bírnak a beszédhelyzet feldolgozásakor. Ilyen lehet egy telefonbeszélgetésben a háttérzaj vagy egyéb környezeti hatások megjelenése.

A jelen tanulmány középpontjában a fenti modulok közül a második működésének vizsgálata áll, vagyis azon információk elemzése, amelyek a beszélő hangjában kódolva rejtőznek, nyelviileg nem feldolgozhatók. Feltételezhető, hogy ez a modul több egymáshoz kapcsolódó, különböző funkciókat ellátó egységből áll (Gocsál 1998).

## 2. A beszélő életkora

Az agy eszköztárának fentebb leírt működései közül egyik az életkorbecslő képesség, amelynek bemenete a hallott beszéd, kimenete pedig a beszélőnek tulajdonított életkor. Az életkorbecslés pontosságának vizsgálatával a magyar beszédre vonatkozóan Gocsál (1998), Gósy és Nikléczy (2000), Gósy (2001), valamint Bóna (2013) foglalkozott.

Gocsál (1998) vizsgálati anyagában 19 férfi beszélő szerepelt, feladatuk egy novella felolvasása volt. A felolvasásokat rögzítő felvételeket 28 egyetemi hallgatónak játszották le, akik azt a feladatot kapták, hogy próbálják megbecsülni a hallott beszélő életkorát, illetve jelezzék, ha felismerni vélik valamelyik beszélőt. Az értékelésnél azokat a becsléseket számították „jó” becslésnek, amelyek legfeljebb öt évet tévedtek a beszélő életkorának meghatározásában. Az eredmények kiértékelése során azt találták, hogy a becslések nagyon változatosak, és jelentősen szórnak, továbbá az ítéletek alkotása során jellemző tendencia volt a kerek vagy „félkerek” számok (nullára és ötre végződő) megjelenése. Voltak olyan beszélők, akiknek az életkorát jellegzetesen pontosabban vagy pontatlan-

nabbul becsülték meg a hallgatók, mint a többi felolvasóét. A kísérlet legfontosabb következtetése szerint az életkor nem pontosan becsülhető, de ebben a pontatlanságban megfigyelhető egyfajta következetesség. A szerző ezt a jelenséget egyrészt a beszélő hangjában potenciálisan megjelenő megtévesztő vagy megszokottól eltérő jegyekkel, másrészt pedig az egyes értékelők egyéni tapasztalatainak sokféleségével magyarázta.

Gósy (2001) szintén foglalkozott az életkorbecslés pontosságával, az erre vonatkozó eredményeit a következő alfejezetben, a beszélő testalkatával kapcsolatban ismertetjük.

Egy másik kutatás során (Bóna 2013) azt vizsgálták, hogy hogyan ítélik meg a 60 év feletti beszélők életkorát a fiatalabb hallgatók. A hangmintákat két csoportban (külön a férfiak és nők mintái) használták fel a percepció teszthez, amelyben egyetemisták hallás után próbálták meghatározni a beszélők életkorát. A hallgatók nem tudták, hogy a hallott mintákban csak 60 éves vagy idősebb beszélők szerepelnek. A kutatás eredményei rámutattak arra, hogy az életkor becslése oly mértékben pontatlan, hogy (függetlenül a számszerűsített konkrét életkortól) még az idős korú beszélők idős életkorának azonosítása sem működik pontosan. Egyes 60 éves kor feletti beszélők hangja a hormonális változások ellenére olyan fiatalos marad, hogy az őket hallgatók rendszeresen alábecsülik az életkorukat, amennyiben csak az auditív csatornára támaszkodhatnak. Az eredmények további oka a szerző szerint az, hogy vannak életkoruknak megfelelően tipikus és atipikus beszélők. Míg előbbiek az életkorukra és az azzal együtt járó változásokra jellemző beszédprodukciónal rendelkeznek, mások ettől eltérően beszélnek. A pontatlanságok okai lehetnek még a becslést végzők egyéni tapasztalatai vagy azok hiánya, illetve elvárásaik, sztereotípiáik is befolyásolhatják őket az idősek beszédével kapcsolatban.

Az életkorról alkotott hallgatói benyomások hátterében a különböző életkorú személyek beszédére jellemző akusztikai sajátosságok állnak. Az idősek hangképzésében szignifikáns különbségeket igazoltak Bóna (2009, 2013) kísérletei. A korábbi kutatás során fiatal felnőttek hangképzését hasonlították össze az idősek által képzett hangokkal. Az elemzésben az idősek zöngéje zörejeesebbnek és szabálytalanabbnak mutatkozott a fiatal korosztállyéval szemben, továbbá az általuk képzett hangok kevésbé voltak elkülöníthetők, illetőleg és hátrébb képződtek, mint a fiatalok produkciójában vizsgált hangok.

A beszédre továbbá gyermekkortól időskorig a teljes életszakasz alatt hatással vannak a fejlődéssel járó változások. A kognitív képességek, a gondolkodás sebessége, a finommozgások koordinációja, az izomműködés mind befolyásolja a beszédprodukción, az időseknél emellett egyéb hatások is előfordulhatnak: gyengül a légzőrendszer (hangerő csökkenése), valamint hallásproblémák is felléphetnek.

Egy másik kutatás (Gósy–Nikléczy 2000) azonos beszélők azonos körülmények között, de 25 év különbséggel készített hangfelvételeit vetették össze a kutatók. Az eredmények azt mutatták, hogy a beszéd nagy egyéni variabilitása mellett egyes formánsértékek statisztikailag igazolható mértékben változtak, alátámasztva ezzel az életkorral járó változások meglétét.

A fentínél rövidebb időtartamokon mérte a beszéd egyéni változatosságát két további kutatás (Németh 2013b; Gósy–Krepsz 2014). Mindkét kísérlet eredményei azt mutatták, hogy a megfigyelt paraméterek mentén (beszédtempó, alaphangmagasság) a beszédben néhány hetes eltéréssel is statisztikailag szignifikáns változások tapasztalhatók. (Az életkor beszédbeli jellemzőiről lásd részletesebben ugyanezen kötetben Balázs és Bóna (2016) tanulmányát).

### 3. A beszélő testalkata

A beszélő hangjában megjelenő rejtett információk a testalkatáról is árulkodhatnak. Bizonyára mindenki találkozott már azzal a hétköznapiakban előforduló jelenséggel, amelynek során egy adott személy – akit eleinte csak szemlélünk – egyszer csak megszólal, hangja pedig az előzetes várakozásunkhoz képest nagymértékben eltérő. A jelenség tudományos megközelítésével hazánkban Gósy (2001) foglalkozott. A vizsgálat kiindulópontja a hétköznapi tapasztalat azon formája, amelyben egy bizonyos embertípushoz egy meghatározott beszéd-típust képzelünk el. Gondoljunk csak a filmek szinkronizálására! Nem véletlen, hogy sok esetben hasonló az eredeti személy alkata, megjelenése a magyar hangját kölcsönző személyével. A szerző példaként hozza erre Bud Spencer és magyar hangja, Bujtor István hasonlóságát. A szinkronizálás másik példája, amikor egy jól ismert színész szinkronhangja megváltozik, a nézőnek kellemetlen, bizonytalan érzése támad. A filmvásznon megjelenő, állatokat szinkronizáló színészek megválasztásában is döntő szempont a testalkat. Az egérke beszédét kis termetű színész, általában nő valósítja meg, míg a medve hangját mély hangú és erőteljesebb testalkatú színész kölcsönzi. Nem véletlen tehát, hogy a rajzfilmek szereplőinél is döntő szempont az alkati harmónia a hangok megválasztása során. A kutatás fontossága azonban nemcsak a filmipar fenti jelenségében rejlik, hiszen fontos tapasztalatokkal szolgál a beszélőfelismerés, illetve a kriminallisztika vizsgálataihoz is.

A beszélő testalkatának percepció vizsgálatához tíz magyar anyanyelvű pedagógustól vettek fel hangmintákat. Öt férfi és öt nő szerepelt a kutatásban, életkoruk 23 és 55 év között alakult. Kiválasztásuk testméreteik alapján történt. A választás során a kiindulás a normál magasság és súly volt, azaz a sem alacsonynak, sem magasnak, sem pedig soványnak vagy kövérnek tartott személyek képezték a referenciát (egy férfi és egy nő). A továbbiakban mindkét nem

képviselőiből egy alacsony kövér, egy alacsony sovány, egy magas kövér és egy magas sovány személyt választottak ki a felvételhez. A beszélőktől rövid, olvasott szöveget rögzítettek. A percepciók kísérletben 32 fő vett részt, ugyancsak pedagógusok, 25 és 35 év közötti életkori szórással. Feladatuk az volt, hogy egy előre elkészített tesztlapon becsüljék meg a beszélők testalkatát, súlyát, valamint életkorát. A kutatás hipotézise szerint a három paraméter közül csak az életkor az, amely relatív biztonsággal megbecsülhető. A kísérlet eredményei ezt megerősítették, bár az életkorra és a testalkatra vonatkozó becslések igen változatos eredményeket mutattak. A valóságnak leginkább megfelelő ítéletek az életkor megítélésében keletkeztek, míg a legalacsonyabbak a súly meghatározásában voltak adatolhatóak. A természetre vonatkozó becslések a két említett paraméter között voltak megtalálhatóak, a beszélők magasságát tehát pontosabban becsülték meg a kísérletben részt vevő hallgatók, mint a beszélők súlyát.

Az eredmények a következőképp magyarázhatók: az emberek jobban tudatosítják magukban a beszélő személy életkorát, mint más egyéb tulajdonságait. Ebből következtetni lehet arra, hogy az agy neurális spektrogramjai tárolnak adatokat az életkorról. Továbbá Gósy szerint a mindennapi kommunikációnk során jelentősége van annak, hogy milyen életkorú személlyel beszélünk; annak azonban, hogy milyen alkatú vagy súlyú, kommunikációs szempontból nincsen jelentősége. Ebből következően jóval kisebb az esélye annak, hogy a beszélő személy alkatával kapcsolatban információkat dolgozzunk fel.

Létezhetnek továbbá olyan öntudatlanul lezajló feldolgozási folyamatok, amelyek a testalkattal kapcsolatosak (alkati harmónia). A beszédpercepció során tehát nem tudatosítjuk azt, hogy az illető milyen alkati tulajdonságokkal rendelkezik, mégis furcsa érzésünk támad, ha hangja szélsőségesen eltér az általunk elvárttól. A szerző szerint nem tudjuk megmondani, hogy egy alacsony kövér ember vagy egy magas sovány ember milyen hangját érintő sajátosságokkal rendelkezik. Ha azonban a látvány és az akusztikum ellentmond egymásnak, arra rendszerint felfigyelünk.

#### **4. A beszélő iskolázottsága, műveltsége**

A beszélő nyelvhasználata általában összefügg az iskolázottságával, műveltségével is. Bizonyára mindenki találkozott már olyan személyekkel, akik köztudottan nagy műveltséggel rendelkeztek, mindez pedig nyelvhasználatukban is érződött választékos kifejezőmódjukon és mondanivalójuk gondos megformáltságán keresztül. Nyelvhasználatunk nemcsak gondolkodásunkról, de műveltségünkéről is információkkal szolgál kommunikációs partnereink számára. Magában a nyelvben is megtalálhatóak azok a kifejezések, amelyekkel egyéneket minősítünk megnyilvánulásaik alapján, ugyanakkor szorosan összefonódnak az illető előze-

tes fejlődésével, tanultságával. Ilyenek például a *kifinomult*, *művelt*, *kulturált* fogalmak, vagy ezek ellentétéként a *bárdolatlan*, *műveletlen*, *kulturálatlan*, de gondolhatunk akár a *paraszt* szó pejoratív jelentésére is.

A fenti jelenség nem kizárólag a közvetített nyelvi viselkedésben vagy kódban, de a vele szoros összefüggésben lévő beszédben (és annak hangzásában) is megnyilvánulhat. Barra Mária (2010) kísérlete ennek megismerésére vállalkozott: azt elemezte, hogy a hallgatókban milyen benyomások alakulnak ki a beszélő hangja alapján az iskolázottságra vonatkozóan. Vizsgálatának kiindulópontja a szociálpszichológiából ismert holdudvar-hatás (Forgács 2007) volt. A jelenség lényege, hogy ha egy személy pozitív tulajdonságait észleljük, hajlamosak vagyunk az illetőhöz további pozitív tulajdonságokat társítani még abban az esetben is, ha ez egyébként nem lenne indokolt. A jelenséget pszichológiai kísérletek során igazolták, amelyek szerint a jó megjelenésű emberekhez olyan tulajdonságokat is hozzárendeltek a kísérletben részt vevők, amelyek nem következhetek az illető jó megjelenéséből. Ezek között szerepeltek a magasabb társadalmi státusz, magas iskolai végzettség, illetve kiemelkedő intelligenciaszint. Az is kimutatható volt a vizsgálatok során, hogy a vonzó külsejű egyének negatív tulajdonságait sokkal enyhébben ítélték meg a többiekénél (Forgács 2007).

Barra Mária szerint a beszéd is egyfajta külső megjelenés, amely képes lehet pozitív vagy negatív irányban befolyásolni azokat a hallgatókat, akik a beszélőt nem látják, csupán a hangját hallják. Tanulmányában egy hallgatók által kitöltött percepcióos teszt segítségével értékeltette különböző személyek felolvasását. Az értékelők feladata a beszélők iskolázottságának és munkahelyi pozíciójának megítélése volt.

A kutatáshoz szükséges hangfelvételeket öt személy közreműködésével készítették el, akik mindannyian ugyanazt az egy mondatból álló szöveget olvasták fel. A hangját kölcsönző öt személyt úgy választották ki, hogy beszédzavarral, hangzóképzési hibával ne rendelkezzenek, ugyanakkor ne legyenek képzett beszélők sem. További kiválasztási szempont volt, hogy az iskolai végzettségüket és a munkahelyi pozíciójukat tekintve különbözőek, valamint sorba állíthatóak legyenek. A felolvasók között elrejtettek egy „kivételt” is: egy olyan női beszélőt, aki az átlagnál jelentősen szebben beszélt (ugyanakkor ő sem volt képzett beszélő, munkahelyi pozícióját és iskolai végzettségét tekintve azonban messze elmaradt a csoport többi tagjától). Az ő kiválasztásával a szerző célja az volt, hogy kontrollszemélyként vizsgálva őt megfigyelhetővé és összehasonlíthatóvá váljon a szép beszéd hatásának erőssége a benyomások kialakulásában.

A percepcióos kísérletben szereplő adatközlőknek azonos értékelőlapon, fokozatmentes skálán kellett jelölniük, hogy az öt felolvasó személyt mennyire érzik iskolázottnak, illetve benyomásaikra támaszkodva milyen munkahelyi pozíciót társítanak hozzájuk. Az eredményeket ezt követően csoportonként feldolgozták és összesítették, majd statisztikai elemzést végeztek rajtuk. A hangmintákat to-

vábbi vizsgálatnak vetették alá, amelynek során három gyakorlott beszédoktató elemezte a hanganyagokat.

Az eredmények kiértékelésekor a kutatók azt találták, hogy nagyon nagy eltérések vannak a vizsgált mintában a beszélőkről kialakult benyomásokban, mind az életkort, mind pedig az iskolázottságot tekintve. Az iskolázottságra vonatkozó adatok azt mutatták, hogy a legtöbb pozitív értéket (a legmagasabb iskolai végzettségre vonatkozó véleményt) a középfokú végzettséggel rendelkező édesanya kapta (ő volt a „kivétel”), míg a legkevésbé iskolázottnak az egyik felsővezető férfit tartották. A hallás utáni vizsgálat mellett elemezték az egyes beszédprodukciók jellemzőit is. A hallás utáni vizsgálat során – az előzőekkel megegyezően – a kontrollszemély szerepét betöltő édesanya kapta a legjobb minősítést.

A bemutatott kutatási eredmények alátámasztották, hogy a megvizsgált szupraszegmentális elemek (beszédtempó, beszéddallam és hangerő) befolyásolták a válaszadók benyomásainak kialakulását az iskolázottság és a munkahelyi pozíció meghatározásában. A holdudvar-hatás a beszédben is érvényesült, hiszen a „szép” beszédű felolvasó értékelésében jelentős pozitív, míg a beszédprodukciójukat alacsonyabb szinten megvalósító beszélők esetében jelentős negatív torzítást tapasztaltak a hallás utáni megítélésben.

## **5. A beszélő foglalkozása**

Egy adott foglalkozás üzése nagyfokú ráhangolódást, gyakorlottságot eredményez a huzamosabb ideig azt végzők esetében. Ez nemcsak a munkafolyamatokhoz tartozó cselekvések és szakszavak alapos ismeretét jelenti, hanem kihathat a nyelvhasználatra, a beszédviselkedésre, alapvető emberi megnyilvánulásokra is. A műszaki területen dolgozók, különösen a mérnökök, programozók, gazdasági szakemberek vagy természettudományokkal foglalkozók esetében az egyértelmű, rövid és logikus közlések az irányadóak. Ezzel szemben a politikusok, ügyvédek, pedagógusok, írók és költők, színészek számára alapvető fontosságú a változatos kifejezésmód, az idézések, visszaulások, és nem egyszer hasznos eszköz a kétértelműség. Ebből adódóan az sem meglepő, ha egyes foglalkozások olyannyira kihatnak az adott személy beszédére, hogy az a hétköznapi megfigyelő számára is felfedezhető.

Azzal kapcsolatban, hogy kikövetkeztethető-e a beszélő foglalkozása a beszédből, Auszmann (2012) végzett vizsgálatokat. Míg a beszédtudományban eddig lezajlott vizsgálatok szinte kizárólag átlagos beszélők produkcióját vizsgálták, kevés tanulmány foglalkozott a színészek beszédével. A szerző ezen hiány pótlásának érdekében hasonlította össze színészek és hétköznapi beszélők beszédét. A kutatás egyik kiinduló hipotézise szerint a beszélőt nem látó, de hangját halló személyek képesek megkülönböztetni a színészeket az átlagos beszélők-

től. A további hipotézisek szerint a színészek magánhangzóinak formánsértékei nagyobb térben realizálódnak, azaz ejtésük pontosabb, kevésbé jellemző rájuk az „egymásba csúszó”, kevésbé variábilis magánhangzóejtés. A fentiekén túl a harmadik hipotézis szerint a színészek *s* és *sz* hangjainak ejtése is kevésbé variábilis a beszédet nem hivatásszerűen űző társaikénál.

A kísérlet alapjául szolgáló hangfelvételekhez 4 átlagos beszélő és 4 színész felolvasásait használták fel. A beszélők kiválasztásánál fontos szempont volt, hogy beszédhibával ne rendelkezzenek, valamint korban és nemben megfeleltethető párokat alkossanak egymással a színész – nem színész beszélők, ily módon összehasonlítható legyen az adott párok produkciója. A páronként megfeleltetett hangfelvételekből egy-egy mondatot használtak fel a percepciós vizsgálathoz.

A percepciós tesztben részt vevők feladata az volt, hogy a hangok páronkénti meghallgatása során válasszák ki, hogy a két hallott hangminta közül melyik a színész hangja, döntésüket pedig szövegesen is indokolniuk kellett.

A percepciós teszt igazolta az előzetes hipotéziseket: a színészek és az átlagos beszélők felolvasása szubjektíven, hallás alapján egyértelműen elkülöníthető. Vannak mérhető egyéni különbségek is, például a magánhangzók formánszerkezete a képzési jegyeiktől függetlenül jól elkülöníthető a két csoport esetén. A kutatás utat nyit a két csoport felolvasása közötti különbségek további vizsgálatára (légzéstechnika, artikuláció stb.), illetve tanulsággal szolgál mind a hivatásos, mind az átlagos beszélők beszédtechnika-oktatására nézve is.

## 6. A beszélő vonzó tulajdonságai

A beszélő hangja nemcsak a fentiekben említett külső és belső tulajdonságokról, de arról is információkat közvetít a hallgatónak, hogy az adott beszélő mennyire bizonyul vonzónak. A jelenséget a médiában lépten-nyomon tetten érhetjük: ilyen például a rádióban hallott vagy a már említett szinkronhangok esete. A vonzónak, kellemesnek hallott hanghoz sok esetben kellemes arcot társítunk, amely a valóságban messze állhat az általunk elképzelttől. Ezt használják fel a színészek vagy zenei előadók is, akik hangszínezetüket aszerint változtatják, hogy éppen dühösnek, határozottnak, vonzónak stb. szeretnék feltüntetni magukat. A hivatásos beszélők mellett a hétköznapi életben is éppúgy találhatjuk vonzónak vagy ellenszenvesnek egy adott személy hangját.

Ennek tudományos megközelítését Gocsál Ákos és Huszár Ágnes (2003) egy percepciós vizsgálat során kísérte meg. Kutatásukban a beszélőkről előre megalkotott tulajdonságok (pl.: nyugodtság, természetesség, okosság, szexisség stb.) becsléseit hasonlították össze az objektíven mérhető beszédjellemzőkkel. A kísérlet alapfeltevése az a tény volt, amely szerint a hallott beszéd akkor is emocionális reakciót vált ki a hallgatókból, ha nem látják a beszélőt. Gondoljunk csak a köztu-

datban is élő jelzőkre, mint a *bársonyos hang*, *meleg hang*, *kemény hang*, *rideg hang* stb. A kutatók szerint az emberekkel való találkozásaink összbenyomásokat alakítanak ki bennünk. Ezek a benyomások eleinte strukturálatlanok, később azonban ún. burkolt személyiségelméletté rétegződnek bennünk egy mentális sablon alapján. George Kelly elmélete (Forgács 2007) szerint kétpólusú konstruktumokat (mérceket) illesztünk az észlelt személyekre. Ilyen konstruktumok lehetnek például a „szeret – nem szeret”, „jópofa fickó – unalmas fráter”, „határozott – határozatlan” stb. A percepciók kísérlet módszere Kelly elméletét használja fel, ugyanis arra a kérdésre keresi a választ, hogy a beszéd mérhető fonetikai sajátosságai mennyiben felelősek az észlelt benyomásokért.

Ennek megfelelően a kutatók arra a kérdésre keresték a választ, hogy milyen személyiségvonásokat tulajdonítanak a hallgatók azoknak a beszélőknek, akiket nem látnak. A vizsgálatban női beszélők hangját ítélték meg férfi hallgatók.

A női beszélők egy érzelmileg semleges témáról beszéltek: egy általuk jól ismert város közterét mutatták be. Az elkészült hangmintákat ezt követően fiatal férfiaknak játszották le. Az ő feladatuk az volt, hogy az egyes felvételek meghallgatása után ítéletet alkossanak a beszélőt jellemző tulajdonságokról. A tesztlapokon hat tulajdonságpár szerepelt, az alábbi sorrendben: fiatal – nem fiatal, okos – nem okos, szexis – nem szexis, természetes – mesterkélt, nyugodt – nyugtalan, érthető – nehezen érthető. Az adott tulajdonságpárok értékelésére ötfokú skála szolgált, attól függően, hogy mennyire tartották jellemzőnek az adott személyre a kérdéses tulajdonságot.

A hangmintákat ezt követően akusztikai fonetikai vizsgálatnak vetették alá, amelyben beszélőnként megállapították a legfontosabb paraméterek átlagértékeit, így az alaphangmagasságot, a beszédtempót, az artikulációs tempót és az artikulációs határfokot (ez a jellel kitöltött részek aránya a beszédben). A méréseket követően a vélt tulajdonságok és a fonetikai paraméterek között kerestek összefüggéseket.

A kutatás során bizonyosságot nyert, hogy a beszélőkről a hallgatók különböző benyomásokat alakítanak ki, a verbális üzenettől függetlenül. A több szünettel, megfontoltabban beszélő személyeket inkább tartották érthetőbbnek és okosabbnak, mint a gyorsan beszélőket. A mélyebb hang az okosság benyomását keltette, az alaphangmagasság és a szexiség pedig szintén összefüggést mutatott egymással.

## 7. A beszélő személyisége

A beszéd egyedi produktum, amely alapján a fentiekben bemutatott számos tényezőre következtethetünk. Hordoz-e azonban a beszéd olyan információt, amelyből alapvető tulajdonságaink, viselkedésünk, mentalitásunk valóban

visszakövetkeztethető? A beszédben megjelenő lehetséges személyiségvonásaink nemcsak hitelességünkre, egyéniségünkre engedhetnek következtetni, de a bűnügyi fonetika számára is segítséget nyújthat egy beszélő profiljának elkészítésekor. Kérdésként merül fel továbbá az is, hogy a beszélőről kialakuló benyomások mennyire jellemzik őt valójában, illetve mennyire félrevezetőek? A fenti témában Gocsál (2010) végzett kutatásokat, amelyekben a személyiségvonások és a beszédakusztikai paraméterek, az akusztikai paraméterek és a benyomások, valamint a személyiségjegyek és a benyomások lehetséges összefüggéseit vizsgálta. Az eredmények szerint a szünettartások összefüggést mutattak a beszélőnek tulajdonított barátságosság mértékével. A férfi és a női adatközlők megítélésében nem volt különbség adatolható. További kérdéseket vizsgált Németh (2013a), valamint Németh és Nagy (2015) a beszélő hangjának és személyiségének lehetséges összefüggéseire vonatkozó kísérleteiben. A vizsgálatok kiindulópontja az volt, hogy összefügg-e az egyén beszéde azzal, hogy milyen személyiséggel rendelkezik. Ehhez a személyiség mérhető adatait és az ennek megfelelő akusztikai fonetikai paramétereket hasonlították össze. A két fenti kutatás közül a korábbi egy iskolai helyzetben vizsgálta meg a diákok véleményét a tanárok személyiségével kapcsolatban, a későbbi pedig különböző életkori csoportok véleményét hasonlította össze a vizsgált egyének személyiségéről.

A mindkét kísérlet során felhasznált személyiségmodell a *Big Five* volt. A modell a vizsgált személyiséget öt faktorra osztja, és a faktorok (tengelyek) mindegyikén elfoglalt értékkel jellemzi. Így nem egy terminussal, hanem öt jellemzővel írja le az adott egyén természetét. A faktorok az alábbiak: extravertió, barátságosság (vagy kellemesség), lelkiismeretesség, emocionalitás (más néven neuroticizmus) és intellektus. Az öt fő faktort több hasonló, egymástól kisebb mértékben eltérő elmélet is használja más-más alárendelt tulajdonságlistákkal. A tulajdonságlisták közül a Goldberg-féle változat rövidített formája szolgált a kísérlet alapjául, amelynek előnye, hogy olyan hétköznapi tulajdonságokat használ, amelyek nem csak pszichológusok számára értelmezhetőek. A tulajdonságlista rövidítését a percepciós tesztet kitöltő adatközlők mentális igénybevétele indokolta, mivel több beszélő hangját is értékelniük kellett.

A tanári beszédet elemző kísérletben (Németh 2013a) hat tanár szakos hallgatót kértek fel arra, hogy tartsanak rövid előadást, tanítási helyzetet imitálva. A felvételek elejéből készült, azonos hosszúságú részleteket egy percepciós kísérlethez használták fel.

A percepciós vizsgálat (megfigyelői értékelés) egy budapesti általános iskola hetedikes évfolyamának két osztályában valósult meg. A diákoknak egy teszt-lapon kellett jelölniük, hogy milyen tulajdonságokat tartanak érvényesnek a beszélőre. Ezzel párhuzamosan a diákokéval megegyező tulajdonságlista alapján minden tanárjelölt értékelte önmagát, illetve két-két, a tanárjelöltet jól ismerő barát is véleményezte őt. A tesztet a jelölteknek saját magukra vonatkozóan (ön-

jellemzés), a barátoknak pedig (megfigyelői értékelés) a tanárhallgatók személyes ismerete alapján, de a hanganyagok meghallgatása nélkül kellett kitölteniük. Mindegyik adatközlői csoport ugyanazt a tesztlapot töltötte ki.

Az értékeléseket a beszéd objektíven megmért paramétereivel is összevetették. A statisztikai vizsgálatok során több esetben szignifikáns összefüggések mutatkoztak nem csak az egyes értékelők (önértékelés, barátok, diákok) ítéletei között, de az ítéletek és egyes személyiségtulajdonságok között is. A beszéd gyorsabb tempóértékei a hallgatókban egyértelműen a magasabb extraverzió és intellektus benyomását keltették. Emellett az összes jellemzés összehasonlítása és több faktorbeli egyezése nyomán igazolódott, hogy a barátságosság, a lelkiismeretesség és az intellektus dimenziója is – helyenként változó eredménnyel, de – viszonylagos biztonsággal helyesen megítélhető pusztán a beszélő hangja alapján.

A második kutatás (Németh–Nagy 2015) során a kutatók azt vizsgálták, hogy másként ítélik-e meg a különböző életkorú hallgatók hallás után a beszélő személyiségét. A percepciók kísérlet anyagai és módszere megegyezett a megelőző kutatás során használtakkal, a megítéléseket végző adatközlőket azonban életkoruk szerint négy csoportra bontották (12–13 évesek, 19–22 évesek, 23–29 évesek, valamint a 30–65 évesek csoportja).

A statisztikai vizsgálatok során az ítéletek átlagait csoportonként összehasonlították egymással, a beszélők saját magukról alkotott jellemzésével, valamint a beszélőt egy jól ismerő barát értékeléseivel is. A kutatók azokra a kérdésekre keresték a választ, hogy mennyire tudják pontosan meghatározni az adott csoportok a valós személyiséget (amely az önjellemezésből és a közeli barát jellemzéséből rajzolódik ki), valamint vannak-e különbségek a személyiségpercepcióban az eltérő életkorokban. Jelentkeznek-e továbbá összefüggések a mérhető beszédjellemezőkkel?

Az eredmények szerint a beszéd egyes mérhető paraméterei valóban mutatnak összefüggést a személyiség egyes jellemzőire (főként az intellektusra) vonatkozóan. A ítéletek összevetése során nem csak a csoportok véleménye egyezett, de az önjellemezésekkel és a baráti ítéletekkel is szignifikáns kapcsolat mutatkozott. A többi faktor esetén kisebb különbségek mutatkoztak, amelyek szerint a csoportok véleménye nem különül el következetesen egymástól, így nem nyert bizonyítást, hogy a személyiségpercepció az életkor előrehaladtával változik.

## **8. A beszélő állapotára vonatkozó információk**

Szintén hétköznapi tapasztalat, hogy a beszélő állapotával kapcsolatban is vonhatók le következtetések a hangja alapján. Másként beszél egy szomorú, letört személy, és másként egy vidám. Scherer (1995) bebizonyította, hogy a

beszélő érzelmi állapota a fonetikai jegyekben is megmutatkozik. A kutatás szerint öröm esetén az alaphang frekvenciájának átlagértéke, variabilitása, energiája növekszik, tartománya pedig szélesebb lesz. Szomorúság esetén ennek fordítottja következik be, valamint az alaphangfrekvencia kontúrjai lefelé görbülnek, azaz a dallam ereszkedő lesz. Egy hazai kutatásban (Földi 1996) is hasonló következtetésre jutottak: a boldog, örömteli megnyilatkozások frekvenciatartománya lényegesen szélesebb, intenzitása erőteljesebb. Ugyanakkor a hallgatók számára nem könnyű a beszéd érzelmi töltetének az azonosítása; a pontos felismerés az érzelemtől függően 15–70% (Lakatos 2015).

A beszélők hangja egyéb fizikai jellemzők mentén is differenciálódhat. Egy kísérletben azt vizsgálták (Braun–Rietveld 1995), hogy a dohányzó és a nem dohányzó beszélők hangjában észlelhető-e különbség. Húsz huzamosabb ideje dohányzó és húsz nem dohányzó személy hangját vették fel, majd játszották le 31 hallgatónak. A vizsgálat során az azonos életkorú dohányzók és nem dohányzók közül a kísérletben részt vevők minden esetben a dohányzó személyeket vélték idősebbnek. A beszédjellemzők közül a szótagszám és a HNR (harmonikus/zaj arány, amely a zöngé tisztségének egyik mérőszáma) mutatott összefüggést a beszélő életkorával.

Az alkohol beszédre gyakorolt hatása a beszéd töredezettségét, a hiba típusú jelenségek számának növekedését és a beszéd gyorsulását hozza magával. A témával a hazai szakirodalomban elsőként Gyarmathy (2007) foglalkozott; ebben a kötetben pedig Németh Szilvia (2016) tanulmányában olvashatunk részletesen az „alkoholos beszédéről” szóló kutatásokról.

## 9. Összefoglalás

A fentiekben bemutatott kutatások haszna nemcsak a beszéd és a beszélő kapcsolatának jobb megértésében, de a beszédprodukción és a beszédpercepción működésének, valamint lehetséges tévedéseinek további tapasztalataival is segíti a beszéd tudományt. Emellett több terület számára gyakorlati haszonnal szolgálnak az itt megismert, valamint a területen zajló további kutatások. Ezek egyik legfontosabb, alapvető felhasználója a retorika tudománya, ahol a szónokok közvetlen érdeke, hogy a hallgatóságuk szemében a lehető legegységesebb színben jelenhessenek meg. Nem mellékes azonban a nyelvészeti hangszakértők ide kapcsolódó tevékenysége sem, akik beszélőazonosítással, beszélőprofil-építéssel, valamint a hanganyagok bizonyítékként való minősítésével, ellenőrzésével foglalkoznak.

## Irodalom

- Auszmann Anita 2012. Máshogy beszélnek a színészek, mint az átlagos beszélők? In Váradi Tamás (szerk.): *VI. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest, 2012. 02. 03. MTA Nyelvtudományi Intézet, Budapest, 3–13. <http://www.nyud.hu/alknyelvdok12/proceedings12/auszmann2012.pdf>
- Balázs Boglárka – Bóna Judit 2016. Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban. In Bóna Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest, 7–18. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu).
- Barra Mária 2010. A felolvasás „mellékhatásai”: Egy attitűdvizsgálat. *Alkalmazott Nyelvtudomány* X/1–2. 82–95.
- Bernstein, Basil 1975. Nyelvi szocializáció és oktathatóság. In Pap Mária – Szépe György (szerk.): *Társadalom és nyelv*. Gondolat Kiadó, Budapest, 393–435.
- Bóna Judit 2009. Az idős életkor tükröződése a magánhangzók ejtésében. *Beszédkutató 2009*. 76–87.
- Bóna Judit 2013. *Az időskori beszéd fonetikai sajátosságai*. ELTE Eötvös Kiadó, Budapest.
- Braun, Angelika – Rietveld, Toni 1995. The influence of smoking habits on perceived age. In: *Proceedings of the 13<sup>th</sup> International Congress of Phonetic Sciences*. Vol. 2. Stockholm, 294–297.
- Carver, Charles S. – Schreier, Michael F. 2006. *Személyiségpszichológia*. Osiris Kiadó, Budapest, 81–86.
- Forgács József 2007. *A társas érintkezés pszichológiája*. Kairosz Kiadó, Budapest.
- Földi Éva 1996. Az érzelemkifejezés szupraszegmentális formái és percepciója. *Egyetemi Fonetikai Füzetek* 20.
- Gocsál Ákos 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutató '98*. 122–134.
- Gocsál Ákos – Huszár Ágnes 2003. Csábító hangok. *Beszédkutató 2003*. 9–18.
- Gocsál Ákos 2010. *A beszédakusztikai paraméterek és a beszélő személyiségjegyei közötti összefüggések vizsgálata*. Doktori disszertáció. ELTE, Budapest.
- Gósy Mária – Nikléczy Péter 2000. Az idő változásának és a beszéd állandóságának paradoxona. *Beszédkutató 2000*. 132–143.
- Gósy Mária – Krepesz Valéria 2014. *Az artikulációs mintázat az idő és a beszédszervek működésének függvényében*. Előadás a XVI. Pszicholingvisztikai Nyári Egyetemen. Balatonalmádi, 2014. május 25–29.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukción. *Beszédkutató 2007*. 108–121.
- Lakatos Boglárka 2015. Másképp észlelik-e az érzelmeket a nők és a férfiak? *Alkalmazott Nyelvtudomány* XV. 115–126.
- Nagy János 2007. Faktoranalitikus személyiség-taxonómiák. In Gyöngyösiné Kiss Enikő – Oláh Attila (szerk.): *Vázlatok a személyiségről*. Új Mandátum Könyvkiadó, Budapest, 110–145.
- Németh Máttyás Ferenc 2013a. A tanári beszéd és személyiség. *Anyanyelv-pedagógia* 2013/1. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=434> (Megjelentve: 2015.02.08)
- Németh Máttyás Ferenc 2013b. *A beszédprodukciónak a variabilitása beszélőn belül és beszélők között*. Előadás a XV. Pszicholingvisztikai Nyári Egyetemen. Balatonalmádi, 2013. május 26–30.

Németh Mátyás Ferenc – Nagy János 2015. Hogyan ítélik meg különböző életkorú hallgatók a beszélő személyiségét? *Alkalmazott Nyelvtudomány XV.* 127–147.

Németh Szilvia 2016. Az alkohol hatása a beszédre. In Bóna Judit (szerk.): *Fonetikai olvasókönyv.* ELTE Fonetikai Tanszék, Budapest, 145–157. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu).

Scherer, Klaus R. 1995. How emotion is expressed in speech and singing. *Proceedings of the XIIIth International Congress of Phonetic Sciences.* Vol. 3. Stockholm, 90–96.

### **Kérdések, feladatok**

1. Milyen tulajdonságokra következtethetünk tévesen a hallott beszédből?
2. Mennyire megbízható az életkor becslése?
3. Mit jelent a beszédben megnyilvánuló holdudvar-hatás?

## Az alkohol hatása a beszédre

Hétköznapi megfigyelések alapján egyértelműnek tűnik, hogy az alkohol hatással van a beszédprodukcóra. A célzott kutatások alapján a beszéd több paramétere valóban informál az ittasságról, azonban a kontrollált körülmények között végzett vizsgálatok azt mutatják, hogy egy ismeretlen személyről csupán a hangja alapján igen nehéz eldönteni, hogy alkoholos befolyásoltág alatt áll-e vagy sem. A következőkben áttekintjük, hogy milyen gyakorlati haszonnal jár a jelenség kutatása, miért hallatszódhat a beszéden, ha valaki alkoholos befolyásoltág alatt áll, hogyan vizsgálható a kérdés tudományos igénnyel, és hogy pontosan mit tudunk eddig az alkohol kiváltotta beszédbeli változásokról.

### 1. Az alkoholos beszéd kutatásának céljai

Az alkohol beszédre gyakorolt hatása mintegy száz éve folyamatosan aktív kutatási téma. Kezdetben (1915–1964) orvosokat és pszichológusokat foglalkoztatott a kérdés: az alkoholos befolyásoltág jelei között a beszédbeli változásokat is vizsgálták. Az ekkor végzett kutatások az alkohol élettani hatásainak feltárását segítették. A második korszakban (1966–1982) mindinkább előtérbe kerültek a szociális és a nyelvészeti szempontok: a szerzők elsősorban a beszéd mennyiségi megváltozására (beszédesség), a beszéd gátlásainak oldódására, a megakadásjelenségekre és a mondattani változatosságra fókuszáltak, de megjelentek már az első fonetikai tanulmányok is (Chin–Pisoni 1997). A harmadik korszakban (1985-től) a digitális jelfeldolgozás fejlődésének köszönhetően az akusztikai fonetika vált az alkoholos beszéd vizsgálatának fő eszközévé, a kutatások száma pedig robbanásszerűen megnőtt (Chin–Pisoni 1997). Az elsődleges cél az alkoholos befolyásoltág hangalapú detektálása lett, az 1990-es években az amerikai Bloomington és a németországi Wiesbaden vált a vizsgálatok fő központjává. Az amerikai kutatócsoport az igazságügyi felhasználásra kezdett koncentrálni, vagyis annak kimutatására, hogy egy cselekmény elkövetésekor a tettes ittas volt-e vagy nem. A másik kutatócsoport ezzel szemben a közlekedésbiztonságra fókuszált: egy olyan fedélzeti számítógépbe illeszthető szoftver létrehozását tűzték ki célul, mely az elhangzó parancsok alapján azonosítja, ha a vezető ittas, és ilyenkor blokkolja a járművet (Chin–Pisoni 1997). 2010 óta a vizsgálatok többsége az automatikus detektálást igyekszik fejleszteni, az elsődleges központ pedig München lett.

A német és az amerikai angol nyelvterületen kívül csak szórványosan jelennek (és jelentek) meg alkoholos beszéddel foglalkozó tanulmányok. Magyarországon 2007-ben (Gyarmathy 2007) készült az első vizsgálat, és bár hazánkban a fő kutatási irány a pszicholingvisztika, minden kutatásban találni akusztikai fonetikai paramétereket is, mint például az artikulációs tempó (Gyarmathy 2009; Szőnyegi 2013; Tisljár-Szabó et al. 2013, Tisljár-Szabó 2014), a beszédtempó (Gyarmathy 2009; Szőnyegi 2013; Tisljár-Szabó et al. 2013, Tisljár-Szabó 2014) vagy az alaphangmagasság (Gyarmathy 2009; Tisljár-Szabó et al. 2013, Tisljár-Szabó 2014). A magyarra vonatkozó vizsgálatok elősegítése érdekében 2013-tól folyamatban van egy magyar alkoholos beszédkorpusz kialakítása (l. Németh 2014; Lados–Németh 2014) is.

## 2. Az alkoholos beszéd biológiai háttere

A köznapi nyelvhasználatban az *alkohol* szó alatt az etanolt (etil-alkoholt) értjük, egy ital etanolkoncentrációját térfogatszázalékban szokás megadni: a sör általában 5%, a bor 11%, a vodka pedig 40% körüli értékkel bír.

Az elfogyasztott etanol egy többlépcsős folyamat végén távozik a szervezetből: a bevitel után már a szájban elkezdi felszívódni, de a folyamat nagyrészt a gyomorban és főleg a vékonybél első két traktusában megy végbe (Chin–Pisoni 1997). Ennek gyorsaságát igen sok tényező befolyásolja (pl. a gyomortartalom, a fizikai aktivitás, a testhőmérséklet – vö. Agarwal–Goedde 1990), ezzel magyarázható az, hogy ugyanaz a személy különböző mértékben lehet részeg két különböző alkalommal, akkor is, ha ugyanannyi és ugyanolyan italt fogyasztott. A felszívódást követően az etanol nagy része a vér útján teritődik, így gyorsan eljut többek között az agyba és a tüdőbe. Az agyba érve a központi idegrendszerre általános nyugtató hatást fejt ki, ennek következtében lassul az idegsejtek működése, fokozatosan gyengülnek az érzékek, és koordinálatlanná válik a mozgás. A folyamat utolsó lépéseként az etanol nagy része enzimek segítségével széndioxidra és vízre bomlik, a maradék pedig vizelettel, izzadsággal, nyállal és kilélegzett levegővel távozik a szervezetből (Agarwal–Goedde 1990).

Az etanol szervezetre gyakorolt hatása két külön tényezőre bontható: az egyik az objektív módszerekkel mérhető véralkoholszint (BAC, vagyis blood alcohol content), ami azt mutatja, hogy egy liter vérnek hány ezreléke etanol. A másik pedig az egyelőre egzakt mérőszámmal nem rendelkező befolyásoltság, ami az etanol kiváltotta válaszreakciók (pl. kipirulás, koordinációzavar, emelkedett hangulat) összességét jelenti. A véralkoholszint vérvétellel állapítható meg a legpontosabban, a gyakorlatban azonban gyakran használják a kilélegzett levegő alapján becsült értéket is (BrAC, vagyis breath alcohol content). Utóbbi azért informál a véralkoholról, mert a tüdőt behálózó érrendszeren keresztül az oxi-

génnel együtt az alkohol párája is átjut a légzsákocskák levegőjébe. Ezt kilélegezve a levegő etanolkoncentrációja alapján becsülhető a vér alkoholtartalma is (Chin–Pisoni 1997). A BrAC és a BAC átváltása azonban nem 100%-osan pontos: az egyéni különbségek és az alkoholszonda minősége is meghatározó, emiatt több országban a jogi gyakorlat csak a vérvételen alapuló véralkohol-meghatározással számol.

A befolyásoltság terén jelentős, elsősorban genetikai okokra visszavezethető varianciát találni, vagyis egyénenként változik, hogy valaki „mennyire bírja” az alkoholos italokat. Vannak magasabb és alacsonyabb toleranciájú személyek is, előbbiek azonos véralkoholszint mellett enyhébb válaszreakciókat mutatnak, mint mások (Agarwal–Goedde 1990). Ezen különbségek miatt a véralkoholszint és a befolyásoltság megfeleltetése csak közelítő érvényű lehet. Kutatások sora (l. Moskowitz–Robinson 1988) mutatja, hogy már alacsony BAC hatására (0,5% alatt) is romolhat a reakcióidő, a finommotoros mozgások, a koordináció és a megosztott figyelmi képesség. Fokozódó alkoholkoncentráció hatására (de még 1% alatt) zavar léphet fel a pszichomotoros képességekben, az információfeldolgozásban és a koncentrált figyelem terén is. Ahogy fokozatosan növekszik a véralkoholszint, véget ér az eufória, majd a nyugtalansági fázis, és lassan bekövetkezik az alkoholemérgezés. Alacsony tolerancia esetén már a 3% is eszméletvesztéssel járhat, halálos szintnek általában az 5%-et tekintik.

Az a feltételezés, hogy az etanol hatással van a beszédre, egyrészt általános megfigyeléseken alapszik, másrészt élettanilag is feltételezhető, hiszen a nyelvi folyamat minden mozzanatát az idegrendszer irányítja. A beszéd önmagában is rendkívül komplex művelet egyrészt azért, mert a hangképző szervek egyikének sem a hangadás az elsődleges funkciója, másrészt azért, mert a pontos artikulációhoz igen precíz mozdulatok szükségesek. A különböző magánhangzók kiejtésakor a nyelv helyzete gyakran mindössze egy milliméterrel tolódik el. Egy hang kimondásakor körülbelül 100 izom mozdul meg, és minden izom működésében körülbelül 100 motorikus egység (egy neuron és egy izomrost) vesz részt. Ha egy másodperc alatt 14 hangot ejtünk ki, akkor ez azt jelenti, hogy egy másodperc alatt körülbelül 140 000 neuromuszkuláris információcsere történik. Ezeket a folyamatokat, vagyis a beszédképző szervek motorikus mozgását a bal agyfélteke harmadik homloki agytekervénye és a Brodmann 44-es terület koordinálja, így lehetséges az, hogy az alkohol homloklebe nyre irányuló, gátló hatása miatt változások keletkeznek az artikulációban. A kérdés csupán az, hogy ezek a változások rendszerszerűek-e: köthetőek-e bizonyos véralkohol- vagy befolyásoltsági szinthez, vagy túl nagy az egyének közötti variancia.

### 3. Az alkoholos beszéd kutatásának módszertana

Az alkoholos beszéd vizsgálatának legfőbb nehézségét a megfelelő hanganyagok rögzítése jelenti. Nem csupán azért, mert a kísérleti protokoll hosszú és költséges, hanem mert különleges biztonsági intézkedéssel és etikai vonatkozásokkal jár. Az alábbiakban egyenként áttekintjük az alkoholos beszéd vizsgálatát célzó kísérletek fontosabb részleteit, majd összefoglaljuk a magyar alkoholos beszéd-korpusz protokollját. Ezek alapján nem csupán a téma kutatásának módszertana, hanem az is kirajzolódik, hogy a témát érintő számos kutatás mely módszertani pontokban különbözik egymástól, vagyis milyen kísérleti helyzetek alapján vizsgálták már az alkohol beszédre gyakorolt hatását.

#### 3.1. A kísérleti személyek kiválasztása

Az alkoholos beszéddel foglalkozó kutatások során az az általános, hogy ugyanannak a beszélőnek rögzítik józan és ittas beszédét is, de egyes kutatók józan kontrollesoporttal dolgoztak (Collins 1980). Még napjainkban is fölülreprezentáltak a fiatal férfi adatközlők, sőt egészen az 1960-as évekig női adatközlők bevonása fel sem merült (Chin–Pisoni 1997). Idővel a kutatásokban részt vevő beszélők száma folyamatosan nőtt, az ezredforduló után pedig megjelent az igény a nagyszámú, nemben, korban és nyelvjárásban kiegyensúlyozott beszéd-korpuszokra. Az eddigi legnagyobb gyűjteményt 2007 és 2009 között rögzítették német nyelvterületen (Schiel et al. 2011).

A vizsgálatok során a hagyományos metaadatok (nem, kor, születési hely stb.) mellett általában szerepel olyan kérdőív is, amely segítségével az adatközlő elhelyezhető egy alkoholizmust mérő skálán. Ennek rögzítése azért fontos, mert a tartósan nagy mennyiségű és az alkalmi alkoholfogyasztás eltérő hatásokat mutathat a beszédben.

#### 3.2. Anyag

A kutatások többsége utánmondást vizsgál (pl. Pisoni et al. 1985) vagy felolvasást szó- (pl. Lester–Skousen 1974), mondat- (pl. Pisoni et al. 1985) vagy szövegszinten (pl. Klingholz et al. 1988), de megnevezéssel (pl. Moskowitz–Roth 1971), félspontán (pl. Künzl et al. 1992) és spontán beszéddel (pl. Fontan et al. 1978) is foglalkoztak már.

#### 3.3. Az interjúk protokoll

Az alkohol beszédre gyakorolt hatásaira irányuló kísérlet tervezésekor először azt kell meghatározni, hogy a józan és ittas állapotban készült felvételek azonos

alkalommal készüljenek-e vagy sem. A kutatások egy részében (pl. Schiel et al. 2011; Tisljár-Szabó et al. 2013) előbb az ittas, majd (általában több hét múlva) a józan hanganyag készült el, így kompenzálták, hogy az interjúk során azonos anyagot kellett felolvasniuk a beszélőknek. Több esetben azonban egy időben készült az összes interjú, így az adatközlő beszédének egyes paraméterei (pl. amelyeket befolyásolhat egy enyhe megfázás vagy a hangulat) biztosan nem/kimondottan az alkohol hatására változtak.

A következő kérdés az alkohol típusának és bevitelének kiválasztása. Ismeretes, hogy az etanol felszívódása gyorsabban végbemegy magasabb koncentrációjú, égetett szeszesitalok esetén, mint az alacsonyabb alkoholtartalmú italok fogyasztásakor (Agarwal–Goedde 1990), így az összehasonlíthatóság miatt már a kísérlet tervezésekor meg kell választani a bevinni kívánt alkoholos italt. Általánosnak számít az az eljárás, hogy az adatközlő meghatározott idő alatt (jellemzően 15 perc) fogyasztja el az italt, ezt követően körülbelül 20 perc várakozás után kezdődik az interjú szakasz. Amennyiben korábban kezdődik az interjú, a véralkoholszint a befejezetlen felszívódás miatt nem lesz állandó a felvétel során.

Több szempontból is fontos már a kutatástervezéskor meghatározni az elvárt véralkoholszintet. Az alkoholos beszéd kutatásának első korszakaiban vizsgáltak igen ittas (pl. detoxikálóban lévő) beszélőket (Fontan et al. 1978), de napjainkban a kutatók többsége 1‰ körüli értékkel dolgozik. Egyrészt mert a komoly mértékű ittaság esetén túlzottan erős változásokat lehet megfigyelni, így gyakorlati szempontból nagyobb jelentősége van az 1‰-es BAC-nak. Ez már hatással van a gépjárművezetői képességekre, de még nem számít súlyos részegségnek. Másrészt az alacsonyabb véralkoholszint biztonsági okokból is kívánatos: az 1‰-es szint még az alacsony toleranciájú személyeknél sem kockázatos. A kutatások többségében a józan és az ittas állapot között egy enyhén ittas (kb. BAC/BrAC: 0,5‰) fázisban is készül felvétel. Ilyenkor az első véralkoholszintméréskor kiderül, ha az adatközlő toleranciája alacsonyabb, és így a kísérletvezető változtathat a második dózison, és nem adagolhatja túl az etanolt. Harmadrészt számolni kell azzal is, hogy az alacsonyabb véralkoholszinttel rendelkező kísérleti személyek közreműködőbbek, mint a magasabb BAC-al rendelkezők. Bár az egyéni tolerancia miatt nem lehet pontosan megállapítani, hogy az adatközlőnek hány gramm etanolt kell meginnia a kívánt véralkoholszint eléréséhez, közelítő becslések végezhetők. A kísérletek többségében két orvosi képletet, a Widmark- (1932) és a Watson-formulát (1980) alkalmazták, ehhez az alkoholos ital etanolkoncentrációjára, az adatközlő nemére, életkorára, magasságára és súlyára van szükség. Egy 25 éves, 160 cm magas, 50 kg-os nőnek kb. 1,5 dl, míg egy 25 éves, 185 cm magas, 80 kg-os férfinak 2,5 dl 37,5%-os vodkát kell meginnia az 1‰ eléréséhez.

A kísérletek nagy részében levegős szondával történt a véralkoholszint ellenőrzése, mivel egy esetleges vérvétel orvosi segítséget igényel, valamint feszültebbé teheti az adatközlőt.

Végezetül, az alkohol hatásaira irányuló kísérletek során folyamatosan szem előtt kell tartani az etikai vonatkozásokat. Ezek közül a biztonsági szempontok a legfontosabbak. Orvosi közreműködés nélkül csak enyhe ittasság idézhető elő, az is csupán akkor, ha az adatközlőre nézve nem veszélyes az alkoholfogyasztás (pl. nem szed speciális gyógyszereket). Amellett, hogy a beszélő bármikor megszakíthatja a kísérletet, a kísérletvezető felelőssége, hogy a véralkoholszint mérése során felismerje, hogy az adatközlő a vártnál alacsonyabb alkoholtoleranciájú, és félbehagyja a kísérletet, vagy kisebb mennyiségű alkoholt adagoljon. További fontos szempont az adatkezelés. Ilyen például, hogy a beleegyező nyilatkozat aláírásának és a beszélőt érintő személyes adatok felvételének az alkoholfogyasztás előtt kell megtörténnie, valamint a felvételek során (amennyiben spontán beszéd is rögzítésre kerül) kerülni kell a személyes témákat. Az adatközlőben tudatosítani kell, hogy az elkészült felvételek felhasználhatóságához adott engedélyét bármikor visszavonhatja, és pontosan tájékoztatni kell a hanganyag tárolási és felhasználási módjáról. Kiemelt figyelmet kell fordítani a felvételek anonimizálására is.

### **3.4. A magyar alkoholos beszédkorpusz felvételi protokollja**

A magyar alkoholos beszédkorpusz jelenleg 51 adatközlőt számlál (32 férfi, 19 nő). A teljes interjú protokoll 4 órát vesz igénybe, egy alkalommal két beszélővel készülnek felvételek józan, 0,5‰-es és 1‰-es állapotban. A felvételek egyenként 15-20 percesek, felolvasást és spontán beszédet is tartalmaznak, így beszélőnként összesen 45-50 percnyi hanganyagot rögzítünk.

Az adatközlők a beleegyező nyilatkozat aláírása után kitöltenek egy háttérinformációkat (l. Németh 2014) tisztázó kérdőívet, ennek része a Michigan Alcohol Screening Test (Selzer 1975) lefordított változata. Ezt követően meghatározzuk a bevinni kívánt alkohol mennyiségét a Widmark- (1932) és a Watson-formula (1980) segítségével: az önkéntes adatközlők vodka, fehér rum és pálinka közül választhatnak. Az önkéntesekkel először józanul készül felvétel, ezt követően három olyan tesztet végeznek, melyek a reakcióidejüket mérik. Ezeket mindkét alkoholos állapotban megismétlik: arra szolgálnak, hogy az alkoholos befolyásoltságot objektív mérőszámokhoz kössük. Ezt követően az adatközlők elfogyasztják a meghatározott alkoholmennyiség felét (egymáshoz képest 20 perces eltéréssel kezdik meg a bevitelt, így egymást váltják a felvételi helyiségben). A 15 perces felszívódási szakasz után következik a második felvétel, melynek megkezdése előtt, közepén és végén alkoholszondával ellenőrizzük a BrAC-ot. Ezt követően az adatközlő elfogyasztja az alkohol második felét

(akinél a fél adag hatására magasabb BrAC jelentkezett, 10-20%-kal csökkentjük a második adagot), és a felszívódási szakaszt követően elkészítjük a harmadik felvételt is, rögzítjük a BrAC-ot és a reakcióidőt.

#### 4. Mit tudunk eddig az alkohol beszédre gyakorolt hatásairól?

##### 4.1. Az alkoholos beszéd észlelése (percepciója)

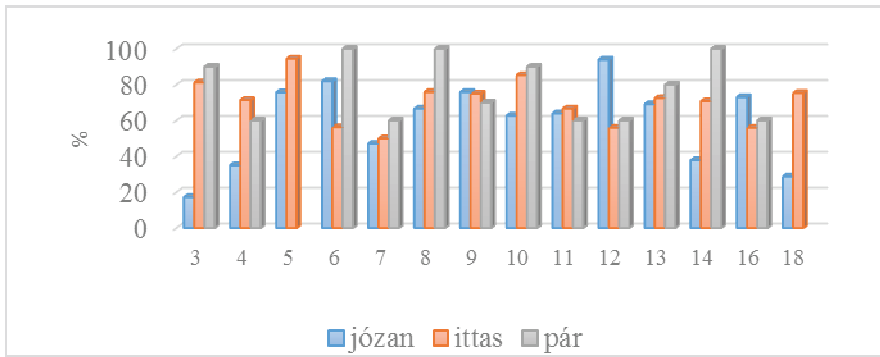
Az általános megfigyelések alapján azt várnánk, hogy az alkoholos beszédet könnyű felismerni. Az eddigi percepciósi (észlelést mérő) tesztek alapján azonban azt látjuk, hogy korántsem egyszerű feladat. Azokban a kísérletekben, amelyekben józan és ittas beszélők hangmintáit hallgatták meg olyanok, akik tudtak róla, hogy a hallott minták között van ittas és józan személyektől származó is, a helyes azonosítások aránya vagy megegyezett a tippelésével (Klingholz et al. 1988), vagy csak kicsivel haladta meg azt (Lados–Németh 2014). Ahogy nőtt a beszélő véralkoholszintje, úgy javult a helyes azonosítások aránya is: 2‰ fölötti BrAC-nál már 80% körüli volt a helyes ítélet (Künzel et al. 1992). Jelentősen javult a felismerés akkor is, amikor a hallgatónak olyan párba állított minták alapján kellett dönteniük, amelyek ugyanazzal a beszélővel józan és ittas állapotban készültek. Ilyenkor a sikeres ítéletek aránya statisztikailag igazolhatóan jobb volt, mint a tippelésé (Klingholz et al. 1988; Künzel et al. 1988; Pisoni et al. 1985; Pisoni–Martin 1989; Schiel 2011; Lados–Németh 2014). Az 1. táblázat összefoglalja a korábbi kutatások eredményeit. A megakadások és a téma nagyban befolyásolta a hallgatók ítéleteit (Klingholz et al. 1988; Lados–Németh 2014), valamint a hallgatók a józan állapotban készült hanganyagokat gyakrabban vélték ittasnak, mint fordítva.

1. táblázat. Az alkoholos beszéd sikeres azonosításának aránya kutatásonként

	Megkülönböztetés			Azonosítás		
	Olvasott	Utánmondás	Spontán	Olvasott	Utánmondás	Spontán
Klingholz et al. (1988)	61,1%			54,0%		
Künzel et al. (1992)				66,8%		
Pisoni et al. (1985)		82,4%				
Pisoni–Martin (1989)					64,7%	
Schiel (2011)	75,7%		68,1%			
Lados–Németh (2014)	77,5%		83,8%	64,2%		63,7%

Amennyiben a hallgatók jelölhették, hogy mennyire biztosak a döntésükben, az ítéletük annál nagyobb arányban bizonyult helyesnek, minél biztosabbak voltak benne (Pisoni et al. 1985; Lados–Németh 2014). Ezen kívül mind a fonetikusok (Pisoni et al. 1985), mind a hivatásos rendőrjárőrök (Pisoni–Martin 1989) jobban teljesítettek a vizsgálatok során, mint a laikus hallgatók.

Ha az egyes beszélőkre vonatkozó helyes azonosítási arányt nézzük, látszik, hogy voltak olyanok, akiken jobban, és olyanok, akiken kevésbé lehetett felismerni, hogy fogyasztottak alkoholt (1. ábra). Elképzelhető, hogy ez a felvétel során elért befolyásoltsági szinttel van összefüggésben.



1. ábra. Az egyes beszélők józan, ittas és párban lehallgatott hangmintáinak helyes azonosítási aránya (%) (Lados–Németh 2014)

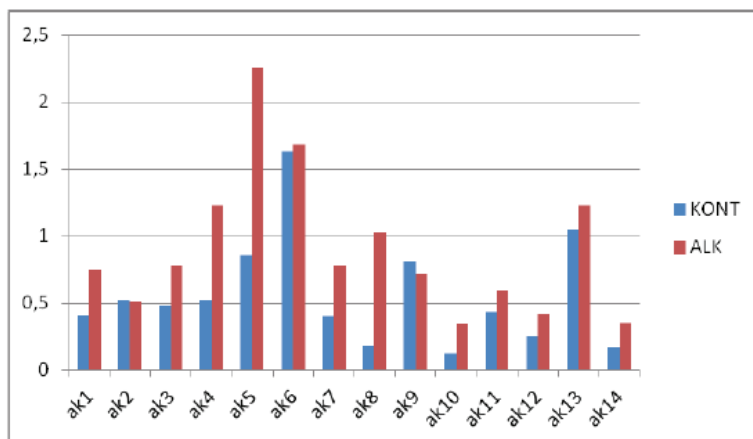
## 4.2. Megakadások

Az alkohol hatására az egyszerű reakcióidő megnő, és ez megfigyelhető nyelvi szempontból is: a kutatások többsége szerint ittas állapotban a tárgymegnevezés ideje növekszik (Moskowitz–Roth 1971), és a felolvasás lassul (Sobell–Sobell 1972).

A szerzők egyetértenek abban, hogy alkohol hatására nő a megakadásjelenségek száma. Elsősorban a hibázások válnak gyakoribbá (pl. Gyarmathy 2011; Szőnyegi 2013): a nyelvi tervezés, az artikulációs tervezés és az artikulációs kivitelezés szintjének hibái is növekvő tendenciát mutatnak (2. ábra) (Szőnyegi 2013).

Számos kísérletben úgy találták, hogy csökken az ismétlés (pl. Oullet 2010; Barffüßer–Schiel 2010), valamint a korrigált nyelvbotlások száma, ami az önmonitorozás romlását mutatja (Szőnyegi 2013). Ezzel szemben Gyarmathy (2007, 2009, 2010, 2011) kutatása éppen az önmonitorozás erősödését észlelte. Az ellentmondást vélhetően a kísérleti protokollok eltérése okozza. Az alkoholos állapotban utánmondott nyelvtörők vizsgálatokor Tisljár–Szabó és munkatársai

(2014) megállapították, hogy a memória romlása nagyban hozzájárul a hibázások arányának növekedéséhez.



2. ábra. A hiba típusú megakadások előfordulási gyakorisága beszélőnként józan (KONT) és ittas (ALK) állapotban (100 szóra jutó hiba típusú megakadások) (Szönyegi 2013: 37).

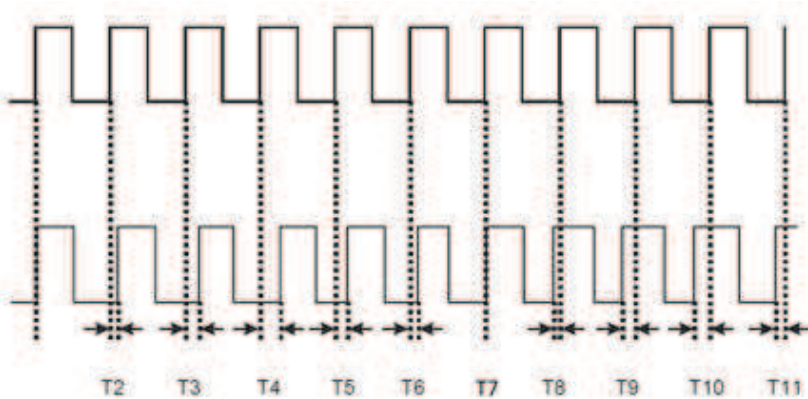
Egyes hangcsoportok esetében gyakran figyelhető meg hibás artikuláció, ilyenek az affrikáták (zár-réshangok), a nazálisok (orrhangok) és a likvidák ([l] és [r]) (Pisoni et al. 1985).

#### 4.3. Akusztikai fonetikai változások

Az eddigi kutatások alapján úgy látszik, hogy az alkohol elsősorban a fonációs (zöngképzési) folyamatokra, az időviszonyokra, valamint az artikuláció korai szakaszára van hatással (Chin–Pisoni 1997).

A legtöbbet kutatott kérdések egyike az alaphangmagassághoz köthető változások mértéke és iránya. Míg az alaphangmagasság ( $f_0$ ) variabilitásának növekedését minden ezt vizsgáló kutatás kimutatta, az  $f_0$  átlaga egyes vizsgálatokban emelkedő (pl. Klingholz et al. 1988; Hollien et al. 2001), másokban (pl. Watanabe et al. 1994) ereszkedő tendenciát mutatott. Baumeister és munkatársai (2012) 75 férfi és 85 nő felolvasása és spontán beszéde alapján mindkét tendenciát megerősítette: a beszélők 79%-ánál  $f_0$ -emelkedés figyelhető meg, a maradék 21%-nál pedig csökkenés. A két csoport közti különbség egyelőre tisztázatlan. Az alaphangmagasság változását vizsgálva Baumeister és társai (2012) megállapították, hogy a beszélők 63,5%-ánál gyorsabban mennek végbe a változások alkoholos állapotban, mint józanul. Szintén a hangszalagműködéshez kapcsolódó érték a

jitter, mely a kváziperiodikus hullámok rövid időtartamú eltolódását jelenti: amikor a hangszalagok zöngékezés közben újra és újra megnyílnak, ezt közel azonos, de nem pontosan azonos hosszúságú időszakosként teszik. Ez a kismértékű időeltolódás a jitter, ami gyakorlatilag a hangszalagműködés rendezetlenségének mérőszáma (3. ábra). Több kutatás (pl. Watanabe et al. 1994; Baumeister et al. 2012) is azt mutatja, hogy a jitter növekszik az alkohol hatására.



3. ábra. A jitter (Mészáros et al. 2010: 4). A felső jel teljesen periodikus, míg az alsó kváziperiodikus. Az alsó jel „rendezetlensége” megfigyelhető a felőhöz képest. A jitter a rendezetlen, T-vel jelölt időszakosok átlaga

A beszéd időviszonyai szintén megváltoznak az alkohol hatására. Az alaphangmagasság mellett ez a másik leginkább kutatott kérdéskör az alkoholos beszéddel kapcsolatban. Kutatások sora (pl. Pisoni et al. 1985; Künzl et al 1992) azt mutatja, hogy mind a **beszédtempó**, mind az **artikulációs tempó** lassul. Pisoni és társai (1985) megfigyelték, hogy az időtartam nyúlása a zöngétlen beszédhangoknál jellemzőbb, vagyis az ittas beszéd különbözik az egyszerű, józan lassú beszédétől.

Az artikuláció pontatlansága több paraméterben is megfigyelhető, e téren azonban kevesebb kutatással számolhatunk. A magánhangzók képzésében csak kisebb eltérések figyelhetők meg: több vizsgálat (pl. Schiel et al. 2010; Németh 2014) találta azt, hogy egyes magánhangzók ([a:] [i]) első formánsának értéke a beszélők egy részénél növekszik. Ezzel szemben a különböző mássalhangzócsoporthoz tartozóknál jellegzetesebb változások figyelhetők meg: az affrikáták (zár-rés hangok) zárszakasza lerövidül, vagy egészen eltűnik, spirantizálódás (réshangúsodás) következik be (Lester–Skousen 1974; Chin–Pisoni 1997). A [s] képzési helyének megváltozását (palatalizációját) számos esetben megfigyelték az amerikai angolban (Lester–Skousen 1974; Chin–Pisoni 1997). Az angolban a [s]

bizonyos hangkörnyezetben józan állapotban is palatalizálódik, azonban az alkohol hatására egyéb helyzetekben (pl. mássalhangzó-torlódás során) is végbe megy a folyamat (Heigl 2013). Künzel és munkatársai (1992) adatolták, hogy a nazalitásban is fellépnek anomáliák: a nazális mássalhangzók (orrhangok) orális hangokként, az orális magánhangzók viszont nazálisan realizálódhatnak.

## 5. További kutatási lehetőségek

Mint láttuk, az alkohol hatására több mérhető változás is bekövetkezik a beszédben, azonban egyes jelenségek nincsenek kimerítően vizsgálva (pl. a nazalitás), másoknál viszont ellentmondó kutatási eredményekkel találkozunk. Ezek részben az egyéni varianciának, részben a különböző kísérleti protokolloknak köszönhetők. Bár az alkoholos beszéd felvételének menete az idők során egyre kiforrottabbá és pontosabbá vált (pl. a kísérleti személyek gondosabb megválasztása, a véralkoholszint pontosabb mérése, az alkohol típusának és bevitelének figyelembevétele stb.), a befolyásoltsággal (vagyis a részegség tulajdonképpeni mértékével) még kevés kutatás számol. A további vizsgálatokban célszerű lenne a véralkoholszinten kívül egyéb objektív mérőszámokat (különböző reakcióidő-adatokat) is rögzíteni a hanganyagok mellett, hogy a befolyásoltság mint további paraméter is felhasználhatóvá váljon, ezáltal a BAC/BrAC mellett új szempontból is csoportosíthatók legyenek az adatközlők.

A kimutatható változások ellenére a tudomány mai állása szerint csupán valószínűsíthető, hogy a beszélő egy adott felvétel elkészültekor fogyasztott-e alkoholt vagy sem. A bizonytalanságban egyrészt szerepet játszik az, hogy az alkohol hatására bekövetkező változások egyenként egyéb esetekben is létrejöhetnek, például az alaphangmagasság nem csak az alkohol hatására, hanem egyes érzelmi állapotok (öröm, düh) során is megemelkedik (Baumeister et al. 2012). Másrészt a jelentős egyéni eltérések is nehezítik az azonosítást, például egyes beszélőknél emelkedik az alaphangmagasság, és lassul az artikulációs tempó, mások viszont éppen az alaphangmagasság csökkenését és a nazalitásban tapasztalható anomáliákat (szabálytalanságokat) produkálják. Ezekre a bizonytalansági tényezőkre megoldást jelenthetne, ha sikerülne az alkohol hatására együtt megjelenő, korreláló (kölcsonösen összefüggő) paramétereket azonosítani.

## Irodalom

Agarwal, Dharam – Goedde, Werner. 1990. *Alcohol metabolism, alcohol intolerance, and alcoholism: Biochemical and pharmacogenetic approaches*. Springer-Verlag, Berlin-Heidelberg-New York-London-Paris-Tokyo-Hong Kong.

- Barfüsser, Sabine – Schiel, Florian 2010. *Disfluencies in alcoholized speech*. Abstracts IAFPA 2010. [http://epub.ub.uni-muenchen.de/13675/1/schiel\\_13675.pdf](http://epub.ub.uni-muenchen.de/13675/1/schiel_13675.pdf) (2015. 01. 30.)
- Baumeister, Barbara – Heinrich, Christian – Schiel, Florian 2012. The influence of alcoholic intoxication on the fundamental frequency of female and male speakers. *Journal of the Acoustical Society of America* 132/1. 442–451.
- Chin, Steven – Pisoni, David 1997. *Alcohol and speech*. Emerald Group Publishing Limited, Madison.
- Collins, Patrick 1980. A comparison of the oral syntactic performance of alcoholic and non-alcoholic adults. *Language and Speech* 23. 281–288.
- Fontan, M. – Bouanna, G. – Piquet, J. M. – Wgeux, F. 1978. Les troubles articulatoires chez l'ethylique. *Lille Medical* 23. 529–542.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukción. *Beszédkutatás 2007*. 108–121.
- Gyarmathy Dorottya 2009. Hogyan hat az alkohol a spontán beszédre? In Kuna Ágnes – Veszelszki Ágnes (szerk.): *FÉLÚTON 3. A harmadik Félúton konferencia (2007) kiadványa*. ELTE BTK Nyelvtudományi Doktori Iskola, Budapest.
- Gyarmathy Dorottya 2010. A beszédellenőrzés működése alkoholos állapotban. In Bárdosi Vilmos (szerk.): *Világkép a nyelvben és a nyelvhasználatban: Az MTA Modern Filológiai Társasága tudományos konferenciájának előadásai*. Tinta Könyvkiadó, Budapest, 125–135.
- Gyarmathy Dorottya 2011. *A megakadások javításának stratégiái a spontán beszédben*. Doktori értekezés. ELTE, Budapest.
- Heigl, Barbara. 2013. *[s] under the influence of alcohol*. [http://www.gmu.edu/org/lingclub/WP/texts/8\\_Heigl2.pdf](http://www.gmu.edu/org/lingclub/WP/texts/8_Heigl2.pdf) (2015. 01. 30.)
- Hollien, Harry – DeJong, Gea – Martin, Camilo – Schwartz, Reva – Liljegren, Kristen 2001. Effects of ethanol intoxication on speech suprasegmentals. *Journal of the Acoustical Society of America* 110/6. 3198–3206.
- Klingholz, F. – Penning, R. – Liebhardt, E. 1988. Recognition of low-level alcohol intoxication from speech signal. *Journal of the Acoustical Society of America* 84. 929–935.
- Künzel, Hermann – Braun, Angelika – Eysholdt, Ulrich 1992. *Einfluss von Alcohol auf Sprache und Stimme*. Kriminalistik-Verlag, Heidelberg.
- Lados Dénes – Németh Szilvia 2014. *Idiolektus vagy ittasság? A részegek beszédének felismerhetősége*. Előadás a XVI. Pszicholingvisztikai Nyári Egyetemen. Balatonalmádi, 2014. május 25–29. Ppt: <http://goo.gl/4V91NQ>
- Lester, Leland – Skousen, Royal 1974. The phonology of drunkenness. In Bruck, Anthony – Fox, Robert – LaGaly, Michael (eds.): *Papers from the parasession on natural phonology*. Chicago Linguistic Society, Chicago, 233–239.
- Mészáros János – Styeták Péter – Tóth László 2010. *Jittermérések optikai szakaszokon*. Mérési útmutató. Óbudai Egyetem, Budapest.
- Moskowitz, H. – Robinson, C. 1988. *Effects of low doses of alcohol on driving-related skills: A review of the evidence*. Washington D.C. <http://ntl.bts.gov/lib/25000/25700/25752/DOT-HS-807-280.pdf> (2015. 01. 30.)
- Moskowitz, H. – Roth, S. 1971. Effect of alcohol on response latency in object naming. *Quarterly Journal of Studies on Alcohol* 32. 969–975.
- Németh Szilvia 2014. Az alkohol hatása a magánhangzók ejtésére. In Várad Tamás (szerk.): *Alknyelvdok8. Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből*. MTA

- Nyelvtudományi Intézet, Budapest, 101–112. <http://www.nytud.hu/alknyelvdok14/proceedings14/nemeth.pdf>
- Ouellett, M. 2010. *Estimating intoxication level from speech*. <http://www.honors.ufl.edu/apps/Thesis.aspx/Details/489> (2015. 01. 30.)
- Pisoni, David – Hathaway, Susan – Yuchtman, Moshoe 1985. *Effects of alcohol on the acoustic-phonetic properties of speech. Final report to GM Research Laboratories*. (SRL Technical Note No. 85-03). Indiana University, Bloomington.
- Pisoni, David – Martin, Christopher 1989. Effects of alcohol on the acoustic-phonetic properties of speech: Perceptual and acoustic analyses. *Alcoholism: Clinical and Experimental Research* 13. 577–587.
- Schiel, Florian – Heinrich, Christian – Neumeyer, Veronika 2010. Rhythm and formant features for automatic alcohol detection. In: *Interspeech 2010*. 458–461.
- Schiel, Florian – Heinrich, Christian – Barfüsser, Sabine 2011. Alcohol Language Corpus: The first public corpus of alcoholized German speech. *Language Resources and Evaluation* 46/3. 503–521.
- Selzer, Melvin 1975. The Michigan Alcoholism Screening Test (MAST): The quest for a new diagnostic instrument. *American Journal of Psychiatry* 3. 176–181.
- Sobell, Linda – Sobell, Mark 1972. Effects of alcohol on the speech of alcoholics. *Journal of Speech and Hearing Research* 15. 861–868.
- Szőnyegi Júlia 2013. *Az alkohol hatása a beszédprodukcóra*. MA szakdolgozat. ELTE, Budapest.
- Tisljár-Szabó, Eszter – Rossu, Renáta – Varga, Veronika – Pléh, Csaba 2013. The effect of alcohol on speech production. *Journal of Psycholinguistic Research* 43. 737–748.
- Tisljár-Szabó Eszter 2014. *Érzelmi állapotok és módosult tudatállapotok beszédjellemzőinek vizsgálata*. Doktori értekezés. BME, Budapest.
- Watanabe, Hiroshi – Shin, Takemoto – Matsuo, Hiromichi – Okuno, Fumio – Tsuji, Tsutomu – Matsuoka, Midori – Fukaura, Junichi – Matsunaga, Hisashi 1994. Studies on vocal fold injection and changes in pitch associated with alcohol intake. *Journal of Voice* 8. 340–346.
- Watson, Patricia 1989. Total body water and blood alcohol levels: Updating the fundamentals. In: Crow, Kathryn E – Batt, Richard D. (eds.): *Human metabolism of alcohol. Vol. I. Pharmacokinetics, medicolegal aspects, and general interest*. CRC Press, Boca Raton, 41–56.
- Widmark, Erik 1932. Die theoretischen Grundlagen und die praktische Verwendbarkeit der gerichtlich-medizinischen Alkoholbestimmung. Urban & Schwarzenberg, Berlin, Wien. (Fordításban megjelent: *Principles and applications of medicolegal alcohol determination*. Biomedical Publications, Davis, 1981.)

### Kérdések, feladatok

1. Milyen változások következnek be az alkohol hatására a beszédben?
2. Megállapítható-e a beszédből, hogy a beszélő ivott-e alkoholt? A választ indokolja meg!
3. Milyen módszertani és etikai problémákat vet fel a beszéd vizsgálata alkoholos állapotban?



## Temporális sajátosságok a beszédben

Ha valaki az átlagosnál gyorsabban vagy lassabban beszél, arra mindig felfigyelünk. A gyorsbeszélőre gyakran rászólunk, hogy *ne hadarj!* (nem tudván, hogy a hadarás és a gyors beszéd nem azonos fogalmak a logopédiában), míg egy lassú beszélőt hallgatva sokszor alig várjuk, hogy végre befejezze a mondatot, és kiderüljön, hogy mi a lényeg. Vannak olyan esetek is, amikor bár a beszélő beszédtempója nem tér el az átlagostól, mi mégis gyorsnak, nehezen követhetőnek tartjuk, vagy épp lassúnak, unalmasnak ítéljük, amit mond.

Milyen tényezőktől függ, hogy valaki gyorsan beszél, vagy lassan beszél, sok szünetet tart, avagy keveset? Mi befolyásolja azt, hogy hallgatóként milyennek ítéljük a beszéd sebességét? Tudunk-e szándékosan változtatni a beszédünk tempóján? Változik-e egy adott társadalom tagjainak a beszédtempója a felgyorsult életritmus következtében?

### 1. Beszédtempó, artikulációs tempó, szünettartás

Mielőtt a fenti kérdéseket megválaszolnánk, érdemes tisztázni a beszéd temporális jellemzőinek leírására szolgáló fogalmakat. A beszéd sebességének meghatározásakor kétféle tempóértéket szokás elkülöníteni. A beszédtempó az időegység alatt elhangzó beszédjelek számát jelenti, mégpedig úgy, hogy a beszéd idejébe a szünetek is beleszámítanak. Az artikulációs tempó a beszédjelek létrehozásának a sebességét mutatja meg. Ilyenkor nem vesszük figyelembe a szünettartást, csak az artikuláció idejét. A beszédtempó kiszámításakor meg kell számolni a beszédhangok vagy a szótagok vagy a szavak számát, majd elosztani a beszéd teljes időtartamával. Az artikulációs tempó meghatározásakor a jelidővel osztjuk a beszédjelek számát; például a teljes beszédidőből kivonjuk a szünetek összesített idejét, avagy külön-külön megmérjük a beszédszakaszok (szünettől szünetig tartó nyelvi jelek) időtartamát, és azokat összegezzük. Mivel ekkor kisebb értékkel osztjuk ugyanazon beszédjelek számát, az artikulációs tempó értéke mindig nagyobb, mint a beszédtempóé (Gósy 2004). A tempóértékeket megadhatjuk szó/percben, szótag/percben, szótag/másodpercben és hang/másodpercben is (Gósy 2004).

A tempóértékek meghatározásakor megkülönböztetjük a globális és a lokális tempót. A globális tempó az egész beszédprodukciónak az átlagos tempóját mutatja meg. A lokális tempó kisebb egységeket vizsgál, és a szótagok vagy a hangok

időtartamainak átlagát veszi figyelembe. Az artikulációs tempó ugyanis változik a beszéd során ugyanazon beszédhelyzetben is, egyes szótagokat (pl. hangsúlyosakat, szakaszvégieket vagy nyújtást mint megakadást tartalmazókat) lassabban ejtünk, mint másokat (Fletcher 2010).

A nyelvi jelek folyamatos sorozatát a beszédben gyakran szünetek és más megakadásjelenségek szakítják meg. A beszédszünet egyrészt fiziológiai szükséglet (a levegővételhez biztosít időt), másrészt lehet akaratlagos (a beszélő tagolni akarja a mondanivalóját, illetve hatásszünetet tart), illetve jelezheti a beszélő beszédtervezési nehézségeit is (Gósy 2004). Ezek mellett a fő szerepek és okok mellett a szakirodalomban számos más szünetfunkcióról is olvashatunk, például gondolkodási szünetről, a beszédtervezési bizonytalanságokat „feloldó” szünetről és hatásszünetről, de a szünet kaphat például nyomatékosító funkciót is (összefoglalásként lásd Bóna 2013). A jelkimaradás nem mindig tekinthető funkcionális szempontból beszédszünetnek: a „csend” megjelenhet 1. egy folyamatban levő egységen belül, 2. társalgási egységek között, illetve 3. jelezheti a társalgás végét (Levelt 1989). A beszédjelnek a társalgási egységek közötti kimaradását hallgatásnak nevezzük, a társalgást vagy beszédeseményt megelőző és követő jelkimaradás pedig a csend (Markó 2005). A szünet, a csend és a hallgatás tehát nem azonos fogalmak a fonetikában. A csend és a hallgatás során a beszélő abbahagyja a beszédet. Ennek több oka lehet: nincs több mondanivalója, vagy egyéb okból nem alkalmas, hogy folytassa a beszédet. A hallgatás általában hosszabb időtartamú, mint a beszédszünet (Markó 2005, 2015).

A beszédszünetet a következőképpen definiálhatjuk (Gósy 2004: 207): „a beszéd folyamatban jelentkező szünet olyan kismértékben akaratlagos beszédkimaradás, amely néma vagy jellel kitöltött, de független a beszédhang képzésétől. Funkcióját tekintve a beszédprodukcióban 1. biztosítja az artikulációt lehetővé tevő légáramot, 2. elősegíti a közlés értelmi tagolását, 3. a beszédtervezés során az ún. ellentmondások, téves utak stb. feloldására szolgál, 4. a mentális lexikonban történő keresési idő kitöltését biztosítja, illetőleg lehetőséget nyújt a nyelvi kódolás módosítására. Funkciói a beszédmegértésben: 1. az elhangzottak könnyebb feldolgozása, 2. az entrópia [rendezetlenség, bizonytalanság] csökkentése és 3. a megértés és az értelmezés működési folyamatainak biztosítása”.

A szünet nem feltétlenül néma jelkimaradás, a jellel kitöltött szüneteket hezitálásnak hívjuk. Ezek leggyakoribb formája a magyarban az *öö* és a *mm* (Gósy 2004), de megjelenhetnek például *öm*, *öh*, *ömh*, *höh*, *eh*, *eö* stb. formában is (Horváth 2010). A néma szünet és a hezitálás különböző variációkban állhat egymás mellett is két beszédszakasz között (Markó 2005; Horváth 2010), az ilyen szünetet kombinált szünetnek hívjuk (Horváth 2014). A szünettartás egyénre jellemző sajátosság, ha egy beszélőre hosszabb néma szünetek a jellemzők, akkor valószínűleg hosszabb kitöltött szüneteket tart (Gósy–Beke–Horváth 2011). A hezitálások fő funkciói a következők lehetnek: a beszédszándék kifeje-

zése, a bizonytalanság jelzése, idő biztosítása a hibajavításhoz, illetve a társalgásban a beszélő kitöltött szünettel jelezheti a beszédszakasz végén a beszédpartnernek, hogy át kívánja adni a szót, vagy épp ellenkezőleg, magánál akarja tartani (Horváth 2010).

## 2. A temporális sajátosságokat befolyásoló tényezők a beszédprodukciónban

Azt, hogy egy beszélő gyorsan vagy lassan beszél, számos tényező befolyásolja. A beszédtempó függ például a beszélő életkorától, a személyiségétől, a nemétől, a fizikai és pszichés állapotától; befolyásolják a beszélőre jellemző szünettartási sajátosságok és megakadások; de fontos tényező a beszédtema, a beszédhelyzet és a beszéd típus is (Gósy 2004).

Az életkor hatásával a tempóra számos tanulmány foglalkozik. Ezek megállapítják, hogy a beszédtempó az anyanyelv-elsajátítás korai szakaszaiban lassabb, mint később; majd időskorban ismét lelassul, a gyermekekéhez hasonló lesz. Bár az egyes szerzők különböző mértékegységekkel adják meg a tempóértékeket, ami megnehezíti az összevetést, közlünk néhány a magyarra jellemző tempóértéket különböző életkorú beszélőktől (1. táblázat).

1. táblázat. A beszédtempó és az artikulációs tempó átlaga különböző életkorokban

Életkor	Beszédtempó (hang/s)	Artikulációs tempó (hang/s)
3 éves (1 fő, Gósy 1997)	nincs adat	6
6 éves (1 fő, Gósy 1997)	nincs adat	8
21–24 évesek (Markó 2014)	10,94	12,82
39–46 évesek (Markó 2014)	10,09	12,25
60–65 évesek (Bóna–Vakula 2013)	9,4	12,3
75–85 évesek (Bóna–Vakula 2013)	8,8	11,4
90 évesek (Bóna–Vakula 2013)	6,4	9,5

	Beszédtempó (szó/perc)	Artikulációs tempó (szó/perc)
6 évesek (Neuberger 2013)	82,1	nincs adat
7 évesek (Neuberger 2013)	81,3	nincs adat
9 évesek (Neuberger 2013)	88,9	nincs adat
9 évesek (Horváth 2013)	75,2	nincs adat
11 évesek (Neuberger 2013)	90,1	nincs adat
13 évesek (Neuberger 2013)	99,6	nincs adat
Felnőttek (Gósy 2004)	135–140	nincs adat

Több kutatás szerint a nők és a férfiak hasonló tempóval beszélnek (pl. Gocsál 2001), míg más vizsgálatok szerint a férfiak artikulációs tempója gyorsabb, mint a nőké (vö. Váradi 2009). A beszélő személyisége is hatással van a tempóra: az extrovertált egyének gyorsabban beszélnek, mint az introvertáltak (vö. Gocsál–Huszár 2003), és a hallgatói benyomások is nyíltabbnak ítélik a gyorsabb beszélő személyeket (Németh 2013).

Hétköznapi tapasztalat, hogy a beszélő pszichés állapota is hatással van a beszédtempóra: az izgalom legtöbbször gyorsítja az artikulációs tempót, míg a szünetek, főként a hezitálások száma növekszik a beszédben. Hasonlóan árulkodik a beszédtempó az örömről és a haragról, ezek kifejezésekor gyorsabb a beszéd, illetve a szomorúságról, amely esetén általában lassabban beszélünk (Scherer 1995).

A beszédtempó megváltozása különböző betegségeknek a jelzése, illetve velejárója is lehet. A szorongás, a depresszió és a gyomorfekély a beszédtempó lassulásával, a pajzsmirigy-túlműködés és a neurózis gyors beszédtempóval jár együtt, míg a pszichopáták beszédsebessége átlagos marad (Vékássy 2004). A kisebb-nagyobb rekedtséggel társuló ritmus- vagy tempóváltozás már a betegség kezdetén felhívhatja az orvos figyelmét a hangszalagműködés kóros elváltozásaira (Balázs 2002); de például a beszéd lassulásával jár együtt az Alzheimer-kór is (Hoffmann et al. 2005).

A különböző beszédhelyzetek más-más tervezési nehézséget jelentenek (Krepsz 2016), ami a temporális sajátosságokban is megjelenik. A különböző beszéd típusok közül a kutatások többsége a felolvasást mérte a leggyorsabb tempójúnak (Váradi 2010; Bóna 2013; Markó 2014). Markó (2005) spontán narratívákat, képleírást és társalgást hasonlított össze. Az eredményei szerint a képleírásban lassabb, a társalgásban gyorsabb tempóértékek voltak adatolhatók a spontán narratívákhoz képest. Bóna (2013) kutatásában a leglassabb tempójú beszéd típus a tartalomösszegzés volt, míg a leggyorsabb artikulációs tempójú a felolvasás. A leggyakoribb és a legritkább szünettartás is ezt a két beszéd típusot jellemezte. A társalgás a szünetezést tekintve jobban hasonlított a felolvasásra, mint a spontán narratívára.

Végezetül fontos megjegyezni a tempót befolyásoló tényezők között azt is, hogy a különböző nyelvek hangzó változataira más-más átlagos tempóértékek jellemzőek. Bár a különböző szakirodalmi adatok nehezen összevethetőek a módszertani eltérések miatt, a 2. táblázatban megadtuk néhány tájékoztató adatot különböző nyelvekre vonatkozóan (forrás: Gósy 2004: 205).

2. táblázat. Különböző nyelvek tempója

Nyelv	Tempóértékek szóródása (szótag/s)
holland	5,5–9,3
francia	4,7–6,8
spanyol	4,6–7,0
arab	4,6–7,0
olasz	5,3–8,9
brit angol	4,5–8,9

### 3. A temporális sajátosságok és a beszédfeldolgozás

Az, hogy egy közlést gyorsnak, lassúnak vagy normál tempójúnak ítélünk, elsősorban az elhangzó beszéd artikulációs tempójától, a szünetek számától és hosszától függ (Grosjean–Lane 1976). Bár az elhangzó beszéd sebességének közel objektív megítélésére az átlagos hallgató is képes, és akár az általa ismeretlen nyelv tempóját is meg tudja ítélni úgy, hogy az észlelési ítéletek hasonlítsanak az objektíven mért adatokhoz (den Os 1985), a beszédtempó megítélésében nagy egyéni különbségek lehetnek (Gósy 1991; Gocsál 1999). Az átlagos beszélők percepcióis folyamatai átlagos tempójúak, míg a szélsőségesen gyors vagy szélsőségesen lassú beszélők beszédfeldolgozási folyamatai szélsőségesen gyorsak vagy szélsőségesen lassúak (Gósy 2004). Az átlagos beszélők/hallgatók között is vannak lassabb és vannak gyorsabb beszédtempót kedvelő személyek, illetve olyanok, akiknek percepcióis viselkedése „extrém” (Gocsál 1999).

A tempóészlelésre hatással van az intenzitásváltozás (a hangerő megváltozása) és az alaphfrekvencia-változás is. A magasabb alaphfrekvencia gyorsabb, az alacsonyabb alaphfrekvencia lassabb tempó benyomását kelti. A monoton dallam hatása nem független az objektív tempótól: „gyorsító hatása” csak egy bizonyos átlagérték után jelentkezik a beszédészlelésben (Gósy 1991). A tempóészlelést befolyásolják még a szintaktikai határok, a szótaghangsúly, a szótagszerkezet és a ritmikai szerkezet is (Gósy 1991). A beszédtempó változásában éppen észlelhető különbség a megnyilatkozás átlagos tempójának 5%-nyi gyorsulását vagy lassulását jelenti (Quené 2007). A beszédpercepció során hozott „szubjektív” tempóítéletek függenek még a beszédhelyzettől, az adott feladat megértési nehézségeitől, a beszédrészlet terjedelmétől, illetve a hallott beszédprodukciónak artikulációs megformáltságától stb. (Gósy 1997).

A tempóészlelést alapvetően befolyásolja az általános beszédmegértés is; amelyre hatással van maga a beszédtempó is. A gyorsabb beszéd általában nehezebben érthető, mint az átlagos tempójú; a nagyon gyorsan elhangzó mondatok feldolgozása igen gyors beszédértést eredményez, ám ilyenkor több az értési hiba is. A nagyon lassú mondatok feldolgozását nehezíti, hogy ilyenkor több idő

jut az asszociációs működésekre (Gósy 2004). A gyorsabb tempóhoz a hallgató adaptálódik, ami a válaszadás során gyorsabb beszédképzésben is megnyilvánul (Gósy–Laczkó 1987). Az, hogy a beszédtempó hogyan hat a beszédmegértésre, függ a hallgató életkorától is. Mondatisméltési, mondatértési, illetve szövegértési tesztek tanúsága szerint a gyermekek és az idősek a lassabb beszédet preferálják (Gósy–Laczkó 1987; Rákli 2004).

A felgyorsított tempójú mondatok és szövegek észlelését és megértését több kutatásban is vizsgálták. Ezek tanúsága szerint a beszédtempó bizonyos mértékű gyorsítása által előidézhető egy olyan állapot, amikor a hallgató már képtelen megérteni az elhangzott üzenetet, de az észlelési szint még működik (Gósy 1991; Bóna 2009). Egy magyar nyelvű kísérletben (Bóna 2009) tizenöt különböző hosszúságú és szerkesztettséggű, illetve jelentéstartalmú mondatot rögzítettek normál tempóban, majd az időtartam-zsugorítás módszerével 60%-ra, 50%-ra és 45%-ra nyomták össze őket úgy, hogy nem változott a beszédhangok frekvenciaszerkezete és intenzitása. A természetes tempó esetén a mondatok átlagos beszédtempója 14,2 hang/s, a 60%-os időtartam-zsugorítás esetén 23,7 hang/s (ez a tempóérték még előfordulhat a hétköznapi kommunikáció során is, ha gyorsbeszélőkkel beszélgetünk), 50%-os időtartam-zsugorítás esetén 28,4 hang/s, 45%-os időtartam-zsugorításkor pedig 31,6 hang/s volt. Az eredmények szerint a 60%-os időtartam-zsugorítás esetén nem volt jelentős különbség a természetes tempójú mondatok ismétléséhez képest, a leggyorsabb tempó esetében azonban a helyesen visszamondott szótagok aránya 50,1% lett. A 60%-ra és az 50%-ra zsugorított időtartamú mondatokra kapott válaszok arról tanúskodnak, hogy egy darabig valóban segíti a megértés szintje a beszédészlelést. A közlések ugyanis nagyon gyorsan hangzottak el, a kísérleti személyek mégis képesek voltak – bizonyos mondatok esetében akár 100%-os eredménnyel is – megismételni az elhangzottakat; a nem megfelelően észlelt szavakat, hangokat a magasabb szintű beszédfeldolgozási folyamataik, nyelvi tudásuk segítségével rekonstruálták. Ezzel szemben a mindössze plusz 5 százalékponttal (45%-ra) zsugorított mondatoknál már drasztikusan lecsökkent a helyesen visszamondott szótagok száma. Ebből arra következtethetünk, hogy a beszédmegértés részfolyamata ebben az esetben már nem működött. Az észlelés azonban tetten érhető maradt még ekkor is, erről azok az értelmetlen mondatok tanúskodnak, amelyeket ennél a tempónál visszamondtak a kísérleti személyek. Ezekben az eredetihez nagyon hasonló magán- és mássalhangzókat tartalmazó szavakat találunk. Azt, hogy a gyorsabb tempó nehezíti a beszédfeldolgozást, nemcsak a helyesen visszamondott szótagok aránya, hanem a reakcióidők is mutatják. A leggyorsabb mondatok esetében mért reakcióidők szignifikánsan hosszabbak voltak a lassabb tempójú mondatoknál mértéknél.

A tempógyorsítással kapcsolatos kísérletek számos módszertani kérdést is felvetnek. Ezek egyike az, hogy hogyan gyorsítsuk a beszédet. Létrehozhatunk

gyors tempójú beszédet 1. természetesen úton, amikor arra kérjük a beszélőket, hogy gyorsítsanak a beszédük tempóján; avagy mesterségesen, 2. a fentebb már említett időtartam-zsugorítás módszerével, illetve 3. a szünetek kiiktatásával. A kísérletek tanúsága szerint a mesterségesen felgyorsított beszéd jobban érthető, mint a természetes gyors beszéd – legalábbis a szavak szintjén –, mert az időviszonyai az átlagos beszédhez hasonlíthatnak. Habár a természetes gyors beszéd is jól érthető, a feldolgozása lassabb, mint a mesterségesen felgyorsítotté. Ráadásul a hallgatók a mesterségesen felgyorsított beszédet preferálják (Janse–Nootboom–Quené 2003). A szövegértési teljesítmény a tempó gyorsításával csökken, és akkor a legrosszabb, ha a beszédtempó-változást a szünetek kiiktatásával, és nem csak az artikulációs tempó gyorsításával érjük el (Bóna 2009).

Az tehát, hogy milyen tempóval beszélünk, nagymértékben meghatározza a beszélő észlelési és megértési mechanizmusainak sikerességét. Ezért társalgás közben érdemes tekintetbe venni a hallgató sajátosságait, és amennyire lehetséges, ehhez igazítani a beszédünket is (a társalgópartnerek önkéntelenül is alkalmazkodnak egymáshoz, vö. Grácz 2009). A következőkben megvizsgáljuk, hogy mennyire lehetséges a beszédtempónk szándékos gyorsítása, avagy lassítása.

#### 4. Tudunk-e változtatni a beszédünk tempóján?

A szakirodalom szerint a beszédtempó az egyik legnehezebben megváltoztatható szupraszegmentális eszköz (Gósy 2004). Akár lassítani, akár gyorsítani szeretnénk rajta, mindenképpen nehézségekbe ütközünk. A tempó szándékos lassítása során hamar kellemetlen érzésünk lesz, és beszédünket visszagyorsítjuk a számunkra természetes tempóra; míg a szándékos gyorsításnak vannak biológiai és beszédtervezésbeli korlátai is.

A beszédtempó biológiai meghatározottságát több kísérlettel is vizsgálták. Az egyik vizsgálatban a felső és alsó ajkak, az állkapocs és a nyelv mozgásának elemzésével próbálták felderíteni a lassú- és a gyorsbeszélés motoros okait. Ehhez egy rövid hangsort (*a bad daba*) mondattak be többször az adatközlőkkel normál, lassú és gyors tempóban. Az eredmények szerint a lassú beszédtempónál a normál értékekhez képest a sebességek csúcserképei csökkentek, míg gyors tempó esetén a beszédszervek sebességének növekedését és csökkenését is észlelték. A felső és alsó ajkak sebességei között mindig pozitív volt a korreláció; de az állkapocs és más szervek sebességeinek korrelációjánál az előjelek eltértek az egyes beszélők esetében (McClean 2000). Ez azt jelenti, hogy a felső és az alsó ajkak működésének sebessége azonos módon gyorsult vagy lassult a beszéd közben; ugyanakkor az állkapocs és más beszédszervek sebességét összevetve előfordult az is, hogy az egyik szerv gyorsabban működött, míg a másik lassult.

Egy másik vizsgálatban (Tsao–Weismer 1997) egy rövid szöveget olvastattak fel kétszer lassú és gyors beszédű kísérleti személyekkel. Az utasításnak megfelelően az adatközlők először a megszokott egyéni beszédtempójukban olvastak, majd a lehető leggyorsabb tempóval, amelyre képesek voltak. Az eredmények azt mutatták, hogy a) szignifikáns összefüggés van a természetes szokásos és a maximális beszédtempó-értékek között, b) szignifikáns különbség van a lassú és a gyorsbeszélők maximális beszédsebessége között, illetve c) nagyjából hasonló az eltérés a természetes és a maximális beszédsebesség között mind a lassú, mind a gyorsbeszélők esetén. Néhány kivétel is előfordult: voltak olyan lassú beszélők, akik nagyon gyors maximális beszédtempót tudtak létrehozni. Ez azt bizonyítja, hogy nem önmagában az ideg-izomrendszeri tényező befolyásolja az egyéni beszédtempó értékét, hanem pl. szociolingvisztikai tényezők is.

A beszédtempó tudatos megváltoztatását a magyar nyelvben is vizsgálták (Laczkó 1993). Egy kísérletben például a hangos olvasás temporális sajátosságait elemezték normál, lassított és felgyorsított tempóban. Az eredmények azt mutatták, hogy a beszédtempó megváltoztatásában elsősorban a szünettartás játszik szerepet. A vizsgálatban részt vevők könnyebben tudták gyorsítani, mint lassítani olvasási tempójukat, a gyorsítás mértéke átlagosan elérte a 4,1–3,03 hang/s-ot, s általában a szünetek időtartamának csökkentése jellemezte. Ezzel szemben a felolvasás tempójának lassítása során a gyakorlott beszélők mindössze 2,23 hang/s-ot, míg az átlagos beszélők csak 1,6 hang/s-ot tudtak lassítani beszédtempójukon. A lassítást a szünetek számának növekedése jellemezte (Laczkó 1993).

A felolvasott szövegek tempójának megváltoztatása nehéz, de egyszerűbb, mint a spontán közléseké. A felolvasás és a spontán beszéd produkciója között ugyanis nagy különbség van: a felolvasás során csak az artikulációs tervezésre kell összpontosítanunk, míg a spontán beszédben a szöveg tartalmi és formai megformálására is. A spontán beszéd tempójának szándékos megváltoztatását nehezebb kísérleti úton vizsgálni, mint a felolvasását, ugyanis ugyanaz a beszédprodukció nem megismételhető, a különböző spontán megnyilatkozások összevetése pedig problémás, mivel létrehozásukkor számtalan tényező befolyásolja a tempót (lásd fentebb).

Egy kísérletben úgy próbálták megoldani a nehézségeket, hogy irányított spontán beszédet elemeztek többféle tempóban (Bóna 2009: 91–101). Az adatközlőknek egy képsort mutattak, amelyen egy közismert logikai feladvány megoldásának lépései voltak összekeverve. A logikai feladvány ismertsége egyrészt segítség volt a történetmeséléshez, másrészt a lépések összekeverése miatt a beszélőnek folyamatosan gondolkodnia kellett a történet megfelelő folytatásán. Így nem volt lehetőség rá, hogy az egyszer már elmondott történetet rutinból ismétlje meg. A kísérleti személyeket arra kérték, hogy meséljék el a képsoron látható történetet először természetes tempóban, majd kétszer egymás után olyan gyorsan, amilyen gyorsan csak tudják. Az eredmények azt mutatták, hogy az

adatközlőknek nem sikerült jelentősen fokozni a beszédsebességüket. A beszédtempót átlagosan mindössze 3,6 hang/s-mal (ez mintegy 31 szót jelent percenként), az artikulációs tempót 2,9 hang/s-mal tudták gyorsítani, és a gyorsítás során a szünettartás is nagymértékben csökkent. A tempóváltoztatás mellett az adatközlők úgy is megpróbálták „gyorsítani”, hogy a normál tempójú elmondásnál kevesebb szóval, mondattal fogalmazták meg gondolataikat, így is rövidítve a beszédidőt. Amíg a természetes tempóban elmondott történethez átlagosan 128 szót használtak, addig az első gyorsításnál ez átlagosan 84, a másodiknál 73 szóra csökkent. Megfigyelhető volt az is, hogy a kísérlet résztvevőinek a fele a feladatot is belefoglalmazta a gyorsított történetmesélésbe (például *a révész nagyon gyors nagyon nehéz feladatot kapott; gyorsan visszarakta a kecskét; az a révész feladata, hogy gyorsan átvigye a kecskét [...] átrohan vele a révész a túloldalra; gyorsan át kell vinni; a kecskét máris visszük vissza; a révésznek az a feladata, hogy átvigye minél gyorsabban a kecskét*). Jellegzetes volt az is, hogy a tempógyorsulással együtt a megakadások előfordulási gyakorisága is nőtt.

A spontán beszéd szándékos lelassítását hadarók és gyorsbeszélők beszédében vizsgálták (Bóna 2009). Az eredmények azt mutatták, hogy egyetlen vizsgált személynek sem sikerült megfelelően lassítania beszédét: vagy továbbra is gyors maradt az artikuláció, vagy a beszélők csak túllassítással, önmaguk karikírozásával, hangszínezet-változtatással tudták megoldani a feladatot. Volt köztük olyan hadaró adatközlő is, akinek a szándékos lassításkor a beszédtempója és az artikulációs tempója 5 hang/s-mal csökkent a természetes spontán beszédéhez képest, de ehhez a lassításhoz a beszélőnek jóval nagyobb mértékű szünettartásra volt szüksége. Amíg a természetes szituációban kb. 24% volt a szünetek aránya, addig ez az arány a félszöveg beszédhelyzetben 38%-ra nőtt, s minden ötödik szünet időtartama meghaladta az 1 másodpercet (Bóna 2009).

Összegezve tehát az eredményeket elmondható, hogy a tempó megváltoztatása nehéz; és elsősorban a szünettartás arányának és gyakoriságának változásával függ össze.

## 5. Változik-e a beszéd tempója társadalmi szinten?

A fentiekben láthattuk, hogy az egyén beszédtempóját és artikulációs tempóját számos tényező befolyásolja egyetlen napon belül is, és ezektől függetlenül is változik az élete során. Kérdés azonban, hogy az egyéntől függetlenül, az évszázadok során változott-e az egy nyelvet beszélők átlagos beszédtempója, azaz gyorsul-e a beszéd. Ha igazolható, hogy a beszédtempó a különböző korszakokban gyorsult, akkor ezzel több nyelvtörténeti változás (hangkiesések, időtartamok rövidülése, bizonyos minőségi változások stb.) is magyarázható (Gósy 1998).

A kérdés megválaszolása azonban nem egyszerű, és a szakirodalomban is ellentmondó adatok olvashatók. Hangfelvételek hiányában ugyanis a korábbi tempóváltozásokra csak következtetni tudunk. Vértes O. András különböző korokban, gyorsírással lejegyzett parlamenti beszédeket elemzett, és matematikai számítások segítségével arra jutott, hogy a beszéde sebesség felgyorsult. A számításai alapján a feltételezett beszédtempó 1869-ben 2–2,3 szótag/másodperc (kb. 5–5,75 hang/s); 1889-ben 3,6 szótag/másodperc (kb. 9 hang/s); 1924-ben pedig 4 szótag/másodperc (kb. 10 hang/s) lehetett (Vértes O. 1989; Kassai 1993). Ezek a számítások azonban bizonytalanok több okból is (maga Vértes O. András is annak tartja őket): 1. a szótag/másodperces érték a különböző szótaghosszúságok miatt csak megközelítőleg képes tükrözni a beszédtempót; 2. a gyorsírók nem kiejtés szerint, hanem a helyesírásnak megfelelően alakítják a sajátos jeleiket, így a gyorsírásból nem derül ki, hogy valójában hány hang hangzott el a beszéd során; 3. a gyorsírással lejegyzett szövegekből nem derül ki, hányszor és milyen időtartamú szünetet tartott a beszélő, ami pedig alapvetően befolyásolja a beszédtempó alakulását.

Mérésekkel a 20. század első felétől lehetséges vizsgálni a tempóváltozást. Gombocz Zoltán 1909-es kísérletében a *tát, tátog, tátogat, tátogatók, tátogatónak* szósorozat magánhangzóinak időtartamát hasonlította össze. Kísérletét 1965-ben megismételte Tarnóczy Tamás, majd 1992-ben Kassai Iлона (1993). Kassai a három kísérlet adatainak összevető elemzése alapján megállapítja, hogy a huszadik század elejéhez képest a magyar beszéd tempója felgyorsult. A gyorsulás okának a hangidőtartamok csökkenését, illetve az időzítési monotóniát tartja, amely a szó belseji időtartamviszonyok átrendeződésének nyomán alakul ki. Ugyanakkor ez az összehasonlítás számtalan kérdést vet fel, mivel az adatközlők kiválasztása nagymértékben befolyásolja a mérési eredményeket, és mindegyik méréskor (vö. Kassai 1993) mindössze egyetlen bemondó ejtésében elemezték az időtartamértékeket.

Gósy (1988) egy tempóészlelést és beszédmegértést vizsgáló kísérletében arra kereste a választ, hogy a beszéd legalábbis az észlelés szintjén felgyorsult-e. Az eredményei szerint a kísérleti személyek a tempóészlelés során a 11 hang/s-os értéket ítélték átlagosnak, ugyanakkor a mondatértési tesztben nagyrészt helyes válaszokat adtak a 13,98 hang/s-os normálnak tekintett tempó esetén is (Gósy 1988). A felnőttek szövegértéskor lassúnak tartják a 7,8 hang/s-os beszédtempót; lassúnak, illetve közepesnek a 8,1–9,1 hang/s-os tempót; közepesnek a 9–10 hang/s-os, kissé gyorsnak a 13 hang/s-os, igen gyorsnak a 14,3 hang/s-os beszédtempót (Gósy 1997).

Hegedűs Lajos 1957-es adatai szerint a hírolvasás tempója átlagosan 12 hang/s, a sportközvetítésé 14 hang/s volt (Hegedűs 1957). Fónagy Iván és Magdics Klára 1960-as mérései szerint az akkori beszédtempó középértéke 11,35 hang/s volt, a versmondást 9,4 hang/s-os, a beszélgetést 12,89 hang/s-os, a

sportközvetítést pedig 13,83 hang/s-os tempóérték jellemezte (1960). Ezek az átlagértékek nem maradnak el a mai tempóértékektől, hiszen a felnőttek átlagos artikulációs tempója (ahogyan az fentebb is olvasható) 12,5–14 hang/s (Gósy 2004); ami különböző helyzetekben lassabbá vagy gyorsabbá válhat. Imre 2005-ben a versmondás artikulációs sebességére 9,9 hang/s-ot, a hírolvasásra 14,4 hang/s-ot, a sportközvetítésre 12,5 hang/s-ot mért. Bóna 2009-es adatai szerint a hírolvasás artikulációs tempója a közszolgálati rádióban 12,9–13,7 hang/s volt, a beszédtempója 10,5–11,6 hang/s.

A Hegedűs-archívum (Nikléczy–Horváth 2007) beszédmintái és mai beszélőkkel készült hangfelvételek összevetése alapján Menyhárt (2010) azt találta, hogy a mai budapesti beszélők gyorsabban beszélnek, mint a mai falusi beszélők, és náluk is lassabban beszéltek az archívumi falusi beszélők. Ugyanakkor gyermekekkel készült felvételeket összevetve Imre és Menyhárt (2014) nem talált szignifikáns eltérést a 60 évvel ezelőtti és a mai gyermekek beszédének tempójában. Az artikulációs tempó is csak a mesemondáskor volt szignifikánsan gyorsabb a mai városi gyermekeknél az archívumi falusi gyermekhez képest. A szerzők azonban azt is megállapítják, hogy egyértelműen megjelennek a beszédkultúra változásai a 60 évvel ezelőtti beszélőkhöz képest a mai gyerekeknél, ezek a beszéd tagolásában, a szünettartásban és a megakadások típusainak változásaiban jelentkeznek. Az, hogy az egyéni különbségek nagyobb mértékben meghatározóak, mint akár 55 év időkülönbség, Vallent 2005-ös eredményeiben is jól megfigyelhető. Esettanulmányában ugyanis az archívumi beszélők (szintén a Hegedűs-archívumból, az 1950-es évekből) gyorsabban beszéltek, mint a 2005-ben felvett beszélők (Vallent 2005).

70–80 év távlatában elemezte a tempóváltozást Mártonfi (2007) és Gocsál (2014). Ők régi filmeket vizsgáltak a tempóértékek tekintetében. Mártonfi (2007) két film eredeti és újraforgatott változatát vetette össze, míg Gocsál (2014) négy színész 1937-es filmszerepeiben mért artikulációs tempókat. Mind a két vizsgálatból arra lehet következtetni, hogy – legalábbis a fiatal színészek filmbeli beszéde esetében – a beszéd tempója nem gyorsult hét-nyolc évtized alatt.

## 6. Kitekintés

A tanulmányban körbejártuk a tempót befolyásoló tényezőket mind produkciós, mint percepciósi szempontból, és megfigyelhettük, hogy számos tényező befolyásolja a tempóindexek alakulását. Emellett a szakirodalom alapján választ kerestünk arra is, hogy változtatható-e a beszédtempó, illetve változik-e társadalmi szinten a különböző korokban.

Kérdés azonban az, hogy mindez, amit a tempóról tudunk, hogyan hasznosítható a gyakorlatban is. A tempóval kapcsolatos vizsgálatoknak számtalan társa-

dalmi hasznossága van a hétköznapi kommunikációs szituációktól kezdve a mesterséges intelligenciáig. A mindennapi kommunikáció során – amellet, hogy az észlelést és a megértést nagymértékben befolyásolja – a tempóból és a szünetekből következtethetünk a beszélgetőpartnerünk érzelmi állapotára, fáradságára, a témában való jártasságára; és a beszédtempója alapján (ha csak a hang alapján kell ítéletet hozni) rokonszenvesnek vagy ellenszenvesnek (vö. Gocsál–Huszár 2003), szavahihetőnek, avagy megbízhatatlannak (Bata 2005) tartjuk; de a tempó befolyásolja a beszélő kompetenciájával, hitelességével vagy akár a szexiságával kapcsolatos ítéleteinket is (Gocsál–Huszár 2003). A tempóvizsgálatok eredményeit hasznosíthatja továbbá a tömegkommunikáció, a média, a retorika (például milyen közönséghez milyen tempóval érdemes/kell szólni, mennyi szünetet tartsunk a jobb megértés vagy a hatás kedvéért), avagy a pedagógia is (a tanítók, a tanárok mintát adnak a diákoknak, ezért különösen fontos, hogy jól beszéljenek; emellet ha a tanító gyorsan beszél, a kisgyermek a beszédfeldolgozási nehézségek miatt lemarad a munkában). A logopédia, gyógypedagógia esetében is kitüntetett szerepük van a tempóvizsgálatoknak, hiszen több beszédzavar és nyelvi zavar jár együtt a tempó megváltozásával. Kiemelendő a hadarás mint nyelvi zavar, amelynek vezető tünete a beszéd temporális sajátosságainak a tipikustól való eltérése. A mesterséges intelligencia, illetve a különböző beszédtechnológiai alkalmazások is kiemelten foglalkoznak a tempóval, hiszen ahhoz, hogy a mesterséges beszéd a természetes beszédhez hasonló hangzású legyen, elengedhetetlen az időzítési sajátosságok pontos megismerése.

## Irodalom

- Balázs Boglárka 2002. Időzítési zavarok mint a kóros hangképzés rendellenes következményei. *Beszédgyógyítás* 2002/1-2. 3–7.
- Bata Sarolta 2005. *Kinek hiszünk? A közléstartalom megítélése a szupraszegmentumok függvényében*. OTDK-dolgozat.
- Bóna Judit 2009. *A gyors beszéd. Produkciós és percepciós sajátosságok*. MTA Könyvtára – Lexica Kiadó, Budapest.
- Bóna Judit 2013. A beszédszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszéd kutatás* 2013. 60–75.
- Bóna Judit – Vakula Tímea 2013. Temporális sajátosságok 90 évesek spontán beszédében. *Alkalmazott Nyelvtudomány* 13/1–2. 59–73.
- Fletcher, Janet 2010. The prosody of speech: Timing and rhythm. In Hardcastle, William J. – Laver, John – Gibbon, Fiona E. (eds.): *The handbook of phonetic sciences. Second edition*. Wiley–Blackwell, Malden–Oxford–Chichester.
- Fónagy Iván – Magdics Klára 1960a. Beszédsebesség, szólam, ritmusérvzés. *Magyar Nyelv* 56. 450–458.
- Gocsál Ákos 1999. Egyéni különbségek az artikulációs tempó percepciójában. *Beszéd kutatás* '99. 19–29.

- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszédkutató 2001*. 61–72.
- Gocsál Ákos 2014. Az artikulációs tempó vizsgálata régi magyar filmekben. *Beszédkutató 2014*. 144–154.
- Gocsál Ákos – Huszár Ágnes 2003. Csábító hangok. *Beszédkutató 2003*. 9–18.
- Gósy Mária 1988. Tempóészlelés és beszédmegértés. *Műhelymunkák V*. 87–122.
- Gósy, Mária 1991. The perception of tempo. In Gósy, Mária (ed.): *Temporal factors in speech. A collection of papers*. HAS Research Institute for Linguistics, Budapest, 63–107.
- Gósy Mária 1997. A magyar beszéd tempója és a beszédmegértés. *Magyar Nyelvőr* 121. 129–139.
- Gósy Mária 1998. Hangtörténeti változások feltételezett okairól. *Magyar Nyelv* 96. 276–283.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutató 2011*. 5–30.
- Gósy Mária – Laczkó Mária 1987. A beszédmegértési teljesítmény fejlődése óvodáskortól felnőttkorig. *Magyar Nyelvőr* 111. 444–451.
- Gráczai Tekla Etelka 2009. Temporális jellemzők a beszédpartnerek ismeretségének függvényében. *Beszédkutató 2009*. 121–133.
- Grosjean, Francois – Lane, Harlan 1976. How the listener integrates the components of speaking rate. *Journal of Experimental Psychology: Human Perception and Performance* 2. No. 4. 538–543.
- Hegedűs Lajos 1957. Beszédtempó-elemzések. *Magyar Nyelvőr* 81. 223–227.
- Hoffmann Ildikó – Németh Dezső – Kálmán János 2005. A beszéd időviszonyai Alzheimer-kórban. *Beszédgyógyítás* 15. 21–31.
- Horváth, Viktória 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57/2–3. 288–306.
- Horváth Viktória 2013. Temporális szerveződés kilencéves gyermekek spontán beszédben. *Beszédkutató 2013*. 144–159.
- Horváth Viktória 2014. *Hezitációs jelenségek a magyar beszédben*. Beszéd – Kutatás – Alkalmazás 3. ELTE Eötvös Kiadó, Budapest.
- Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- Imre Angéla – Menyhárt Krisztina 2014. Különböző műfajú szövegek temporális sajátosságai mai és 60 évvel ezelőtti gyermek beszélőknél. *Anyanyelv-pedagógia* 2014/1. <http://www.anyp.hu/cikkek.php?id=496>
- Janse, Esther – Nootboom, Sieb – Quené, Hugo 2003. Word-level intelligibility of time-compressed speech: Prosodic and segmental factors. *Speech Communication* 41. 287–301.
- Kassai Ilona 1993. Gyorsult-e a magyar beszéd tempója az elmúlt 100-120 évben? *Beszédkutató '93*. 62–69.
- Krepsz Valéria 2016. Fonetikai hasonlóságok és különbségek a beszéd típusokban. In Bóna Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest. 175–188. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu).
- Laczkó Mária 1993. A tempó és a szünet viszonya a hangos olvasásban. *Beszédkutató '93*. 185–193.
- Levitt, Willem J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).

- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hümmögés vizsgálata*. PhD-értekezés. ELTE, Budapest.
- Markó Alexandra 2014. A beszéd temporális szerkezete a beszédmód és a beszédhelyzet függvényében. In Bányi Szilvia – Navracsics Judit – Vigh-Szabó Melinda (szerk.): *Nyelvelsajátítási, nyelvtanulási és beszéd kutatások*. Pszicholingvisztikai tanulmányok IV. Gondolat Kiadó – Pannon Egyetem MFTK, Budapest – Veszprém, 33–45.
- Markó Alexandra 2015. Jelkimaradások a beszédben: szünet és hallgatás. In Bárh M. János – Bodó Csanád – Kocsis Zsuzsanna (szerk.): *A nyelv dimenziói. Tanulmányok Juhász Dezső tiszteletére*. ELTE Magyar Nyelvtudományi és Finnugor Intézet – Magyar Nyelvtudományi Társaság, Budapest, 515–524.
- Mártonfi Attila 2007. *Beszédtempó-változás a filmremake-ek tükrében*. Előadás a *Beszéd kutatás 2007* konferencián. Budapest, 2007. november 15–16.
- McClellan, M.D. 2000. Patterns of orofacial movement velocity across variations in speech rate. *Journal of Speech, Language, and Hearing Research* 43. 205–216.
- Menyhárt Krisztina 2010. A beszédsebesség objektív mérési és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszéd kutatás 2010*. 110–124.
- Németh Máttyás Ferenc 2013. A tanári beszéd és személyiség. *Anyanyelv-pedagógia* 2013/1. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=434>
- Neuberger Tilda 2013. A spontán beszéd temporális sajátosságai 6–14 év közötti gyermekeknél. *Anyanyelv-pedagógia* 2013/2. <http://www.anyp.hu/cikkek.php?id=451>
- Nikléczy Péter – Horváth Viktória 2007. Nyelvjárási hangarchívum az interneten. *Beszéd kutatás 2007*. 173–178.
- den Os, Els 1985. Perception of speech rate of Dutch and Italian utterances. *Phonetica* 42. 124–134.
- Quené, Hugo 2007. On the just noticeable difference for tempo in speech. *Journal of Phonetics* 35/3. 353–362.
- Rákli Veronika 2004. Rövid szövegek hallás alapú feldolgozása – rádióhírek megértése. *Alkalmazott Nyelvtudomány* IV/2. 33–48.
- Scherer, Klaus R. 1995. How emotion is expressed in speech and singing. In Elenius, Kjell – Branderud, Peter (eds.): *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Vol 2. Stockholm, 85–89.
- Tsao, Ying-Chiao – Weismer, Gary 1997. Interspeaker variation in habitual speaking rate: Evidence for a neuromuscular component. *Journal of Speech, Language, and Hearing Research* 40. 858–866.
- Vallent Brigitta 2005. A spontán beszéd ötven éve és ma. Esettanulmány. *Beszéd kutatás 2005*. 99–111.
- Váradí Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszéd kutatás 2010*. 100–109.
- Vékássy László 2004. Betegségek beszédjellemezői. *Beszédgyógyítás* 15/1. 40–47.
- Vértes O. András 1989. Gedanken über die geschichtliche Veränderung des Sprachtempos. In: *Proceedings of the Speech Research '89 International Conference*. MFF/HPP 21. 381–384.

### **Kérdések, feladatok**

1. Olvasson fel egy rövid szöveget természetes tempóban. Ezután olvassa fel nagyon lassan, majd a leggyorsabb tempóban, amire képes. Figyelje meg, milyen stratégiákat használ a tempóváltoztatáshoz (szünettartás, túlartikulálás stb.)! Milyen érzései voltak a tempóváltoztatás során?
2. Képzeld el, hogy gyermekeknek, időseknek, a barátainak, a szüleinek mesél el egy történetet! Hogyan változik meg a beszédtempója a különböző hallgatók esetében?
3. Figyelje meg, hogyan változik a riporterek beszédtempója és artikulációs tempója a különböző sportágak eseményeinek közvetítése során!
4. Keressen olyan rádióadót (a neten), amelynek nem érti a nyelvét! Próbálja meg kitalálni, éppen milyen műsor megy! Figyelje meg, hogyan segítik ebben a temporális sajátosságok!

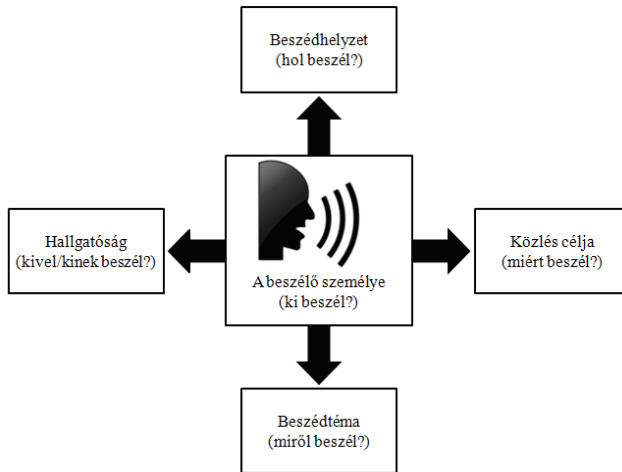


## Fonetikai hasonlóságok és különbözőségek a beszéd-típusokban

A beszéd a kommunikáció legáltalánosabb eszköze. Ahogyan a gyermek bekerül az óvodába, majd az iskolába, egyre több beszédhelyzetben kell helytállnia, nő a kommunikációs tapasztalata. Az életkor előrehaladtával fokozatosan gyarapszik a beszédrutinja, egyre több beszéd-típust ismer meg, tud maga is létrehozni. Felnőtt beszélőként mi magunk is számos szituációban veszünk részt, különböző kommunikációs szerepekben kell megszólalnunk, különböző beszéd-típusokat kell létrehozunk. Kérdés, azonban hogy ezek a beszéd-típusok mennyiben azonosak, és milyen szempontok szerint térnek el egymástól. Milyen jellemzők alapján tekinthetők állandónak az egyes megnyilatkozások, illetve hogyan és milyen tényezők hatására változik beszédünk a mindennapi kommunikáció során?

Szöbeli közléseink igen sokfélék lehetnek, számos tényező befolyásolja a megnyilatkozások minőségét (1. ábra). Ezek közül a következőkben a teljesség igénye nélkül mutatunk néhány olyan paramétert, amelyek korábban fonetikai elemzések tárgyát képezték. Noha az egyes jellemzők számos különböző módon, egymással szoros, gyakran egyedi összefüggésekben alakítják közléseinket, a következőkben külön-külön mutatjuk be a kiválasztott paraméterek beszédprodukcióna gyakorolt hatását.

A szóbeli közlések megvalósulását meghatározza a **beszélő személye**. A megnyilatkozásokra egyaránt hatással van a beszélő genetikai adottsága és a tanult tényezők is. Befolyásolja a beszélő személy életkora, neme, iskolázottsága, testalkata, személyisége, hogy nyelvjárási beszélő-e, aktuális lelki és fizikai állapota, a környezet, ahol nevelkedett és ahol aktuálisan él stb. Egyetlen személy szóbeli megnyilatkozása is nagy különbségeket mutat izgatott és nyugodt lelkiállapotban, kipihenten és fáradtan, egészségesen és betegen, vagy aszerint, hogy épp fogyasztott-e alkoholt (vö. pl. Gocsál 1998; Gósy 2001; Gocsál–Huszár 2003; Gyarmathy 2007; Bóna 2009; Gósy–Krepsz 2015). Amikor mások beszédét hallgatjuk, bizonyos tulajdonságairól relatíve helyes következtetéseket tudunk levonni a hangja alapján, míg más tulajdonságainak azonosítása kevésbé sikeres. Amíg a beszélők nemét és életkorát gyakran műszeres vizsgálat nélkül, csupán a szubjektív feldolgozás segítségével, a korábbi tapasztalatainkra hagyatkozva is képesek vagyunk megítélni (akár telefonon keresztül is), addig a testalkat és a testmagasság becslése nehezebben megítélhető csupán a beszéd alapján.



1. ábra. A beszéd minőségét befolyásoló főbb tényezők

A beszéd minőségére hatással van az aktuális **beszédtema**. Könnyen belátható, hogy saját magunkról, hobbijainkról, tanulmányainkról nehézségek nélkül tudunk összefüggően nyilatkozni. Egy összetettebb, kevésbé hétköznapi témáról előzetes felkészülés nélkül azonban kevesen képesek azonnal folyamatos beszédprodukciónak létrehozni. A számunkra kellemetlen témákról legtöbbször sok szünetet tartva, a pontos szóválasztásra és a megfelelő nyelvi szerkezetekre fokozottan ügyelve beszélünk, az örömhíreket gyakran nagyobb hangerővel, gyorsabb tempóval, kevesebb szünettartással közöljük, míg egy titokról vagy pletykáról halkán, suttogva beszélünk (Gósy 2004).

Meghatározó a **beszédhelyzet**, amelyben a megnyilatkozás létrejön. Másképp beszélünk egy formális és egy informális szituációban, négy szemkört és nagy közönség előtt (előadóként és felszólalóként is). A beszédhelyzet nem kizárólag a megnyilatkozás stílusára vagy a szókincsre van hatással, egyidejűleg befolyásolja az artikuláció minőségét, pontosságát, továbbá a különböző prozódiai eszközök használatát, a hangerőt, tempóértékeket stb.

A beszédhelyzettel szoros összefüggésben befolyásolja a megnyilatkozást a **hallgatóság**. A résztvevők száma, a beszélő(k) és a hallgató(k) között fennálló viszony (főnök–beosztott, barát–barátnő, anya–gyermek) mind hatással van a beszéd minőségére. A társalgások, párbeszéd során a beszélő igyekszik alkalmazkodni a beszédpartner(ek) vélt igényeihez, ami meghatározza a megnyilatkozás szerkesztettségét, a szóátvétel módját és a beszédleptések hosszát. Meghatározó beszédpartnerünk életkora is. Az idős emberek esetében a beszédfeldolgozás folyamata lelassul részben a klinikai hallásromlás, részben a beszédmegértés nehezítetté válása miatt. Beszédünket, elsősorban a szupraszegmentális jegyeket

igyekszünk ennek mentén alakítani, lassabban, pontosabban artikulálva, hangsabban beszélni. A nemzetközi szakirodalomban számos modell létezik a beszédpartnerhez való alkalmazkodás tekintetében. Bell 1984-es modellje szerint azok is hatással vannak az üzenet megvalósítására, akik nem címzettjei az üzenetnek, ám valamilyen módon mégis meghallgatják azt.

A kiejtést befolyásoló tényezők nagy számából adódóan a beszédhangokat és a hangkapcsolatokat sohasem ejtjük egyformán. Az artikulációt vizsgáló kutatások közül napjaink egyik legelfogadottabb felfogása Lindblom (1990) H&H-elmélete (hyper and hypo speech – „túlartikulált és alulartikulált beszéd”). A beszélő két különböző célt próbál elérni a kommunikációs helyzetben. Célja részben a feszes artikuláció a jól érthetőség érdekében, másrészt a gazdaságos beszédprodukciónak. Az artikuláció működése ilyen módon egy skálán képzelhető el, amelynek két végpontja a pontos-pontatlan vagy feszes-laza artikuláció.

A beszédprodukciónak befolyásoló paraméterek igen nagy számából adódóan a szóbeli megnyilatkozásoknak számos típusát, azaz sok különböző beszédtypust különítünk el. Az így létrejött megnyilatkozásokat különböző szempontok alapján csoportosíthatjuk. A kategorizáció alapja lehet a spontaneitás mértéke, amelynek mentén négy kategóriába (spontán, félreproduktív, reproduktív, felolvasás) sorolhatjuk a beszédtypusokat (Wacha 1974). A kategóriák határai egyértelműen meghatározhatók, mégsem merevek.

## 1. A spontán beszéd

A spontán beszéd a hétköznapi kommunikáció legjellemzőbb formája, amely a résztvevők számától függően megvalósulhat monológ, párbeszéd vagy több fős társalgás formájában. A spontán beszéd produkciója a megszólalás szándékától a kiejtésig tartó folyamat (Gósy 2005). A spontán megnyilatkozások jellemzője, hogy a beszélő nem tervezi meg előre sem a mondanivalója tartalmát, sem pedig annak nyelvi megvalósítását. A gondolkodás, a nyelvi tervezés, a szöveg megszerkesztése és annak kiejtése, azaz meghangosítása egy időben történik.

A beszélő az adott pillanatban válogatja ki a közlésre szánt gondolatokat, és tervezi meg a nyelvi formákat, vagyis a pontos nyelvi megvalósítást is. Amíg a közölni kívánt gondolatok és képek aktiválódnak, párhuzamosan történik a nyelvi tervezés is. Ezt a jelenséget nevezik **a beszédtervezés és -kivitelezés paradoxonának**, ugyanis a két különböző folyamat egyidejűleg korlátozza is a másik működését (Gósy 1998).

A spontán beszéd fogalma nem azonos a **természetes beszéddel**. A spontán beszéd létrehozása során a megszólalás szándékát nem előzi meg felkészülés. Természetes beszédnek azt tekintjük, ha a beszélőnek nem okoz észrevehető, feltűnő nehézséget a folyamat működtetése; noha a beszélőspecifikus jellemzők-

ből adódóan a fogalom nem határozható meg objektív fonetikai paraméterek segítségével, számos tényező figyelembevétele szükséges. A beszéd természetességének nagy jelentősége van az előadó-művészi és a retorikai gyakorlatban, ahol a beszélő személy verbális viselkedése alapozza meg a beszélő személy hitelességét (Gósy–Gyarmathy–Horváth 2009).

A mondanivaló tervezésének és meghangosításának egyidejűsége gyakran megnehezíti a beszélő és bizonyos esetekben a hallgató dolgát is. Abból adódóan, hogy a tervezés és a kivitelezés szinkrón tevékenység, a közlés nem minden esetben mintaszerű. Amennyiben a tervezés és a kivitelezés nincs tökéletes összhangban, a megnyilatkozásokban előfordulhat redundancia, azaz túlbiztosítás; hiány, azaz nem kimondott, legtöbbször a közös előismeretből vagy a beszédhelyzetből adódó információ; közbeékelés vagy váratlan bővítmény; alanyváltás; az időrendi és/vagy a logikai sorrendiség be nem tartása; tévesen kiejtett névelő; egyeztetési tévedések; ismétlés stb.

A tartalmilag vagy formailag **hiányos** közlések eredményei lehetnek agrammatikus, azaz nyelvtanilag nem megfelelően kivitelezett, vagy elliptikus, azaz kihagyásos, nem teljes nyelvi formák. Ezek a beszédpartnerek közös nyelvi és fogalmi előismerete nyomán nagy valószínűséggel jól kiegészíthetők, ritkán okoznak megértési problémákat. Például (a példák forrása a BEA beszédadatbázis, vö. Gósy et al. 2012):

- (1) szeretek pingpongozni, korcsolyázni, hát szóval ezeket mondjuk Ő nyilván, aki – én nem – hegyvidéken van az nyilván jobban mehet
- (2) aztán a kertben felmáztam a létra lét rá ra na \*nevetés\*

A spontán beszéd jellemzőiből adódóan azonban nem kizárólag a hiány jelenik meg a megnyilatkozásokban, ugyanúgy jelen van **túlbiztosítás** is, amely során a beszélő mondanivalóját (azonos gondolatot) több, különböző, ám azonos értékű nyelvi formával is kifejez (vö. Gósy 2004). Például:

- (3) én gyógyszerészeti pályát választottam és a végén nem bántam meg hogy ezt a pályát választottam
- (4) környezetvédelmi cégnél dolgozom, de nem csak ott képviselem ezeket az elveket mármint a munkahelyemen hanem a szabadidőmben is
- (5) három gyermekem van két lány és egy fiú aki focistának készül mármint a fiam Dávid

Minden beszéd típusban előfordulhatnak, ám elsősorban a spontán beszéd jellemzői a **megakadásjelenségek**, amelyek a beszédprodukció és beszédpercepció folyamatosságát megakasztó, a beszélő saját nyelvi normájának ellentmondó jelenségek (lásd bővebben: Gósy 2005; Beke et al. 2015; Gyarmathy

2015). Például:

- (6) a nagyapám édesanyámnak a \*néma szünet\* a az és e ööö édesapja
- (7) szabadidőmben kertészkedek a kertemben rengeteg az ööö aze ööö na azálea van meg mindjé mondom mi az sosem jut eszembe a neve pö pö pö pet petúnia van még az erkély alá ültetve

A beszédtervezés során gyakran sorrendiségi hibákból adódóan jelenik meg az egyes részletek előretörése az aktuális megvalósításban, amelyek eredményei lehetnek a közbeékelések, mondatátszövődések vagy a váratlan bővítmények is. Például:

- (8) én úgy gondolkoztam én ööö \*sóhajtás\* gyógyszerészetin de hála a jó istennek hogy ööö nem de akkor ööö ez az időszak volt egy kicsit tessék utánaszámolni
- (9) ennek ellenére én sokat mákszá mák mászkálok az Astoria környékén és akkor hirtelen nem érttem, hogy ellette vagy mellene vannak-e a többiek
- (10) ott kattints rá hogy vizsgálat folytatás nélkül vagyis folytatás vizsgálat nélkül

A kommunikáció során a beszélő és a hallgató ösztönösen betartják az **együttműködési elvet**. A beszélő igyekszik úgy beszélni, hogy a lehető legnagyobb mértékben biztosítsa a hallgatót számára az elhangzottak feldolgozhatóságát, a hallgató pedig az elhangzott közlés legpontosabb, az eredeti közlési szándéknak leginkább megfelelő dekódolására törekszik (Grice 1975/2001).

Az elhangzó közlések pontos feldolgozásának alapvető feltétele a megfelelő **tagolás**. A folyamatos beszédet egységekre bontjuk olyan módon, hogy a beszélő számára fiziológiásan (egy levegővétellel) létrehozható, egyidejűleg a hallgató számára könnyen feldolgozható legyen. A spontán beszéd alapegysége, amelyekre a beszélő saját beszédprodukciónak tagolja, az úgynevezett virtuális mondat. A hallgató számára a különböző egységek határát részben a szünettartás, részben a szintaktikai-szemantikai sajátosságok jelölik (Gósy 2003; Váradí 2013).

A beszédprodukciónak tartott **szünetek** megjelenését befolyásolja a beszéd típusa, noha a kötöttebb beszéd típusokban kevésbé jelentkeznek az egyéni sajátosságok, míg más műfajokban nagyobb mértékű a szünettartás egyéni variabilitása (Bóna 2013). Spontán beszédben gyakrabban és hosszabb időtartamú szüneteket tartanak a beszélők, mint felolvasásban, azonban a szünetek valódi időtartamát elsősorban az egyéni sajátosságok határozzák meg. A spontán beszédre a néma és kitöltött (valamilyen hangjelenséggel kísért), illetve kombinált (néma és jellel kitöltött) szünet egyaránt jellemző (Váradí 2010).

A korábbi kutatások azt igazolták, hogy a szünettartás elsősorban személyspecifikus változást mutat, a különböző típusú monológokban „szabad” (a beszélő számára ismert és kedvelt téma szabad ismertetése) és „kontrollált” (vizsgaszerű szituáció, amelyben a téma nem természetes a beszélő számára) narratíva, illetve történetmondás képsorozat alapján) nem volt megfigyelhető a szövegtípusokra jellemző változás. Ezzel szemben a társalgásban a beszélők kevesebb és rövidebb szüneteket tartottak. A társalgás tehát a legkönnyebb feladatot jelenti a beszélők számára a több résztvevőből adódóan (függetlenül attól, hogy a beszédhelyzetben egymást segítik a beszélők, vagy versenyhelyzet alakul ki), hiszen több idő áll rendelkezésére közléseik megtervezésére, a szókincs bizonyos része már korábban aktiválódik (Markó 2005).

A különböző beszélők spontán megnyilatkozásai között igen nagy egyéni eltérések tapasztalhatók a **temporális szerveződést** tekintve (Markó 2005). A személyspecifikus tényezők sokszor előre nem is jósolhatók, így például a vizsga okozta szorongás kiválthat lassabb beszédtempót hosszabb szünettartással és gyors, szinte hadaró beszédtempót is.

A spontán beszédben megjelenő beszédhangok időtartama rövidebb, mint a felolvasásban vagy izolált ejtésben megjelenő hangoké (Gósy 2004), ami azzal magyarázható, hogy felolvasás során gyorsabb a beszélők artikulációs és beszédtempója, mint spontán beszédben, amely valószínűsíthetően szintén a tervezési folyamatok egyidejű működéséből adódik (Váradi 2010; Bóna 2013). Ezzel szemben nincs jelentős eltérés a „szabad” és „kontrollált” narratívák tempóértékei között, míg a társalgás beszédtempójában – az egyéni különbségek ellenére is – nagymértékű gyorsulás volt kimutatható. Mindez magyarázható a több résztvevő jelenléte miatti rövidülő szünettartással, az artikulációs tempó gyorsulásával is (Markó 2005).

## 2. A félreproduktív (vagy félspontán, félinterpretatív) beszéd

A félreproduktív beszéd jellemzője, hogy a beszélő átgondolja mondanivalóját, tartalmilag felkészül a megszólalásra, ám a megfelelő nyelvi formát csak az adott beszédhelyzetben rendeli hozzá. A megnyilatkozás során nem betanult szöveget mond el, hanem spontán módon hozza létre a beszédet. A szövegalkotás folyamata hol megelőzi a meghangosítást, hol a spontán beszédnek megfelelően egy időben zajlik a nyelvi tervezés és a kiejtés. Félreproduktívnak tekintjük például a tanári megnyilatkozást, ahol az oktató már begyakorolt szöveget, magyarázatot mond el, ám az aktuális megvalósítást (a stílust, a szókincset, az adott grammatikai szerkezeteket) a beszélő, jelen esetben a tanár az adott pillanatban választja meg, és amelyet meghatároz többek között a témakör, az adott tanóra témája, a diákok életkora, az iskola típusa stb.

### 3. A reproductív beszéd

A beszélő egy előre elkészített, mások vagy önmaga által korábban megfogalmazott szöveget tanul meg, majd mond el fejből, azaz a közlés erősen függ a memória működésétől. A megnyilatkozás célja lehet a gyönyörködtetés vagy a spontán beszéd illúziójának keltése, például színészek vagy üzletkötők esetében. Ide tartoznak a különböző előadói megnyilatkozások, amikor a beszélő valamilyen irodalmi művet vagy memoritert, verset, mesét, mondókát, monológot vagy köszöntőt hangosít meg.

### 4. A felolvasás

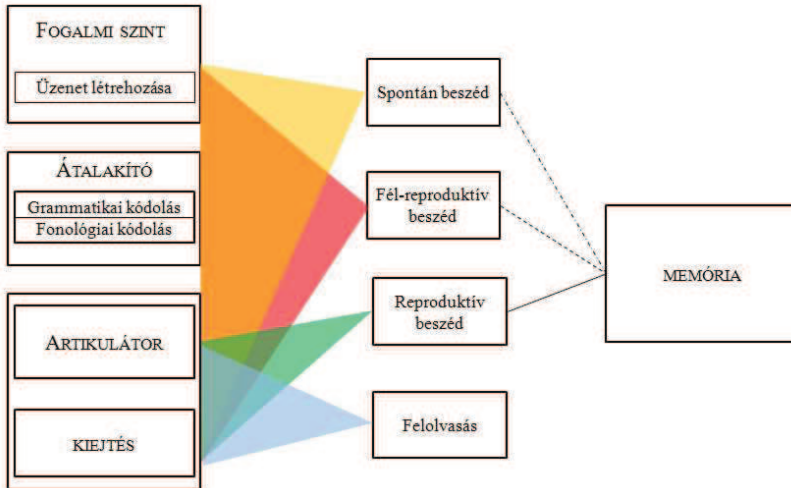
A felolvasás során a beszélő egy előre rögzített, megtervezett, grammatikailag és szókincsében elkészített szöveget hangosít meg. A reproductív beszédhez hasonlóan, itt sincs szükség a gondolatok megfogalmazására (a szöveg megalkotása és megszólaltatása nem egy időben történik), illetve a nyelvi tervezésre (a szöveg nyelvtanilag és szókincsében is előre adott). A felolvasás során a tagolást és a szünetezést is a központozás irányítja, segíti. A spontán beszéddel szemben a szupraszegmentális elemek, például a szünetezés, beszédhang vagy hangsúly megvalósítása nagymértékben tudatos, kisebb mértékben tekinthető automatikusnak. Például:

- (11) a permetszerek \*néma szünet\* hatóanyagainak jelentős része \*néma szünet\* nem a kezelni kívánt növényekre és kártevőkre kerül hanem a környezetbe jutva károsítja a környező területek élővilágát \*néma szünet\* szennyezheti a talajt és a vizeket \*néma szünet\* olvasható a levegő munkacsoport tanulmányában

### 5. A beszédprodukciónak működése

A beszédprodukciónak működését korábban számos módon modellálták. Napjaink egyik legelfogadottabb felfogása a Willem Levelt, holland pszicholingvista által (1989) kidolgozott beszédprodukciónak modell, amely a megnyilatkozás szándékától a kiejtésig mutatja be a beszéd létrehozásának működését. A modell (2. ábra) szerint a megszólalást a közlésre szánt gondolatok megtervezése (fogalmi szint), az aktuális nyelvi forma hozzárendelése (átalakító), így a közlési szándéknak megfelelő szavak, szókapcsolatok, kifejezések, grammatikai szerkezetek kiválasztása előzi meg. A beszédtervezés során létrehozott fonetikai terv az artikulátor segítségével válik hallható beszéddé az artikulációs gesztusok megva-

lósításával. A folyamatok egymással párhuzamosan, automatikusan és rendkívül gyorsan zajlanak.

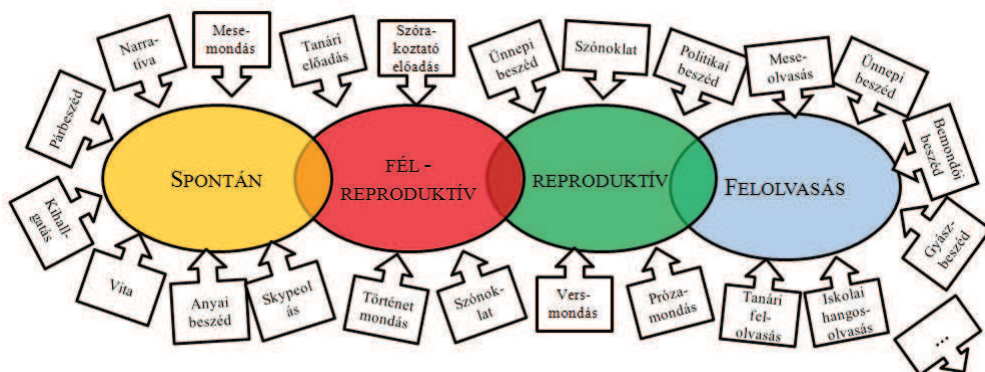


2. ábra. A levelti beszédprodukción model alkalmazása a tárgyalt beszéd típusokra

A különböző beszéd típusok a tervezés különböző szintjeit veszik igénybe. Spontán és félreproduktív beszéd létrehozása esetén a beszédprodukción mind a négy szintje aktiválódik, hiszen az üzenet létrehozásától a kiejtésig minden az adott pillanatban zajlik. A reproduktív beszéd és a felolvasás kizárólag az artikulátor és a kiejtés szintjét érinti, a beszélő mindkét esetben egy előre elkészített szöveget hangosít meg, így a felsőbb tervezési szintek nem érintettek. A memória működése is szorosan hozzájárul az egyes beszéd típusok létrehozásához. A spontán és a félreproduktív beszédben az előismeretek, a közölni kívánt tartalmi és nyelvi egységek felidézését végzi, segít a megfelelő nyelvi formák, így például szavak, kifejezések előhívásában. A reproduktív beszéd esetén a tartalmi elemek és az aktuális nyelvi forma felelevenítésében végez központi feladatot.

A spontaneitás mértéke alapján meghatározott kategóriákhoz különböző beszéd típusok rendelhetők. A 3. ábrán a teljesség igénye nélkül mutatunk be néhány beszéd típust, ahogyan a három pont is jelzi, a felsorolás folytatható. Egyes beszéd típusok több kategóriához is sorolhatók az adott típus egyéni megvalósításának függvényében. Így például a mesemondás történhet teljesen spontán módon, amikor az adott pillanatban találjuk ki a cselekményt, fogalmazzuk meg az eseményeket, és hangosítjuk meg azt. Történhet félreproduktív módon abban az esetben, ha a történet váza, a szereplők, a főbb cselekmények ismertek, ám az adott nyelvi megvalósítást például a hallgatóság korának és előismeretének meg-

felelően választjuk meg. A mesemondás történhet reprodukzív módon abban az esetben, ha például egy gyermek olyan sokszor hallott már egy mesét, hogy ő maga is „kívülről” tudja azt, illetve felolvashatjuk könyvből is.



3. ábra. Beszéd típusok lehetséges csoportosítása

## 6. A beszéd típusok jellemzése

A következőkben nagy mennyiségű anyagon készült vizsgálatok eredményeire építve mutatjuk be néhány beszéd típus fonetikai sajátosságait. A felsorolás szubjektív válogatás eredménye. Noha a különböző beszéd típusok jellemzőire közel pontos meghatározások adhatók, az aktuális megvalósulást minden esetben a beszélő/olvasó egyéni jellemzői határozzák meg.

### 6.1. A színészi beszéd

A korábbi évszázadokban úgy tartották, hogy a színház a szép beszéd temploma, a színészek a beszéd mesterei, akik a nyelv használatának közvetítői, a nyelvi esztétika őrzői. Az előadás típusától függően a színészi beszéd paradoxona, hogy az előadóknak spontán beszédként kell meghangosítani interpretatív beszédet, azaz a köznapi beszéd látszatát kell kelteni. Kivétel ez alól a klasszikus művek bemutatása, ahol egyértelmű az irodalmi idézettség. Kérdőíves vizsgálat eredményei szerint a nézők napjainkban is úgy gondolják, a színészeknek példát kell mutatniuk a beszédük tisztaságával (Ausmann 2011).

A színészek és a hétköznapi beszélők beszédprodukcója közötti különbség egyaránt igazolható a feldolgozás, azaz a percepció és a beszéd létrehozása, azaz a produkció oldaláról. Az átlagos hallgató képes különbséget tenni a professzió-

nális és a hétköznapi beszélők megnyilatkozásai között. Kérdés, hogy milyen paraméterekben mutatható ki különbség a két csoport között. Eltérés igazolódott többek között a magánhangzók ejtésében és a műfaj tekintetében is. Míg a színészek magánhangzói jól elkülönültek egymástól, addig az átlagos beszélőknél átfedés volt kimutatható az egyes hangok között a formánsértékek vizsgálata alapján. A prózafelolvasásnál nagyobb, a versolvasásnál kisebb különbség volt kimutatható a két beszélői csoport között. Az egyik lehetséges magyarázat szerint a lírai műfaj előadásmódja erősen kötött, az átlagos beszélők is nagyobb figyelmet fordítottak az artikuláció pontosságára (Auszmann 2012).

## 6.2. A tanári beszéd

A tanári beszéd olyan multifunkciós eszköz, amelynek segítségével a tanár nevel, tudást közvetít, fejleszti a diákok nyelvi képességeit, és hat a személyiségük alakulására is (Antalné Szabó 2011). Az órai munka során a tanár kérdez, magyaráz, vezeti a tanár–diák interakciót, értékeli a diákok munkáját, megszervezi és vezeti az órai egyéni, csoportos és osztálymunkát stb. Kiemelten hangsúlyos, hogy a tanár kommunikációja hiteles legyen, megfelelő módon tudjon élni a beszéd verbális és nem verbális eszközeivel (Antalné Szabó 2006).

A számos funkciót kizárólag akkor képes a tanár a verbális kommunikációja segítségével ellátni, ha megnyilatkozásai jól érthetők, feldolgozhatók a különböző életkorú csoportok számára. A pedagógusok nyelvi kifejezőmódjával kapcsolatos követelmény, hogy alkalmazkodjon a hallgatóságához. A tanulói csoport életkorának, képességeinek, a tananyag nehézségének megfelelő legyen a kifejezőmód, közlésének grammatikai komplexitása, stílusa, szókinccse, beszédtempója, szünettartási jellemzői, artikulációjának pontossága, ritmusa, hangereje stb.

## 6.3. Beszédprodukción a médiában

A tömegkommunikáció főbb jellemzői, hogy egyirányú és közvetett, azaz a megnyilatkozó (például műsorvezető, bemondó) közlésére nem érkezik közvetlen módon válasz. Különbség van azonban a rádió és a televízió kommunikációs jellemzői között. Míg a rádióban mindent hanggal kell kifejezni, ezért a beszéd tempója és a hang érzelmi modulációja kiemelt szerepet kap, addig a televízióban a kép és a hang egyszerre közvetít információt. A normával, így a médianormával kapcsolatos kérdéskör számos tényezője máig megválaszolatlan maradt. Elmondható azonban, hogy alapkövetelményként van jelen a tiszta hangképzés és az értelem- és érzelemtükröző hangsúlyozás megléte. A rádiós beszéd normájának alapvető követelményei a hangtalan légzés, természetes, egészséges, jó rezonanciájú beszédhang, amely alapján rokonszenvesnek ítéljük meg a beszélőt. Elvárt a megfelelő hangerő, a tiszta hangképzés, a helyes beszédtempó, beszédritmus, hang-

súly, hanglejtés alkalmazása, a szöveg helyes tagolása, a megfelelő, témához és beszédhelyzethez illő beszédstílus megválasztása stb. (Balázs 2000).

A mindennapos tapasztalatok szerint részben a korábban felsorolt tényezők, részben a tartalom, szóhasználat, stílus, „mondatszerkesztés” figyelembevételével – azok tudatosítása nélkül is – rövid idő alatt eldönthető egy adott szövegrészlet alapján annak műfaja. A különböző műfajoknak eltérő a „hangzásvilága”, többek között a tartalomnak megfelelően változik a beszéd tempója, ritmusa, változik a beszélő felhasznált hangterjedelme, a szünetek hossza és minősége (Imre 2005).

#### **6.4. Társalgás, diskurzus, konverzáció, párbeszéd**

A hétköznapi megnyilatkozások jellemző beszédtypusa a párbeszéd vagy társalgás (más néven diskurzus vagy konverzáció) a mindennapi érintkezés legfőbb formája. Két résztvevő esetén párbeszédnek, több résztvevő esetén társalgásnak nevezzük az interakciót. Narratíván azt értjük, amikor a beszélő és hallgató szerepe nincs egyensúlyban, az egyik résztvevő egy vagy több beszédtemáról hosszabban közöl valamit. Párbeszédes társalgási forma esetén a két résztvevő felváltva közli gondolatait, egymás megnyilatkozásait is figyelembe véve. Gyakran a párbeszédes formát, a folyamatos beszélőváltásokat hosszabb-rövidebb narratívák szakítják meg, amikor az egyik fél hosszabban fejt ki gondolatait egy-egy témáról. Ide sorolhatók az irányított interjú helyzetben létrejött párbeszéd, a műsorvezető megnyilatkozásai vagy a terápiás beszélgetés is. Párbeszéd vagy társalgás során fokozottan van jelen a beszédpartnerhez való alkalmazkodás szükségessége (Iványi 2001).

A spontán beszédben létrejövő társalgásokat sokáig strukturálatlannak tekintették, később azonban számos kutatás készült az interakciós szabályok, a forduló, a beszédlépések, a társalgásokat felépítő szomszédsági párok (pl. kérdés-válasz) megjelenésének vizsgálatára, amely a folyamatot irányítja (vö. Grice 1975). A beszélőváltás elősegíti, hogy a résztvevők valamilyen rendszer szerint beszéljenek. A szó átvételére alkalmas helyet általában a beszédpartner jelzi verbális (pl. dallammenet, tempóváltozás, szünettartás) és nem verbális eszközökkel (Boronkai 2008). A társalgások leírása, jellemzése megkezdődött az elmúlt évtizedben, így leírás született például – monologikus szövegekkel összevetve – a szünetezés, a hümmögések mintázatairól és a beszélőváltások típusairól felnőtt és gyermek beszélőknél is (Markó 2005, 2006, 2007; Szabó–Tóth 2012).

#### **6.5. Anya–gyermek kommunikáció**

Az anya–gyermek kommunikáció egy speciális beszédhelyzet, a beszédalkalmazkodás egyik legkutatottabb területe. Dajkanyelvnek nevezik az édesanyának

vagy gondozónak a gyermekéhez intézett sajátos közléseit, amelyek akusztikai fonetikai tekintetben is eltérnek a felnőttek verbális kommunikációjától. Ez egyfajta sajátos nyelvhasználat, amely fonetikai szempontok szerint is eltér a felnőttekhez intézett megnyilatkozásokétól. A dajkanyelv jellemzője a felnőttekhez szóló beszédhez képesti magasabb hangfekvés, szélesebb frekvenciatartomány, változatos dallammoduláció, emelkedő végű intonációjú közlések, lassú beszédtempó. Ez a beszédmód elsősorban a nyelvelsajátítás első periódusa során meghatározó. Más kutatások azt igazolták, hogy a csecsemőkhöz, illetve kisgyermekhez intézett közlések számos kicsinyítő kifejezést, ismétlést és rövid közléseket tartalmaznak, jellemző kifejezési formájuk az utasítás, az erős expresszivitás, amelyet széles gesztusok és mimika kísérnek (Kátainé Koós 2006). A dajkanyelv egyidejűleg szolgálja az anyának arra irányuló törekvését is, hogy gyermekét bevezesse az adott kultúrába. Ilyen módon az anya és a kisgyermek között fennálló kapcsolat az antropológiai nyelvészeti vizsgálatok eredményei szerint kulturálisan erősen determinált, azaz az eltérő kultúrákban különböző formákban valósulhat meg (Réger 2002).

## 7. Összegzés

A beszéd típusok leírása és jellemzése nem teljes. A különböző beszéd típusok a főbb paraméterek mentén jól jellemezhetők, ám a beszéd megvalósítását és megvalósulását befolyásoló tényezők igen nagy számából adódóan gyakran egyedi az adott megnyilatkozás. A beszéd típusok jellemzőinek megismerése hozzájárulhat a beszédazonosítás pontosabb működéséhez, továbbá a mesterséges beszédkeltés több területén való felhasználásához. A beszélő személye, a téma, a hallgatóság, a szituáció és a közlés célja mentén a beszédprodukcióban adatolt akusztikai fonetikai különbségek felhívják a figyelmet az összehasonlító vizsgálatok (pl. életkori sajátosságok, genderkutatások, beszédpatológiás kutatások) esetén a beszéd típus szempontjának figyelembevételére is. A különböző beszéd típusok megismerése hozzásegítheti az olvasót azok tudatos használatához.

## Irodalom

- Antalné Szabó Ágnes 2006. *A tanári beszéd empirikus kutatások tükrében*. Magyar Nyelvtudományi Társaság, Budapest.
- Antalné Szabó Ágnes 2011. A tanári értékelő megnyilatkozások. *Beszéd kutatás 2011*. 121–133.
- Auszmann Anita 2011. *A művészi beszéd*. Szakdolgozat. ELTE BTK, Budapest.
- Auszmann Anita 2012. Máshogy beszélnek a színészek, mint az átlagos beszélők? In Váradi Tamás (szerk.): *VI. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest, 2012. 02.

03. MTA Nyelvtudományi Intézet, Budapest, 3–13. <http://www.nytud.hu/alknyelvdok12/proceedings12/auszmann2012.pdf>
- Balázs Géza 2000. A média nyelvi normája. *Magyar Nyelvőr* 124. 5–24.
- Beke András – Gósy Mária – Horváth Viktória – Gyarmathy Dorottya – Neuberger Tilda – Auszmann Anita 2015. Megakadások a beszédstílus függvényében: spontán narratívákban és társalgásokban. In Gósy Mária (szerk.): *Diszharmóniás jelenségek a beszédben*. MTA Nyelvtudományi Intézet, Budapest, 85–100.
- Bóna Judit 2009. Az idős életkor tükröződése a magánhangzók ejtésében. *Beszédkutatás* 2009. 76–87.
- Bóna Judit 2013. A beszédszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszédkutatás* 2013. 60–75.
- Boronkai Dóra 2008. Konverzációelemzés és anyanyelvtanítás I. *Anyanyelv-pedagógia* 2008/2. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=60> [Letöltés: 2012. február 19.]
- Gocsál Ákos – Huszár Ágnes 2003. Csábító hangok. *Beszédkutatás* 2003. 9–18.
- Gocsál Ákos 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutatás* '98. 122–35.
- Gósy Mária – Beke András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória 2009. A beszéd természetességéről alkalmazott fonetikai szempontból. *Beszédkutatás* 2009. 170–181.
- Gósy Mária – Krepsz Valéria 2015. Magánhangzók temporális jellemzői az idő múlásának függvényében. *Beszédkutatás* 2015. 53–65.
- Gósy Mária 1998. A beszédtervezés és beszédkivitelezés paradoxona. *Magyar Nyelvőr* 122. 3–15.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125. 478–488.
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszédkutatás* 2003. 19–44.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. In Prószéky Gábor – Váradi Tamás (szerk.): *Általános Nyelvészeti Tanulmányok XXIV. Nyelvtchnológiai kutatások*. Akadémiai Kiadó, Budapest, 329–349.
- Grice, Paul 1975/2001. A társalgás logikája. In Pléh Csaba – Síklaki István – Terestyéni Tamás (szerk.): *Nyelv – kommunikáció – cselekvés*. Osiris Kiadó, Budapest, 213–227.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukcóra. *Beszédkutatás* 2007. 108–121.
- Gyarmathy Dorottya 2015. Diszharmóniás jelenségek, megakadások a beszédben. In Gósy Mária (szerk.): *Diszharmóniás jelenségek a beszédben*. MTA Nyelvtudományi Intézet, Budapest, 9–48.
- Imre Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- Iványi Zsuzsanna 2001. A nyelvészeti konverzációelemzés. *Magyar Nyelvőr* 125. 74–93.
- Kátainé Koós Ildikó 2001. Az anya személyiségjegyeinek hatása a csecsemő preverbális kommunikációjára. *Beszédkutatás* 2001. 101–113.
- Kátainé Koós Ildikó 2010. A dajkanyelvi tényezők szerepe az anyanyelv-elsajátítás csecsemőkori szakaszában. In: Balaskó Mária – Kovács László (szerk.): *Hálózat kutatás*. Tinta Kiadó, Budapest, 184–194.

- Laczkó Mária 2000. Miért nem szeretnek felelni a diákok? Egy kísérlet tapasztalatai és tanulságai. *Magyartanítás* 2. 28–34.
- Laczkó Mária 2008. Hogyan minősítik a tizenévesek beszédét a diákok és a tanárok? *Anyanyelv-pedagógia* 2008/3–4. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=87>
- Levelt, Willem J. M. 1989. *Speaking: From intention to articulation*. A Bradford Book. The MIT Press, Cambridge (Massachusetts)–London (England).
- Lindblom, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In Hardcastle, William – Marchal, Alain (eds.): *Speech production and speech modeling*. Kluwer, Dordrecht, 403–439.
- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata*. Doktori disszertáció. ELTE, Budapest.
- Markó Alexandra 2006. Nonverbális vokális jelek a társalgásban. *Beszédkutató* 2006. 57–68.
- Markó Alexandra 2007. A társalgás hangtana. *Magyar Tudomány* 164/1. 54–57.
- Réger Zita 2002. *Utak a nyelvhez. Nyelvi szocializáció – nyelvi hátrány*. Soros Alapítvány – MTA Nyelvtudományi Intézet, Budapest.
- Szabó Ágnes – Tóth Andrea 2012. A beszélőváltások sajátosságai óvodáskorú gyermekek társalgásában. *Beszédkutató* 2012. 234–245.
- Váradí Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutató* 2010. 100–109.
- Váradí Viola 2013. *A spontán beszéd szegmentálása produkciós és percepciós szempontból*. Doktori disszertáció. ELTE, Budapest.
- Wacha Imre 1974. Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti Tanulmányok* X. 203–16.

### Kérdések, feladatok

1. Hogyan csoportosíthatók az egyes beszéd típusok?
2. Hogyan jellemezhető a spontán beszéd?
3. Mely tényezők határozzák meg az egyes beszéd típusokat?

## Szociofonetika

A szociofonetika a szociolingvisztika és a fonetika határterülete, amely szociolingvisztikai tényezők (pl. nem, társadalmi rétegződés) beszédre gyakorolt hatásait vizsgálja. Tárgykörébe beletartozik a beszédprodukción és a beszédpercepció vizsgálata is. Alapjait Labov és munkatársai fektették le, akik főként az amerikai angol magánhangzó-variánsait vizsgálták különböző dialektusokban (Foulkes et al. 2010). A szociofonetikához számos kutatási terület tartozik: a **kétnyelvűség** fonetikai vetületei; az **akcentusjelenségek**; a **nyelvi variabilitás** (a beszélőn belüli és a beszélők közötti) fonetikai vetületei; a **társadalmi rétegek** sajátosságai, a kisebbségi nyelvhasználat; a **nyelvjárások** hangtani vonatkozásai; a beszédben megjelenő **nemi** és életkori különbségek; a nyelvi szocializáció, beszéd típusok stb. (Foulkes et al. 2010) (1. ábra). A vizsgálatok a beszéd szegmentális és szupraszegmentális szintjére is kiterjednek. Szegmentális szinten elemzik pl. az egyes nyelvjárásokban megjelenő mássalhangzók, illetve magánhangzók ejtősváltozatait. Szupraszegmentális szinten ide tartozik pl. az alaphangmagasság vizsgálata a beszélő nemének vagy szexuális identitásának kapcsán, illetve a beszédtempó vizsgálata különböző életkorokban.



1. ábra. A szociolingvisztika és a fonetika határterülete

A percepció vizsgálatok célja a fonetikai változók és a beszélő/nyelv szociolingvisztikai sajátosságai közötti kapcsolatok vizsgálata, a hallgató szempontjából (Baranowski 2013). Ide tartozik pl. egy beszélő dialektusának azonosítása a beszéde alapján vagy a gyermekek nemének meghatározása a hangjukból.

A tanulmányban többek közt a következőkről olvashat: Lehetséges-e azonosítani a beszélő anyanyelvét az akcentusa alapján? Hogyan hat a kétnyelvűség a beszédre? Milyen ejtőingadozások figyelhetők meg különböző nyelvhasználók-

nál? Milyen hatással van az írásképe a kiejtésre? Milyen hangzásbeli különbségek jellemzik a nyelvjárásokat? Milyen különbségek fedezhetők fel a férfiak és nők beszédében?

### 1. A többnyelvűek kiejtése

A kétnyelvűség – illetve többnyelvűség – sokféleképpen hat a beszédre. Befolyásolja a beszéd szegmentális és szupraszegmentális szintjét: gondoljunk csak arra, hogy amikor idegen nyelven beszélünk, bizonyos hangokat másképp ejtünk ki, mint az anyanyelvi beszélők; vagy másképp hangsúlyozunk, más dallamot használunk pl. a kérdő mondatokban – anyanyelvünkhöz hasonlóan. Ezek a különbségek a második nyelv és anyanyelvünk különbségeiből adódnak. A nyelvek artikulációs bázisa (egy nyelvre jellemző hangképzésbeli sajátosságok összessége) ugyanis különbözik, és ez a beszédben akcentushoz vezet. (Az akcentus mértéke természetesen függ attól, hogy mikor kezdődik a második nyelv elsajátítása/tanulása, befolyásoló tényező a beszélő életkora. A gyermekek még képesek arra, hogy akcentus nélkül sajátítsanak el egy idegen nyelvet, ugyanakkor felnőttkorban ez már kevésbé jellemző.)

Akcentusnak nevezzük azokat a kiejtésbeli sajátosságokat, amelyek egy nyelv használatakor annak normatív kiejtéséhez képest eltérésként jelentkeznek egy nem a nyelvközösséghez tartozó, de annak nyelvén megnyilatkozó egyén beszédében (Kassai 1995). Vizsgáljuk meg egy kicsit közelebbről, mit is jelent ez a magyar mint idegen nyelv szempontjából. Akcentusjelenségekkel a magyart idegen nyelvként beszélők esetében is találkozhatunk. Azonban arra a kérdésre, hogy pusztán az akcentusa alapján azonosítható-e a beszélő nemzetisége, nem tudunk biztos választ adni. Bóna (2004) vizsgálatában 12, különböző nyelvcsaládokból származó beszélő (pl. japán, horvát, francia, német) felolvasását rögzítette, majd egy percepciósi tesztben arra kérte a résztvevőket, hogy azonosítsák a beszélő anyanyelvét, indokolják meg döntésüket, illetve ítélik meg, mennyire akcentusos az illető beszéde. A tesztet kitöltőket két csoportba osztotta: az egyik csoport semmilyen segítséget nem kapott, a másik ugyanakkor megkapta a lehetséges nyelvek felsorolását. Az anyanyelv felismerése szempontjából a helyes válaszok száma alacsony volt. Az a csoport teljesített jobban, amelyik előre megkapta a lehetséges válaszokat. A válaszadók 33%-a ismerte fel helyesen a japán beszélőt, 25%-a a németet, 21%-a pedig az oroszot. A legkevésbé a lengyel beszélő anyanyelvét azonosították helyesen (1%). A döntések hátterében a megkérdezettek szerint a következők álltak: a japán esetében a magas alaphang, a szótagoló és éneklő beszédmód és a jellegzetes hanglejtés árulkodott a beszélő nemzetiségéről. A szláv nyelveknél többen felismerték a „szláv” voltot, a „darabos” ejtés és a mássalhangzó-torlódások miatt. A lengyel beszélő

anyanyelvét a finnugor nyelvekkel keverték össze, ennek oka a beszélő jó kiejtése lehetett. Az olasz és francia beszélőket szlávként azonosították, részben a gyors tempó, részben a szövegben előforduló *szláv* szó miatt (a szemantika ebben az esetben fontosabbnak bizonyult a kiejtésnél). A német beszélőt többen franciának vélték, a nem elvárt hangsúlyozás, hehezetes ejtés és a szóvégek „elharapása” miatt. Bóna az eredményekből arra következtetett, hogy a magyarban pusztán az akcentus alapján nem állapítható meg az idegenek anyanyelve. A kutatásba olyan tanárokat is bevont, akik magyar mint idegen nyelvet tanítanak, eredményeik azonban hasonlóak voltak a percepciók tesztet kitöltőkhöz (kivételet képeztek ez alól az olyan anyanyelvű beszélők esetei, akik nagyobb számban jönnek magyart tanulni, pl. a japánok, őket jobb arányban azonosították).

A beszéd szegmentális szintjét – ezen belül a magánhangzókat – érintő akcentusjelenségekkel foglalkozott Menyhárt (1999, 2002b). Spontán beszéden végzett kutatásokat szláv, főként bolgár anyanyelvű beszélőkkel. A bolgár nyelv és a magyar artikulációs bázisa eltér egymástól, előbbiben pl. nincsenek rövid és hosszú magánhangzók, ezért a bolgár anyanyelvű adatközlők a magyarban a hosszú magánhangzókat is rövidebben ejtették. Olyan hangok esetében, mint az [a:] és [ɛ], amelyeknek nincs ekvivalense a bolgárban, artikulációs nehézségekbe ütköztek, beszédükben ezáltal több hang összemosódott. Menyhárt kutatásában – szintén a magánhangzók vizsgálata kapcsán – azt is megfigyelte, hogy az artikuláció nem állandó bizonyos magánhangzók ejtésekor, ami szintén eltérésekhez vezethet.

Menyhárt (2002a) azt is elemezte, hogy mi jellemzi az egynyelvű és kétynyelvű gyermekek beszédtempóját. 60 gyermek (közülük 20 egynyelvű magyar, illetve 20 egynyelvű bolgár, valamint 20 kétynyelvű magyar-bolgár) beszédét elemezte egy képleírási feladat segítségével. Hipotézise igazolódott, a kétynyelvű gyermekek artikulációs tempója lassabb volt, mint az egynyelvűeké. Ugyanakkor a szünettartásuk nem különbözött. Menyhárt szerint ez a kétynyelvűsége vezethető vissza, ugyanis a kétynyelvű gyermek két egymással szimultán nyelv birtokosa, ami beszédében lassabb tervezési és kivitelezési folyamatokhoz vezet.

Németh (2012) kutatásában arra kereste a választ, hogy meg lehet-e különböztetni a valódi és az imitált akcentust. Orosz és francia beszélők felolvasott mondatait játszotta le a hallgatónak, és arra kérte őket, becsüljék meg a beszélő anyanyelvét, illetve azt, hogy eredeti vagy imitált akcentusról van szó. Az eredmények szerint az eredeti akcentus felismerése 10%-kal sikeresebb volt az imitálténál. Az is bebizonyosodott, hogy azok a hallgatók, akik tudnak az adott idegen nyelven, jobban azonosítják a beszélő anyanyelvét, illetve az eredetiséget. Németh (2013) azt is elemezte, hogy milyen különbségek vannak a magánhangzók ejtésében az imitált és az eredeti akcentus között. Az orosz anyanyelvű beszélők a magánhangzók nagy részét zártabban ejtették, mint az

imitátorok, illetve a magánhangzók redukciója jellemző eleme volt az orosz akcentusnak.

A kétnyelvűség a beszédben megjelenő megakadásokat is érintheti. Navracsi (2007) 22 kétnyelvű felnőtt beszélővel készített vizsgálatot, amelyben megkérte őket, hogy meséljenek el képsorozatokat mindkét nyelvükön. A beszélők között voltak, akik korai életkorban sajátították el a nyelvet, és voltak, akik később. Eredményei szerint a korai kétnyelvűeknél mind a két nyelv úgy működik, mint az egynyelvűeknél az első nyelv. Ennek megfelelően más-más megakadások jelentkeztek a csoportok között. A korai kétnyelvűeknél pl. gyakoribbak voltak a szerkesztési és grammatikai hibák.

Az akcentusjelenségek vizsgálata az alkalmazott nyelvészet számára több szempontból is hasznos. Egyrészt pedagógiai célok mentén: a magyar mint idegen nyelv oktatásában, ahhoz kapcsolódóan a kiejtés tanításában. A beszédtechnológia számára is fontos, hiszen a beszélőfelismerő és nyelvfelismerő rendszerek pontosabb működése megkívánja az akcentusos beszéd felismerését és hatékony kezelését. Ugyanakkor a kriminalisztikában is hasznosítható a profilalkotás kapcsán (pl. a valódi és imitált akcentus meghatározásakor). Mindehhez azonban szükség van a magyarban jelentkező akcentusjelenségek további átfogó és széles körű vizsgálatára.

## 2. Nyelvi variabilitás és nyelvi változás

A nyelv egyik legérdekesebb jelensége a variabilitás, vagyis az egyénen belüli variációk, amelyeket használni képesek (Gecső 2004). Ezek közé tartozik például, hogy kétszer nem tudunk kiejteni „ugyanúgy” egy mondatot (Gósy 2004).

Gósy Mária (2009) az *akkor* szó akusztikai fonetikai sajátosságait figyelte meg az adott funkciójától (határozószó, töltelékszó) függetlenül. 5 fiatal férfi és 5 fiatal nő spontán beszédét elemezte (255 perc), mely 286 *akkor* szót tartalmazott. Fő kérdése az volt, hogy mely paraméterekben lesz eltérés a kiejtés variabilitása következtében. Az eredmények nagy artikulációs különbségeket igazoltak. A gyakori szóelőfordulás a kiejtés nagymértékű változatosságához vezet a beszédben. Eredményei felhasználhatók a beszédfelismerő rendszerek számára is.

A nyelvi tudatosságot, normát és rendszerkényszert vizsgálta Gósy Mária és Kovács Magdolna (2003) a *h* hang esetében 43 adatközlővel. A kísérlet célja az volt, hogy az ejtésingadozásokról, illetve a nyelvhasználati jellemzőkről képet kapjanak. A kísérleti adatok igazolták, hogy a magyar anyanyelvű adatközlők nyelvhasználati tapasztalata és a „norma” szigorúbb vagy elnézőbb alkalmazásai hatással vannak a *h* viselkedésére. Következtetesként megfogalmazták, hogy a *h* végű szavak esetén a gyakoriság a vártnál jobban meghatározza a vizsgált

szavak produkciós és észlelési sajátosságait. Más az ejtésben és a percepcióban a *h* végű tőszók, a toldalékolt szavak és a szóösszetételek megjelenése. Kísérletük-ből kiderül, hogy az észlelés a konkrét akusztikai-fonetikai adatok alapján történik, azonban a beszélők nyelvi tudatossága is hatással van döntéseikre.

Nyelvi változás következik be akkor, amikor egy közösség tagjai átvesznek és normaként elfogadnak egy új nyelvi formát, melyet korábban csak egy másik kisebb csoport használt egy beszélőközösségen belül (Coates 1992). Wardhaugh (1995) szerint a nyelvi változásban a két legfontosabb tényező a változások megléte és a társadalmi nyomás, illetve ezek kölcsönhatása. A változásokat éppen ezért gyakran hibának gondolja a szinkrón nyelvhasználat. Crystal (1998) megfogalmazza, hogy nyelvtől függetlenül a szókincs és a kiejtés változásai a legfeltűnőbbek a kortárs nyelvhasználók számára.

Gósy Mária és Gyarmathy Dorottya (2008) a megakadásjelenségeket tanulmányozták bizonyos időtávlatból. Az 1940-es és 1950-es évekből származó, spontán beszédet rögzítő Hegedűs-archívum anyagát vetették össze mai beszélőktől rögzített saját spontánbeszéd-anyagukkal. Arra a kérdésre keresték a választ, hogy a spontán beszédtervezés és kivitelezés folyamataiban igazolható-e eltérések az ötven-hatvan évvel ezelőtti és a mai beszélők között. A kutatásban 28 archívumi beszélő anyagát és 28 mai beszélő anyagát elemezték. Eredményeik szerint a mai beszélőkre jellemzőbbek a megakadásjelenségek. A töltelék-szavak, nyújtások és szó belseji szünetek hasonlóképpen jelen vannak a régi és a mai beszélőknél. Szignifikáns eltérést a hezitálások, az ismétlések és a hiba típusú megakadásjelenségek esetében adatoltak, amelyek a mai beszélőknél fordultak elő lényegesen nagyobb arányban. A hibajelenségek száma háromszorosá-ra növekedett, amely arra utal, hogy a beszédtervezés folyamatában több problémával szembesülnek a mai beszélők, s ezeket különféle stratégiákkal igyekeznek feloldani. A mentális lexikon aktiválásából adódó nehézségek (téves kezdés, téves szótalálás, lexikális előhívás nehézségei) az archívumi anyagban fordultak elő nagyobb arányban. Erre utalnak a régi beszélőknél adatolt gyakori nyújtások is. A mai beszélőknél ezzel szemben a megfelelő nyelvi forma megtalálása okozott több gondot, melyet a grammatikai hibák relatíve nagy száma, valamint a hezitálás erőteljes növekedése igazolt. A különbségek okaként a szerzők megnevezik a beszédtempó gyorsulását, a rádió és a televízió hatását, valamint a produkciós gazdaságosságra való törekvést.

Gósy Mária és Horváth Viktória (2009) kutatásukban arra a kérdésre keresték a választ, hogy hogyan tükrözi a kiejtés a nyelvi funkció változását. 17 beszélő spontán beszédében elemezték a *tehát* és az *ilyen* szavak előfordulási arányát és funkciófüggő kiejtését. A 459 perces anyag 461 *tehát* és 154 *ilyen* szót tartalmazott eltérő alakváltozatokban. Eredményeik szerint napjainkban végbemegy a *tehát* kötőszó és az *ilyen* névmás funkciókör-bővülése, mindkettő gyakran realizálódik töltelék-szóként, ami a beszédtervezés problémáira utalhat. A funkcióbő-

vülés a lexémák akusztikai-fonetikai szerkezetében is nyomon követhető. Nyelvi funkciótól függően statisztikailag szignifikáns különbséget a szavak időtartamában, valamint a magánhangzók második formánsainak értékében kaptak. Töltelékszóként a *tehát* rövidebb, míg az *ilyen* hosszabb időtartamban valósul meg. Az eredmények alapján arra a következtetésre jutottak, hogy az újabb nyelvi eszközökre azért van szüksége a beszélőnek, hogy áthidalhassa a gyakori hezitálást (őzést). A töltelékszók a közlésekbe illeszkednek, így a hallgatók számára nem tűnnek fel olyan mértékben, mint a hezitálások. Eredményeik hozzájárulnak a spontán beszédben zajló szinkrón változások bemutatásához.

### 3. A különböző társadalmi rétegek beszéde

G. Varga Györgyi (1972) a budapesti nyelvhasználat társadalmi rétegződését elemezte. 47 nővel és 53 férfivel végzett kutatásában vizsgálta az egyedi alakváltozásokat, az íráskép hatását a kiejtésre, a felső nyelvállású magánhangzók rövidülési tendenciáját, valamint a nyelvi magatartást. Eredményeiből kiderül, hogy az életkor csupán kismértékben befolyásolja a nyelvhasználatot, azonban az iskolai végzettség egyértelműen hatással van a kiejtésre, a beszédre. A nyelvi magatartást (a nyelvi jelenségek megítélése) az anyanyelvi és általános műveltség, a nyelvi neveltetés, a nyelvi előélet, a szűkebb közösség (család, iskola, munkahely, baráti kör) szokásai irányítják.

Egy attitűdvizsgálat eredménye szerint az iskolázottságra vonatkozó benyomásra hatással van az artikuláció minősége (Barra 2010), bár korántsem olyan erős a következtetés a hangzóejtés minőségéből az iskolázottságra, mint az angolban.

### 4. A nyelvjárási beszéd

A különböző társadalmi rétegek, illetve a társadalmi státuszok hatással vannak beszédünkre. A regionális köznyelv az utóbbi fél évszázad nagy társadalmi változása, átmenet a köznyelv és a tájnyelv között. Olyan nyelvváltozat, amelyet tájnyelvi elemek színeznék, például a tájnyelvi ejtémód, illetve ritkán nyelvjárási szavak. A köznyelv választékos formáját mindennapi nyelvhasználatban csak kevesen beszélik, az emberek többsége a regionális köznyelvet használja (Kiss 1995).

Nyelvjárások alatt a területi alapon elkülönülő nyelvváltozatokat értjük, melyek elsősorban a falvakban élők beszédére jellemzőek, és a köznyelvtől eltérő formákat tartalmaznak. A nyelvjárások különféleképpen hatnak a beszédre, gazdagabbá tehetik nyelvhasználatunkat, de gátolhatják is azt. A nyelvjárási nyelv-

használat egyéneenként variálódhat, hiszen ugyanaz az ember másként beszél eltérő beszédhelyzetekben (Kiss 2003).

A szociofonetika tárgykörébe a nyelvjárások hangtani vonatkozásai tartoznak. A magyar nyelvre jellemző, hogy nyelvjárásai között nincs nagy különbség. A magyarban tíz nyelvjárási régiót különböztetünk meg: a nyugat-dunántúli, közép-dunántúli-kisalföldi, dél-dunántúli, dél-alföldi, palóc, Tisza–Körös-vidéki, északkeleti, mezőségi, székely, moldvai nyelvjárási régió (2. ábra).



2. ábra. A mai magyar nyelvjárásterületek

[http://www.bihalbocs.hu/dialektologia\\_demo/nyelvjaraster.html](http://www.bihalbocs.hu/dialektologia_demo/nyelvjaraster.html)

A következőkben néhány hangtani jellegzetességet mutatunk be (Kiss 2011).

A nyugat-dunántúli nyelvjárási régió beszélőire jellemző, hogy fonémaként megmaradt a zárt *ĕ* (*embĕr*), továbbá erős az *l*-ezés (*illen* 'ilyen'). A köznyelvi *ú*, *ű*, *í* általában rövid (*kut*, *tűz*, *viz*).

A közép-dunántúli-kisalföldi nyelvjárási régió beszélőire szintén jellemző a zárt *ĕ* használata. Kiejtésük közelebb áll a köznyelvhez, mint a nyugat-dunántúli nyelvjárási régió beszélőie.

A dél-dunántúli régiót erős *ö*-zés jellemzi (*löhet* 'lehet').

A dél-alföldi nyelvjárási régió fő jellemzője szintén az *ö*-zés (*embör*, *mögövette* 'megvette'). A köznyelvi *ó*, *ő* helyén gyakori az *ú*, *ű* (*rúzsza* 'rózsza').

A palóc nyelvjárási régió beszélői az ajakkerekítéses köznyelvi *a* ajakkerekítés nélküli változatát ejtik, a köznyelvi *á*-val szemben pedig hosszú *a*-t használnak. Ezen nyelvjárásban még ma is ejtik az *ly*-t.

A Tisza–Körös-vidéki nyelvjárási régióra jellemzők a záródó kettőshangzók,

és erős az *l*, *r*, *j* nyújtó hatása (*bórnyu* 'borjú'), valamint jelen van az *i*-zés (*kerítés* 'kerítés', *níz*).

Az északkeleti nyelvjárási régióban nincs meg a zárt *ë*. Igen erős az *l*, *r*, *j* nyújtó hatása (*körte*).

A mezőszéki nyelvjárási régió beszélőinél gyakran megrövidülnek a hosszú magánhangzók (*apro*, *kesztyü*). Jellemző a nyílt *a*-zás (*hasszú* 'hosszú', *sak* 'sok').

A székely nyelvjárási régió zárt *ë*-ző, az *ó*, *ő*, *é*-nek kettőshangzós realizációi jelennek meg, valamint jellemző az *ö*-zés és az *o*-zás (*körösz* 'kereszt', *havas* 'havas').

A moldvai nyelvjárási régió többségében zárt *ë*-ző, és több mássalhangzót jésítetten ejtenek. Csángó sajátosság az *sz*-elés is, vagyis a köznyelvi *s*-nek *sz*-ként való ejtése (*maszt* 'most') (Kiss 2011).

## 5. Nemek közötti különbségek a beszédben

A beszéd egyes jellemzőit a beszélő anatómiai sajátosságai határozzák meg. Fiziológiailag az emberi hangképzés alapja a kilégzés okozta levegőáramlás, a hangszalagok fonációs beállítottsága és rezgése, illetve a toldalékcső rezonátor-tevékenysége (Gósy 2004). A beszédképzésben részt vevő szervek (tüdő, gége, toldalékcső) méretét meghatározza a beszélő testalkata, neme és életkora is. A szervek az életkor előrehaladtával változnak, ezért a beszéd akusztikai lenyomata is változik. A férfiak és nők beszédének egyik legfeltűnőbb különbsége az alaphangmagasság eltérése: a felnőtt férfiak alaphangmagassága 80–140 Hz, a nőké 160–260 Hz közötti, míg a gyermekeké az 500 Hz-et is elérheti (Kassai 2006). Az alaphangmagasság ugyanakkor a beszélő életkorától is függ.

A gyermekek beszédével foglalkozó kutatások ellentmondásosak abban a tekintetben, hogy van-e különbség például a fiúk és a lányok alaphangmagasságában, magánhangzók formánsértékeiben (formáns: a zöngé rezonátorüregben felerősített felharmonikusa), vagy bármilyen más tényezőben, amely lehetővé tenné a hallgató számára a beszélő nemének azonosítását. Több kutatás szerint (Lee et al. 1999; Vorperian et al. 2009) a pubertáskor előtt nincs különbség a toldalékcső hosszában a nemek között. Perry és munkatársai (2001) ezzel szemben azt találták, hogy a magánhangzók formánsértékei már négyéves korban is mutatnak nembeli eltéréseket. Egy, a magyar nyelvre végzett kutatás az [ɔ], [a:], [ɛ] hangok esetében talált eltérést a fiúk és lányok ejtésében 5 és 10 éves korban (Tóth 2014).

Az alaphangmagasság tekintetében 5–6 éves korban nem találtak különbségeket a nemek között a kutatók (Hasek 1980), míg 7 és 10 éves kor között már megkezdődött az alaphangmagasság csökkenése, amely a két nemnél valószínű-

síthetően eltérő ütemben zajlik. Whiteside és Hodgson szerint (1999) a lányoknál az alaphangmagasság csökkenése 6 és 10 éves kor között, míg a fiúk esetében később, 8 és 10 éves kor között jelentkeznek. Az eltérő ütemű változással magyarázható az, hogy több vizsgálatban találtak a lányokénál magasabb alaphangot a 7 és 10 éves fiúknál (Deme 2012; Tóth 2014).

A gyermekek nemének beszédalapú azonosítása nem könnyű feladat. Traunmüller és Bezooijen (1994) 5 és 11 éves kor között vizsgálta a nemek felismerését. Míg az 5 éveseknél mindössze 19%-ban azonosították sikeresen a beszélő nemét, 11 éves korra a duplájára nőtt a helyes felismerések aránya. Az azonosításban a szerző szerint a formánsok és az alaphangmagasság mellett egyes szupraszegmentumok is szerepet játszhatnak. Amir és munkatársai (2012) 8 és 10 éves gyermekek beszédének elemzésekor azt találták, hogy a nemek felismerési aránya 70%-ról 80%-ra nő az életkor előrehaladtával. Egy magyar nyelvű kutatásban az 5 éveseknél a válaszadók 69%-ban, a 10 éveseknél pedig 85%-ban azonosították a beszélő nemét (Tóth 2014).

A beszéd alaphangja és a magánhangzóformánsok mellett elemezték még a beszéd tempóját is a férfiaknál és a nőknél. 1994-ben Byrd kutatásában azt találta, hogy a felnőtt nők lassabban beszélnek, mint a férfiak. Ennek ellentmondó eredményt kapott Gocsál (2001), aki szerint a nők és a férfiak sem a beszédtempó, sem az artikulációs tempó értékeiben nem különböztek egymástól. Mivel a nők szünettartása egységesebb volt, és kevesebb szünetet is tartottak, a szerző feltételezte, hogy azonos tempó esetén a kevesebb szünetet tartót a hallgatók gyorsabb beszédűnek ítélik meg, mint a több szünetet tartót. Markó (2014) eredményei megerősítették, hogy a férfiak hosszabb szüneteket tartanak beszédükben, mint a nők, illetve azt is, hogy az életkor előrehaladtával a szünetek egyre hosszabbá válnak. Mint látható, ebben a kérdésben az eltérő vizsgálati módszerek különböző eredményeket hoztak, ezért további, nagyszámú adatközlővel készült kutatások adhatnak választ erre a kérdésre.

A nők és a férfiak beszédét elemezve az is felmerülhet kutatási kérdésként, hogy vajon mennyi és milyen megakadások tarkítják a beszédüket. A nők beszédében kevesebb megakadás (a beszédet megszakító elem, amely nem járul hozzá a közlés tartalmához) fordult elő (Horváth 2007). Voltak ugyanakkor olyan megakadásjelenségek, amelyek csak a férfiakra voltak jellemzők: az újraindítás és a grammatikai összehangolatlanság. Bóna (2014) vizsgálatában az életkor, a nem és a beszéd típus függvényében elemezte a megakadásokat. Tendenciaszerűen a fiataloknál a nők, az időseknél pedig a férfiak csoportjában talált kevesebb megakadást.

A férfi és női beszédhangok eltéréseiben az anatómiai különbségek mellett természetesen egyéb tényezők is szerepet játszhatnak. Néhány példa: tanult genderspecifikus viselkedésminták; az artikuláció pontossága (Diehl et al. 1996);

a testalkat (Xue et al. 2010); a gége leszorításával elért mélyebb alaphang a fiatal fiúknál (Huszár 2009).

## 6. Összegzés

Mint ahogy azt a tanulmányban láthattuk, a szociofonetika számos kutatással gazdagítja a beszéd tudományát. A nemek közötti különbségek, a többnyelvűség vizsgálata, a nyelvi variabilitás, a különböző társadalmi rétegek beszéde, illetve a nyelvjárási beszéd adják fő kutatási területeit. Ezek az elemzések amellet, hogy hozzájárulnak a vizsgált jelenségek empirikus leírásához, más szempontból is fontosak lehetnek: pedagógiai, beszédtechnológiai vonatkozásban. A nyelvjárási beszéd kapcsán például összekapcsolhatók az anyanyelvi neveléssel, a nemek beszédbeli különbözőségeinek vizsgálatakor pedig a beszédtechnológiával.

## Irodalom

- A mai magyar nyelvjárásterületek.* [http://www.bihalbocs.hu/dialektologia\\_demo/nyelvjaraster.html](http://www.bihalbocs.hu/dialektologia_demo/nyelvjaraster.html). (A letöltés ideje: 2016. január 1.)
- Amir, Ofer – Engel, Merav – Shabtai, Esther – Amir, Noam 2012. Identification of children's gender and age by listeners. *Journal of Voice* 26/3. 313–321.
- Baranowski, Maciej 2013. Sociophonetics. In: Bayley, Robert – Cameron Richard – Lucas, Ceil (eds.): *The Oxford Handbook of Sociolinguistics*. Oxford University Press, Oxford, 403–424.
- Barra Mária 2010. A felolvasás „mellékhatásai”: egy attitűdvizsgálat. *Alkalmazott Nyelvtudomány* 1–2. 81–96.
- Bóna Judit 2004. Az akcentus az előfeltevéseink tükrében. *Alkalmazott Nyelvtudomány* 2004/1. 31–39.
- Bóna Judit 2014. Megakadások az életkor, a nem és a beszédtypus függvényében. *Beszédkutatás 2014*. 123–143.
- Byrd, Dani 1994. Relations of sex and dialect to reduction. *Speech Communication* 15. 39–54.
- Coates, Jennifer. 1992. *Women, Men and Language*. Longman, London.
- Crystal, David 1998. *A nyelv enciklopédiája*. Osiris Kiadó, Budapest.
- Deme Andrea 2012. Óvodások magánhangzóinak akusztikai jellemzői. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngeszűrésig*. ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest, 77–99.
- Diehl, Randy L. – Lindblom, Björn – Hoemeke, Kathryn A. – Fahey, Richard P. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24. 187–208.
- Foulkes, Paul – Scobbie, James M. – Watt, Dominic J. L. 2010. Sociophonetics. In Hardcastle, William – Laver, John – Gibbon, Fiona (eds.): *Handbook of phonetic sciences*. Blackwell, Oxford, 703–754.

- Gecső Tamás 2004. A hová? és a honnan? különbségének pszicholingvisztikai természetéről a magyarban. In Gecső Tamás (szerk.): *Variabilitás és nyelvhasználat*. Segédkönyvek a Nyelvészet Tanulmányozásához XXXIV. Tinta Könyvkiadó, Budapest, 114–123.
- Gocsál Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszédkutatás 2001*. 61–72.
- Gósy Mária – Gyarmathy Dorottya 2008. A nyelvhasználati változás egy jelensége. *Magyar Nyelvőr* 132. 206–222.
- Gósy Mária – Horváth Viktória 2009. Hogyan tükrözi a kiejtés a nyelvi funkció változását? In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában – grammatika a diskurzusban*. Tinta Könyvkiadó, Budapest, 37–45.
- Gósy Mária – Kovács Magdolna 2003. Nyelvi tudatosság, norma és rendszerkényszer a *h* esetében. *Magyar Nyelvőr* 127/4. 439–455.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2009. Szóejtés és szóészlelés: változatosság és adaptálódás. *Beszédkutatás 2009*. 46–76.
- Hasek, Carol S. – Singh, Sadanand – Murry, T. 1980. Acoustic attributes of preadolescent voices. *Journal of the Acoustical Society of America* 68. 1262–1265.
- Horváth Viktória 2007. Vannak-e „női” és „férfi” megakadásjelenségek a spontán beszédben? *Magyar Nyelvőr* 131/3. 315–323.
- Huszár Ágnes 2009. *Bevezetés a gendernyelvészetbe*. Tinta Könyvkiadó, Budapest.
- Kassai Ilona 1995. Pszicho-szociolingvisztikai jegyzetek az akcentusról. *Általános Nyelvészeti Tanulmányok* 18. 103–115.
- Kassai Ilona 2006. Fonetika. In Kiefer Ferenc (főszerk.): *Magyar nyelv*. Akadémiai Kiadó, Budapest, 789–835.
- Kiss Jenő 1995. *Társadalom és nyelvhasználat*. Nemzeti Tankönyvkiadó, Budapest.
- Kiss Jenő 2003. *Magyar dialektológia*. Osiris Kiadó, Budapest.
- Kiss Jenő 2011. Nyelvjárások, regionális nyelvváltozatok. In Kiefer Ferenc (főszerk.): *Magyar nyelv*. Akadémiai Kiadó, Budapest, 517–549.
- Lee, Sungbok – Potamianos, Alexandros – Narayanan, Shrikanth 1999. Acoustics of children’s speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America* 105/3. 1455–1468.
- Markó Alexandra 2014. A beszéd temporális szerkezete a beszédmód és a beszédhelyzet függvényében. In Bátyi Szilvia – Navracsics Judit – Vígh-Szabó Melinda (szerk.): *Nyelvelsajátítási-, nyelvtanulási- és beszédkutatások*. Gondolat Kiadó – Pannon Egyetem, Budapest–Veszprém, 33–45.
- Menyhárt Krisztina 1999. Szláv anyanyelvűek magyar beszédének fonetikai sajátosságai. *Beszédkutatás 1999*. 30–44.
- Menyhárt Krisztina 2002a. The temporal organisation of speech in monolingual and bilingual children. *Acta Linguistica Hungarica* 49/3–4. 347–362.
- Menyhárt Krisztina 2002b. Az akcentus a magánhangzók tükrében. *Beszédkutatás 2002*. 38–52.
- Menyhárt Krisztina 2003. A spontán beszéd megakadásjelenségei az életkor függvényében. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 125–138.
- Navracsics Judit 2007. *A kényelvű mentális lexikon*. Balassi Kiadó, Budapest.

- Németh Szilvia 2012. *Igazi vagy hamis? Megkülönböztethető-e a valódi és az imitált akcentus?* Előadás. XV. Pszocholingvisztikai Nyári Egyetem, 2012. 05. 21.
- Németh Szilvia 2013. *Hogyan álcázzuk magunkat oroszok? – az imitált és eredeti orosz akcentus egybevetése.* Előadás. XVI. Pszicholingvisztikai Nyári Egyetem, 2013. 05. 28.
- Perry, Theodore L. – Ohde, Ralph N. – Ashmead, Danieal H. 2001. The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America* 109/6. 2988–2998.
- Tóth Andrea 2014. Gyermek nemének és életkorának meghatározása a beszédük alapján. *Beszéd kutatás 2014.* 98–111.
- Traunmüller, Hartmut – Bezooijen, Renée van 1994. The auditory perception of children's age and sex. In: *Proceedings ICSLIP.* 1171–1174. <http://www2.ling.su.se/staff/hartmut/barn.pdf>. (A letöltés ideje: 2015. 01. 03.)
- G. Varga Györgyi 1972. A budapesti nyelvhasználat társadalmi rétegződése. *Általános Nyelvészeti Tanulmányok* 8. 213–230.
- Vorperian, Houri K. – Wang, Shubing – Chung, Moo K. – Schimek, E. Michael – Durtschi, Reid B. – Kent, Ray D. 2009. Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of the Acoustical Society of America* 125/3. 1666–1678.
- Wardhaugh, Ronald 1995. *Szociolingvisztika.* Osiris–Századvég, Budapest.
- Whiteside, Sandra P. – Hodgson, Carolyn 1999. Acoustic characteristics in 6–10-year-old children's voices: Some preliminary findings. *Logopedics Phoniatrics Vocology* 24. 6–13.
- Xue, Steve A. – Wing Chi Cheng, Regine – Manwa Ng, Lawrence 2010. Vocal tract dimensional development of adolescents: An acoustic reflection study. *International Journal of Pediatric Otorhinolaryngology* 74. 907–912.

### Kérdések, feladatok

1. Határozza meg a szociofonetika fogalmát!
2. Mutassa be 1-1 példán keresztül az egyes kutatási területek szociofonetikai hozadékait!

## Milyen egységekre tagolható a beszéd?

A nyelvhasználat két alapvető formája a beszéd és az írás. A beszéd a nyelvhasználat hangzó formája, fizikai paraméterekkel (idő, intenzitás, frekvencia) jellemezhető levegőrezgés, akusztikai jel (Gósy 2004). Mivel a beszédképző szerveink működése következtében folyamatosan változik az akusztikai szerkezet, ezért a beszédet folyamatos, más szóval analóg jelnek kell tekintenünk. Ezzel szemben az írás – ami elsősorban a beszéd rögzítésére jött létre – véges számú, egymástól jól elkülöníthető, diszkrét elemből (betűből) épül fel, vagyis digitális jelnek kell tekintenünk (Nádasdy 2006). Ebből következik, hogy az írásban egyértelműen el tudjuk különíteni egymástól a szöveget felépítő egységeket (betű, szó, mondat stb.), míg a beszéd elemzése során meg kell birkóznunk azzal a nehéz feladattal, hogy egy analóg jelet átalakítsunk diszkrét jelek sorozatává.

A beszéd vizsgálata során, a fonetikai kutatásokban gyakran felmerül a kérdés, hogy milyen egységekre tagolhatjuk a beszéd folyamatot, illetve hogy milyen jellemzők mentén történhet meg ez a tagolás. A jelen tanulmány ezt a kérdéskört járja körül, összegezve a magyar beszédre vonatkozó legfontosabb eredményeket.

Először röviden bemutatjuk a beszéd tagolásának problémáját, majd áttekintjük a leíró nyelvészetből ismert egységeket (bekezdés, mondat, szó, szótag, beszédhang) fonetikai szempontból, végül a prozódiai egységeket (intonációs frázis, prozódiai frázis) vesszük górcső alá.

### 1. A tagolás folyamata

Az átlagos hallgatónak nem okoz nehézséget, hogy a hallott közlésekben azonosítsa a szavakat, mondatokat vagy akár nagyobb egységeket. Vannak azonban különleges helyzetek, amikor szembesülhetünk a tagolás nehézségeivel. Például idegen nyelven elhangzó beszédben vagy nagyon zajos környezetben nehezebben különítjük el egymástól a szavakat. Nehézséget okozhat továbbá, ha a beszélő gyakorlatlan, és fiziológiai tényezők – például a tüdő kapacitása – szervezik a közlésének tagolását. A kisgyermek gyakran nem tudja a tüdőkapacitásukhoz igazítani a tagolást, ezért akár szó közepén is „kifogyhatnak” a levegőből, és lélegzetvétellel kényszerülhetnek.

Az elhangzó hosszabb közlések megértésének és értelmezésének egyik alapfeltétele, hogy a hallgató megfelelően tagolja azt (Gósy 2003). A befogadónak azonban nincs szüksége minden esetben a szegmentálásra, például az írott beszéd esetén karakterhiány, központosítás jelzi az olvasó számára az egyes egységek – szavak, mondatok stb. – határát. A folyamatos beszédben ugyanakkor nem ilyen egyértelmű a határjelzés, nem jelzik minden esetben például szünethelyek a szemantikai, szintaktikai, prozódiai egységek kezdetét vagy befejezését (vö. Gósy 2005).

Először hasonlítsuk össze az írott szöveg és annak felolvasott változatának tagolása során tapasztalható eltéréseket. A magyar helyesírás szabályrendszeréből egyértelmű, hogy mit hogyan kell leírunk, hogyan használjuk az írásjeleket. Szövegfelolvasáskor azonban sokkal lazább szabályok határozzák meg a megvalósítást, vagyis többféle „helyes” megoldás létezik egy adott szöveg felolvasására (ezt nagymértékben meghatározza a felolvasó értelmezése is). Alapvetően az írás vezérli a felolvasást is, figyelembe kell vennünk például az írásjeleket a meghangosítás során, de másfajta tagolási elveket kell követnünk. Például nem minden vesszőnél kell szünetet tartanunk felolvasáskor, és nemcsak vesszőnél tarthatunk szünetet (Olaszy 2010a).

## 2. A leíró nyelvészeti egységeken alapuló fonetikai kutatások

Már kisiskoláskorban megtanuljuk azokat a metanyelvi kifejezéseket és jelentésüket, amelyek mentén feltételezhetően a beszédpercepció folyamatában feldolgozzuk az elhangzott közléseket. Ilyen fogalmak például a bekezdés, a mondat, a szó és a szótag. Több fonetikai kutatás foglalkozott ezeknek az egységeknek a feldolgozásával, a beszélt nyelvre való alkalmazhatóságával.

### 2.1. Bekezdések a magyar beszédben

A mondatnál hosszabb egységeket a nyelvészeti szakirodalomban *bekezdés*-nek nevezik, a terminust azonban általánosan az írott szövegek bizonyos egységeinek tipográfiai elkülönítésére használják. A verbális kommunikáció megnyilatkozásainak vizsgálatában félrevezető lehet (vö. Lehiste 1979), ezért Váradi (2012) a gondolategység terminust vezette be a beszélt nyelv bekezdés szintű egységeire. Tolcsvai Nagy (2001) az írott szövegek bekezdéseinek következő fontos tulajdonságait sorolja fel: szemantikai és funkcionális egység; egy kiemelkedő téma fogja össze; van belső szerkezete; a témáját megnevező csúcspont a bekezdés elején vagy végén található. A bekezdés határait a beszédben különböző nyelvi és egyéb eszközök jelölik, például szünet és habozás; az idő, a tér, a szereposztás, a nézőpont változásainak jelölői. Szikszai-

né Nagy (1999) meghatározásában a bekezdés vagy paragrafus a szöveg szemantikai-logikai, szintaktikai-intonációs-tipográfiai egysége a szövegben, funkciója a téma tagolódásának vizuális jelölése.

Felmerül a kérdés, hogy a hallgatók percepció mechanizmusa mennyire egyöntetűen működik a bekezdésszintű tagolás során; illetve a hallgatók milyen akusztikai-fonetikai, szintaktikai és pragmatikai jellemzőket használnak fel a spontán beszéd bekezdésszintű tagolása során. Várad (2012) vizsgálatában a spontán beszéd bekezdésszintű tagolását, a gondolategységeket elemezte a beszédfeldolgozás szempontjából. A kutatás adatközlői azt a feladatot kapták, hogy egy spontán közlés hallgatása során jelöljék meg a bekezdéshatárokat a központosítás nélküli, helyesírásban lejegyzett átiratban. A vizsgálat egyik fontos tanulsága, hogy az adatközlőknek különbözhetnek a kódolt ideáik a spontán beszéd bekezdéseiről, a gondolategységekről. Úgy tűnik, hogy nem egy általános stratégia alapján tagolják a hallgatók a spontán beszédet gondolat-egységekre, hanem többféle stratégia létezhet, akárcsak írásbeli szövegalkotás során. A bekezdésszintű tagolásban sem figyelhetünk meg általános tendenciát, hanem nagymértékben hallgatófüggő, hogy mekkora gondolat-egységekre tagolják a spontán beszédet. A tagoló pontok általában szünetet megelőzően fordultak elő, de az elemzések szerint a szünet időtartama és a jelölések száma között nincs összefüggés. A szünetek valamilyen mértékben meghatározók a bekezdésszintű tagolásban is, de úgy tűnik, hogy önmagukban nem elegendők, főképp a rövidebb időtartamú szüneteknél jellemző a bizonytalanság. A nem szünethez kapcsolódó jelölésekre nagyfokú bizonytalanság volt jellemző, csak egyetlen jelölésnél fordult elő, hogy az adatközlők egyötöde gondolat-egység-határt jelölt. Ez az eredmény alátámasztja a szünet erőteljes határjelző szerepét a gondolat-egységek szintjén is.

A határjelölések nagymértékű összefüggést mutattak a szintaktikai szerkezet befejezettségével, 98%-ban szintaktikai szerkezet határán jelöltek egység-határt az adatközlők. Az irreguláris zöngékepzés – amely percepció szempontból a hangszínezetváltás hatását kelti – közlészáró funkciója ebben a kutatásban is igazolódott. Nem volt elvárás az adatközlők részéről, hogy a mondatnál nagyobb egységek zárlatainál – normál zöngékepzés esetén – ereszkedő legyen a dallammenet. Ezt a megfigyelést támasztja alá továbbá, hogy az ereszkedő dallammenetknél az adatközlők bizonytalansága volt jellemző, általában csak egy-két adatközlő jelölt gondolat-egység-határt. Markó (2009) az emelkedő/szökő dallammenet spontán beszéd-*beli* előfordulásának vizsgálata során arra a megállapításra jutott, hogy az emelkedő dallam közlészáró helyzetben gyakori a spontán beszédben. Azokra a jelölésekre, ahol csak egy adatközlő észlelt gondolat-egység-határt, a lebegő dallammenet volt legnagyobb arányban jellemző. A lebegő dallammenet a gondolat lezártatlanságát jelentheti, amit a hallgatók is észleltek, ezért volt jellemző a nagyfokú bizonytalanság. A hallgatói ítéletek azonosságá-

nak csökkenésével nőtt az általuk határnak ítélt szöveghelyek száma. Az egyöntetű jelöléseket a szemantikai-szintaktikai, pragmatikai és szupraszegmentális jelölők együtt járása jellemezte: témaváltás, szintaktikai szerkezethatár, hosszabb szünet, irreguláris zöngékezés.

A gondolategységekre való szegmentálás nem problémamentes, nem találtak olyan jegyeket, amelyek egyértelműen, az adatközlők többsége számára jelzik az egység határát. A kapott eredmények alapján úgy tűnik, hogy a mondatnál nagyobb összefüggő szövegegységek tematikusan szerveződnek, önmagukban a szintaktikai és a szupraszegmentális jellemzők ennél kisebb egységeket határoznak meg.

Beszélt nyelvi anyagon nemcsak a bekezdésnyi egységek, hanem a tematikus tagolás működését is elemezték. Görög nyelvű híradásokban vizsgálták a tematikai szegmentálás sajátosságait. Főtémákra és altémákra tagolták a híradásokat, és elemezték ezek prozódiai megvalósulását. Eredményeik szerint a tematikus tagolás és a témakategorizáció elsődleges prozódiai jelölői a zárlat alaphangmagasság-értéke és a szünet időtartama (Botiniset al.2003).

Gósy és munkatársai (2011) a tematikus egységeket abban az esetben különítették el egymástól, ha a kísérletvezető újabb kérdéssel újabb témát indított. Elemzéseik során arra a következtetésre jutottak, hogy az intenzitás a beszélők 90%-ánál nagyobb, az alaphangmagasság pedig a beszélők 70%-ánál magasabb a tematikus egység elején, mint a végén. A tematikus egységek időtartamait vizsgálva megállapították, hogy a beszélők relatíve azonos időtartamú tematikus egységeket hoztak létre narratíváik során.

## 2.2. Mondatok a magyar beszédben

A mondatnak többféle meghatározása, megközelítése – logikai, formális, lélektani, funkcionális – létezik a nyelvészeti szakirodalomban, a *Magyar grammatika* a következő definíciót adja: „A mondat egy vagy több szóból áll, zárt intonációs szerkezet jellemzi. A mondat a legnagyobb nyelvi egység, amely a nyelv szabályai, mintái szerint nyelvi elemekből megszerkeszthető. Egyben a legkisebb nyelvi egység, amely egy adott beszédhelyzetben közlésegyiséggé válhat” (Kugler 2000: 371). Az írott beszédre alkalmazható ez a meghatározás, de például a zárt intonációs szerkezet nem feltétlen velejárója a spontán beszéd mondatainak. Gósy (2003) bevezette a virtuális mondat fogalmát, a virtuális jelzőt a ’benne rejlő’ értelmében használva. Ezek az egységek többé-kevésbé azonos paraméterekkel jellemezhetők, melyeket a hallgatók képesek tagoló funkcióban használni spontán közlések hallgatásakor. A kutatásában alkalmazott módszert több kutató átvette, a továbbiakban ezeket fogjuk áttekinteni. A kísérletben részt vevő adatközlők megkapták az elhangzó közlés helyesírásban, központosítás nélkül lejegyzett változatát, és miközben hallgatták a beszélt nyelvi

anyagot, a tesztalapon függőleges vonallal jelölték az általuk észlelt mondathatárokat.

A Gósy (2003) kutatásában részt vevő 58 adatközlő nagy arányban (70%) egyöntetűen észlelte a mondathatárokat spontán beszédben. Úgy tűnik, hogy a hallgatók nagymértékben hasonló paraméterek alapján hozzák meg döntéseiket a mondathatárokat illetően. Váradí (2009) két beszédmódot, a spontán beszédet és a felolvasást, hasonlított össze a mondatszintű tagolás tekintetében 26 adatközlő bevonásával. Az ő eredményei szerint a spontán beszédben 64%-os, míg a felolvasásban 91%-os egyöntetűséget mutatott a mondathatár-jelölés. Ezt a különbséget az eltérő nehézségű tervezési mechanizmusok magyarázhatják.

Nemcsak felnőttek, hanem gyermekek spontán beszédben is elemezték a virtuális mondatokat (Neuberger 2012). Ebben a vizsgálatban két gyermek hanganyagát használták fel a percepció tesztben, az egyik szöveg nyelvtanilag egyszerűbb, a másik bonyolultabb szerkezetű volt. Az eredmények azt mutatják, hogy a gyermekek beszédének szegmentálása kevésbé egyöntetű, mint a felnőtteké.

A potenciális akusztikai fonetikai jelölők közül az egyik leggyakrabban elemzett jelenség a szünet. Felolvasott és spontán angol beszédben is igazolták a néma szünet határjelző szerepét (Goldman-Eisler 1972). Egy későbbi kutatás is megerősítette Goldman-Eisler eredményeit. Az angol és japán dajkanyelvben a szünetek 96%-a együtt járt a megnyilatkozásvégi határokkal. Az angolban a szünetek 59%-ánál, a japánban 69%-ánál adatoltak határt (Fisher–Tokura 1996). Kemler és munkatársai (1989) gyermekeknek és felnőtteknek szóló közléseket játszottak le 8 és fél hónapos csecsemőknek. Tagmondatokba és tagmondatok közé illesztettek be egy perces néma szüneteket a közlésekbe. A csecsemők a gyermekeknek szóló közléseket és a tagmondathatáron lévő szüneteket preferálták a felnőtteknek szólókéval és a tagmondaton belüli szünetekkel szemben.

Magyar nyelven is elemezték felolvasásokban és spontán beszédben a szünetek határjelző szerepét. Gósy (2003) eredményei magyar spontán beszédben alátámasztották, hogy a szünet időtartama meghatározó a virtuális mondatok észlelésben. Minél hosszabb szünetet tartott a beszélő, annál nagyobb arányban jelölték mondathatárt a hallgatók.

Váradí (2009) kutatásában összehasonlította a felolvasásokban és a spontán közlésekben a szünet határjelző szerepét is. A kutatás eredményei szerint a felolvasásban sokkal erőteljesebb volt a szünet határjelző szerepe. A felolvasásban a szünetek 70%-ánál, míg a spontán közlésekben csak 23%-ánál jelölt legalább egy adatközlő mondathatárt. Spontán közlések létrehozásakor a beszélő többféle okból tarthat szünetet. Ha nem találja a megfelelő szót, vagy más okból nehézségei támadtak a közlésének létrehozásában, akkor gyakran szünettartással biztosítja az időt a produkciós folyamatoknak. Felolvasáskor azonban egy ké-

szén kapott szöveget kell meghangosítania a beszélőnek, ezért kevesebb nehézsége van a beszédtervezés során és ebben az esetben a szünetek szinte kizárólag tagoló funkcióban fordulnak elő. Úgy tűnik, hogy a hallgatók is képesek megkülönböztetni az eltérő funkciójú szüneteket.

Gyermekek beszédében a hallgatók nem minden szünetnél jelöltek mondathatárt, a mondathatár-jelöléseik nagymértékben szünetnél fordultak elő. A nyelvtanilag egyszerűbb szerkezetű szövegben a jelölések 91%-a, a bonyolultabb szerkezetűben a jelölések 78%-a szünetet követően volt adatolható (Neuberger 2012).

Egy holland anyanyelvűekkel végzett vizsgálat eredményei szerint a beszélő használja az alaphangváltozást tagoló funkcióban, és ezt a hallgató is észleli (Swerts–Collier 1992). Gósy (2003) kutatásának eredményei szerint a magyar anyanyelvűek is felhasználják döntéseikben az alaphangmagasság változását vagy változatlanágát. Legnagyobb arányban ereszkedő dallammenetnél, legkevésbé emelkedő dallammenetnél észleltek mondatvéget a hallgatók.

Váradi (2009) kutatásában jellegzetes eltéréseket talált a felolvasás és a spontán beszéd között. Az utóbbiban az ereszkedő, míg az előbbiben a lebegő dallammenet bírt erőteljesebb határjelző funkcióval. Megvizsgálta azokat a dallammeneteket is, amelyekben irreguláris volt a zöngképzés. (Ez olyan zöngképzési mód, amelynek az eredményét érdes, rekedtes hangként észleljük, vö. Markó 2005). A felolvasásokban erőteljes az irreguláris zöngé határjelző szerepe, a mondathatárjelölések közel 40%-ánál adatolt irreguláris zöngképzést, míg a spontán narratívákban egyetlen esetben sem. Az irreguláris zöngé határjelző szerepét több nemzetközi és magyar nyelvű kutatás is igazolta.

Neuberger (2012) kutatásában az egyszerűbb és a bonyolultabb nyelvtani szerkezetű spontán, gyermeknyelvi közlésekben is – Gósy (2003) felnőttnyelvi eredményeihez hasonlóan – az ereszkedő dallammenet bírt a legerőteljesebb határjelző funkcióval, az emelkedő és a lebegő dallammenet kevésbé keltette a befejezettség érzetét.

Markó (2010) a prozódia szerepét vizsgálta a spontán beszéd tagolásában. Kutatásában háromféle teszhelyzetben elemezte a mondathatár-észleléseket. Az első tesztben Gósy (2003) módszertanát alkalmazta, vagyis a hallgatók megkapták a helyesírásban lejegyzett, de központosítás nélküli szöveget, és a beszédanyag meghallgatása során kellett jelölniük a szövegben a mondathatárokat. A második tesztben csak a központosítás nélküli szövegben kellett mondathatárokat jelölni, míg a harmadikban egy módosított hangfelvételben (a hanganyag értelme nem volt kivethető, csak a prozódiai jellemzők voltak észlelhetőek) kellett mondathatárokat jelölniük. A hallgatók a legeggyöntetűbben az első teszhelyzetben jelöltek mondathatárokat, és a legnagyobb bizonytalanságot a harmadik teszhelyzetben tapasztalta a szerző. A kutatás eredményei alapján elmondhatjuk,

hogy bár vannak akusztikai fonetikai jelölői a virtuális mondatok határának, de ezek nem egyértelműek és nem minden kétséget kizáróak.

### 2.3. Szavak a magyar beszédben

A *szó* terminus önmagában nem definiálható. Különbséget kell tennünk a nyelv szavai, vagyis a lexémák és a beszéd szavai, vagyis a szóelőfordulások között. A *Magyar grammatikában* olvasható meghatározás szerint: „A lexéma a közös alaki, szerkezet- és jelentéstani tulajdonságokkal rendelkező szóelőfordulásokból elvont nyelvi egység. A nyelv elemkészletének része, szótári szó, melyet a rá jellemző morfológiai, szintaktikai és szemantikai információkkal együtt raktározunk el tudatunkban” (Lengyel 2000: 27).

A beszédben előforduló szavak a szóelőfordulások, melyek általában önálló fonetikai, szerkezeti, jelentéstani egységeknek tekinthetők. Ezek építik fel a beszéd mondatait, és az adott mondat határozza meg konkrét jelentésüket, alakjukat és szerepüket (Lengyel 2000). A továbbiakban a szóelőfordulásokkal foglalkozunk, ezért a *szó* terminust ’szóelőfordulás’ értelemben használjuk.

A fonetikai kutatások során a szavakkal kapcsolatban a legtöbb nehézséget a szószintű tagolás jelenti. Feltehetőleg a beszédpercepció folyamatában a szószintű tagolás elsősorban szemantikai és nem fonetikai alapon történik. Ez azt jelenti, hogy a hallgató felismeri azokat a hangsorokat, amelyekhez jelentés kapcsolható, és ez alapján tagolja szavakra a hallott közlést. Írásban szóközzel különítjük el egymástól a szavakat, ezért az olvasó könnyen szavakra tudja tagolni az olvasott szöveget. Beszédben azonban ilyen egyértelmű határjelölő híján nehézségekbe ütközhetünk a szószintű tagolás során. Megoldást jelenthet, ha mindkét szóhoz tartozónak vesszük, az első szó záró hangjaként, illetve a második szó kezdőhangjaként (Olaszy 2010b).

A szavak megvalósulásának időtartama a beszédben nem állandó. Több tényező befolyásolja, hogy például a *kelkáposzta* szó mekkora időtartamban realizálódik. Nem mindegy, hogy gyorsan vagy lassan beszélünk-e (beszédtempó). A gyorsabb tempó következtében a szóidőtartamok rövidülhetnek. A beszélők között is jelentős különbségeket adathozhatunk ebben a tekintetben, illetve a beszédhelyzet is nagymértékben meghatározza a szavak időtartamát. Ez azt jelenti, hogy ha megmérjük a *kelkáposzta* szó időtartamát két beszélő spontán, nem előre tervezett közlésében (például megkérjük, hogy meséljen az ebédjéről, ami történesen kelkáposzta-főzelék volt), akkor két – akár nagyon – eltérő időtartamadatot fogunk kapni. Ha ugyanezt a vizsgálatot elvégezzük azzal a különbséggel, hogy ugyanazt a beszélőt kérjük meg két különböző időpontban és eltérő beszédhelyzetben ugyanerre a feladatra, akkor is eltérő időtartamadatokat fogunk kapni. Vagyis bár egy adott szó általában ugyanannyi beszédhanggal valósul meg beszélőtől, beszédtempótól, beszédhelyzettől függetlenül, mégis

eltérő időtartamban realizálódnak. Egy kutatásban ugyanazt a mondatot többször is bemondták a beszélők, és a mondat szavaira kapott átlagidőtartamok nagy változatosságot mutattak, 500 ms-mal, 590 ms-mal, 608 ms-mal, illetve 696 ms-mal realizálódtak (Gósy 1998).

A szótagszám is befolyásolja a szó időtartamát: minél több szótagból áll egy szó, annál hosszabb ideig tart a kiejtése. Azonban nem arányosan növekedik a szó időtartama a szótagszám növekedésével. Az egy és két szótagból álló szavak közti különbség csak mintegy 40 ms, de a két-három szótagos szavak között (az átlagot tekintve 143 ms), illetve a négy és öt szótagos szavak között (az átlagot tekintve 133 ms) ugrásszerűen megnövekszik ez a különbség. A három és a négy szótagosak, valamint az öt és hat szótagosak között azonban ezekhez viszonyítva kisebb, mintegy 100 ms körüli az időtartam-különbség. Ennek a különbségnek a hátterében az eltérő, időigényes beszédtervezési folyamatok állhatnak. A három vagy több szótagból álló szavak esetében nagyon gyakran toldalékolt alakok fordulnak elő, vagyis valamilyen időigényes grammatikai művelet sor eredményeként jön létre a szóalak (Gósy 2004).

A szótagszám mellett a szógyakoriság is befolyásolja a szóidőtartamok alakulását. Feltehetőleg a szóelőhívás nehezítettsége okozza, hogy a ritkábban használt szavak tendenciaszerűen hosszabb időtartamban realizálódnak, mint a gyakrabban használatosak. Például a *lányomé* szó egy női adatközlő ejtésében 511 ms volt, míg a *langyos* szó ejtéséhez 721 ms-ra volt szüksége az adatközlőnek (Gósy 2004).

## 2.4. Szótagok a magyar beszédben

Nemcsak a szavak, hanem a szótagok definiálása is nehézségekbe ütközik, mibenléte napjainkban is vitatott kérdés. Egyes kutatók nem is tekintik a szótagot önálló hangtani kategóriának (Kassai 2005). A beszélők azonban tapasztalati úton vesznek tudomást a szótag létezéséről, a beszéd folyamat legkisebb (önálló) időbeli egységének érzik (Siptár 2003). Amit biztosan állíthatunk a szótagról, hogy a beszédnek a beszédhangtól magasabb, a szónál alacsonyabb szintű szerveződési egysége. Általában a beszédhangnál nagyobb, a szónál pedig kisebb terjedelmű. De fontos megjegyezni, hogy egyetlen beszédhang is alkothat önálló szótagot (például a *Kla-u-di-a* szó második és utolsó szótagja), illetve egy szó is állhat egyetlen szótagból (például *tó*). A meghatározásbeli nehézségek ellenére a naiv beszélők is képesek a szótaggal különböző műveleteket elvégezni. Például meg tudják számolni, hány szótagból áll egy szó, sőt szótagokra is tudják bontani azt. De ki is tudják cserélni egy szó szótagjait egymással (például *ka-lap* helyett *lap-ka*) (Kassai 2005).

Bár a szótag kérdéskörével elsősorban a fonológia foglalkozik, azért a fonetika tudományában is használatos egység. Például egy beszélő beszédtempóját

meghatározhatjuk az alapján is, hogy hány szótagot ejt ki adott időegység alatt, például percenként.

A magyarban megkülönböztethetünk nyílt és zárt szótagokat. A nyílt szótagok magánhangzóra végződnek (például a *he-ve-rő* szó szótagjai), a zárt szótagok pedig – egy vagy több – mássalhangzóra végződnek (például a *hang-fal-szett* szó szótagjai) (Gósy 2004).

A szótagok időtartamát elsősorban a szerkezetük és az őket alkotó beszédhangok sajátosságai határozzák meg. A különböző szerkezetű szótagok időtartamai jellegzetes eltéréseket mutatnak. Átlagosan a legrövidebb időtartamban a magánhangzó + mássalhangzó (VC) típusú zárt szótagok realizálódnak (átlag: 165 ms), ennél valamivel hosszabbak a mássalhangzó + magánhangzó (CV) típusú nyílt szótagok (átlag: 194 ms), és leghosszabban a mássalhangzó + magánhangzó + mássalhangzó (CVC) típusú zárt szótagok realizálódnak (átlag: 219 ms) (Gósy 2004).

## 2.5. Beszédhangok a magyar beszédben

Mindenekelőtt meg kell különböztetnünk egymástól a beszédhangot és a fonémát. Az előbbieket a szegmentális hangszerkezet legkisebb egységei, egy adott nyelvre jellemző hangképzési mechanizmussal hozzuk őket létre. A beszédhangok hangkapcsolatokká, szótagokká, hangsorokká szerveződve alkotják a szegmentális hangszerkezetet. „A fonéma elvont nyelvi entitás, egy adott nyelvben jelentésmegkülönböztető erővel bíró egység” (Gósy 2004:245). A beszédhang fizikai jelenség, amelyet beszédképző szerveinkkel hozunk létre, ezzel szemben a fonéma elvont, fizikai tulajdonságokkal nem jellemezhető egység (Gósy 2004).

Úgy tűnhet, hogy az artikuláció során létrehozott beszédhangok eleve adva vannak, és ezekből vonatkoztatjuk el a fonémákat. Valójában azonban éppen az ellenkezőjéről van szó. A beszédflowam ugyanis nem önálló, egymástól független beszédhangok sorozata, hanem egymásból jönnek létre, az egyik hang artikulációs beállításából fokozatosan jutunk el a követő hang artikulációs beállításáig. Feltehetőleg nem a beszédflowamat, hanem a nyelv, a beszélők nyelvtudása szerveződik beszédhangnyi egységekbe (valójában fonémákba) (Siptár 2003). Tipikus beszédfejlődés esetén a gyermekek legkorábban hatéves korukra sajátítják el azt az olvasástanuláshoz elengedhetetlen képességet, hogy a beszédflowamot beszédhangokra tagolják (Kassai 2005).

A beszédhangokkal foglalkozó fonetikai kutatásokat nagymértékben megnehezíti a beszédhangok két fontos tulajdonsága. Az egyik nehézség, hogy egy adott beszédhang többféleképpen realizálódhat (más a *király k*-ja és a *katona k*-ja). A másik nehézség, hogy a beszédhangok hangátmenettel kapcsolódnak egymáshoz. Akusztikai szempontból a beszédflowam egy folyamatosan változó

akusztikai rezgésnek tekinthető, amelyben beszédhangok és az őket összekapcsoló hangátmeneti részek találhatók (Olaszy 2010b).

A fonetikai elemzésekhez gyakran nélkülözhetetlen, hogy a beszéd folyamatot szegmentáljuk beszédhangszinten. A hanghatárok kijelölésére használhatunk kézi és gépi módszereket, illetve ezek kombinációját. Az előbbi lényege, hogy a kutató nagy pontossággal kijelöli a hanghatár helyét, általában nem pusztán hallás alapján, hanem a hullámforma és a hangszínkép segítségével „láthatóvá” teszi a beszéd folyamatot. Ennek a módszernek a legfőbb hátránya, hogy meglehetősen lassú és megerőltető folyamat, illetve szubjektív. Ezzel szemben a gépi hanghatárjelölés során egy erre a célra kifejlesztett algoritmus állapítja meg a hanghatárokat, tehát objektívebb, mint a kézi jelölés, nagy hátránya azonban a pontatlansága. Ezt a pontatlanságot azonban lehet javítani utófeldolgozással, kézi módszerrel (Olaszy–Bartalis 2008).

A kézi hanghatárjelölés pontosságával foglalkozott egy korábbi kutatás (Váradi 2013). Ebben arra a kérdésre keresték a választ, hogy hét fonetikailag képzett lejegyző között mekkora különbségek adatolhatók a beszédhangszintű tagolásban (lejegyzők közötti variancia), illetve hogy a lejegyzők önmagukhoz képest mennyire konzekvenssek ugyanannak a hanganyagnak eltérő időpontokban való tagolása során (lejegyzőn belüli variancia). Az első lejegyzésekkor a CV (mássalhangzó + magánhangzó) hangkapcsolatok tagolása során átlagosan mintegy 16 ms volt az eltérés a hét adatközlő jelölése között, míg a VC (magánhangzó + mássalhangzó) hangkapcsolatok esetén kicsit több mint 20 ms (viszonyításképpen Gósy (2004) alapján az egyik leggyakoribb magyar magánhangzó, az *e* időtartama hangsúlyos helyzetben 78–220 ms között változik). A második lejegyzéseknél mind a CV, mind a VC hangkapcsolatok esetén kisebb eltéréseket adatoltak, mint az első lejegyzéseknél, és az adatközlők közötti variancia a VC hangkapcsolatoknál kisebb volt (átlagosan 12 ms), mint a CV hangkapcsolatoknál (átlagosan 17 ms).

A lejegyzőn belüli variancia elemzése során összehasonlították egy adott adatközlő első és második lejegyzésében adatolt hanghatárokat. Az összes beszélő esetében a 10 ms-nál kisebb eltérés volt a jellemző (80%-nál nagyobb arányban). A lejegyzők között azonban jelentős eltéréseket lehetett megfigyelni a konzekvencia tekintetében.

### 3. Prozódiai egységek a magyar beszédben

Az előző fejezetben bemutattuk, hogy az elsősorban a leíró nyelvészetben használatos egységek alkalmazása a fonetikai kutatásokban nehézségekkel jár, sok esetben nem egyértelmű, ezért a fonetikusok gyakran prozódiai egységeket különítenek el a beszéd folyamatban.

A beszélt nyelv prozódiai szerkezetének leírására a legelterjedtebb elképzelés mind a hazai, mind a nemzetközi szakirodalomban a hierarchikus felépítés (Gussenhoven 2004; Roca–Johnson 2005; Varga 1994; Hunyadi 2002). Felülről lefelé haladva a következő szintek különíthetők el: megnyilatkozás, intonációs frázis, fonológiai frázis, fonológiai szó, láb, szótag. A magyar nyelven végzett vizsgálatokban azonban a terminusok sokfélesége figyelhető meg.

A felolvasott szövegek esetén Elekfi (1962) a beszédütem terminust használja azokra a nagyobb értelmi, illetve kritikai egységekre, amelyekből a mondat felépíthető. „A beszédütem nagyjából a klitikumos egységnek (egy szóhang-súllyal ejthető szósorozat) feleltethető meg. Bolla Kálmán (1992: 10) más meghatározását adta a szupraszegmentális alapegységnek: „A szegmentális szerkezet struktúráképző alapeleme a beszédhang, míg a szupraszegmentális hangszövet legkisebb szerkezeti építőblokkját szupraszegmentális hangszerkezetnek nevezük”. A *szupraszegmentális hangszerkezettel* szinonim kifejezés az *intonációs szerkezet* vagy *intonációs frázis*, más szerzőknél ezzel a kifejezéssel találkozhatunk.

Olaszy (2006) vizsgálatában a legnagyobb szövegegység a mondat volt, a mondatokon belül pedig prozódiai egységeket különített el a szünettartás alapján. Ezeket a prozódiai egységeket az alaphangmagasság alapján bontotta kisebb egységekre, intonációs frázisokra. A fonetikai szakirodalomban azonban a szünettől szünetig terjedő egységekre a beszédszakasz terminus használata terjedt el, ami megegyezik az Olaszy-féle prozódiai egységnek nevezett egységgel.

Wacha (1988: 152) az élőbeszédben a megnyilatkozást tekintette alapegységnek, és a következőképpen definiálta: „Megnyilatkozáson az élőszóbeli, a (spontán) beszélt nyelvi közlésegesznek (szövegnek, szövegegésznek, beszédműnek) azt a – pontosan többé-kevésbé – elkülöníthető kisebb részét/egységét értem, melyet az írott nyelvhasználatról szólva a mondat, szövegmondat terminussal szokás megnevezni. A megnyilatkozás a beszélt nyelvnek-nyelvhasználatnak olyan mondat értékű része tehát, melynek határait (kezdetét és végét) utólag – az elhangzó szöveg lejegyzésekor (átírásakor) – állapítottuk meg és jelöltük meg írásjelekkel, figyelembe véve az írásbeliség alapján kialakult mondatfelfogást (konvenciót) is.” A szünettől szünetig terjedő egységekre, vagyis a beszédszakaszokra pedig a *fonemikus frázis* kifejezést használja.

Szaszák és Beke (2012) olvasott és spontán beszédben elemezte statisztikai módszerekkel a beszédjel szintaktikai és prozódiai szerkezetét. Kutatásuk fő kérdése az volt, hogy lehetséges-e a szintaktikai szerkezet feltárása a beszédjel prozódiai elemzése alapján. Olvasott beszédben a prozódia megbízhatóan jelezte a szintaktikai határokat: elkülönítette a tagmondathatárokat a szószerkezetek határaitól. Spontán beszéd esetén a prozódiai szerkezet alapján kevésbé volt megbízható a szintaktikai tagolás.

Olaszy (2006) elemezte a prozódiai szerkezeteket a hírfelolvasásban, a mese-mondásban, a novella és a reklámok felolvasásában. A felolvasások időszerkezetét, alapfrekvencia-szerkezetét, intenzitás-szerkezetét és hangszínezetét konkrét számadatokkal határozta meg, hogy leírja az egyes szövegtípusokhoz tartozó felolvasási stílust. Vizsgálatának eredményei azt mutatták, hogy a felolvasók követik a szöveg tartalmi és műfaji vonatkozásait, amely megmutatkozott a felolvasások hangsúlyozási, dallamviteli, ritmikai, dinamikai megformáltságában.

Váradí és Beke (2013) elemezte a prozódiai egységek, a beszédszakaszok és az intonációs frázisok szupraszegmentális szerkezetének akusztikai-fonetikai megvalósulását nyolc adatközlő felolvasásában és spontán beszédében. Az eredmények azt mutatják, hogy a két prozódiai egység akusztikai-fonetikai realizációja eltérő. A beszédszakaszokra lassabb artikulációs tempó és tágabb hangterjedelem volt jellemző, mint az intonációs frázisokra. Az artikulációstempó-értékek a beszédszakaszokban kisebb variabilitást mutattak, mint az intonációs frázisokban. A felolvasásokban szignifikánsan gyorsabb volt az adatközlők beszédsebessége a beszédszakaszokban és az intonációs frázisokban egyaránt. Az artikulációs tempó variabilitása a kisebb prozódiai egységekben, a hangsúlyközi szakaszokban figyelhető meg, amely variabilitás a nagyobb egységekben kiegyenlítődni látszik.

A felolvasás és a spontán beszéd eltérő akusztikai fonetikai megvalósulása a prozódiai egységek szintjén is adatolható volt. A beszédszakaszok időtartama szignifikánsan rövidebb volt a spontán közlésekben, mint a felolvasásokban. Mind a nyolc adatközlő magasabb arányban tartott szünetet spontán beszédében, mint felolvasásában. A felolvasások és a spontán közlések intonációs frázisainak átlagos időtartamában azonban nem volt szignifikáns eltérés. Az adatközlők közötti variabilitás is kisebb mértékű volt, mint a beszédszakaszok időtartamában. A beszédmód nem befolyásolta az intonációs frázisok időtartamát. A beszédszakaszok azonban szignifikánsan rövidebb időtartamban realizálódtak a spontán közlésekben, mint a felolvasásokban.

A felolvasások artikulációstempó-mintázatait, illetve a beszélők artikulációs tempójának összefüggéseit elemezve azt a következtetést vonták le, hogy a szavak artikulációs tempója az intonációs frázison belül nem önkényesen, hanem rendszerszerűen változik. Ez azzal magyarázható, hogy a beszéd temporális szerveződésének az intonációs frázis lehet az egyik alapegysége a felolvasás során. Mindhárom típusú intonációs frázis esetén a *rallentando* (fokozatos lassulás az egység vége felé) tendencia volt a legjellemzőbb. Az intonációs frázisok többségénél adatolható volt az egység vége felé a lassulás, amelyet a hallgatók felhasználhatnak a beszéd tagolása során. A spontán közlésekben azonban nem volt jellemző ez a lassulási tendencia, sokkal dominánsabbak voltak az egyéni sajátosságok. Az intonációs frázisok artikulációstempó-struktúráiban adatolt

egyéni különbségek arra utalnak, hogy más „időzítési stratégiák” mentén szerveződhetnek a spontán és a nem spontán közlések.

#### 4. Összefoglalás

A jelen tanulmányban a magyar beszéd szegmentálásával foglalkozó kutatások bemutatása révén betekintést nyerhettünk ebbe a komplex folyamatba. A bekezdések, a tematikus egységek, a mondatok, a szavak, a szótagok, a beszédhangok és a prozódiai egységek elemzése bepillantást enged a beszélők és a hallgatók tagolási stratégiáiba és azok fonetikai megvalósulásába. A humán mechanizmus rugalmassága azonban megnehezíti a tagolás folyamatának leírását. Hallgatóként képesek vagyunk arra, hogy alkalmazkodjunk a beszélő egyéni sajátosságaihoz – például jellemzően lebegteti alaphangmagasságát a közlésvégeken –, beszélőként pedig meglehetősen változatosan valósítjuk meg a tagolást. A mesterséges beszédfelismerő rendszerek számára nagy kihívás a beszéd értelmes egységekre való tagolása, mert nem léteznek egyértelmű határjelölők (Honbolygó 2011). Vizsgálatok igazolják, hogy a humán mechanizmus hatékonyabb és pontosabb, mint a gépi a bekezdéseknél kisebb egységek felismerésében. Rosszabb körülmények között (zajos környezet, spontán beszéd) kisebb hibaszázalékon belül teljesít, mint a számítógép (Lippmann 1997). A humán mechanizmus meglehetősen rugalmas, sokkal inkább viszonyok, mint konkrét értékek alapján dönt, és alkalmazkodik a beszélő tagolási sajátosságaihoz, illetve a kommunikációs körülményekhez.

#### Irodalom

- Bolla Kálmán 1992. Szupraszegmentális elemzések. Egyetemi Fonetikai Füzetek 8. ELTE Fonetikai Tanszék, Budapest.
- Botinis, Antonis – Gawronska, Barbara – Katsika, Argyro – Panagopoulou, Dionisia 2003. Prosodic speech production and thematic segmentation. *PHONUM* 9. 113–116.
- Elekfi László 1962. *Vizsgálatok a hanglejtés megfigyelésének módjaihoz*. Nyelvtudományi Értekezések 34. Akadémiai Kiadó, Budapest.
- Fisher, Cynthia – Tokura, Hisayo 1996. Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development* 67/6. 3192–3218.
- Goldman-Eisler, Frieda 1972. Pauses, clauses, sentences. *Language and Speech* 15. 103–113.
- Gósy Mária 1998. Szavak hangalakjának változása a gyermeknyelvben. *Beszédkutatás '98*. 1–38.
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszédkutatás 2003*. 19–43.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.

- Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutatás 2011.* 5–30.
- Gussenhoven, Carlos 2004. *The phonology of tone and intonation.* Cambridge University Press, Cambridge.
- Honbolygó Ferenc 2011. *A beszéd prozódiai jellemzőinek észlelése. A hangsúly pszicholingvisztikai és agyi háttere.* Akadémiai Kiadó, Budapest.
- Hunyadi László 2002. *Hungarian Sentence Prosody and Universal Grammar.* Peter Lang, Frankfurt am Main–Berlin–Bern–Bruxelles–New York–Oxford–Wien.
- Kassai Ilona 2005. *Fonetika.* Nemzeti Tankönyvkiadó, Budapest.
- Kemler Nelson, Deborah G. – Hirsch-Pasek, Kathy – Jusczyk, Peter W. – Cassidy, Kimberly Wright 1989. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 16/1. 55–68.
- Kugler Nóra 2000. A mondattan általános kérdései. In Keszler Borbála (szerk.): *Magyar grammatika.* Nemzeti Tankönyvkiadó, Budapest, 369–393.
- Lehiste, Ilse 1979. Perception of sentence and paragraph boundaries. In Lindblom, Björn – Öhman, Sven (eds.): *Frontiers of speech communication research.* Academic Press, London–New York–San Francisco, 191–201.
- Lengyel Klára 2000. A nyelvi egységek szinteződése. In Keszler Borbála (szerk.): *Magyar grammatika.* Nemzeti Tankönyvkiadó, Budapest, 24–33.
- Lippmann, Richard P. 1997. Speech recognition by machines and humans. *Speech Communication* 22. 1–15.
- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége. Monologikus és dialogikus szövegek összevetése, valamint a hűmmögés vizsgálata.* PhD-disszertáció. ELTE, Budapest.
- Markó Alexandra 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás 2009.* 88–106.
- Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutatás 2010.* 82–99.
- Nádasdy Ádám 2006. Nyelv és írás. In Kiefer Ferenc (főszerk.): *Magyar Nyelv.* Akadémiai Kiadó, Budapest, 907–931.
- Neuberger Tilda 2012. A spontán beszéd grammatikai fejlődése – a KFM módszer alapján. In Markó Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngékezdési időig.* ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest, 116–128.
- Olaszy Gábor 2006. Prozódiai szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás 2006.* 21–50.
- Olaszy Gábor 2010a. Tagolási különbözőségek. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd.* Akadémiai Kiadó, Budapest. 79–81.
- Olaszy Gábor 2010b. Hang- és szóhatárok kijelölése a beszéd hullámformáján. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd.* Akadémiai Kiadó, Budapest. 83–86.
- Olaszy Gábor – Bartalis Mátyás 2008. Jelfeldolgozási algoritmusok kombinációja a gépi hanghatárjelölés javítására. *Beszédkutatás 2008.* 208–220.
- Roca, Iggy – Johnson, Wyn 2005. *A course in phonology.* Blackwell, Oxford.
- Siptár Péter 2003. Hangtan. In É. Kiss Katalin – Kiefer Ferenc – Siptár Péter (szerk.): *Új magyar nyelvtan.* Osiris Kiadó, Budapest, 285–384.

- Swerts, Marc – Collier, René 1992. On the controlled elicitation of spontaneous speech. *Speech Communication* 11/4–5. 35–48.
- Szaszák György – Beke András 2012. Statisztikai módszerek alkalmazása a szintaktikai szerkezet és a beszédjel prozódiai szerkezetének feltérképezéséhez olvasott és spontán beszédben. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest, 236–250.
- Szikszaíné Nagy Irma 1999. *Leíró magyar szövegtan*. Osiris Kiadó, Budapest.
- Tolcsvai Nagy Gábor 2001. *A magyar nyelv szövegtana*. Nemzeti Tankönyvkiadó, Budapest.
- Váradí Viola 2009. Határjelzés a felolvasásban és a spontán beszédben. In: Geccsó Tamás – Sárdi Csilla (szerk.): *A kommunikáció nyelvészeti aspektusai*. Tinta Könyvkiadó, Budapest, 278–283.
- Váradí Viola 2012. Bekezdések a spontán beszédben. In Markó Alexandra (szerk.): *Az anya-nyelv-elsajátítástól a zöngeszédésig*. ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest, 304–315.
- Váradí Viola 2013. *Phonemic segmentation of Hungarian spontaneous speech*. Előadás, XV. Pszicholingvisztikai Nyári Egyetem, Balatonalmádi.
- Váradí Viola – Beke András 2013. Az artikulációs tempó variabilitása felolvasásban. *Beszédkutatás* 2013. 26–41.
- Varga László 1994. A hanglejtés. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan 2. Fonológia*. Akadémiai Kiadó, Budapest, 468–549.
- Wacha Imre 1988. Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest, 102–158.

### Kérdések, feladatok

1. Melyek a legjellegzetesebb különbségek a beszéd és az írás, illetve azok tagolása között?
2. Milyen nehézségekkel szembesülhetünk hallgatóként a beszéd tagolása során?
3. Milyen nehézségekkel szembesülhetünk a szószintű tagolás során?
4. Milyen problémákkal szembesülünk a szótag definiálása során?
5. Melyek a legjellemzőbb különbségek a kézi és a gépi hanghatárjelölés között?



## Beszél(get)ünk a számítógéppel? A beszéd mesterséges előállítása, számítógépes beszéd- és beszélőfelismerés

### 1. Bevezetés

A 21. század egyik új kutatási és alkalmazási területe a **beszédtechnológia**. A beszédtechnológia a mesterséges intelligencián belül a beszédalapú gyakorlati alkalmazások kifejlesztésével foglalkozik. A beszédtechnológia mindehhez a beszéddel kapcsolatos kutatások (például a fonetika), valamint az információtechnológia eredményeit használja fel (Németh–Olaszy szerk. 2010).

Jóllehet nem figyelünk rá, nem is vesszük észre, vagy legalábbis nem tudatosul bennünk, de mindannyian kommunikálunk számítógépekkel, és ez a kommunikáció sokszor beszéd segítségével zajlik. Gondoljunk arra, hogy amikor egy telefonközpont irányít minket a kívánt információhoz vagy a hívott félhez, az a telefonközpont egy „beszélő számítógép”. Amikor a telefonszámlánk egyenlegét szeretnénk megtudni, azt akár egy számítógép bemondásában is meghallgathatjuk. Ha a tudakozót hívjuk fel, a kívánt telefonszámot ugyancsak egy számítógép közli velünk. De akkor is egy számítógép beszédére hagyatkozunk, amikor például az autóban navigációs rendszerre bízunk a tájékozódást térképolvasás helyett. Arra is számos példát találhatunk, amikor a „gép leiratozza a beszédünket”, mint amikor parancsszavakkal irányítunk valamilyen operációs rendszert, konzolt, vagy akár a név szerinti tárcsázást kívánjuk használni a telefonunkon. Vannak olyan diktáló rendszerek, ahol a gépnek felolvasott szöveg leiratozásra kerül (például orvosi leletezők). Ezeken kívül kereshetünk kulcsszavakat rövid hanganyagokban, híradós adásokban is. Az olvasók közül azonban a legtöbben valószínűleg a hangvezérelt okostelefonokkal, a videómegosztó webhely és a keresőfelület, illetve a fordító programok hangvezérelt keresésével találkozhattak.

Láthatjuk, hogy ahogyan az ember-ember közötti kommunikációt is beszédprodukcióna (a beszéd létrehozása) és beszédpercepcióna (a beszéd feldolgozása) osztja a tudomány, ugyanígy sorolhatók be az ember és gép közötti beszédkommunikáció alkalmazásai a számítógépes beszéd-előállítás, valamint a számítógépes beszédfelismerés nagy területeire. De természetesen – akárcsak az emberi kommunikáció, amelynek lényege a kölcsönösség – ezek a fejlesztések dialógusrendszerekben is működnek, vagyis vannak olyan alkalmazások, amelyekben az ember beszélve fordul a számítógéphez, amely ugyancsak beszéddel válaszol (pl. mesterséges intelligencia, robottechnológia, okosépületek).

A beszédtechnológia létrejötte és dinamikus fejlődése számos okra vezethető vissza. Ezek az alkalmazások részben a kényelmünket szolgálják, egyszerűsítik a munkát. Másrészt a számítógép sok műveletet gyorsabban végez el, mint az ember; nem téved; nem zavarja a monotonia (vagyis ugyanazt a feladatot akár több tízezer-szer is el tudja végezni egymás után ugyanúgy, anélkül, hogy belefáradna, és nem hibázik); ötvözni tudja többféle alkalmazás előnyeit (például a navigációs rendszer a helymeghatározást és a térbeli tájékozódást az információközléssel). A beszédtechnológiai alkalmazások fontos előnye, hogy a valamilyen sajátos igénnyel élő emberek számára nagy segítséget nyújtanak. Például meghangosítják a számítógép képernyőjét a vakok és gyengénlátók számára; súlyos beszédzavar esetén közvetítik a beszélő szándékát; a siketek és nagyothallók számára a beszédjelet a száj mozgását ábrázoló vizuális jellé alakítják.

A következőkben néhány olyan alkalmazást és ezek működési módját mutatjuk be röviden, amelyeknél az ember és a számítógép közötti kommunikáció beszéd segítségével zajlik.

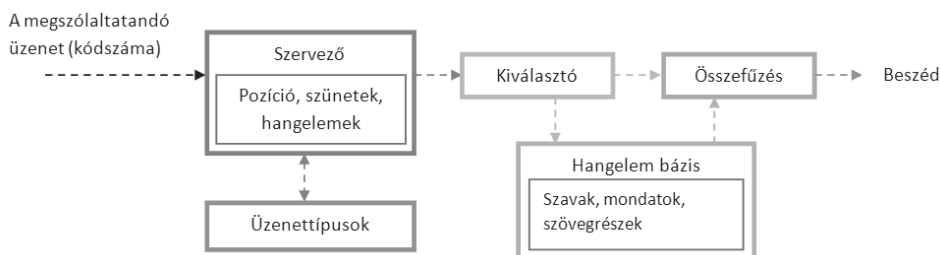
## 2. Számítógépes beszéd-előállítás

A számítógépes beszéd-előállításon – más szóval: beszéd-szintézisen – belül az alkalmazások két nagy csoportját különíthetjük el: az ún. kötött szótáras rendszereket és a szövegfelolvasókat. A kötött szótáras rendszerek zárt (és általában nem túl nagy) szókészlettel működnek, néhány üzenettípust tudnak meghangosítani, amelyeknek a szerkezete is kötött. A szövegfelolvasók ezzel szemben (elvből) bármilyen üzenet meghangosítására képes, rugalmas rendszerek.

### 2.1. Kötött szótáras rendszerek

A kötött szótáras rendszerek olyan alkalmazások, amelyekkel nap mint nap találkozunk, ha tömegközlekedési eszközökön utazunk, GPS-es útvonaltervezést használunk, hivatalokkal vagy vállalatokkal kell telefonhívást lebonyolítanunk, vagy olyan kényelmi szolgáltatásokat veszünk igénybe, mint a telefonébresztő vagy a telebank. Ezek a rendszerek emberi beszédből előre felvett rövid szövegrészleteket (egyedi beszédelemeket, vö. Olaszky 1999) fűznek össze egy meghatározott logika szerint. Ezekben a felhasznált szövegrészletek száma nem túlságosan nagy (néhány tíz), és ugyancsak nem túl nagyszámú kombinációkban hangoznak el. Gondoljunk például egy számítógép vezérelte telefonközpontra, vagy akár a járműveken szóló utastájékoztatókra! Ezeknek a létrehozásához először megtervezik azt a nyelvi anyagot, amelyet rögzíteni érdemes. Az adott vállalat kiválaszt egy beszélőt, akinek a bemondásában felveszik ezeket a részleteket. Az így összeállított hangelemtárból fogja az összefűző algoritmus kiválasztani a szükséges elemet. Az ismétlődő része-

ket csak egyszer veszik fel, de az algoritmus több különböző körben többször ki tudja ezeket választani. Az 1. ábra mutatja be a kötött szótáras szintézis folyamatát és szükséges moduljait. A hangelembázis az üzenetek, üzenetrészek hullámformáit tárolja. A kiválasztó algoritmus meghatározza, hogy az elemtárból mely összefűzendő hangelemeket (üzenetrészeket) kell kiválogatni a bejövő parancs alapján, majd az összefűző modul egymáshoz kapcsolja ezeket.



1. ábra. Az általános kötött szótáras beszédszintetizátor blokksémája  
(Forrás: Pandur 2011: 10)

Ha a kötött szótáras rendszer csak változatlan mondatokat tartalmaz (pl. *A Kálvin-tér következik. – Kálvin tér. – Az Astoria következik. – Astoria.*), akkor a hangzása jó minőségű lesz, hiszen csak az ember által bemondott szerkezeteket szólaltatja meg változtatás nélkül.

Ha az alkalmazás változó elemeket is tartalmaz, a hangzás minősége nagyban függ attól, hogy a tervezéskor mennyire jártak el körültekintően. Például egy egyenlegközlő vagy az üzenetrögzítőnk lehallgatásakor értelemzavaró lehet, ha a számokat nem olyan bontásban hangosítja meg a rendszer, amelyhez szokva vagyunk; vagy ha a számkapcsolatok hangzása nehezíti a megértést. Képzeljük el, hogy a bankszámlánk egyenlegét szeretnénk lekérdezni, amely 25 974 forint, és a gép a következő információt közli, a | helyén rendre szünetet tartva: *Az ön egyenlege | huszon | öt | ezer | kilenc | száz | hetven | négy |forint.* A bemondás még nehezebben értelmezhető, ha a gép nemcsak szünetet tart az elemek között, de mindegyiket külön is „hangsúlyozza”. Ez könnyen megtörténhet, ha a számelemek felvételekor nem figyeltek az összetartozó számelemek hangzására, csak felvették a magyar nyelvű számalakok főbb elemeit külön-külön. Ahhoz, hogy jól hangzó számfelolvasót állítsunk elő, többek között olyan fonetikai ismeretekkel kell rendelkezniünk, mint hogy milyen a különféle magyar szerkezetek hangsúlyozása, hanglejtése; a beszédhangok hogyan hatnak egymásra; hol tarthatunk szünetet, és hol lenne értelemzavaró a szünettartás, stb. A 2. ábra arra mutat példát, hogy mindezek az ismeretek hogyan hasznosulnak egy számfelolvasó kötött szótáras rendszer tervezésekor.

Eredeti alapelem	Új alapelem + időszerkezet + alapfrekvencia	Új alapelem + koartikuláció	Példa	Példaszám	
egy	kezdő	EGY	EGY	1	
		E(TY)	E(TY)SZÁZ	124	
	belső	egy	egy	...harmincegy...	531468
		e(ty)	e(ty)	...e(ty)száz...	5129
		(n)egy	(n)egy	...ötven(n)egy...	451689
	záró	egy.	egy.	...százegety.	5301
		(n)egy.	(n)egy.	...kilencven(n)egy.	5091
		egy,	egy,	...harmincegety,	631-22-22
		(n)egy,	(n)egy,	...huszon(n)egy,	521-22-22

2. ábra. Az egy számelem optimális változatai egy fonetikailag jól megtervezett számfelolvasó esetén (Forrás: Olasz 2010a: 288)

Az ábrán az *egy* számelem különféle szükséges bemondatait látjuk. A tervezéskor figyelembe vették, hogy az *egy* elem kezdő (pl. **egy**ezer...), belső (pl. **kétezer-egy**százhat) vagy záró helyzetű (pl. **négyszázkilencvenegy**) is lehet. A kezdő helyzetben az *egy* főhangsúlyosan kell, hogy elhangozzon, magasról induló dallammal; míg a záró helyzetben hangsúlytalanul, alacsony hangmagasságon. Az *egy* esetében a mássalhangzó hosszúsága is függ a számelem helyzetétől, illetve az őt követő számelem kezdő hangzójának minőségétől. Ha az *egy* szerkezet végén áll, a *gy*-t hosszán ejtjük, akár csak két magánhangzó között, mint például az *egy*ezer esetében. Ha azonban kezdő vagy belső helyzetű, és mássalhangzó követi (pl. *egy*millió), a *gy* ejtése rövid lesz. Mint a példából látható, a jó minőségű hangzás szempontjából az sem mindegy, hogy figyelembe vesszük-e, milyen számelem követi az épp aktuálisat. Az *egy* olyan zöngés mássalhangzóra végződik, amely részt vesz a zöngességi hasonulás fonológiai folyamatában. Vagyis ha ezt a számelemet zöngétlen mássalhangzó követi, az *egy* *gy*-je *ty*-vé zöngétlenül. Ez történik az *egyszáz* → *ety száz* kiejtésekor. A modell azt is figyelembe veszi, ha a belső vagy záró helyzetben álló *egy* előtt nazális mássalhangzó van, hiszen ebben az esetben az *egy* magánhangzója nazalizálódhat. Mindezek alapján a jól megtervezett számfelolvasóban az *egy*-nek 9 változatát kell rögzíteni és eltárolni.

## 2.2. Automatikus szövegfelolvasás

A szövegfelolvasók (angol rövidítésük alapján TTS-nek is nevezik, mint text-to-speech, szöveg-beszéd átalakító) a kötött szótárasoknál sokkal bonyolultabb rendszerek. Hiszen míg a kötött szótáras alkalmazások előre megadott szövegelemeket használnak, addig a szövegfelolvasóknak az a céljuk, hogy (elvben) bármilyen témájú, műfajú szöveget képesek legyenek az emberi beszédhez hasonló hangzással

meghangosítani. Ahhoz, hogy ezt a célt el tudják érni, nyelvi és prozódiai modelleket és elemzőket használnak, amelyek a megszólaltatáshoz szükséges információkat kinyerik az írott szövegből (Olaszy 2010b). Például megadják, hogy a gépi felolvasó hol tartson szünetet, hová tegyen hangsúlyt stb. Egyes szintézismódszerekben ezeknek nagyobb a jelentőségük, másokban kisebb.

Ahhoz, hogy a számítógép meg tudja hangosítani az írott szöveget, mindeneke-lőtt olyan formára kell hozni az írást, hogy azt a gép fel tudja dolgozni, és abból hangzó anyagot tudjon előállítani. Ehhez olyan módosításokra van szükség, mint például a számoknak kiejthető betűsorrá való átalakítása (pl. 8.25 → *nyolc óra huszonöt perc, 1945. 05. 09. → ezerkilencszáznegyvenöt május kilenc*). A betűszavakat, rövidítéseket, mértékegységeket fel kell oldani (pl. *stb.* → *satöbbi, kg* → *kilógram, SMS* → *esemes*), az idegen vagy hagyományos írásmódú szavakat ugyan-csak a számítógép által feldolgozható, egységes írásmódra kell alakítani (pl. *e-mail* → *ímél, New York* → *nyújork, Batthyány* → *battyányi*). SMS-ek, e-mailek estében az is gyakori, hogy a karakterek számának csökkentése céljából a szöveg írója betű-szám kombinációkat vagy nem szokványos rövidítéseket alkalmaz, ezeket is normál formára kell hozni (pl. *5let* → *ötlet, Lmegy* → *elmegy, vok* → *vagyok*). Az emotikonok kezelése is az előfeldolgozás része.

Ha a szöveg már a kívánt betűkarakterek sorozataként áll rendelkezésre, ezt kö-vetően lehet létrehozni a kiejtési modellt a magyar nyelv fonológiai-fonetikai szabá-lyai alapján. Vagyis a betűsört beszédhangokká alakítják. Itt azokat az ismereteket kell alkalmazni, hogy például mi történik két mássalhangzó vagy két magánhangzó találkozásakor a folyamatos beszédben (pl. az előbbi esetében hasonulások, rövidü-lések, kiesések, az utóbbi estében hiátustöltés jelentkezhethetnek). Ezeket általában szabályok vezérlik, hiszen zöngés és zöngétlen mássalhangzók találkozásakor min-dig történik zöngésségi hasonulás. Ugyanakkor szükség van egy ún. kivételszótárra is, amely azokat az eseteket tartalmazza, amikor valamilyen elvárt alkalmazkodási folyamat mégsem történik meg. Például az *átjön* szót nem ejtjük *tty*-vel, pedig a *tj* kapcsolatokat általában igen (*botja, látja, mutatja*).

A nyelvi elemzés és átalakítás után kerül sor az előállított beszéd meghangosítá-sára, vagyis az akusztikai szerkezet megvalósítására. A fejlesztők erre többféle módszert dolgoztak és dolgoznak ki, és ezek összekapcsolása (ún. hibrid módszerek fejlesztése) is gyakori megoldás. A létező módszertanok közül most csak a formáns-szintézist, a diád-triád alapú szintézist és az ún. korpuszos (változó elemhosszúsá-gú) szintézist mutatjuk be. Ezek a beszéd akusztikai szerkezetéből indulnak ki, de vannak olyan módszerek is, amelyek az emberi beszédkeltés artikulációs sajátossá-gait modellezik és másolják le. (A további tájékozódáshoz ajánljuk a Németh és Olaszy által szerkesztett kötetet (2010) és a hozzá tartozó honlapot: <http://magyarbeszed.tmit.bme.hu/index.php?p=home>.)

A **formánszintézis** a számítógépes beszéd-előállításra elsőként kidolgozott módszer. Az USA és Svédország jártak legelől a módszer alkalmazásában, és Ma-gyarországon volt a harmadik olyan kutatóhely, ahol létrehoztak ilyen rendszert.

A HUNGAROVOX-ot 1982-ben mutatták be. A formánsszintézis módszer lényege, hogy a számítógéppel leutánozzák az emberi beszédet, oly módon, hogy az emberi beszédből elemzett és kivont sok-sok paramétert betáplálják a számítógépbe egy beszédmodell segítségével. A beszédet előállító szabályok összeállítása hosszadalmas, és mély fonetikai szaktudást igényel. A legismertebb és legszélesebb körben alkalmazott magyar formánsszintetizátor a MULTIVOX (1990 óta), amely szabadon hozzáférhető és használható ([http://magyarbeszed.tmit.bme.hu/index.php?p=multivox\\_letoltes](http://magyarbeszed.tmit.bme.hu/index.php?p=multivox_letoltes)). A honlapon megszólaltatható hangminták alapján jól hallható, hogy ennek a hangzása robotos, gépies, fémes, tehát nem közelíti meg az emberi hangszínezetet, ugyanakkor nagyon jól érthető. A formánsszintézisnek nagy előnye, hogy kis tárkapacitást igényel, ezért bármilyen (akár régi) számítástechnikai eszközön, telefonon futtatható. Mivel (hozzáértéssel) nagyon könnyen lehet módosítani a hangot, jól alkalmazható olyan fonetikai kísérletekben, amelyeknél egy-egy paraméter szisztematikus módosítása a cél (amire egy humán beszélő nem volna képes).

A számítógépes beszéd-előállításban ma már sokkal jellemzőbb, hogy emberi beszédből hoznak létre adatbázisokat, és a szövegfelolvasók ezekből állítják össze a hangzó beszédet. Hogyan történik mindez?

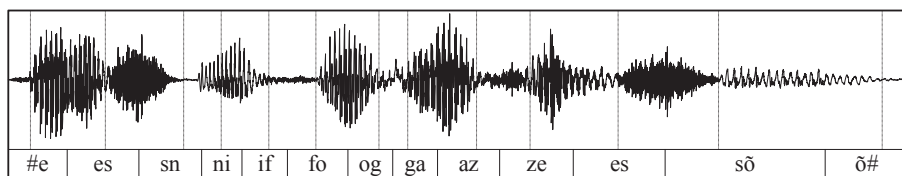
Egy fonetikában járatlan személy előállhatna egy olyan ötlettel, hogy állítsunk elő magyar beszédet a számítógép segítségével úgy, hogy felvesszük az összes (a magyar esetében kb. 40 db) beszédhangot egy valaki bemondásában, és létrehozunk egy algoritmust, amelynek nem volna más dolga, mint hogy ezeket az elemeket az írott forma alapján egymás után, szünet nélkül lejátszsa. Milyen lenne ez a beszéd? Természetellenes lenne a ritmusa, a hangsúlyozása, a dallama, a tagolása. Valószínűleg annyira eltérne az emberi beszédétől, hogy első hallásra meg sem értenénk, de biztosan nagyon fárasztó lenne hallgatni.

Egy fonetikai ismeretekkel rendelkező szakembernek azonban vannak ismeretei arról, hogy a beszédhangokat folyamatosan formáljuk, és a közöttük lévő hangátmenetek igen fontosak mind az artikuláció, mind pedig a beszédfeldolgozás szempontjából (lásd Gósy Mária (2016) írását a jelen kötetben). Valamint azt is tudja, hogy milyen nagy a beszédben a szupraszegmentumok jelentősége. Ezért ő feltehetőleg módosítaná az ötletet. Két javaslata lenne: 1. ne a beszédhangokat vegyük fel önmagukban, hanem rögzítsük a hangátmeneteket; 2. hozzunk létre egy prozódiai modellt, amely gondoskodik arról, hogy a beszéd szupraszegmentális szerkezete megfelelő legyen!

Ezzel a logikával hozták létre az ún. **diádos** szintézist. A diádok két fél beszédhangnyi hullámformarészletet tartalmaznak (lásd 3. ábra). Ha azt szeretnénk, hogy a számítógép a *Jó napot!* hangsort mondja ki, ehhez a következő diádokra lesz szükségünk:

1. diád: [szünet] + a *j* beszédhang első fele,
2. diád: a *j* beszédhang második fele + az *ó* beszédhang első fele,
3. diád: az *ó* beszédhang második fele + a *n* beszédhang első fele,
4. diád: a *n* beszédhang második fele + az *a* beszédhang első fele,

5. diád: az *a* beszédhang második fele + a *p* beszédhang első fele és így tovább.

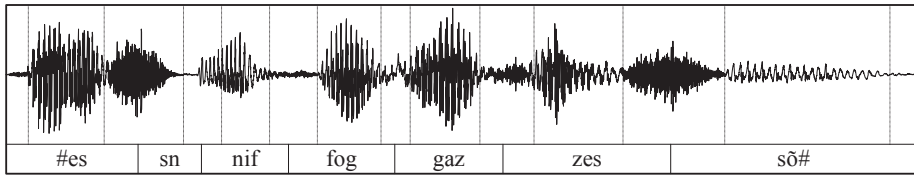


3. ábra. Az *Eszni fog az eső* mondat szintetizált hullámformája 13 diádból összefüvve (Forrás: Olaszky 2010: 293)

40 beszédhanggal számolva 1600 diád nagyjából lefedi az egy nyelv hangzásában megjelenhető beszédhangkapcsolatokat. Ez azért előnyös, mert viszonylag kis méretű adatbázist kell kezelni. Ugyanakkor nem mindegy, hogy ezeket a diádokat hogyan vesszük fel, ez gondos tervezést igényel. Például a diádokat értelmetlen hordozó hangsorokban (álszavakban, logatomokban), monoton kiejtéssel kell felvenni, állandó tempóval, mert csak így lehet biztosítani, hogy a hangzásuk igen hasonló legyen. A diádos adatbázisból behívott, majd egymás után fűzött elemek azonban önmagukban még nem adnak jól hangzó beszédet. Ehhez szükséges a megfelelő prozódia ráültetése, az akusztikai paraméterek módosításával. Ilyen például az, hogy eldöntendő kérdő mondat esetében az utolsó előtti szótagon frekvenciacsúcsot kell elhelyezni (ennek akusztikai paraméterei természetesen aprólékos tervezést kívánnak meg), vagy az, hogy vesszőnél szünetet tart, és új dallamívet indít a szövegfelolvasó.

Ilyen diádos módszerrel hozták létre a magyar ProfiVox (1995 óta) beszédsszintetizátor egyik első változatát ([http://magyarbeszed.tmit.bme.hu/index.php?p=profivox\\_letoltes](http://magyarbeszed.tmit.bme.hu/index.php?p=profivox_letoltes)).

A diádos szintézisnek sok előnye (az adatbázis kis tárigénye, egyszerű kezelhetősége, kis hibaaránya, könnyű javíthatósága) mellett hallható hátránya volt, hogy az elemek összefüvési helyein az illesztés gyakran hallatszott, és ez torz, gépies hangzást idézett elő, még annak ellenére is, hogy az adatbázist emberi beszédből hozták létre. Ennek kiküszöbölésére merült fel az az ötlet, hogy ne legyen illesztés nagy energiájú helyeken (ahol ez a legjobban hallatszik), azaz a magánhangzókat ne két elemből illesszék össze. Ennek érdekében a ProfiVox adatbázisát ún. **triádokkal** egészítették ki (vö. 4. ábra), ezeknek a szerkezete „mássalhangzó második fele + magánhangzó + mássalhangzó első fele” volt. A magánhangzó-magánhangzó és a mássalhangzó-mássalhangzó kapcsolatokra továbbra is diád elemeket használtak.



4. ábra. Az *Esni fog az eső* mondat szintetizált hullámformája 6 triádból és 1 diádból összefűzve (Forrás: Olaszky 2010a: 299)

A számítógépes technológia fejlődésével egyre kevésbé volt annak jelentősége, hogy a beszédszintézishez használt adatbázisok kis tárhelyet foglaljanak. Ennek megfelelően a magyar fejlesztések is a nagy beszédatadbázisok irányában indultak el, elsődlegesen a változó elemhosszúságú egységek kiválasztásának módszertana, később pedig a gépitanulás-alapú módszerek felé. Összefoglalónkban az előbbit tárgyaljuk részletesen.

Az **elemkiválasztásos technológia** a beszédatadbázisban (korpuszban – ezért gyakran korpuszos szintézisnek is nevezik) való közvetlen keresésen és a talált hullámformarészek közvetlen összefűzésén alapul (Olaszky 2010a: 300). Az adatbázis ez esetben többórányi, célzottan felcímkézett beszédet tartalmaz (lásd Varjasi Gergely (2016) írását a jelen kötetben). Ellentétben a diád-triád alapú adatbázisokkal, ez bővíthető, és nem kötött elemeket, hanem mondatokat tartalmaz. A bemondó szempontjából ez azt jelenti, hogy míg a diádok és triádok felvétele néhány órás igénybevétel, a korpuszos szintézishez több ülésben, alkalmanként több órán keresztül kell felvenni a hanganyagot, ráadásul úgy, hogy a beszélő hangszínezete eközben, illetve alkalomról alkalomra nem változhat jelentősen. Hiszen ha így lenne, az a szintézisben komoly minőségi romlást okozna (akár szavanként, szintagmáknak eltérő hangzást). Ez egyrészt azt jelenti, hogy a bemondónak professzionális beszélőnek kell lennie, másrészt azt is, hogy egy adatbázishoz csak egy személy hangja használható fel.

Míg a diádos-triádos vagy a formánszintézissel bármilyen szöveg meghangosítható, a korpuszos szintézis témaspecifikus. Ahhoz ugyanis, hogy jól hangzó beszédet valósíthassunk meg vele, valamilyen témára kell korlátoznunk a felolvasható szöveget. Mindenki számára elérhető a metnet.hu oldalon működő elemkiválasztásos szintézis, amely a napi időjárás-előrejelzést hangosítja meg, de ezzel a módszerrel működik a Keleti pályaudvar új utastájékoztató rendszere is (mindkettő a BME TMIT fejlesztése).

Az elemkiválasztásos szintézis előállítására röviden: felmerül az igény valamilyen témában arra, hogy a számítógép hangosítson meg szövegeket. A fejlesztők az adott témában gyakran elhangzó közléseket összegyűjtik (pl. kivonatolnak időjárás-jelentéseket), és különféle szempontok alapján összeállítanak 5-10 ezer mondatot. A szempontok között szerepel például, hogy a gyakori kifejezések a mondatban többféle helyzetben jelenjenek meg (a mondat elején, közepén, végén, tagmondat-

határon stb.). Erre azért van szükség, mert az elemkiválasztásos módszerben nincs külön prozódiaárúltetés (vagyis nem utólag kapja meg a szöveg a hangsúly-, hanglejtés- stb. mintázatokat), hanem már az elem kiválasztásakor igyekeznek a válogató algoritmus olyan részletet találni, amely a mondatbeli helyzetét tekintve is hasonlít a felolvasandó szerkezet mondatbeli helyzetéhez.

Nézzünk erre egy példát! A szintetizátornak a következő mondatot kell meghangosítania: *A szélcsendes délnyugati völgyekben néhol ködfoltok képződhetnek.* A válogató algoritmus a lehető legnagyobb mértékben egyező részt keresi az adatbázisban. Ha talál olyan mondatot, amely teljes egészében megfelel a felolvasandónak, akkor azt választja be. Ha ilyet nem talál, akkor próbál minél nagyobb egyezést találni. Feltételezhetjük, hogy megtalálja a *néhol ködfoltok képződhetnek* szerkezetet. Önmagában a szöveg szerinti egyezés azonban nem elegendő, mert ha ezt az egybeálló részt egy mondat elején találja, a *néhol* sokkal magasabb dallammal indulna ahhoz képest, mint amelyet a fenti, előállítandó mondatban elvárnánk, ahol a mondat végén szerepel ez a szerkezet. Így nem választja ki az algoritmus ezt a szerkezetet, hanem tovább keres, míg prozódiaileg megfelelőt nem talál. Ha nem talál megfelelőt, akkor a *néhol*, a *ködfoltok* és a *képződhetnek* szavakat esetleg külön-külön fogja kiválasztani az adatbázisból, olyan helyekről, ahol ezek mondat belsejében vagy végén szerepelnek.

Természetesen így is előfordulhat, hogy az algoritmus csak olyan elemeket talál az adatbázisban, amelyek nem felelnek meg prozódiai szempontból. Ekkor egy ilyet fog kiválasztani, a fejlesztők pedig a minőség-ellenőrzés során – hallva az eltérést – korrigálják ezt a hibát oly módon, hogy az adatbázist bővítik az adott szerkezetet megfelelő prozódiaival tartalmazó mondattal. (Az ilyen hibákat folyamatosan gyűjtik, és időről időre újabb megtervezett mondatcsoportokat olvastatnak fel a bemondóval, akinek ezen alkalmakkor újra rá kell hangolódnia a korábbi beszédminták beszédajátosságaira.)

Mindemellett az is lehetséges, hogy egy-egy kisebb-nagyobb nyelvi egység nem áll rendelkezésre az adatbázisban. Ekkor a ProfiVox a diád-triados elemtárból pótolja ki a hiányt.

A gépi tanulási módszerek (ma Magyarországon ez a rejtett Markov-modell alapú, HMM-szintézist jelenti) ugyanilyen aprólékosan címkézett beszédatadatbázisokat használnak. A gépi tanulás előnye egyebek mellett az, hogy kisebb induló adatbázis is elegendő, mivel a gépi tanuló algoritmusok fel tudják venni a bemondó beszédajátosságait, így a bővítéshez sem szükséges a bemondó jelenléte. Ebből az is következik, hogy sem a téma, sem a bemondás hangzása nem korlátozott, bármilyen beszélő hangjára lehet adaptálni ezeket a rendszereket. Ezeknek a fejlesztése előtt azonban még hosszabb út áll.

### 3. Számítógépes beszéd felismerés

Az automatikus beszéd felismerés szintén egy nagyobb csoportot alkot, hiszen számos részterületet foglal magában. A beszéd felismerés legismertebb célja, hogy az ember által kimondott szavakat, szövegeket automatikusan leiratozza, vagyis az ember által gerjesztett hullámformát szöveges karakterekké alakítsa. Emellett azonban számos más részterület is ide tartozik, hiszen a tartalmi leiratozáson kívül az is fontos lehet, hogy kitől származik az elhangzott beszéd, milyen érzelmi, egészségügyi állapotban van az illető. A „Ki beszél?” kérdéssel az automatikus beszélő felismerés foglalkozik. Az egyén hangulatának automatikus felismerésével az érzelme felismerés, az egészségügyi állapot meghatározásával pedig a klinikai beszéd felismerés.

#### 3.1. Beszéd felismerés

A beszéd komplex folyamat, ahol az információ akusztikai formában közvetítődik, azonban nyelvi tartalmat hordoz. Ezért a beszéd felismerésben pusztán csak az akusztikai jel felől közelíteni nem elégséges, hanem valamiféle nyelvtant is létre kell hoznunk, amely azt adja meg, hogy az egyes modellezett beszédegységek (lehetnek ezek beszédhangok, szótagok, szavak stb.) milyen valószínűséggel követik egymást. A beszéd felismerő tehát két nagyobb részből tevődik össze, egy akusztikai modellezésből és egy nyelvtanból. A mai beszéd felismerők szinte kizárólag rejtett Markov-modellt használnak.

A gépi beszéd felismerésnek többféle változata van az artikuláció, a beszélő, az akusztikai környezet és a szótárméret függvényében. Az artikuláció szerint lehet izolált szavas beszéd felismerő, amely szavak felismerésére alkalmas, illetve folyamatos, amely képes folyamatos beszéd felismerésére, így ez áll a legközelebb az emberi beszéd felismeréshez. A beszéd felismerő lehet beszélőfüggő, illetve beszélőfüggetlen. A beszélőfüggetlen felismerő alapvető célja, hogy olyan modelleket alkosson, hogy a beszéd felismerőt bármely felhasználó használhassa. Az akusztikai környezet szintén jelentősen befolyásolja a beszéd felismerő működését. Csendes körülmények között elhangzott beszéden a beszéd felismerő pontos eredményt tud adni, míg zajban az eredmények jelentősen romlanak. Kiemelt feladat a telefonos beszélgetések automatikus leiratozása. Ez a feladat abban tér el a fentiektől, hogy a telefon (telefontípustól függően) más-más frekvenciasávot ereszt át (jellegzetesen 300–3800 Hz közötti frekvenciatartományt). A beszéd felismerésben használt szótár mérete is fontos szempont. Léteznek kicsi (< 100 szó), közepes (100–1000 szó), nagy (> 10 000 szó) és kötetlen szótáros beszéd felismerők. Ez azt adja meg, hogy hány szó felismerésére képes a rendszer.

A beszéd felismerő rendszer eredményességét számos körülmény nehezítheti. Az egyik ilyen a beszéd stílus, hiszen a felolvasott beszédet a rendszer közel 80-90%-os pontossággal ismeri fel, a spontán beszédet csupán 50-60%-os pontossággal, hiszen

ennek akusztikai változatossága nagyobb, a beszédhangok kiejtése pontatlanabb, illetve a nyelvtana is kevésbé feltérképezett. A másik nehézség a nyelvi adaptálhatóság. A morfológiailag gazdag nyelvekre, mint amilyen a magyar is, a felismerés eredménye rosszabb, mivel igen nagy a ritka szavak száma (itt a *kutya*, *kutyának*, *kutyái* stb. mind külön szóként reprezentálódik a szótárban), ezért nagyon nagy szótárra lenne szükség, illetve igen nagy a szótáron kívüli elemek száma is.

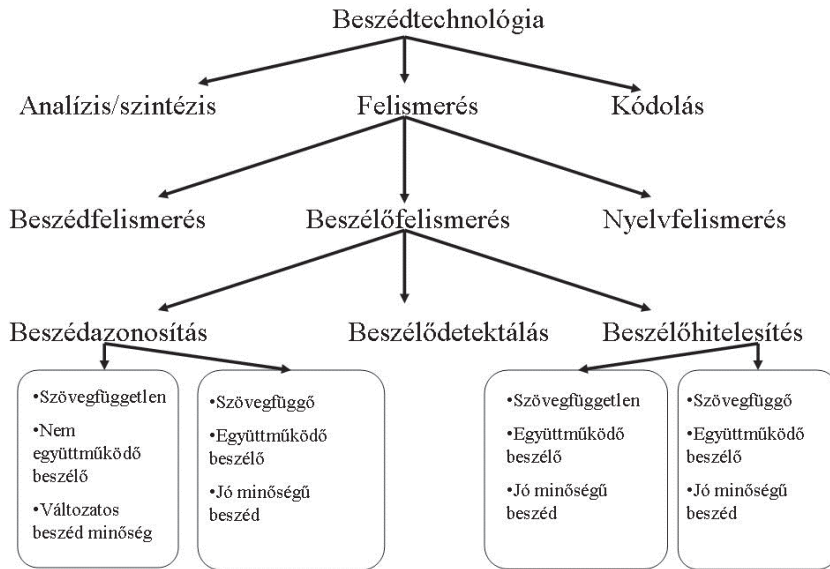
A beszédfelismerő rendszert számos helyen alkalmazzák a már eddig említetteken kívül, mint az orvosi lelet automatikus leiratozása vagy a beszédterápia, az olvasásfejlesztés (pl. Beszédmester). További alkalmazás az audiovizuális számítógépes beszédfejlesztő program beszédhibás gyermekek részére (Varázsdoboz).

A beszédfelismerésről részletesebb leírás olvasható (Németh–Olaszy szerk. 2010; Mihajlik 2013).

### 3.2. Beszélőfelismerés

A mindennapi életben képesek vagyunk akár néhány másodperces hangmintából azonosítani az általunk ismert személyeket. Ez azért lehetséges, mert a beszédhang olyan akusztikai jegyeket tartalmaz, amelyek jól reprezentálják az adott egyént (Böhm 2007). Kutatások kimutatták, hogy a hangfelismerésért, akárcsak az arcfelismerésért, egy külön agyi terület felelős. Képpalkotó eljárások ugyanis bizonyították, hogy más-más agyterület aktiválódott az ismert és nem ismert személy beszédének hallgatása során (Belin et al. 2004, idézi Böhm 2007). Az ismert személyek felismerése mellett képesek vagyunk a nem ismert személyekről is profilt készíteni, vagyis általános információkat adni például a nemre (Lass et al. 1976), az életkorra (Ptacek–Sander 1966; Gocsál 1998), testalkatra (Dommelen–Moxness 1995; Gósy 2001) vagy hangulatra (Scherer–Banse–Wallnott 2001) stb. vonatkozóan.

A gépi beszélőfelismerés alapvetően három területre osztható (vö. 5. ábra). Megkülönböztetünk beszélőazonosítást (speaker identification), beszélőhitelesítést (speaker verification) és beszélődetektálást (speaker diarization) (Bimbot et al. 2004). A beszélőhitelesítés célja, hogy a rendszer egy személyről eldöntse, hogy ő az, akinek állítja magát. Ez a cél megegyezik a többi biometrikus személyazonosítás (pl. ujjlenyomat, íriszvizsgálat) céljával. Ebben a feladatban bináris döntést kell hoznia a gépnek: elfogadás/elutasítás. Ekkor a beszélőnek érdeke, hogy a gép felismerje a hangját, ezért a beszédminőség igen jó. Ezzel ellentétben a beszélőazonosítás célja, hogy a beszélők egy lehetséges köréből kiválasszuk az aktuálisan beszélőt. Lehetséges azonban az is, hogy a lehetséges beszélők halmaza nyílt, vagyis a beszélő nincs benne a halmazban, ekkor a rendszer ismeretlen személyként kell, hogy azonosítsa a beszélőt. A beszélődetektáláskor két- vagy több-beszélős társalgásokban kell azonosítani azt, hogy ki mikor beszél. A beszélőazonosítás és a beszélőhitelesítés lehet szövegfüggő vagy szövegfüggetlen. A kutatók általában a szövegfüggetlen osztályozásra törekkenek, mivel ekkor tetszőleges tartalmú beszédminta alapján történhet a beszélő azonosítása vagy hitelesítése.



5. ábra. A gépi beszédfeldolgozás területei

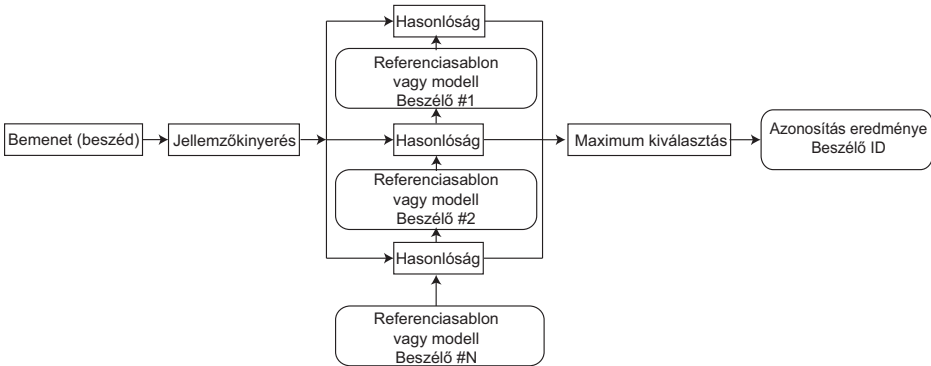
A beszélőhitelesítés napjainkban egyre inkább megoldottnak tűnik, mivel közel 98-99%-os eredménnyel működik. A beszélőazonosítás eredményei ehhez képest jóval változatosabbak. Az eredmények nagyban függenek a felvétel minőségétől, azaz hogy milyen zajos a felvétel, és a beszédminta hosszától stb. A gyakorlatban legtöbbször igen rövid akusztikailag feldolgozható minta áll rendelkezésre az azonosításhoz (Nikléczy 2001, Nikléczy–Gósy 2008). Kutatások szerint a legrövidebb beszédminta hossza, ami még alkalmas az azonosításra, 16 másodperc (Nikléczy–Gósy 2008).

A beszélőfelismerés öt lépésből áll: a beszédjel tisztítása, jellemzőkinyerés, beszélőmodellek létrehozása, mintaillesztés, döntés (6. ábra).

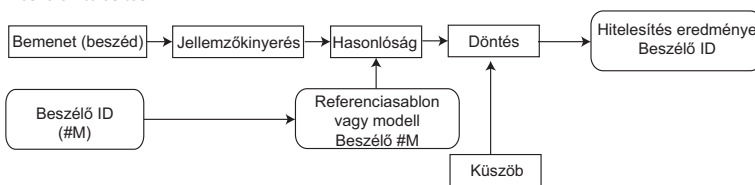
A bemeneti beszédjelből eltávolítjuk azokat a részeket, amelyek nem járulnak hozzá a beszélő személy felismeréséhez vagy nehezítik azt. Ilyen tipikus eljárás a zajszűrés, beszédjeltisztítás, amely során a beszédből eltávolítjuk a zaj minél nagyobb részét, javítva ezzel a jel/zaj viszonyt. A másik eljárás a beszédetektálás, amely során csak azokat a részeket tároljuk el, ahol a beszélő valóban beszél, kiszűrve ezzel a szüneteket, hosszabb légvételeket, zajos részeket. A beszédjel megtisztítása után számítjuk ki az akusztikai jellemzőket. Az akusztikai jellemzők igen sokfélék lehetnek. A jellemzőkinyerés célja az, hogy megtaláljuk azon akusztikai tulajdonságokat, amelyek mentén az egyes beszélők elkülöníthetők, azaz amelyek beszélőszemély-specifikusak. Az akusztikai jellemzőknek ugyanakkor egyszerűen mérhetőnek, minden beszélőnél jól mérhetőnek, érzelmi állapottól függetlennek kell

lenniük. A feladatra használt akusztikai jellemzők száma igen nagy, azonban továbbra is kérdés marad, hogy létezik-e, és ha igen, akkor mely akusztikai paraméterben mutatható ki az egyéni hangszínezet.

#### Beszélőazonosítás



#### Beszélőhitelesítés



6. ábra. A beszélőazonosítás (fent) és a beszélőhitelesítés (lent) folyamatábrája

A jellemzőkinyerés után előállnak az úgynevezett jellemzővektorok, amelyek alapján elvégezhető az osztályozás. Az osztályozáshoz a beszédfelismerésben is használt algoritmusokat szokás alkalmazni (pl. kevert Gauss-modell, rejtett Markov-modell, neurális hálózatok, szupport vektor gépek és ezek kombinációi). A beszédfelismeréshez képest azonban a beszélőhitelesítéskor (6. ábra lent) a modellek közötti hasonlóság mérését végezzük, ami a referencia-adatbázisban található személyek modelljei és az aktuálisan azonosításra kerülő személy modellje közötti hasonlóság mérését jelenti.

### 3.3. Érzelemfelismerés

Az érzelmi töltet felismerése viszonylag fiatal ága a beszédfelismerésnek (Sztahó 2014). Napjaink célkitűzése e területen az, hogy 4 ún. alapérzelmet (haragos, örömteli, semleges, bánatos) különítsenek el akusztikai jellemzők alapján gépi osztályozó módszerekkel. A számos felhasználási lehetőségen túl a felhasználó érzelmeinek

követése sokat segíthet a dialógusok dinamikus felépítésében, a beszélő érzelmeire adekvát gépi válasz kiválasztásában, ily módon az ember-gép kommunikáció teljesebbé tételében.

#### 4. Kitekintés

Napjainkban egyre több munkát és anyagi erőforrást összpontosítanak arra, hogy az ember minél természetesebben tudjon érintkezni az őt támogató gépekkel. Ez nem elsősorban a kényelmünket szolgálja, hanem a minket körülvevő nagymennyiségű információ feldolgozásában lehet segítségünkre. Emellett igen nagy szerepe van a valamilyen kommunikációs hátránnyal élők életminőségének javításában is; továbbá azon betegségek diagnosztizálásában is jelentőségük van a beszédtechnológiai alkalmazásoknak, amelyek a beszédben a betegség korai stádiumában produkálnak tüneteket (pl. a gége eltérései hallhatók a zöngeminőségben; az Alzheimer-kór a beszéd időzítésében korán tetten érhető). A mesterséges intelligencia kifejlesztése közben a humán gondolkodás sajátosságait is mélyebben megismerhetjük.

#### Irodalom

- Belin, Pascal – Fecteau, Shirley – Bédard, Catherine 2004. Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences* 8/3. 129–135.
- Bimbot, Frédéric – Bonastre, Jean-François – Fredouille, Corinne – Gravier, Guillaume – Chagnolleau, Magrin Ivan – Meignier, Sylvain – Merlin, Teva – Garcia, Ortega Javier – Petrovska-Delacrétaz, Dijana – Reynolds, Douglas A. 2004. Tutorial on text-independent speaker verification. In: *Proceedings of EURASIP, Journal on Applied Signal Processing* 4. New York, USA. 430–451.
- Bóhm Tamás 2007. Beszélőfelismerés – neurológiai háttér és pszichológiai modellek. *Magyar Pszichológiai Szemle* 62/4. 541–563.
- Dommelen van, Wim A. – Moxness, Bente H. 1995. Acoustic parameters in speaker height and weight identification: Sex-specific behaviour. *Language and Speech* 38. 267–287.
- Gocsál Ákos 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutató* 1998. 122–134.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125/4. 478–487.
- Gósy Mária 2016. Beszédhangok viselkedése a spontán beszédben. In Bóna Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest, 19–31. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu).
- Lass, Norman J. – Hughes, Karen R. – Bowyer, Melanie D. – Waters, Lucille T. – Bourne, Victoria T. 1976. Speaker sex identification from voiced, whispered and filtered isolated vowels. *Journal of the Acoustical Society of America* 59. 675–678.
- Németh Géza – Olasz Gábor szerk. 2010. *A magyar beszéd. Beszédkutató, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest.
- Nikléczy Péter 2001. A műszeres személyazonosítás lehetőségei rövid időtartamú beszédminták alapján. *Beszédkutató* 2000. 154–172.

- Nikléczy Péter – Gósy Mária 2008. A személyazonosítás lehetősége a beszédanyag időtartamának függvényében. *Beszédkutató 2008*. 172–181.
- Olaszy Gábor 1999. Beszédadatbázisok készítése gépi beszéd-előállításához. *Beszédkutató '99*. 68–89.
- Olaszy Gábor 2010a. Beszédből készített elembázisok beszédszintézishez. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutató, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 283–310.
- Olaszy Gábor 2010b. Automatikus szövegfelolvasás. In Németh Géza – Olaszy Gábor (szerk.): *A magyar beszéd. Beszédkutató, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 429–445.
- Pandur Balázs 2011. *Üzenetkezelő rendszer vak és látássérült felhasználók részére mobil eszközön*. TDK-dolgozat. BME, Budapest.
- Ptacek, Paul H. – Sander, Eric K. 1966. Age recognition from voice. *Journal of Speech and Hearing Research* 9/2. 273–277.
- Scherer, Klaus R. – Banse, Rainer – Wallbott, Harald 2001. Emotional inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology* 32/1. 76–92.
- Sztahó Dávid 2014. Automatikus érzelem-felismerés akusztikai paraméterek alapján. PhD értekezés. BME TMIT, Budapest.
- Varjasi Gergely 2016. Beszédadatbázisok. In Bóna Judit (szerk.): *Fonetikai olvasókönyv*. ELTE Fonetikai Tanszék, Budapest, 233–244. [www.fonetikaitanszek.hu](http://www.fonetikaitanszek.hu).

### Kérdések, feladatok

1. Gyűjtsön az irodalomból, a filmtörténetből olyan alkotásokat, amelyekben beszélő számítógépek szerepelnek. Hogyan kommunikál ezekben az esetekben a gép és az ember? Milyen hasznuk van ezeknek a gépeknek?
2. Képzeld el, hogy egy vállalat Önt kéri fel, hogy a számítógép vezérelt telefonközpontjához válasszon beszélőt! Milyen szempontokat venne figyelembe, amikor kiválasztja az illetőt?
3. Folytassa a *Jó napot!* hangsor bemondatásához szükséges diád elemek felsorolását! Ugyanezt a hangsort milyen triád elemekből lehetne létrehozni?
4. Hallgassa meg a napi időjárás-előrejelzést a [metnet.hu](http://metnet.hu) oldalán! Mely pontokon tér el a bemondás a természetes magyar beszédétől, hol hangzik furcsán? Mi lehet ennek az oka?
5. Tesztelje az okostelefonján vagy személyi számítógépén a magyarra is létező beszédfelismerőt! Milyen minőségben működik eltérő háttérzajban, illetve különböző beszédtempó, hangerő esetén?



## Beszédatbázisok

### 1. Mik azok a beszédatbázisok?

A beszédatbázisok hanganyagokat és azok átiratait, lejegyzett változatait tartalmazzák valamilyen tudományos célból. A nyelvészek ilyen adatbázisokat használnak a vizsgálataik anyagaként. Léteznek beszédtechnológiai célú beszédatbázisok: ezeket olyan szempontok szerint hozzák létre, hogy minél hatékonyabban tudják modellezni a beszédet; ennek eredményeképpen alkotják meg a beszéd felismerőket vagy beszéd szintetizátorokat. Gósy Mária 2008-ban írja, hogy a beszédatbázisok célja, hogy „olyan mennyiségű adatot gyűjthessünk, amely kellőképpen reprezentálja egy közösség nyelvi jellemzőit, illetve nyelvhasználatát, és ezen az anyagon végezzék a különféle elemzéseket” (194) a nyelvészek.

A beszédatbázis egyfajta specifikus beszéd korpusz: létrehozásának körülményei és metodológiája egységes. A különböző típusú adatok (az adatközlőhöz kapcsolható adatok: nem, kor, iskolai végzettség stb.) egységes formában vannak kódolva, így az ezekhez való hozzáférés hatékony. Az adatbázisokban megtaláljuk a hangzó anyagokat, azok átiratait, az anyagokhoz kapcsolódó úgynevezett metaadatokat (például az adatközlők anonimizált adatait), valamint az adatbázis létrehozásához és bővítéséhez használatos módszertani leírást is.

A beszédatbázisokat többféleképpen csoportosíthatjuk: leggyakrabban tartalom alapú különbségtétellel találkozhatunk – így vannak például olvasott szövegek felvételeit vagy éppen spontán nyelvi megnyilatkozásokat tartalmazó adatbázisok is. Fontos tudnunk azonban, hogy vannak olyan típusú beszédatbázisok, amelyek a beszéd típusától függetlenül tartalmazznak anyagokat.

A beszédatbázisok létrehozása során mindig meg kell válaszolnunk bizonyos alapvető kérdéseket: egy-egy szempont különböző súllyal esik latba a korpusz céljának megfelelően. Az adatbázisok tulajdonságai:

- cél,
- nyelv,
- adatközlők (száma és jellemzői),
- felvételi körülmények,
- protokoll, feladatok,
- etikai kódex,
- annotáció (az anyag címkézése),
- tárolás.

A beszédatadtbázisok létrehozásának többféle célja lehet, hiszen beszélhetünk egy apró nyelvhasználati jelenség leírásáról, például egy betegség beszédre gyakorolt hatásának a vizsgálatáról vagy akár egy nyelv szinkron metszetének rögzítéséről is. A célunk meghatározhatja az anyagot, az adatközlőket vagy akár a metaadatainkat is: belátható, hogy egy nyelvjárási beszédet vizsgáló adatbázis tervezése során gondolnunk kell arra, hogy a metaadatoknál tároljunk szociolingvisztikai és dialektológiai információkat is.

Amennyiben egy nagyobb területet szeretnénk vizsgálni (például a budapesti felnőttek beszédét), úgy fontos az adatközlők összetételének kérdése is. Gondoljunk itt arra, hogyha szeretnénk összevetni a férfiak és a nők beszédének egy-egy jellemzőjét, akkor érdemes úgy tervezni az anyagot, hogy a nemek közel egyenlő arányban szerepeljenek a kutatás adatközlői között. A korpusz célja meghatározhatja a felvétel körülményeit is, hiszen például részletes fonetikai elemzéseket csak megfelelő minőségű hanganyagon tudunk végezni.

A korpusz céljának meghatározása során válaszolnunk kell a következő kérdésekre:

- Mit akarunk vizsgálni?
- Kiket akarunk vizsgálni?
- Milyen változók (nyelv, állapot, kor, nem, iskolai végzettség, memória stb.) alapján?
- Milyen körülmények között?
- Milyen jelenségeket jegyezzünk le? Hogyan kódoljuk azokat (annotáció)?
- Milyen feladatokkal hívhatjuk elő a számunkra fontos nyelvi jelenségeket?

Természetesen ezen szempontok dinamikusak; minden adatbázis más-más választ ad a fenti kérdésekre – a célnak megfelelően egy-egy kérdés lehet kiemelten vagy kevésbé fontos is. Például egy gyermeknyelvi (óvodáskori beszédet leíró) adatbázisban célszerű tekintettel lenni a gyermekekre: gondolunk kell arra, hogy a gyermekek mennyire lesznek kommunikatívok, képesek lesznek-e egyáltalán a vizsgálat végéig figyelni. A gyermekek általában a saját környezetükben beszédesebbek, mint egy idegen helyen, ezért a gyermeknyelvi adatbázisoknak nagy részét nem állandó helyszínen, hanem az adott gyermeknek ismerős óvodai vagy iskolai teremben veszik fel – láthatjuk, hogy ebben az esetben a gyermekekkel való sikeres kommunikáció szempontja erősebb, mint a hely állandósága.

A beszédfelvétel készítése megkerülhetetlen problémakört vet fel: milyen etikai szabályokat kell betartani egy kutatónak? Gyermekek esetében ez kiegészül egy további szemponttal is, hiszen ők még nem dönthetnek arról a kérdésről, hogy részt vesznek-e egy kísérletben vagy sem, ezért minden esetben írásos szülői hozzájárulási engedélyt kell kérni. Iránymutatásként érdemes lehet az MTA etikai kódexébe beleolvasni ([http://mta.hu/data/dokumentumok/hatteranyagok/tudomanyetikai\\_bizottsag/tudomanyetikai\\_kodex\\_kgy\\_20100504.pdf](http://mta.hu/data/dokumentumok/hatteranyagok/tudomanyetikai_bizottsag/tudomanyetikai_kodex_kgy_20100504.pdf)).

Egy fonetikai célú beszédkorpusz létrehozása során fontos, hogy a felvétel körülményeit megfelelően válasszuk ki. A környezeti zaj miatt általában kerülni szoktuk a hangos termeket, utcafronti szobákat – lehetőség szerint csendesített helyiségben készítjük a felvételeket. A hely kiválasztásánál a berendezésre is érdemes figyelni, hiszen például a székek nyikorgása vagy a visszhang nemcsak az adatközlőt zavarhatja, hanem az eredményeinket is befolyásolhatja.

A lejegyzendő jelenségek kiválasztása során az elsődleges szempontunk a kutatási kérdések megválaszolása – csak a megfelelően kiválasztott jelenségek vizsgálatával kaphatjuk meg a kérdéseinkre a pontos választ. A nagyobb adatbázisoknál a kódolás egységesítése és konzekvens használata teszi lehetővé a kereshetőséget: az anyagok számítógépes feldolgozása során ezeket a kódokat használjuk.

A feladatok meghatározása módszertani kérdés: kutatásismertetőket olvasva gyakran találkozunk különböző elicitációs módszerekkel, szerepjátékokkal is. A módszertani ismeretek fontossága nemcsak a kutatások sikerességének a mérceje, hanem a megismételhetőségnek is az alapja. A megfelelően megválasztott és kidolgozott feladatrendszer nemcsak megkönnyíti a kutató munkáját, hanem az eredmények hitelességét is alátámasztja.

## 2. Az adatbázisok különböző típusai

A nyelvi adatbázisok egyik legfontosabb célja az, hogy a nyelvi adatokat egységesen kódolva, könnyen használhatóan érhesse el a nyelvész. Amennyiben létezik egy nagy mennyiségű adatot tartalmazó, megfelelő módszertannal kiépített adatbázis, úgy a nyelvészek munkája is jóval hatékonyabb lesz.

A nyelvi adatbázisok másik fontos célja a nyelvi adatok rögzítése az utókor számára. A nyelv folyamatosan változik, így fontos, hogy bizonyos időközönként megvizsgáljuk, hogy milyen módon történik ez. Az ilyen változások következménye a szótárak, a helyesírás folyamatos frissítése is. Bár a magyar nyelvet ez nem érinti, megemlítenéd, hogy egyes, már nagyon kevés beszélő által használt nyelvek esetén (például a lív nyelv esetében), a beszédkorpusz létrehozása a nyelv fennmaradásának a kulcsát képezi.

### 2.1. Beszédkorpuszok nyelvtechnológiai és gyakorlati alkalmazások számára

A beszéd- és nyelvi korpuszokat nemcsak nyelvészek, hanem más szakemberek is gyakorta használják. A beszéddel vagy emberi hanggal kapcsolatos alkalmazások nagy része valamilyen valós emberi hanganyagon alapszik. A nyelvtechnológiának sokféle alkalmazási területe van: gondoljunk csak a

beszédfelismerő programokra, a hangalapú azonosításra, a GPS-re vagy akár egy szövegfelolvasóra.

## **2.2. Fonetikai célú beszédatbázisok**

Ahogy korábban említettük, a beszédatbázisok egy részét fonetikai vizsgálatok támogatására hozzák létre. Ezeken belül is megkülönböztethetünk spontán beszédet és olvasott beszédet tartalmazó adatbázist. A BEA és a GABI is alapvetően fonetikai kutatások számára készül, ezek bemutatását lásd később. Az összehasonlító kutatások alapja, hogy legyenek olyan eredmények, amelyekre viszonyítási alapként tudunk tekinteni. Például a „Milyen a különböző hallásérültek beszéde?” kérdésre úgy lehet választ adni, hogy az ép hallásúak (korban és nemben egyeztetett) beszédéhez viszonyítjuk azt. Az ép beszéd sajátosságainak megismerésén túl tehát az is célja ezeknek az adatbázisoknak, hogy a nem tipikus beszédet fel tudjuk ismerni.

A gyermeknyelvi beszédatbázisok létrehozásának egyik célja a nyelvelsajátítás folyamatának megismerése, továbbá az ehhez képest atipikusnak mondható fejlődés meghatározása. A gyermekek különböző ütemben tanulják meg anyanyelvüket – van, aki folyamatosan, van, aki nagyobb ugrásokban. Vannak azonban olyanok is, akik valamilyen szempontból lassabban haladnak: az artikulációjuk pontatlanabb vagy például hosszabb a nyelvfékük. A lassúság nem feltétlenül probléma, hiszen nagyon sok út van a nyelv elsajátításához; tudnunk kell azonban, hogy egy-egy ilyen probléma akár összetettebb zavart is jelezhet.

## **2.3. Orvosi célú beszédatbázisok**

A beszédatbázisok között vannak olyanok, amelyek valamilyen betegségben szenvedő emberek hangját rögzítik, feltételezve azt, hogy a hanganyag vizsgálata során válaszokat kaphatunk a betegséggel kapcsolatban. A technológia fejlődésével ma már egyre több olyan eszközzel találkozunk, amely az orvosi munkát segíti: ezek lehetnek otthoni mérőműszerek vagy diagnosztizáló segítő telefonos alkalmazások is. Ilyen például a Pocket Doctor nevű alkalmazás is, amely egyebek között az Alzheimer-kór diagnosztizálása során használ orvosi célú beszédkorpuszt. 99%-os pontossággal képesek meghatározni a Parkinson-kórt is, akár olyan beszédhangjellemzők segítségével, mint az alaphangmagasság változása vagy a magánhangzók minőségének torzulása.

A nagy mennyiségű, jó minőségű, hanggal kapcsolatos adatok nemcsak a betegség típusának felismerését és meghatározását, hanem annak kezelési módját is támogathatják. Például egy beteg levegős hangszínezete utalhat a hangszalagok normáltól eltérő állapotára. Ha rendelkezünk egy olyan adatbázissal, ahol

nagy mennyiségű normál és atipikus hangot is találunk, úgy jó pontossággal meghatározhatjuk, hogy a beteg hangja problémát jelez-e vagy sem.

Orvosi célú korpuszon alapul a Donateacry nevű kezdeményezés is. A célja az, hogy a gyermekek sírását azonosítva segítse a szülő-kisgyermek kommunikációját. A projekt lényege, hogy nagyon sok csecsemő különféle sírását felvesszük úgy, hogy a szülő megmondja, hogy most mi miatt sír a gyermek. A felhasználó a saját gyermekének a sírását felvéve, a program alkalmazásával az esetek döntő többségében meg fogja tudni határozni, hogy miért is sír a baba. Az alkalmazás neve Lullabond – The Newborn Translator.

### 3. Protokoll

Az adatbázisok létrehozásának az egyik leglényegesebb részfolyamata a felvételek elkészítése. Az adatbázis tervezése során nagy figyelmet kell fordítani ezért a felvétel forgatókönyvére, a protokollra. A nagy adatbázisok esetében mindenképpen szükséges, de még a szakdolgozathoz létrehozandó kisebb korpuszokhoz is ajánlott a protokoll, a módszertan tesztelése, hiszen a valóságban számos olyan jelenséggel találkozunk, amelyekre az első pillanatban nem gondoltunk.

A protokoll megszabja, hogy a különböző feladatokat milyen módon végezzük vagy végeztetjük el az adatközlővel. Jellemzője, hogy minden esetben nagyon szigorú, csak a legkritikább esetben térhetünk el tőle. Ennek oka az, hogy az esetleges változtatások eltérést okoznak az adatainkban, így a kapott eredményeink sem lesznek helytállóak. A protokoll bizonyos esetekben megengedhet valamilyen mértékű rugalmasságot a felvétel készítőjének, azonban a feladatok felvétel közbeni megváltoztatását, kiegészítését mindenképpen kerüljük el.

A forgatókönyv létrehozásakor figyeljünk arra, hogy az adatközlő számára teljesíthető legyen a feladat. Többnyelvűek nyelvhasználata esetében fontos a feladat instrukciójának a nyelve: ilyen esetben érdemes tájékozódni a nemzetközi szakirodalomban. A protokoll kidolgozásánál lényeges, hogy az ne öncélú legyen, hanem feleljen meg a kutatási céljainknak és az adatközlőknek is. Tekintettel arra, hogy a korpusz létrehozása igencsak forrásigényes feladat, ezért érdemes úgy összeállítani a nyelvi anyagot, hogy az hosszú távon, akár több kísérlet elvégzéséhez is biztosítson anyagokat.

A protokoll részének tekinthető az adatfelvételi lap kitöltése is. Ezekben olyan adatokat kérünk az adatközlőtől, amelyek segítségével elemezni fogjuk az anyagokat (független változók). Például a nemre, az iskolai végzettségre vagy a lakóhelyre is gondolhatunk. A kutatások során anonimizálnunk (azaz megszüntetnünk a személy és az adat közötti kapcsolatot) kell ezeket a dokumentumokat, valamint meg is kell őriznünk ezeket. A későbbi ellenőrzés miatt fontos, hogy

azonosítószámmal lássuk el az adatlapokat, így a digitalizálás során egyszerűbb lesz az adatbevitel, valamint a felvétel is visszakereshető lesz.

#### 4. Lejegyzés, annotálás

Az annotálás a vizsgálandó jelenségek bejelölését és felcímkézését jelenti. A beszédatadbázisok nemcsak hanganyagokból és metaadatokból állnak, hanem a kutató számára érdekes nyelvi jelenségek lejegyzéséből is. Az adatbázis tervezésekor ezért meg kell válaszolnunk a lejegyzéssel kapcsolatban felmerülő kérdéseket:

- Milyen jelenségeket és milyen szinten annotáljunk?
- A vizsgált jelenségeken pontosan mit értünk?
- Milyen módon annotáljunk?
- Milyen címkékkal annotáljunk?
- Milyen lejegyzési rendszerrel?
- Milyen pontossággal/hibahatárral?

Egy hangfájlban például jelölhetjük a szüneteket SIL címkével (lásd 1. ábra).



1. ábra. Példa a mondat szintű annotálásra – a kép felső részén az oszcillogramot látjuk, középen a spektrogramot, alul pedig az annotált jelenségeket: jelen esetben a szüneteket (SIL) valamint a mondatokat (PRAAT program)

Az annotálandó jelenségek meghatározását minden esetben a lehető legprecízebben végezzük el, hiszen ez egyrészt a lejegyzési munkát is segíti, valamint a kísérletek értékelésénél is nagyon fontos. Általában véve három annotációs szintről beszélhetünk: mondat szint, szó szint és hang szint. Amennyiben a hangok időtartamát szeretnénk vizsgálni, úgy a hang szintű lejegyzést végezzünk, ha a

kitöltött szünetekre vagyunk kíváncsiak, úgy a szószint is elég; mondat szinten elemezhetjük például a beszélőváltásokat is.

Ma általában háromféle módon készítünk lejegyzéseket: kézzel, automatikusan, illetve vegyesen. A kézi lejegyzés nagyon precíz, viszont általában nagyobb időigényű, az automatikus viszont gyorsabb. A két módszer ötvözését gyakran használják gépi annotáció, majd kézi ellenőrzés formájában.

Ma már a lejegyzés szinte mindenhol digitálisan történik, szoftveresen támogatva: vannak olyan szoftverek, mint a PRAAT, amely nyílt szótáras lejegyzési keretet kínál; viszont vannak olyanok is (ELAN), amely lehetőséget ad zárt szótárak használatára is. A két típus közti különbség, hogy a címkék soraiban (ún. tierekben vagy layerekben) milyen címkéket használhatunk. A zárt szótáras tierekben csak az előre meghatározott címkékből választhatunk (így némileg gyorsabb), a nyílt szótáras tierbe viszont bármilyen címke kerülhet.

Minden lejegyzési módnak van valamilyen hibahatára: tudnunk kell, hogy akár a gép, akár az ember hozza létre a nyelvi lejegyzést, mindig találhatunk majd benne valamilyen módon rosszul címkézett anyagot. Ennek a kontrollálása kiemelten fontos, ezért az adatbázist ebből a szempontból is ellenőriznünk szükséges.

## 5. Példák adatbázisokra

### 5.1. A BEszélt nyelvi Adatbázis: BEA

A BEA adatbázist az MTA Nyelvtudományi Intézetének Fonetikai Osztálya 2007 ősze óta fejleszti. Célja, hogy rögzítse a magyar köznyelv jelenlegi állapotát, továbbá lehetőséget adjon a nyelvészek számára, hogy jó minőségű hanganyagokon alapuló kutatásokat végezhesenek. A BEA anyagot kínál a különböző nyelvtechnológiai és egyéb gyakorlati alkalmazások számára is (Gyarmathy–Neuberger 2011).

A BEA tervezési, infrastrukturális-fejlesztési és képzési munkálatainak elvégzése, valamint a próbafelvételek elemzése után, 2008-ban kezdődtek el a beszédadatbázis hanganyagainak a felvételei (Gósy 2008). A felvételek általában 40-45 percesek, 8 különböző részből állnak. Az adatok összehasonlítása érdekében nemcsak spontán beszéd, hanem ismétléses és felolvasásos részfeladat is található a protokollban.

A felvételi protokoll állandó:

1. Spontán beszéd:

- a) narratíva,
- b) véleménykifejtés: aktuális téma véleményezése,
- c) tartalomösszegzés, két monológ létrehozása:

- i) egy tudománynpszerűsítő cikk alapján,
- ii) egy történet meghallgatása alapján.

2. Mondatisméltés: 25 különböző hosszúságú és változatos struktúrájú mondat utánmondása, egyszeri meghallgatás után.

3. Felolvasás:

- a) a 25, korábban megismételt mondat felolvasása,
- b) egy (az előbbtől eltérő) tudománynpszerűsítő cikk felolvasása.

4. Háromfős társalgás. Az interjúkészítőn és az adatközlőn kívül egy másik nyelvész kolléga vesz részt a beszélgetésben.

A BEA adatközlői 20 és 90 közötti, budapesti felnőttek; hallásállapotuk jó, köznyelvet beszélnek. Az adatbázis felvételei mindig ugyanazon a helyszínen készülnek, az MTA Nyelvtudományi Intézet Fonetikai Osztályának csendesített szobájában; a felvételek technikája is állandó (Gósy 2008). Az adatközlőről a következő adatokat ismerjük: kor, végzettség, magasság, súly, foglalkozás, beszédhibája (ha van), dohányzik-e. Az anonimizálás a felvételt követően azonnal megtörténik, így a beszédanyag semmilyen módon nem kapcsolható össze a beszélő nevével. A beszédanyagokat természetesen csak kutatási és oktatási célokra használhatják fel (Gósy 2013). Az adatok tárolása DVD-n és külső winchesteren történik. Az anyag annotációja háromszintes: beszédszakasz-, szó- és hangszint (Neuberger 2009), és olyan nonverbális hangokat is jelöl, mint például a tüszentes, a nevetés, a köhögés. Annotálják továbbá a különböző típusú szüneteket is.

A BEA mára számos kutatás anyagát biztosította, például a beszéd szegmentális és szuprasegmentális elemzéséhez; a koartikulációs folyamatok vizsgálatához; a beszéd hangzásának elemzéséhez; a spontán beszéd megismeréséhez (lásd Gósy szerk. 2012).

A BEA jelenleg rögzített anyagát 305 beszélő körülbelül 265 órányi felvétele képezi. A legrövidebb felvétel 24 perc 27 másodperc, a leghosszabb pedig 2 óra 24 perc és 47 másodperc. A BEA hangfelvételei tudományos célokra hozzáférhetők; hozzáférés a Nyelvtudományi Intézet Fonetikai Osztályának osztályvezetőjétől, prof. dr. Gósy Máriától kérhető.

## 5.2. Gyermeknyelvi beszédAdatBázis és Információtár: GABI

A GABI az ELTE BTK Fonetikai Tanszékéhez köthető gyermeknyelvi adatbázis. Az előzetes munkák (Imre 2011) elvégzése után, 2013-tól kezdődően gondozza a tanszék Gyermeknyelvi Kutatócsoportja, dr. Bóna Judit vezetésével. A korpusz fejlesztés alatt áll, jelenleg mintegy 250 felvétel készült el. A létrehozás során kiemelt szempont volt a fonetikai megalapozottság és az, hogy olyan, jó minőségű spontánbeszéd-adatbázis legyen, amelyet különböző tudományterületek is hasznosíthatnak. Az adatbázis célja, hogy rögzítse a köznyelvet beszélő

gyermek és kamasz beszédét (Bóna et al. 2014). Az adatközlők 3–18 éves kor közötti magyar anyanyelvű gyermekek, fiatalok. A tervek között az szerepel, hogy minden egyes életkorban 50-50 főtől legyen hanganyag. Ez az első olyan beszédadatbázis, amely a gyermekek hangját ilyen módszertannal rögzíti. Az adatbázis protokolljának létrehozása a BEA fejlesztőivel konzultálva zajlott.

A GABI kettős protokollal rendelkezik: van egy 3–9 éves kor közötti, és van egy 9–18 éves kor közötti változat is. A felvétel ideje gyermekenként 30-40 perc között változik. A kisgyermek számára a következő felvételi eljárást hozták létre (Bóna et al. 2014):

1. 15 mondat utánmondása.
2. Spontán narratíva.
3. 20 szó jelentésének meghatározása.
4. Egy hallott szöveg tartalmi összegzése. Részkorcsoportonként eltér a gyermekek érdeklődési körének megfelelően.
5. Mondatfelolvasás (második osztályos kortól).
6. Történetmesélés képek alapján.

A nagyobbak számára létrehozott protokoll a BEA protokolljának és a 3–9 évesek számára készített változatnak az ötvözése. A 9–18 éveseknél egy hetedik feladat is található: ez vita két gyermek között, amelyet a felvételvezető moderál.

A felvételek annotálása elkezdődött, jelenleg .trs (Transcriber) formátumban vannak az annotált a fájlok.

### 5.3. Budapesti Szociolingvisztikai Interjú: BUSZI

A Budapesti Szociolingvisztikai Interjú egy, az MTA Nyelvtudományi Intézet által gondozott, nagyszabású vizsgálat, amelynek célja az, hogy a budapesti lakosok nyelvhasználatát rögzítse (Kontra – Váradi é. n.). A minta életkorra és iskolai végzettségre reprezentatív. Az anyaggyűjtés 1987-ben kezdődött 50 darab próbainterjú rögzítésével. A felvételek a kor technológiájának megfelelően magnetofonnal történtek. A BUSZI adatai is anonimizáltak; a felvételekhez való hozzáférést a gondozó intézetben kérhetünk. Nem árt tudnunk, hogy bár a felvételek nyelvészeti szempontból igen hasznosak, de a rögzítés technológiai korlátai miatt fonetikai elemzésre nem kimondottan alkalmasak. Hangszintű elemzésre nem, de bizonyos szó- és mondatszintűekre megfelelő anyagot biztosít a BUSZI (a beszédtempó szociolingvisztikai különbségeinek a vizsgálata például egy érdekes kutatási terület lehet).

A BUSZI célja, hogy rögzítse és tárolja a budapesti beszéd szociolingvisztikai és stilisztika változatosságát. Az interjúk mellett különböző minőségben felvett beszédet és kérdőíveket is tartalmaz a korpusz. A próbafelvételekben (kvótamin-tás megoldással) 10-10 fő vett részt: 50 év feletti tanárok, egyetemisták, szakmunkások, üzletemberek, 15–16 éves tanulók. Ezt a kutatást hívják BUSZI 2-

nek. A BUSZI 3 felvételeire 1988-89-ben került sor. Ekkor 200 felvételt készítettek el az Intézet munkatársai.

A korpuszról és az adatai felhasználásával készült publikációk listája megtekinthető a <http://www.nytud.hu/buszi/ril/socio/osztbib3.htm> weboldalon. A korpusz protokolljáról, a kódolásról lásd Reményi (1989).

#### 5.4. TALKBANK

A TALKBANK egy világméretű összefogás; a kezdeményezés célja, hogy a lehető legtöbb ingyenesen hozzáférhető adatbázist egy keretrendszerben tegyen elérhetővé. Az adatbankból a következő adatbázisok érhetők el:

- CHILDES – az egyik legrégebbi és legnagyobb nyelvfüggetlen gyermeknyelvi adatbázis.

- BilingBank – bilingvális adatközlők hanganyagai.

- SLABank – második nyelvet tanulók anyagai (Second Language Acquisition).

- CABank – konverzációelemzés számára előkészített adatokat tartalmaz.

- PHONBank – a fonológiai fejlődés vizsgálatához létrehozott nyilvános adatbázis.

- Danish SamtaleBank – dán beszélt nyelvi adatbázis.

Ezek az adatbázisokon kívül több, úgynevezett speciális korpusz is megtalálható. Van például afáziás, demenciás betegek hanganyagait tartalmazó rész is, de találunk TBI (Traumatic Brain Injury – 'traumás agysérülés') és különböző audiológiai problémával rendelkező emberektől származó hanganyagokat is.

Az adatbázis anyagai a Creative Commons Licence hatálya alá tartoznak, így szabadon terjeszthetők és adaptálhatók. Az adatok felhasználása során hivatkozni kell az adatok forrására, meg kell adni egy Creative Commons Licence-re irányuló linket, valamint jelölni kell azt is, ha az adatokban valamilyen változtatást (például összevonást) alkalmaztunk. A CC licenc alá tartozó adatok ingyenesen használhatók bármilyen nem üzleti célból. További fontos szabály, hogy az adatok átrendezésével, átalakításával kapott adatok is a CC licenc alá fognak tartozni.

A TALKBANK honlapja: <http://talkbank.org/>; koordinátora: Brian MacWhinney (Carnegie Mellon University). A honlapon új adatbázis hozzáadásához is van lehetőségünk. A szerveren megtalálhatóak az adatbázisokhoz kapcsolódó legfontosabb információk és egy videó- és hanganyag-kezelési útmutató is. A CHILDES projekthez kapcsolódóan megtalálhatjuk az adat- és metaadatkezelő rendszerét is, a CLAN-t. Ennek a leírása és az útmutatója letölthető innen: <http://childes.talkbank.org/clan> és innen: <http://childes.talkbank.org/manuals/clan.pdf> (MacWhinney 2000).

A TALKBANK honlapjáról letölthetjük a beszédkorpuszokhoz tartozó használati és kezelési útmutatókat is. Ezek formátuma .pdf, illetve .doc és .docx. A kezdeményezés fenntartásához, valamint az adatok elemzéséhez számos olyan programot felhasználnak, amelyet a nyelvészek gyakran alkalmaznak, például: PRAAT, CLAN, ELAN, Phon, PhonTalk, Chatter, EXMARaLDA. A programok listája elérhető: <http://talkbank.org/software/>. A fenti oldalon megtalálhatók azok az adatelemző szoftverek is, amelyeket kifejezetten nagy mennyiségű adatok feldolgozására fejlesztettek ki. Ezek a programok elég drágák (általában több száz dollár), azonban sok esetben ingyenesen kipróbálhatók 1–4 hétig.

### Adatbázisok internetes elérhetőségei

<http://www.nytud.hu/buszi/ril/socio/osztbib3.htm>  
<http://talkbank.org/>  
<http://childes.talkbank.org/manuals/clan.pdf>  
<http://childes.talkbank.org/clan>  
<http://talkbank.org/software>  
[http://speechlab.tmit.bme.hu/beszinfo/beszedadatbazis\\_info.pdf](http://speechlab.tmit.bme.hu/beszinfo/beszedadatbazis_info.pdf)  
[http://fonetika.nytud.hu/letoltesek/bk13/Olaszy\\_Beszedkutatatas2013\\_261-270.pdf](http://fonetika.nytud.hu/letoltesek/bk13/Olaszy_Beszedkutatatas2013_261-270.pdf)  
<http://childes.psy.cmu.edu/>  
[http://media.dwds.de/clarin/userguide/text/metadata\\_CMDI.xhtml](http://media.dwds.de/clarin/userguide/text/metadata_CMDI.xhtml)  
(Az url-eket 2015. július 1-jén ellenőriztük.)

### Irodalom

- Bóna Judit 2010. Bizonytalansági megakadások idősek és fiatalok spontán beszédében. *Beszédkutatás 2010*. 125–138.
- Bóna Judit – Imre Angéla – Markó Alexandra – Váradi Viola – Gósy Mária 2014. GABI – Gyermeknyelvi BeszédAdatbázis és Információtár. *Beszédkutatás 2014*. 246–251.
- Gósy Mária 2013. A beszédadatbázis jelentősége a beszédkutatásban. *Anyanyelv-pedagógia 2013/4*. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=489>
- Gósy Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutatás 2008*. 194–207.
- Gósy Mária szerk. 2012. *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest.
- Gyarmathy Dorottya – Neuberger Tilda 2011. A BEA adatbázis alkalmazásfüggő lejegyzései. *Beszédkutatás 2011*. 109–120.
- Imre Angéla 2011. Mit árulnak el a gyermeknyelvi felvételek? Előadás a *Beszédkutatás 2011* konferencián. Budapest, 2011. 10. 27–28.
- Johnston, Trevor 2015. *Auslan Corpus Annotation Guidelines*. 2015 május. [https://www.academia.edu/12360442/Auslan\\_Corpus\\_Annotation\\_Guidelines\\_May\\_2015\\_version](https://www.academia.edu/12360442/Auslan_Corpus_Annotation_Guidelines_May_2015_version)
- Kontra Miklós – Váradi Tamás é. n. *The Budapest Sociolinguistic Interview: Version 3*. <http://www.nytud.hu/buszi/wp2/index.html>

- Neuberger Tilda 2009. A spontán beszéd lejegyzése – a BEA adatbázis tapasztalatai alapján. *Beszédkutatás 2009*. 182–195.
- MacWhinney, Brian 2000. *The CHILDES Project: Tools for Analyzing Talk*. 3rd Edition. Lawrence Erlbaum Associates, Mahwah, NJ.
- Markó Alexandra – Bóna Judit 2006. A spontán beszéd lejegyzésének néhány módszertani kérdése. *Beszédkutatás 2006*. 124–133.
- Reményi Andrea Ágnes 1989. A Budapesti Szociolingvisztikai Interjú kódolási rendszeréről. *Hungarológiai Közlemények* 21/3/80. 405–416. [http://epa.oszk.hu/02400/02401/00036/pdf/EPA02401\\_Hungarologiai\\_kozlemenyek\\_1989\\_80\\_405-416.pdf](http://epa.oszk.hu/02400/02401/00036/pdf/EPA02401_Hungarologiai_kozlemenyek_1989_80_405-416.pdf)

### **Kérdések, feladatok**

1. A dadogást szeretné vizsgálni: milyen adatközlői csoporttal végezné el a kutatást, és milyen feladatokat adna a számukra?
2. Milyen metaadatokat venne fel egy óvodáskorú gyermekek hangját vizsgáló beszédatadabázishoz?

## A fonetika története röviden

A fonetika a beszéd sajátosságaival foglalkozó tudomány, mely az elmúlt évszázadokban – más tudományokhoz hasonlóan – nagy fejlődésen ment keresztül. Néhány évszázada a beszédhangok rendszerezése és jellemzése még a grammatikához tartozott, az 1890-es években útnak induló eszközfonetika előtt nem voltak külön szakemberek a beszéd vizsgálatára. A mai értelemben vett fonetika pedig csak a 20. század második felében alakult ki, a produkció mellett ekkor vált a beszéd észlelésének kutatása is a fonetika szerves részévé.

### 1. Általános fonetika

A beszéd hangjai iránti érdeklődés az írás kialakulásának abban a fázisában lép fel először, amikor az emberek a fogalomjelölő írás helyett a hangalakjelölő írás valamelyik válfajára tértek át. Ehhez a változtatáshoz ugyanis legalább kezdetleges ismeretekre volt szükség a nyelvi tények hangalakjáról. Egyes írásoknál, például az óegyiptominál, világosan lehet látni, hogy hogyan alakult át fokról fokra az írásrendszerük: hogyan lett fogalomjelölésből hangalakjelölés, miként lett az eredetileg tagolatlan hangalakjelölésből szótagcsoportokat, szótagokat, végül pedig hangokat jelölő írás.

Bár a hangokról való ismereteik nem olyan szerteágazók, a hangírás teljes alakjában a görögöknél fejlődik ki először, akik a mássalhangzók jelölése mellett már a magánhangzókat is jelölik, illetve osztályozni is megkísérlik őket a képzési sajátosságaik alapján. Ismereteiket a rómaiak öröklik meg, akik az egyes képzési mozzanatokra vonatkozó megállapításokat a latin nyelv megfelelő hangjaira alkalmazzák, ezen kívül az egyes hangcsoportok neveit is lefordítják (vocalis, consonans, liquida stb.).

Hagyatékek a középkoron át kevésbé gyarapodik, bár a XVI. században már akadnak olyan szerzők, akik a saját anyanyelvük vagy a latin nyelv írási és ejtési kérdéseivel kapcsolatban megfigyeléseket tesznek. Közülük kiemelkedik az aarhusi Jacobus M. Madsen, aki 1586-ban megjelent munkájában<sup>1</sup> (*De literis libri duo*) a hangokat a kilenc képzőszerv működése szerint tárgyalva természetes rendszerbe szervezi. A dán tudós sok helyes megállapítást tesz a hangok képzé-

---

1. Az ebben a fejezetben olvasható, a 20. századot megelőző hivatkozások többségénél a dátum a tájékoztatást szolgálja, a művek azonban nem kerültek be az irodalomjegyzékbe. Valamennyit lásd Laziczius (1944), illetve Vértes O. (1980) munkájában.

sével kapcsolatban: hangsúlyozza például, hogy a magánhangzók képzésénél a szájüreg alakja fontosabb, mint a zöngé, illetve nyelvierekre és ajakképzésüekre osztja őket a nyelv helyzetének figyelembe vételével együtt is.

A XVII. és XVIII. században különböző indíttatásból ismét nagyobb figyelmet kapnak a beszédhangok. Az egyik legfontosabb indíttatás a siketek oktatásának segítése a beszédhangok képzésének és a zöngéképzésnek egyre jobb megismerésén keresztül, mások egy megalkotandó világnyelv szempontjából kapcsolódnak be a hangok kutatásába, emellett az énekkutatás fejlesztésének lehetősége is inspirálólag hat a kutatókra. A szomszédos tudományterületek művelői fizikusok és fiziológusok is jelentős adalékokkal járulnak hozzá a fonetika fejlődéséhez.

A XVIII. század végén ennek köszönhetően két fontos fonetikai munka is napvilágot lát: Christoph Friedrich Hellweg értekezése (*Dissertatio inauguralis de formatione loquale*, 1781), amelyben először jelenik meg a német magánhangzók elrendezésének jellegzetes háromszögű sémája; a másik a magyar Kemelen Farkas németül megírt könyve (*Mechanismus der menschlichen Sprache*, 1791), amely a légzés folyamatára és a beszélőszervek tevékenységére vonatkozóan tartalmaz rendkívül értékes megfigyeléseket.

A század végén még egy nevezetes esemény gyarapította a fonetikai ismereteket: a kutatók „felfedezték” a szankszrit nyelvet, amelynek gazdag és az európai nyelvekétől eltérő hangrendszerrel olyan addig ismeretlen képzési mozzanatokra is ráirányította a figyelmet, mint például a retroflex ejtés (a hangképzés közben a nyelvhegy visszahajlik).

A fonetikai kutatás első igazi fellendülése a XIX. századra tehető, ekkor már a fiziológusok és az akusztikusok részéről is nagy érdeklődés mutatkozott a beszéd iránt. Czermak Nepomuk János a hangszalagok működéséről szóló írásában először alkalmazza Garcia gégetükrét fiziológiai szempontú vizsgálatokra (*Der Kehlkopfspiegel und seine Verwendung für Physiologie und Medizin*, 1854).

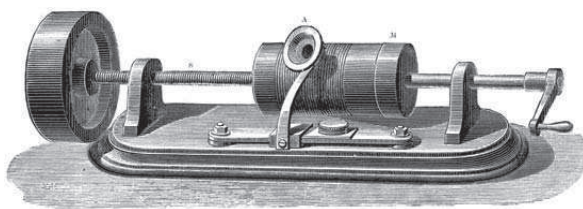
A század második felében a fonetikai kutatások fellendülnek: vizsgálni kezdik az ismertebb nyelvek (angol, dán stb.) hangjait, illetve az ismeretek szélesebb körben való terjedése is megindul. Több olyan folyóirat is létezik, amely hangjelölési, ejtési és más fonetikai kérdésekkel foglalkozik. 1850-től kezdve *Fonetic* (1852–: *Phonetic*) *Journal* néven működik tovább a korábban *Phonographic*, illetve *Phonotypic Journal* címen megjelenő folyóirat. 1889-ben a francia Paul Passy megalakítja a Nemzetközi Fonetikai Társaságot (International Phonetic Association, IPA), melynek hivatalos lapja, a *Le Maître Phonétique* azóta is kizárólag fonetikai tartalommal jelenik meg.

A század végéhez közel még egy tudománytörténetileg fontos eseménysor megy végbe: útjára indul a kísérleti vagy más elnevezéssel az eszközfonetika. Megteremtője Jean Pierre Rousselot abbé, aki az eszköztechnikát először avatja módszerré a fonetikában. Mindez az 1897–1901 között kiadott művében, a

*Principes de phonétique expérimentale*-ban jelenik meg elsőként a maga teljességében.

Az eszközfonetikusok alap gondolata az volt, hogy pusztán az emberi fül nem megbízható sok esetben, így ha biztos ismereteket akarunk szerezni, akkor azt ne szubjektív megfigyelések vagy becslések, hanem gépi regisztrálás és mérés útján tegyük. Kezdetben ez a kutatói társadalom nagy részét meggyőzte, az eszköztan rohamos fejlődése miatt azonban a vizsgálati anyagról egyre inkább a vizsgálati eszközre helyeződött a figyelem. Az egyre több műszaki és fizikai tudás szükségessége miatt egy új fonetikustípus is megszületett: a sokkal inkább technikus vagy fizikus, mint nyelvész típusa.

Fontos megemlíteni, hogy az eszközfonetika működéséhez olyan eredményekre is szükség volt, mint Thomas Alva Edison fonográfja (1. ábra), amely a hangrögzítés problémáját oldotta meg, vagy mint Alexander Graham Bell fonautográfja, amely a beszéd vizuális megjelenítésére volt képes.



1. ábra. Edison első fonográfja 1877-ből

(<http://www.britannica.com/biography/Thomas-Edison/images-videos/First-model-of-Thomas-Alva-Edisons-phonograph-1877/161874>)

Ebben az időben indult meg az Eberhard és Kurt Zwirner nevéhez fűződő fonometria is, amely nem más, mint nyelvészeti rendeltetésű eszközfonetika. Lényege, hogy bár az eszközfonetikushoz hasonló regisztráló és mérő műszerekkel rendelkezik a fonometrikus is, ő nemcsak bánni tud velük, de ismeri is őket műszaki szempontból. Egy ilyen szakember csak egy nyelvessel együtt dolgozva szól hozzá bármilyen beszédhanggal kapcsolatos nyelvi kérdéshez, mert tudja, hogy a hangok nemcsak fizikai jelenségek, hanem nyelvi tények is. Zwirnerék szerint a nyelvész az, aki a kérdést felteszi, és a fonometrikus csak azt mondja meg, hogy meg lehet-e oldani a problémát a rendelkezésre álló technikai eszközökkel, és ha igen, hogyan.

A fonetika XX. századi rohamos léptékű fejlődése azt is szükségessé tette, hogy a tudomány művelői időnként személyesen is találkozzanak, és megvitasák a tudomány szak új eredményeit. Ennek a törekvésnek 1932 óta a Nemzetközi Fonetikai Társaság a fő támogatója, amely időről időre világkongresszusokat (International Congress of Phonetic Sciences, ICPhS) rendez. Az elsőt Amszter-

damban tartották 1932-ben, a második, 1935-ös kongresszus Londonban volt, újabb három évvel később pedig Genf városában került megrendezésre. Az elmúlt évtizedekben ez az esemény vált a fonetikusok legnagyobb és legelismertebb találkozási pontjává, az itt elhangzó előadások szakmai színvonala rendkívül magas, a résztvevők pedig a világ minden tájáról érkeznek. A legutóbbi, 18. Fonetikai Világkongresszus helyszíne Glasgow volt 2015-ben.

## 2. A magyar fonetika története

A magyar fonetika történetében sokszor csak egymásutániság állapítható meg szerves összefüggések helyett. Vértes O. András hatalmas kutatómunka eredményeként 1980-as művében a tudomány szak fejlődését négy időszakra osztotta. E négy korszakot az ő munkássága alapján tekinti át a tanulmány, a Balassa utáni időszakot mi ötödikként jelölve, a modern magyar fonetika koraként tárgyaljuk.

### 2.1. Az első korszak: a kezdetektől az első magyar grammatikáig

A szórványos hangtani megjegyzések korának legrégebb tudományos igényű munkái a XV. századból valók. Galeotto Marzio már feljegyzi, hogy a magyar parasztok és nemesek kiejtése között nincs különbség, valamint tőle származik a magyar magánhangzókra vonatkozó első ismert megfigyelés is. A kor másik fontos műve Temesvári Pelbárt *Rosarium*-a (1503–1508), amelyben Temesvári a nyelv anatómiájának tárgyalása után a feladatait is sorra veszi, és a beszédhibákat is tárgyalja, különbséget téve anatómiai és funkcionális eredetűek között.

### 2.2. A második korszak: fonetikai megállapítások a Bél Mátyás előtti szakirodalomban

A hang első magyar meghatározását az első magyar nyelvtanban találjuk, amelynek szerzője Sylvester János. A *Grammatica Hungarolatina* 1539-ből származik, és a hang definiálásán kívül a magyar nyelv hangjainak rendszerezését is elvégzi.

A beszédhangokra vonatkozó fonetikai megjegyzések ekkor négy típusba sorolhatók: a metaforikus leírások (pl. Dévai Bíró Mátyásnál), az idegen nyelvek kiejtéséhez való hasonlítások (pl. Sylvesternél vagy Szenczi Molnár Albertnél), egy-egy magyar hangnak más magyar hangokhoz való hasonlítása (pl. Komáromi Csipkés Györgynél) és a képzésmód leírásai (pl. Misztótfalusi Kis Miklósnál). Emellett vannak kísérletek a hangkapcsolatok leírására és a magyar kiejtés általános jellemzésre is. A korszak Tsétsi János 1708-ban kiadott szabályaival végződik.

### 2.3. A harmadik korszak: Bél Mátyástól az első magyar nyelvű fonetikáig

A XVIII. században a hazai fonetikai érdeklődés több forrásból táplálkozott, a hangok vizsgálata elmélyültebbé vált. Már nem csak grammatikai, helyesírási és retorikai munkák miatt fordultak a hangok felé, hanem verstani, gyógypedagógiai vagy épp élettani kutatásoknak is köszönhetjük a korszak eredményeit.

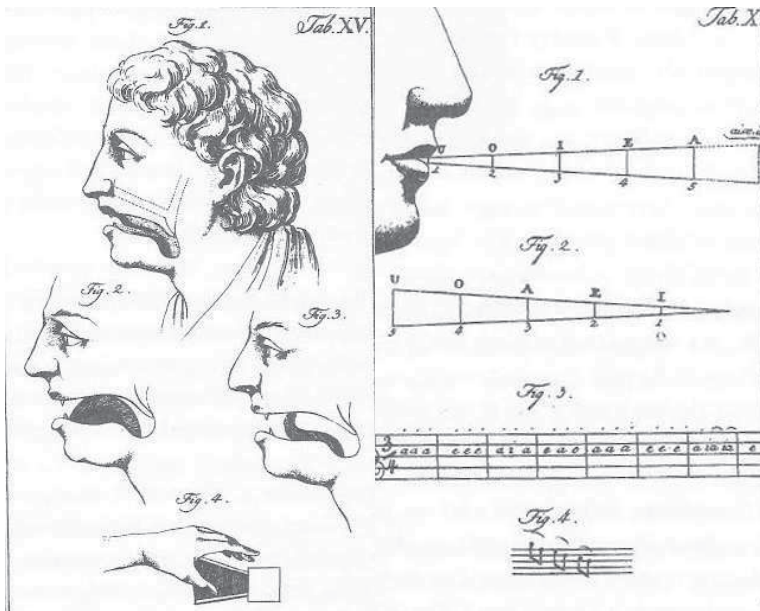
Az első teljes magyar hangtan Bél Mátyás nevéhez kötődik, munkája a székelv rovásírásról szóló nagy tanulmányának részeként maradt fent 1718-ból. A fonetikai rész mindössze 7 lapnyi terjedelmű, mégis jelentős, mert minden magyar hanggal foglalkozik, azokat ábécérendben tárgyalja, harmincnégy pontban. Megállapításainak egy része ma is helyt áll: leírja például, hogy az *atya* szó két *a* hangját különbözően ejtjük, és a rövid *e*-nek is kétféle kvantitását különbözteti meg. Emellett az idegen szavak magyar hangzásával is foglalkozik, illetve más megfigyeléseket is tesz a latin és a magyar kiejtéssel kapcsolatban. Közel 50 évvel később más munkák is foglalkoztak a ma a szociolingvisztikához tartozó problémákkal, például a falusi és a városi ejtés közötti különbségekkel, a műveltek és az iskolázatlanok hangképzésbeli sajátosságaival vagy a férfiak és a nők sajátos kiejtési módjaival.

A kor kiemelkedő alakjának, Kempelen Farkasnak a munkássága is erre az időre esik. Wolfgang von Kempelen 1734-ben Pozsonyban született ír eredetű családban. A modern magyar fonetika előfutára Győr és Bécs mellett szülővárosában járt iskolába. Igazi polihisztorként jogot, filozófiát, matematikát, fizikát és építészetet tanult, de irodalmi alkotásai is jelentősek, és kézügyessége is jó volt, rézkarcokat készített. Több találmány is köthető nevéhez, így a beteg Mária Teréziának készített mozgatható ágy, írógép, a pozsonyi vízrendszer terve és a híres sakkozó automatája is, a sakkozó török. Több nyelven beszélt, a magyar és a latin mellett németül, franciául, olaszul és szlovákul is. Tudott még románul és szerbül, illetve valamennyire angolul is. Nevéhez fűződik az 1788-ban szabadalmaztatott gőzturbina és két kiemelkedő fonetikai munka is.

Ezek közül az egyik könyve a *Mechanismus der menschlichen Sprache*, amely 1791-ben jelent meg. Számos fonetikai tétele a magyar hangokra is érvényes, megállapításai ma is helytállóak. Ebben részletesen írt a beszédhangok képzésének sajátosságairól, a hangátmenetekről (ezzel messze megelőzte korát). Az *A*-val jelölt hang képzésénél leírta például, hogy a hangszalagok rezegnek, az orrüreg felé vezető út el van zárva, a nyelv fekvő helyzetben van, a nyelvcsatorna nyílása harmadfokú, a fogaknak nincs szerepük a képzésben, és hogy az ajkak ötödfokú nyílásban vannak. A hangok artikulációját és az artikuláció fiziológiai hátterét saját maga rajzolta ábrákkal szemléltette is (2. ábra).

Kempelen azért tanulmányozta a beszédet ilyen részletesen, mert célja volt annak mesterséges előállítását. A *Mechanismus*-ban írt a beszélőgépről is (3. ábra), amellyel a világon elsőként szavakat és rövid mondatokat tudott kimondatni,

mindezt több nyelven is. E találmányán 22 évig dolgozott, könyvében a munka több fázisát is leírta, így ezek alapján gépe nemcsak továbbfejleszthető, de rekonstruálható is lett.



2. ábra. Az *l* hang képzése és a magánhangzók zeneiségének szemléltetése Kempelen ábráin (Gósy 2004: 320)



3. ábra. Kempelen beszélőgépe a Deutsches Museumban (<http://www.deutsches-museum.de/ausstellungen/musikinstrumente/objekte/sprechmaschine-1/sprechmaschine-2/sprechmaschine-3/>)

A feljegyzéseknek köszönhetően az MTA Nyelvtudományi Intézetének Kempelen Farkas Beszédkutató Laboratóriumában is elkészült a gép hiteles, működő másolata 2002-ben (Nikléczy–Olaszy 2002). A rekonstrukció célja az volt, hogy működő szerkezetként előállíthatók legyenek vele hangok és hangsorok, így ki lehessen próbálni azokat a hang-előállítási formákat, amelyeket Kempelen leírt. További célkitűzés volt a gép által előállított hangok akusztikus összetevőinek a mai korszerű jelfeldolgozási eljárásokkal való megvizsgálása. Az eredmények igazolták, hogy a gép működött, így Kempelen kétszáz évvel ezelőtt tulajdonképpen egy artikulációs, nyelvfüggetlen beszédszintetizátort alkotott. Gépe és vizsgálatai azért is olyan jelentősek, mert elvégezte a beszédképzésről való addigi ismeretek rendszerezését, és lehetővé tette addig nem ismert tények felfedezését is (Gósy 2004).

A hazai hangtan Kempelen évszázadában került a korabeli irodalmi és tudományos érdeklődés középpontjába. A magyar irodalmi élet kialakulásával két hosszan fennálló irodalmi vita is kirobbant, ezeknek hangtani vonatkozásai tudományszakunkhoz tartoznak. Az egyik a prozódiai harc volt, amely a *h* más-salhangzó létjogosultságát kérdőjelezte meg. Baróti Szabó Dávid már 1777-ben leírta, hogy a *h*-t egyesek mássalhangzónak, mások magánhangzónak tartják, ő köztes felfogást képviselve azt állította, hogy néha mássalhangzó, máskor csak lehellés. Virág Benedek egy fonológiai érveléssel bizonyította, hogy a *h* mássalhangzó, amikor az *ara* és *hara*, *ora* és *hora*, *hegy* és *egy*, *hála* és *ála* stb. szavakat minimálpárokba állította.

A másik polémia a diftongusok köré csoportosult, Rájnis József és Baróti Szabó Dávid között zajlott. A vitát ekkor még Révai Miklós azzal zárta le, hogy a helyes magyar kiejtésben nincsenek diftongusok; a kérdés majd a romantika erősödő népnyelvi kutatásaikor kerül újra előtérbe. Viták voltak még a zöngés affrikáták egyeshang vagy kettőshang voltáról is. A jottista-ipszilonista háború is rendelkezett hangtani vetülettel, és az ebben a perben szereplő hangtani kérdéseknek köszönhetően irányult a figyelem a hangkapcsolatok ejtésére.

Ennek a korszaknak köszönhetjük Révai Miklós történeti nyelvtanát, amely *Elaboratior grammatica Hungarica* címmel jelent meg 1806-ban. E mű már leírja valamennyi hang képzését, így voltaképpen magyar leíró hangtanként olvasható, ugyanakkor a szupraszegmentális elemek leírásában is úttörő volt, a hangsúly és a hanglejtés mellett a tempóról is írt.

Az első magyar nyelvű fonetika, Simon Antal *Igaz mester-e* (1808) a pedagógia tudományának köszönhető, a szerző ugyanis a váci „süketnéma”-intézet első igazgatója volt. Simon Antalt leendő tanár kollégáival együtt Bécsbe küldték, hogy a szükséges szakismereteket elsajátítsa, ezután írta meg könyvét, amelynek tárgya a „betűk” helyes kimondásának tanítása volt. Ezzel magyarázható, hogy a mássalhangzók képzését többször is tárgyalja, először önmagukban,

utána hangkapcsolatban, végül a tanítás rendjébe helyezve. A magánhangzóknál a hosszú és rövid hangok között minőségbeli különbséget is megállapított.

#### 2.4. A negyedik korszak: útban a magyar fonetika felé

A XIX. század elejétől a 80-as évekig számos tényező együttese készíti elő a magyar fonetika kialakulását. Ekkor már számos önálló cikk foglalkozik egy-egy hanggal vagy hangcsoporttal, a tudomány iránt érdeklődők a folyóiratokban, napilapok hasábjain is olvashatnak hangtani kérdésekről. A tudományos élet megindulásával már egyre többen vesznek részt hangtani kérdésekkel kapcsolatos vitákban. A kor fontos jellemzője az is, hogy a tudomány nyelve ekkor már a magyar.

Megindul a köznyelv és a nyelvjárások hangtani leírása, itt Czuczor Gergely, Fogarasi János és Budenz József munkája említendő, valamint Bugát Pálé, aki először osztályozta következetesen a magyar mássalhangzókat a képzés helye szerint.

A szupraszegmentális tényezők vizsgálata is felélénkül. E korig viszonylag kevés szó esett róluk az értekezésekben, mert korábban írásban nem jelölték őket; a romantika korában azonban fontossá vált a beszéd érzelmeinek írásban való jelölése is.

A beszédtempó kutatása is megindul, Beregszászi Nagy Pál összeveti a magyar beszédtempót a franciával, Hunfalvy Pál pedig a finn, az észti és a magyar tempó összehasonlító vizsgálatát végzi el. Ezen túl a magyar kiejtés általános jellemzése is előtérbe kerül a tiszta artikuláció és az érthetőség szempontjából. Az első magyar beszédjavító értekezés Szilágyi János munkája, aki 1835-ben írt *A' Selypségről*. Szilágyi e terminuson a dyslaliát (pöszeség, az anyanyelv beszédhangjainak helytelen képzéséből származó artikulációs zavar) és/vagy paralaliát (olyan beszédhibák, amely során egy hang ejtését a gyermek egy másik hanggal helyettesíti) érti, az *r* hang hibáinak például több fajtáját is leírja. A szigmatizmussal (a sziszegő hangok képzését érintő pöszeség) is részletesen foglalkozik.

A fonetikus átírás igénye is jelentkezik, 1813-ban Kyss Sándor a külföldi kortársakat megelőzve már az átírás kérdésével foglalkozik. Rendszere betűkből, számokból és egyéb jelekből áll össze, már 1818-ban használ „organic” jeleket (olyan jelek, amelyek a képző orgánusok működését is jelölik). Munkája ugyan nem terjedt el széles körben, de hatott kortársaira, akik többek között az egyetemes hangírás megalkotásával foglalkoztak.

1862-ben jelenik meg Regner Tivadar *A magyar nyelv kiejtése* című munkája is, amely a magyar fonetika legjelentősebb alkotásaink egyike. Értekezése nagyrészt kontrasztív hangtan, a magán- és mássalhangzókat több szempontból, többször is csoportosítja. Közel jár a hangnyújtás fogalmához, két csoportra bontva

tárgyalja a kettőshangzókat is (ti. tökéletes és tökéletlen kettőshangzók). Értékesek a laterálisra vonatkozó megjegyzései is, az elöl és a hátul képzett *l*-ről is szól, emellett a *h* hang leírásához is hozzájárul. Részletesen értekezik a hang magasságáról, a hanglejtésről és a hangsúlyról is. Ír még az énekesek sajátos kiejtéséről, a magyar nyelv ritmusáról, a nemzeti hangsajátságáról is.

Az 1800-as évek második felében elindul a beszéd eszközfonetikai elemzése is. Az első kísérleti fonetikai adatok a beszédhangok képzésének műszájpaddal végzett artikulációs vizsgálatából származnak Balassa Józseftől (1877), aki a módszert *stomatoskopiá*-nak nevezi. 1904-ben kiadott *Magyar fonétika*-jával zárul a korszak, mellyel a mássalhangzók osztályozásával az évtizedekig tartó vitába kapcsolódik be, valamint itt közöl először rajzokat a nyelv és a szájpaddalás helyzetéről a magánhangzók ejtésekor. Balassa ezen kívül a Keleti Kereskedelmi Akadémiához tartozó Fonetikai Laboratóriumban az elsők egyikeként folytatott eszközfonetikai vizsgálatokat. Különböző finnugor nyelveken beszélők hanganyagain mérte például a zöngé jelenlétét, a levegő kitódulási erejét vagy a hangképzési energiát.

## 2.5. Az ötödik korszak: a modern magyar fonetika kora

A kísérleti fonetika és az azóta kifejlesztett technikai lehetőségek új utakat nyitottak meg a nyelvészek és más tudományterületek művelői számára, így a fonetika a XX. században óriási lépésekben haladt és halad ma is előre. A kor jelentősebb eredményeit Gósy Mária (2004) írta le, az ő áttekintése alapján követjük végig a század egy-egy jelentősebb eseményét.

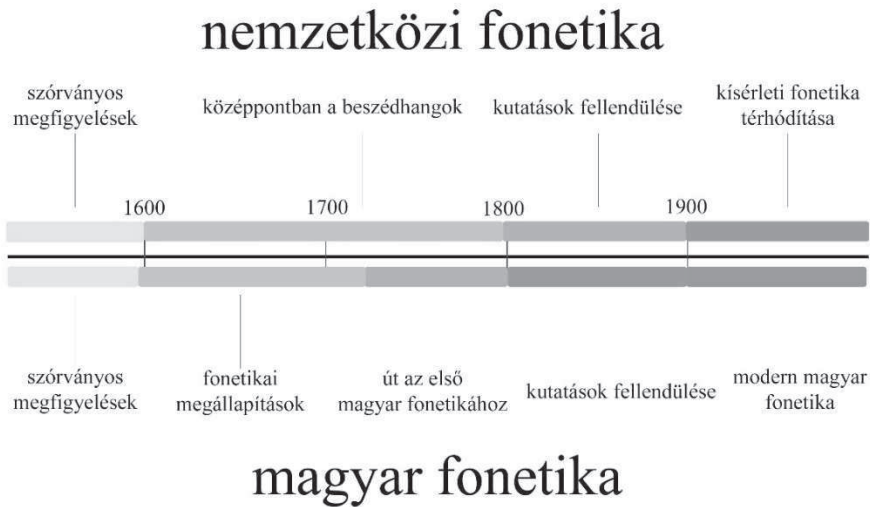
A beszéd eszközfonetikai elemzését még 1887-ben Balassa József kezdte el, a palatogramokkal történő ajakartikulációs vizsgálatokat Hegedűs Lajos (1941–43) és Bakó Elemér folytatta (1941–43), és Csúry Bálint szélesítette ki a nyelvjárási kísérleti-fonetikai vizsgálatokra (1936).

A világháború kitörése előtt már oscillogramok is készültek a magyar magánhangzokról, az elektroakusztikai elemzés 1941-ben indult meg.

Az első „mozgófénykép-felvételek” már a 30-as évek elején jelen voltak, az első fonetikai célú röntgen-hangosfilm (beszéd közben készített röntgenfilm) pedig 1966-ban készült.

A beszéd akusztikai szerkezetének korai kutatása az 1940-es évektől Tarnóczy Tamás, Magdics Klára és Fónagy Iván nehéz kötődik, az azóta eltelt 75 évben pedig számtalan nyelvész, fizikus és mérnök kapcsolódott be a beszéd akusztikai vizsgálatába, de még így is számos kérdés vár megválaszolásra.

A nemzetközi és a magyar fonetika rövid történetét a 4. ábrán látható idővonalon szemléltetjük.



4. ábra. A nemzetközi és a magyar fonetika rövid története

### 3. A fonetikus átírások története

Alexander Graham Bell fonautográfjának megjelenésével egy időben vetődött fel egy az addiginál pontosabb és következetesebb hangjelölési mód szükségessége is, a tudósok így a fonetikai célokra is alkalmas írás kifejlesztésén kezdtek dolgozni. A fonetikus írás megalkotásának kezdeti kísérletei alapvetően két irányban haladtak. Az egyik irány (Erns Wilhelm Brücke kezdeményezésével) elvetette a gyakorlati írások megszokott betűit, és teljesen új jelek segítségével igyekezett az írást megalkotni. A másik irány (Karl Richard Lepsius képviselésével) megtartotta a latin betűket, a tudományos célokhoz kevés elemet pedig a betűalakok módosításával, más rendszerek jeleinek átvételével és segédjelek alkalmazásával duzzasztotta fel.

Brücke és követői egy írásforradalommal akarták megoldani a problémát arra hivatkozva, hogy a megszokott betűjelek használata könnyen tévedésekhez vezet, illetve ez útjában áll a rendszer egyetemességének is (ti. egy átírási rendszernek alkalmasnak kell lennie a legkülönbözőbb nyelvek leírására is). Brücke ennek fényében alkotta meg saját rendszerét, amelyben a betűk képzési receptnek is megfeleltek: pontos előírást adtak arra vonatkozóan, hogy miképpen is kell az adott hangokat képeznünk, a mássalhangzóknál például megadta a képzés helyét, módját, és a zöngességet is jelölte. Bár a kezdeményezés nagy érdeklődést keltett, mégsem terjedt el a tudományos használatban.

Az angolok kísérletei ezzel szemben sikeresebbnek bizonyultak. Bell 1867-es *Visible Speech* című munkájában egy Brücke nyomán készült írást mutatott be, bár betűi jóval egyszerűbbek voltak. A karakterek kevés részletvonalból épültek fel, de így is jól tudták szemléltetni a hangok képzési sajátosságait is, így lehetővé téve a reprodukciójukat. Ez azért fontos, mert a *Visible Speech* eredeti célja a „süketnémák” beszédre való tanításának segítése volt. A mássalhangzóknál egyenes vonal jelölte például a képzőszervek között létrejött zárat, hajlított vonal pedig azt a képzőszervet, amely az adott hang képzésében leginkább részt vesz. A magánhangzóknál függőleges vonal jelölte a zöngét, ehhez pedig pont vagy kampó csatlakozott aszerint, hogy szűk vagy széles ejtésű volt-e a hang, illetve a csatlakozás iránya a hang palatális vagy veláris voltát is megadta. A rendszert Henry Sweet fejlesztette tovább, aki szerint az általa „romic”-nak nevezett latin betűs írás helyett az „organic” írást kell használni, vagyis azt, amelyik a képző orgánusok működését is jelöli a karaktereivel. Az angol tudós közreadta Bell módosított rendszerét, így a Bell–Sweet „organic” rendszerű írás volt az egyetlen, amelyet ily módon szélesebb körben is megismertek, a használata azonban ennek sem volt tartós.

A reformok útján haladó Karl Richard Lepsius Brückével egy időben kezdte meg munkáját. 1855-ös művében a latin betűs írás jeleire épített rendszerét publikálta, amelyet görögből átvett betűjelekkel és mellékjelek használatával egészített ki. Rendszere túl sok diakritikus jelet (a betű fölé, alá vagy mellé helyezett mellékjelet) tartalmazott, így egyetemesen használt rendszerré nem válhatott, ennek ellenére számos, afrikai nyelvekkel foglalkozó tudós átvette.

10 évvel később vált ismertté John August Lundell hangjelölése is, mely a latin nyelv betűkészlete mellett görög és cirill módosított betűket is használt. Jelentős haladást tett lehetővé azzal, hogy csak olyan betűket használt, amelyek folyamatos vonalvezetéssel írhatók, illetve bevezette a monotyp-elmet, amelynek lényege, hogy egy-egy jel csak egy beszédhangot szemléltet. Lundell munkájának megjelentését követően még több évtizedig zűrzavar uralkodott az átírások tekintetében, amire a megoldást végül az 1925-ös koppenhágai értekezlet adta meg, ahol tizenkét nyelvész vett részt. A kutatók elsőként különbséget tettek az átírás és az átbetűzés között. Átírásra ott van szükség, ahol a latin betűs gyakorlati írás nyelvészeti szempontból nem kielégítő, tehát a jelöléseit át kell írni fonetikailag hűen, az ejtést követve. Átbetűzés viszont ott szükséges, ahol a gyakorlati írás nem latin betűs, vagy nem is betűíró. Az előbbinél azt tisztázták, hogy mi az, amit mindenképp jelölni kell, és azt milyen módon (ti. „organic” vagy „romic” rendszerben, és hogy monotyp-elven vagy sem), illetve felhívták a figyelmet a hang és a fonéma közötti különbségre is. A szakemberek végül a latin betűsor alapulvétele mellett döntöttek, a monotyp-elmet azonban csak a hangok egy részénél vitték véghez, így a mellékjelek kérdését nyitva hagyva.

Erre ma érvényes megoldást a Nemzetközi Fonetikai Társaság (IPA) tudott adni, amely fennállása óta többször is körvonalazta a szövetség átírásának elveit. A társaság írása a nemzeti helyesírások jelölésének fonetikai átírását szolgálja, emellett olyan népek nyelvi megnyilatkozásainak lejegyzését, amelyeknek nincs saját írásuk. Fonetikus írásként azonban tudományos célokra is jól használható, ma ez a jelkészlet az egyik legelterjedtebb. A készlet az International Phonetic Alphabet nevet kapta, rövidítése így a társaság nevének rövidítésével megegyezően IPA. Az IPA a kezdetektől fogva latin betűs, a latin antiqua íráson alapszik, de ez két- és hárompontos betűjelöléssel egészül ki, így lehetővé téve, hogy viszonylag kevés jellel az összes ismert nyelv lejegyezhetővé váljon. A nemzetközi fonetikus átírás a mássalhangzók jeleit táblázatban, a magánhangzókét absztrakt, trapéz alakú síkban jeleníti meg. Az előbbinél a képzés helye, a képzés módja és a zöngéesség megléte vagy hiánya derül ki az oszlopokból és sorokból, illetve a táblázatban jelölt az is, ha egy hang ejtése nem lehetséges vagy rendkívül nehezen képezhető. Ugyanígy feltüntetik azt is, ha egy hang elvben létezhet, de még egyetlen ismert nyelvben sem találtak rá példát, illetve a minimális eltéréssel használatos hangok cellái is jelöltek. A magánhangzótrapéz a nyelv vízszintes és függőleges pozícióit írja le, illetve ajakműködés szerint is csoportosítja a vokálisokat. A hangokhoz kapcsolódó mellékjelek számos funkciót kaptak, így a fonáció különböző aspektusai mellett a nyelv, az ajkak vagy a nyelvcsap pozícióját is mutatják. A jelölésrendszer nagy előnye, hogy a fonetikus megközelítésmódú ábrázolás mellett a fonematikus elv alkalmazását is lehetővé teszi, illetve hogy a szupraszegmentumok jelölését is önálló szimbólumkészlettel oldja meg.

Az IPA azóta a kor igényeihez igazodva kidolgozta az extIPA rendszert is, mely a különböző beszédzavarral küzdő egyének beszédének lejegyzésére is alkalmas. A rendszerben a normál ejtéstől való eltéréseket diakritikus jelek szemléltetik, így jelezhető például a denazalizáció. A jelrendszer a beszéd több jellemzőjének visszaadása mellett lehetővé teszi a beszéd tempójának és a szünet-tartásnak a jelölését is.

#### 4. Kitekintés

Az utóbbi két évtizedben a fonetika óriási fejlődésen ment keresztül. Az infokommunikációs technikák rohamos elterjedése, az egyre javuló minőségű rögzített felvételek, az internet, az annotáló és statisztikai programok sokasága olyan utakat nyitottak és nyitnak meg előtte, amelyekre a kutatók korábban talán gondolni sem mertek. Lehetővé válik az egyre pontosabb és matematikailag is igazolható vizsgálatok elvégzése, a korpuszok pedig segítik az esettanulmányokban kapott eredmények nagy mennyiségű adaton alapuló igazolását.

---

## Irodalom

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

Laziczius Gyula 1944. *Fonetika*. Királyi Magyar Egyetemi Nyomda, Budapest.

Nikléczy Péter – Olasz Gábor 2002. Kempelen Farkas beszélőgépezet rekonstrukciója. *Beszédkutatás 2002*. 5–17.

Vértes O. András 1980. *A magyar leíró hangtan története az újgrammatikusokig*. Akadémiai Kiadó, Budapest.

## Kérdések, feladatok

1. Ki volt Kempelen Farkas? Miért jelentős a munkássága?
2. Miért fontos az eszközfonetika megszületése?
3. Miért van szükség az egységes, nemzetközi fonetikus átírássra?

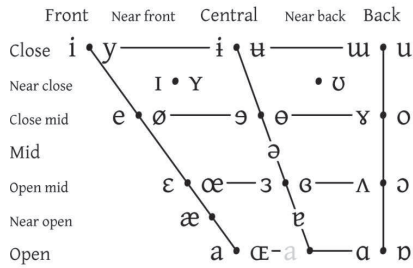


CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

VOWELS



Vowels at right & left of bullets are rounded & unrounded.