

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, nyelőcső-beszédre vagy akár több különböző technikájú éneklésre is (pl. operaéneklés, jóddli, 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 beszédképzés 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 nyelőcsőbeszéd 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 jelrészlet, 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 alaphangfrekvencia-tartományában is tapasztalják, hogy hangjuk hangszínezetet vált, „elcsuklik” akár már kisebb alaphangemeléskor is?

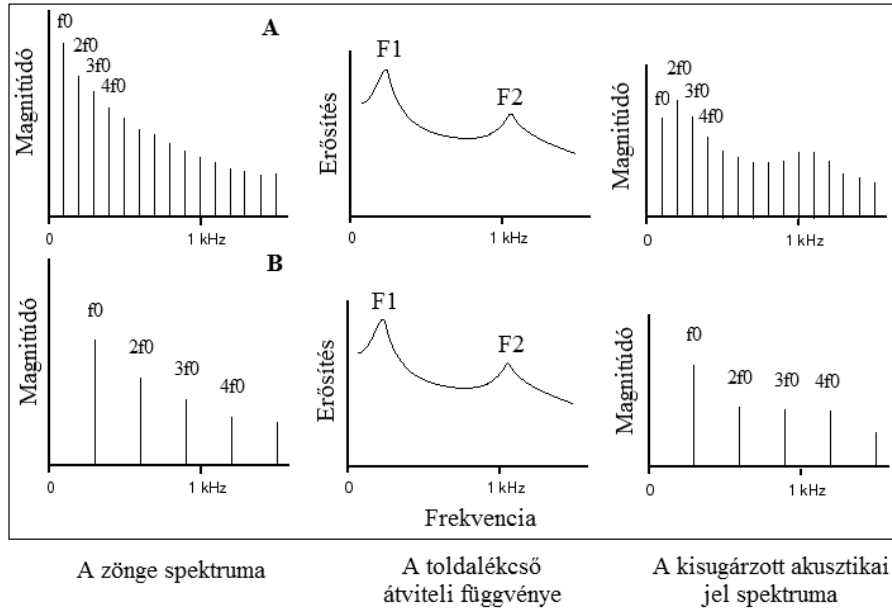
Az énekelt magánhangzókhoz kapcsolatos percepciók vizsgálatainak legfőbb kérdése, hogy miként befolyásolja az alaphangfrekvencia az énekelt magánhangzók azonosítását, ugyanis – mint látni fogjuk – az éneklésben az alaphangfrekvencia emelésével jelentősen megváltozhat a magánhangzók ejtése.

2.1. Az alaphangfrekvencia hatása a magánhangzók akusztikai szerkezetére

A zöngé a természetben előforduló hangok legnagyobb részéhez hasonlóan komplex hang, azaz több összetevőből áll: az alaphangfrekvenciából vagy alaphangból (f_0), valamint annak egész számú többszöröseiből, a felhangokból (egy lehetséges jelölés szerint $2f_0$, $3f_0$ stb.; Vicsi 2010). Az alaphangfrekvencia 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 alaphangfrekvenciája (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 hangfokvásban 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 formánsuk (F_1 és F_2) frekvenciaértéke segítségével (Peterson–Barney 1952). A formánsok értéke



1. ábra. Az alapprofundencia hatása a magánhangzók spektrális szerkezetére. Első oszlop: a gégében 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)

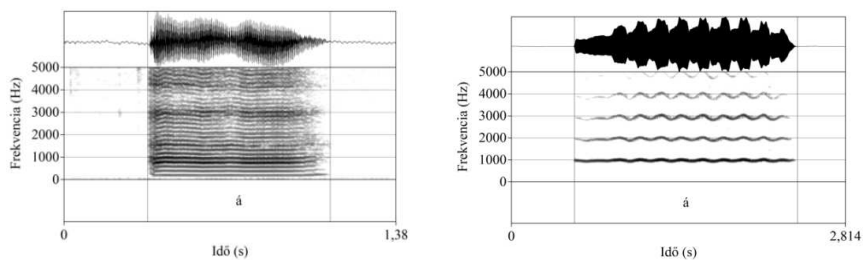
attól függ, hogy az egyes magánhangzók ejtésekor milyen a toldalékcső belső alakja (milyen helyzetben vannak az egyes képzőszervek – ez az artikulációs konfiguráció). A formánsok a toldalékcső ún. sajátrezonanciái (Fant 1960), amelyek az artikulációs konfigurációtól függenek. Az első két formáns (F_1 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 sávzélességébe 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).

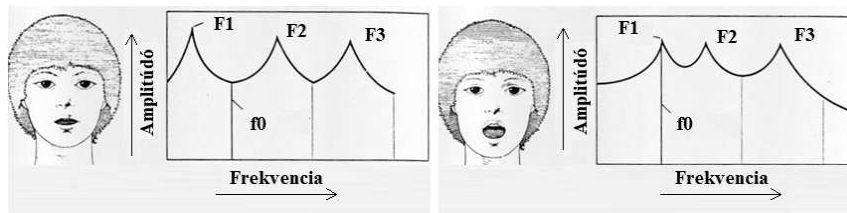
Egy valós beszédmintából származó példát mutat be a 2. ábra. Itt az *á* magánhangzó beszédbeli megvalósítását (bal oldal), valamint a magas alapfrekvenciájú éneklésben rögzített realizációját (jobb oldal) látjuk egy női énekes kitarított ejtésében (állandó alapfrekvencián). A hangszíneképeken (alsó panelek) látható, hogy míg a beszédben vízszintes sötét csíkként jól kivehetők a magánhangzó formánsai, addig az 1 kHz közeli alapfrekvencián már csak az egyes felharmonikusok rajzolódnak ki. Így bár magas alapfrekvencián a felharmonikusok sötétedésük mértéke alapján sejtetni engedik, hogy beleesnek-e valamely formáns sávzélességébe, a formánsok sávzélessége 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-

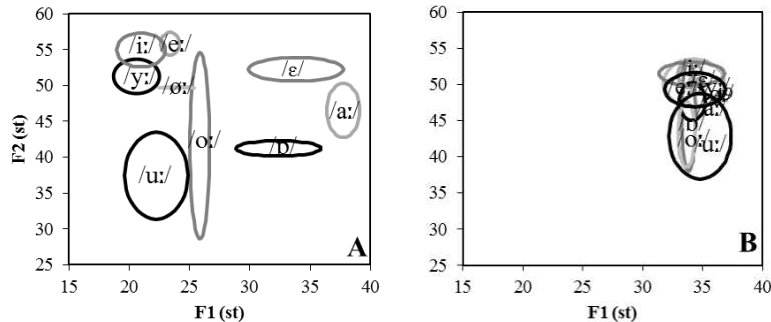
sát módosítják: növelik az állkapocsnyitás szögét, illetve csökkentik a nyelvállás fokát (vö. Sundberg 1987). Ebből az következik, hogy a két cél elérése ellentmondásba kerül: a homogén hangszínű, csengő hangú éneklés és az észlelés számára azonosítható minőségű magánhangzóejtés a zártabb magánhangzók esetében. Ez a konfliktus kivétel nélkül az éneklés javára dől el; a magas alapfrekvenciájú éneklésben (az $f_0 > F_1$ helyzetekben) ugyanis az énekes (az alapfrekvencia 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 alapfrekvencia emelésével egyre növelték az állkapocs nyitásának szögét már közel 300-400 Hz alapfrekvenciától 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 alapfrekvenciákon 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 alapfrekvencia 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 alapfrekvencia 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 alapfrekvencia emelésével fokozatosan lezsugorodik úgy, hogy az *á* pozíciója felé mozdul, mígnem – megközelítőleg 1 kHz alapfrekvencián – 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 alapfrekvencia emelésével egyre jobban uniformizálódó (és egyre ritkább spektrumú) magánhangzók az ejtési szándék szerint.



4. ábra. Az F_1 és F_2 formánsokkal meghatározható akusztikai magánhangzótér (A) a beszédben ($f_0 \sim 200$ Hz) és (B) magas alapfrekvenciájú éneklésben ($f_0 = 698$ Hz) három énekes ejtésében. (Az ellipszisek középpontja: az F_1 és F_2 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 azonosítsák a hallgatók (vö. pl. Scotto di Carlo–Germain 1985). Ezt a feltételezést azonban sem a beszédhangok percepciójának szakirodalma, sem pedig az énekelt magánhangzók egyes újabb kísérletes vizsgálatai nem támasztják alá.

A beszédtudományokban közismert, hogy a magánhangzók percepciójában fontos szerepe van az F_1 és az F_2 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_1 és az f_0 érzeti távolsága (Traunmüller 1981), az elöl és a hátul képzettség érzetét pedig az F_2 és az F_3 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_1 és f_0 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_0 és az F_1 távolsága számít kritikusnak (de ez az alapfrekvencia-tartománytól függően változik), valamint tudjuk, hogy az éneklésben az $F_1 : f_0$ 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 alapfrekvencia emelésével. A vizsgált legmagasabb alapfrekvenciákon (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 alapfrekvencia-tartományon (l. pl. Scotto di Carlo–Germain 1985; Deme 2014b, 2015). 1 kHz alapfrekvencia 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 alapfrekvencián 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).

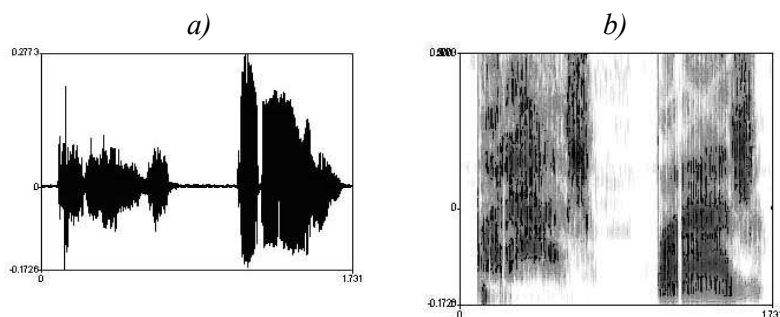
A magas alapfrekvenciájú szoprán éneklés során tehát gyakorlatilag nem ejtethetők és nem is ismerhetők fel a kotta szövegében feltüntetett magánhangzók. Feltehető, hogy ezek kizárólag azért ismerhetők fel mégis az ejtési szándék szerint, mert azokat az énekesek értelmes szavakban, mondatokban ejtik, ez pedig egészen egyszerűen leszűkíti a felismerésben szóba jöhető magánhangzókat, tehát nagyobb eséllyel „tippelünk 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 „prézelés”, 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 zörejessége (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 nyelvéllátásokat tévesztették a hallgatók. Ennek oka az énekelt magánhangzóknál leírt észlelési mintázat: azaz a vokális zártságát az $F_1 : f_0$ viszony alapján azonosítjuk, azonban ez az akusztikai kulcs suttogásnál a hangszalagregzés hiányában nem elérhető az észlelés számára. A jelentés nélküli hangsorok esetében lényegében csak a felső, míg a szavakban a középső nyelvéllátásfokú magánhangzóknál is fordult elő téves azonosítás.

A hangszalagregzés 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

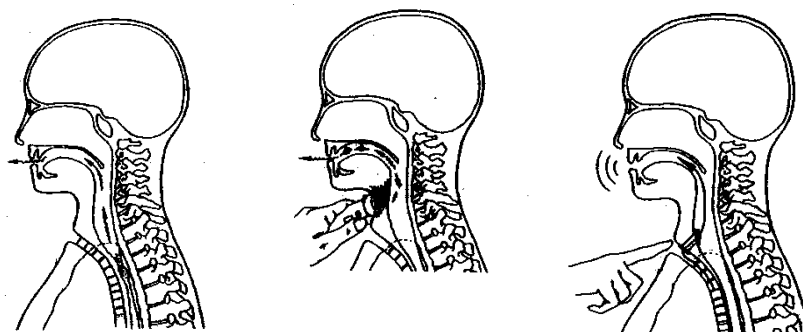
Magyarországon évente több száz esetben végeznek teljes gégeeltávolítást (laryngectomia), általában valamilyen rákos megbetegedés következményeképpen. Ezt követően – hangszalagok hiányában – a beszélő nem tud a megszokott módon zöngét létrehozni, hanem valamilyen alternatív beszédképzési módot kell elsajátítania. 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 beszéd szervei 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 nyelőcsőbeszédről 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 nyelőcsőbeszédrel,

illetőleg hogy a hallgatók ezeket a közléseket kérdésként azonosítják-e (vö. Markó–Grácsi 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 alapprofrekvencia (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 gége nyílást, és a garat is szűkebb, mint a normál beszédprodukción. 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. *Beszédkutató* 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. *Beszédkutató* 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. *Beszédkutató* 1997. 180–187.
- Markó Alexandra 2008. A magyar eldöntendő kérdő mondat megvalósulásai különböző beszédmódokban. *Beszédkutató* 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 nyelvcsöbeszédben. *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 nyelvcsöbeszéd és a hangprotézises beszédproduktió 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. Beszédkutató, 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 alaphékvenciákon a szopránok nyíltabban ejtik a magánhangzókat, amiket a hallgatók is nyíltabb magánhangzókként azonosítanak.
2. Milyen beszédhelyzetekben jellemző a suttogott beszéd?
3. Keressen rá videómegosztó portálokra 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 portálokra, és figyelje meg a hasbeszélés sajátosságait!