

Eötvös Loránd Tudományegyetem  
Bölcsészettudományi Kar

## A DOKTORI ÉRTEKEZÉS TÉZISEI

Deme Andrea

### Az énekelt magánhangzók fonetikai elemzése

Nyelvtudományi Doktori Iskola

vezető: Prof. Dr. Bárdosi Vilmos CSc

Alkalmazott Nyelvészeti Doktori Program

vezető: Prof. Dr. Gósy Mária DSc

A bíráló bizottság tagjai és tudományos fokozatuk:

A bizottság elnöke: Prof. Dr. Gósy Mária DSc

Hivatalosan felkért bírálók: Prof. Dr. Vicsi Klára DSc

Dr. habil. Bóna Judit PhD

A bizottság titkára: Dr. Gyarmathy Dorottya PhD

A bizottság további tagjai: Dr. Szaszák György PhD;

Prof. Dr. Olaszgy Gábor DSc, Prof. Dr. Adamikné Jászó Anna DSc (póttagok)

Témavezető: Dr. habil. Markó Alexandra PhD

Társtémavezető: Prof. emer. dr. Johan Sundberg PhD

Budapest, 2015

## Bevezetés

A jelen értekezésben bemutatott kutatás és kísérletsorozat alapját a szoprán éneklés és a magánhangzóejtés látszólagos ellentmondása képezi. Kísérleteinkben e kérdés produkciós és percepciós aspektusait vizsgáljuk a magyar nyelv magánhangzóin.

Az egyes magánhangzók minőségét és a magánhangzók egymástól való elkülöníthetőségét a magánhangzók létrehozásakor az artikulációs szervek állását jellemző üregi rezonanciák, azaz a formánsok alakítják ki (vö. Fant 1960). A formánsok között kitüntetett jelentőséggel az első és második formáns (F1 és F2) bír, ezek a formáns-frekvenciák ugyanis az egyes magánhangzókra jellemzőek, ráadásul az artikulációs szervek állásával viszonylag egyszerűnek tekinthető összefüggésben állnak. Az első formáns a nyelv függőleges helyzetével fordított arányban változik (a magasabb nyelvállás tehát alacsonyabb frekvenciaértékeket eredményez az első formásra, így pl. az *i* első formánsa alacsonyabb, mint a nála nyitabb *á* első formánsa), míg a második formáns a vízszintes nyelv helyzet függvényében alakul (a hátul képzett, veláris hangzók, pl. az *u* második formánsa alacsonyabb, mint az elől képzett, palatális magánhangzók, pl. az *i* második formánsa, vö. pl. Gósy 2004). Ennek megfelelően a zárt magánhangzók, így például az *i* vagy az *u* első formánsa meglehetősen alacsony frekvenciaértékű. A magas hangosztályok énekesei, elsősorban a szopránok, azonban enél a frekvenciaértéknél általánosan magasabb alapprofrecvencián ( $f_0$ ) énekelnek.

Felmerül tehát a kérdés, mi történik a magánhangzók első (vagy adott esetben akár második) formánsával a magas alapprofrecvenciájú éneklésben akkor, ha az énekelt alapprofrecvencia magasabb, mint a kérdéses formánsok átlagos (beszédben jellemző) frekvenciaértékei. Ha ugyanis ezek a formánsok „eltűnnek” a magánhangzó spektrumából (tehát az énekelt magánhangzókban csak a magasabb sorszámú formánsok rezonálnak), az azt jelenti, hogy csökken az énekhang hangereje, hirtelen drasztikusan megváltozik a hang hangszínezete és az egyes magánhangzó-minőségek is eltűnnek a magas alapprofrecvenciákon. Mivel mindezt nem tapasztaljuk (hiszen a szoprán énekesek általában jól érthetőnek tűnő, de feltétlenül nagy hangerejű hangot, nagy intenzitású magánhangzókat képeznek nagymértékben homogén hangszínezettel), az énekelt magánhangzók ejtésében sajátos képzésjegyek feltételezhetők.

A kérdéssel kapcsolatban a szakirodalomban egy mára igen elterjedt nézet uralkodik. Eszerint azokban a helyzetekben, amikor az F1 értéke alacsonyabb lenne, mint az  $f_0$  frekvenciaértéke (rövid jelölése:  $F1 < f_0$ ), az énekes az F1 értékét a megemelt  $f_0$  értékére vagy valamivel a fölé hangolja (rövid jelölése:  $F : f_0$ ). Az énekesek tehát egyre nyittabban ejtik a magánhangzókat az alapprofrecvencia emelésével (Sundberg 1975, Johansson et al. 1982, Sundberg–Skoog 1995, Bresch–Narayanan 2010, Hertegård–Gauffin 1993, Joliveau et al. 2004, Garnier et al. 2010). A hangolás eredményeképp az énekhang elméletben akár 20-30 dB hangnyomásszint-növekedés (azaz hangerő-növekedés) is elérhető lehet (Sundberg 1979, 1987). A hangolás megjelenése azonban empirikus úton igen nehezen igazolható, hiszen a magas alapprofrecvenciákon ejtett magánhangzók akusztikai lenyomatában a spektrális alul-mintavételezettség jelensége miatt (vö. de Cheveigné–Kawahara 1999) a hagyományos, Fourier-transzformáción

alapuló elemzések segítségével nem határozhatók meg megbízhatóan az egyes formánsok középértékei, így az F1 frekvenciája sem. Az F1 : f0 formánshangolás vizsgálata tehát nem szokványos kísérleti és mérési módszerek alkalmazását követeli meg (artikulációs modellezést, külső gerjesztés használatát a felvételkedzítésnél stb.), melyek használata azonban szükségszerűen problémákkal vagy sokszor akár megbízhatatlan adatokkal is jár (vö. pl. Bresch–Narayanan 2010, Hertegård–Gauffin 1993).

Mivel az F1 : f0 hangolás következtében az éneklés és a magánhangzóejtés konfliktusba kerülhet, felmerül az a kérdés is, hogyan lesznek észlelhetőek azok az énekelte magánhangzók, melyek specifikus rezonanciáit (formánsait) az énekes inkább a megfelelő minőségű énekhangképzés szolgálatába állította az egyes magánhangzók közti különbség megtartása helyett. Ezt a kérdést eddig csak néhány nyelv magánhangzóival kapcsolatban vizsgálták. E vizsgálatok eredményei szerint csökken a magánhangzóknak az ejtés szerinti azonosítása az alulfrekvencia emelésével, és az eddigi kísérletek igazolni látszottak azt az intuitív feltételezést is, hogy a hallgatók egyre nyíltabb magánhangzókként azonosítják az énekelte magánhangzókat (Scotto di Carlo–Germain 1985, Benolken–Swanson 1990, Hollien et al. 2000). Egy további vizsgálat – a beszédpercepció szakirodalmából kiindulva – azt is demonstrálni látszott, hogy lehetséges a magánhangzók „eredeti” minőségének magasabb arányú felismerése akkor, ha a hangzókat mássalhangzó-környezetben ejtik az énekesek (Smith–Scott 1980). Ugyanakkor bizonyos értelemben a fent említette vizsgálatok egyike sem jól kontrollált kísérleteken alapszik, így eredményeik megszorításokkal kezelendők, a kérdések pedig további vizsgálatot érdemelnek.

## **A jelen kutatás célja, kérdések, hipotézisek**

A jelen értekezésben bemutatott vizsgálatok célja a magas alulfrekvenciájú éneklésben megvalósuló magyar magánhangzók akusztikai és percepció jellemzőinek leírása. Kérdéseinket a szakirodalom nyomán felmerült problémákra és felvetésekre alapoztuk, és négy kérdéscsoportba rendeztük. Az énekelte magánhangzók akusztikumát érintő kérdések:

1. Hogyan írható le a magyar magánhangzókat érintő produkciós változások az alulfrekvencia függvényében, különösen akkor, ha az f0 magasabb, mint az adott magánhangzónak a beszédben megjelenő első formánása? Mi jellemzi a spektrális maximumokat, és mi jellemzi a toldalékcsoben az artikuláció során megjelenő (F1 és F2) üregi rezonanciákat?

Mivel a kérdések megválaszolására a hagyományos, Fourier-transzformáción alapuló spektrális elemzésen túl módszertani újítások alkalmazása is szükséges, a kérdéshez kapcsolódó további megoldandó feladat egy megfelelő módszertani eljárás kidolgozása és tesztelése is.

Az énekelte magánhangzók percepcióját érintő kérdések:

2. Mi jellemző az énekelte magyar magánhangzók észlelési tendenciáira? Hogyan változik az ejtési szándék szerinti azonosítás az f0 emelésével? Igazolható-e, hogy a magas alulfrekvenciákon csak a nyíltabb ejtésű magánhangzóknak az ejtésnek megfelelő azonosítása jellemző? Mi jellemzi az ejtési szándéktól eltérő azonosításokat?

Igazolható-e a produkció és a percepció közti lineáris összefüggés, azaz az  $f_0$  emelésével megjelenő egyre nyíltabb magánhangzó-minőségek azonosítása?

3. Miként hat a mássalhangzó-környezet az énekelt magánhangzók azonosítására? Milyen hatása van a koartikulációs formánsátmenetekben rejlő akusztikai információknak a magánhangzók észlelésére? Miként hat (és a formánsátmeneteknél tapasztalt hatástól elkülöníthető-e) a mássalhangzó-környezet akkor, ha a mássalhangzók és a magánhangzó kapcsolatából előálló szekvenciák értelmes szavakat alkotnak (tehát az észlelést a felsőbb nyelvi szintekhez kapcsolódó feldolgozási folyamatok aktivációja is segíti)?

4. Miként hat a zöngéindítás az énekelt magánhangzók elkülönítésére/azonosítására? Alkalmazható-e az énekelt magánhangzók esetében az a hangszerek hangszínének megkülönböztetésében talált eredmény, mely szerint a hangindítás mint akusztikai kulcs jelenléte növeli a hangszín felismerésében nyújtott percepció teljesítményt?

5. Különbözik-e bármilyen módon a naiv, énektechnikailag képzetlen hallgatók és az énekesek percepciója – sikeresebbek-e az énekesek az egyes énekelt hangzóknak az ejtési szándék szerinti azonosításában? A kérdés vizsgálata azt célozza, hogy a lehetőségek szerint legalaposabban körüljárjuk a produkciós változások percepció kompenzációjának lehetséges módozatait.

A vizsgálandó kérdéseinkkel kapcsolatban három hipotézist fogalmaztunk meg:

**H1:** Az énekesek a magas alapprofrekvenciákon énekelt magyar magánhangzók első formánsát a megemelt  $f_0$  értékére (vagy valamivel a fölé) hangolják akkor, ha egyébként az  $f_0$  értéke magasabb lenne, mint a magánhangzóra a beszédben jellemző  $F1$  értéke. Ennek értelmében az  $F1 : f_0$  megjelenése az éneklésben a magánhangzók nyíltsági fokának (illetve beszédben jellemző  $F1$  értékének) függvénye.

**H2:** Mivel magas alapprofrekvenciákon az énekelt magánhangzóknak a frekvenciaszerkezete és a spektrális felbontása is megváltozik, a magas alapprofrekvencián énekelt magyar magánhangzóknak az ejtési szándék szerinti azonosítása az  $f_0$  emelésével törvényszerűen csökken. Az  $f_0$  és  $F1$  csökkenő távolsága miatt a magas  $f_0$ -on ejtett magánhangzók akusztikai megjelenése az ejtési szándéktól eltérő azonosításokban inkább a zárt hangzók felismerését facilitálja.

**H3:** Mivel magas alapprofrekvenciákon az énekelt magánhangzók frekvenciaszerkezete megváltozik, továbbá a magánhangzók spektrális felbontása (a környező mássalhangzókkal alkotott szonoráns hangátmenetek spektrális felbontásával együtt) csökken, a mássalhangzó-környezet a formánsátmenetekben rejlő akusztikai kulcsok révén nem fejt ki pozitív hatást a magánhangzóknak az ejtési szándék szerinti azonosítására. Ugyanez érvényes az izolált ejtésű magánhangzók zöngéindítására is. A mássalhangzó-környezet pusztán abban az esetben segítheti az ejtési szándék szerinti azonosítást, ha az azonosítandó magánhangzót tartalmazó hordozó hangsorok olyan értelmes szavak, melyek száma kisebb, mint a válaszként szóba jöhető összes magyar magánhangzó száma.

## Kísérletek

### A magyar magánhangzók ejtése és észlelése az éneklésben

Első vizsgálatunk egyik célja az volt, hogy feltárjuk a magas alaphangfrekvencián énekelte magánhangzóknak a hagyományos (Fourier-transzformáción alapuló) akusztikai elemzés segítségével kinyerhető jellemzőit, illetve hogy megállapítsuk az alaphangfrekvencia emelésével megfigyelhető változásokat. A második cél az ejteni szándékozott magánhangzók észlelésében az  $f_0$  emelésével megjelenő tendenciák leírása volt.

#### Anyag, módszer és kísérleti személyek

Vizsgálatunk anyagát a hosszú vagy fonémakategória-váltás nélkül hosszban ejtendő 9 magyar magánhangzó /v a: e e: i: o: ø: u: y:/<sup>1</sup> képezte,  $mVn$  szerkezetű hangsorokban, egy szoprán (életkor: 50 év) énekesnő beszédejtésében és énekprodukciónak, négy alaphangfrekvencián (500 Hz, 550 Hz, 650 Hz).

A percepció tesztnél 11 adatközlő (4 férfi és 6 nő) vett részt (21 és 25 év közöttiek, átlag: 22 év). A tesztelők feladata az volt, hogy azonosítsák és egy válaszlapon jelöljék az elhangzó hangsorokban felismert magánhangzókat. A stimulusokat randomizált sorrendben mutattuk be a résztvevőknek. A percepció adatok statisztikai próbákat végeztünk az SPSS 13.0 programban.<sup>2</sup>

A magánhangzóknak a Praat program (Boersma–Weenink 2009) segítségével megállapítottuk meg az énekelte magánhangzóknak akusztikai kimenetében megfigyelhető maximumok frekvenciahelyeit<sup>3</sup>.

#### Eredmények

Az észlelési adatok alakulását az 1. ábra mutatja. Eszerint az egyes magánhangzóknak és az egyes képzési jegyeknek az ejtési szándék szerinti azonosítása bár nem fokozatosan, de – az elvárásoknak megfelelően – csökkent az  $f_0$  emelésével. Ennek értelmében tehát minél magasabb alaphangfrekvencián énekelte az egyes magánhangzókat az énekes, annál kevésbé ismerték fel azokat a hallgatók az ejtési szándék (azaz az énekesnek a produkció felvételkor mutatott stimuluslista) szerint. (Az ábrán egy adatpont az egyes alaphangfrekvenciákon ejtett összes magánhangzóra adott összes válasza utal mint 100%.)<sup>4</sup> Az adatok összevetéséből az is kiderül, hogy a labialitás

---

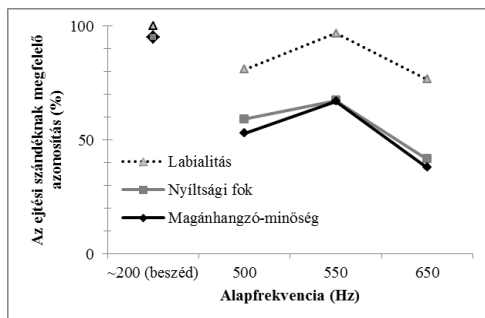
<sup>1</sup> Mivel a rövid magánhangzók ejtése az éneklésben problematikus lehet, a rövid magánhangzókat sem itt, sem másutt nem vizsgáltuk az értekezésben. Ezzel együtt a szakirodalom alapján mégis biztosított, hogy a jelen kísérletben az összes magyar magánhangzó-minőség vizsgálatára módunk volt (vö. pl. Gósy 2004).

<sup>2</sup> Az itt közölt vizsgálatok során előbb minden esetben Shapiro–Wilk-próbával ellenőriztük a normális eloszlás kritériumát, majd parametrikus adatok esetén varianciaanalízist vagy ismételt mérés varianciaanalízist alkalmaztunk (ANOVA, rep. meas. ANOVA,  $t$ -próba), nem parametrikus adatok esetén pedig  $\chi^2$ -próbát vagy Mann–Whitney U-próbát. A korrelációt a Pearson-féle korrelációelemzéssel teszteltük.

<sup>3</sup> A lineáris Hz értékeket érzetileg (Barkba) transzformáltuk.

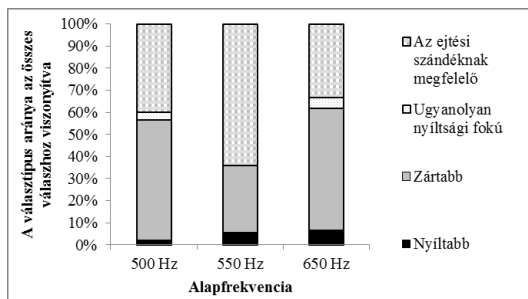
<sup>4</sup> A beszéd módra kapott azonosítási adatok elkülönítésének oka (itt és másutt is), hogy ezek az adatpontok nem csak alaphangfrekvenciájában, hanem produkció módjában is eltérő stimulusokra kapott eredményeket mutatnak.

jegy jóval ellenállóbb az f0 emelésével szemben, mint a magánhangzó nyíltsági foka (ANOVA,  $F(2) = 8,34$ ;  $p = 0,02$ ). Az f0 emelésével a válaszok egyre növekvő szóródását láttuk.



1. ábra: Az egyes magánhangzó-minőségek, valamint az ejteni szándékozott nyíltsági fok és ajakműködés az ejtési szándék szerinti azonosítása

A 2. ábra a második vagy harmadik nyíltsági fokú (azaz középső vagy alsó nyelvállású) /ε e: ø: o:/ magánhangzókra kapott válaszok megoszlását foglalja össze. Ezek a magánhangzók egy fokkal nyíltabb és egy fokkal zártabb szomszédal is rendelkeznek a magyar magánhangzó-rendszerben.

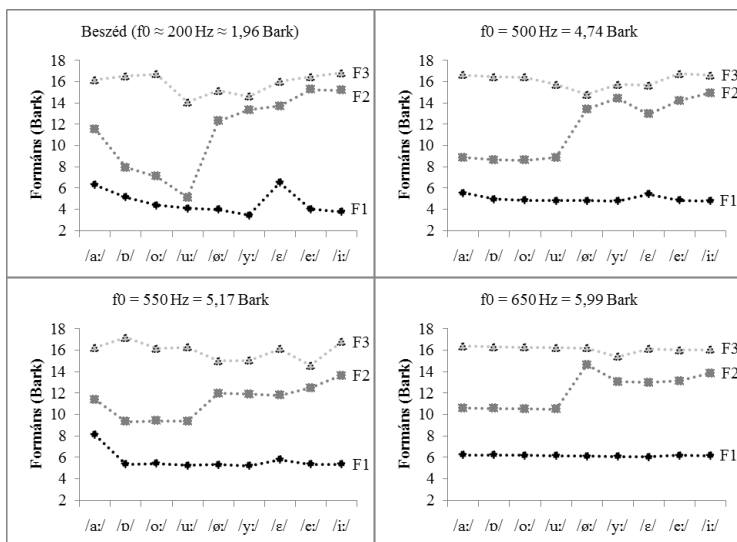


2. ábra: A második és harmadik nyíltsági fokú /ε e: ø: o:/ hangzókra az egyes alapfrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltsági foka szerint

A szakirodalomban olvasható, az artikuláció és akusztikum, valamint a percepció lineáris összefüggéseit feltételező hagyományos nézet szerint az f0 emelésével (az F1 : f0 hangolás miatt) egyre nyíltabban megvalósuló hangzókat egyre nyíltabbakként is azonosítják a hallgatók. A 2. ábrán közölt adatok e felvetés a lehetőségek szerint legjobban kontrollált ellenőrzését teszik lehetővé (korábban ilyen elemzést nem végeztek). Az összevetésből kiderül, hogy a vizsgált alapfrekvenciákon az adatközlők legnagyobb részét az ejteni szándékozott magánhangzónál zártabb magánhangzókat azonosítottak akkor, amikor nem az ejteni kívánt magánhangzókat azonosították a kísérlet során.

A 3. ábra a hangnyomás-idő függvény Fourier-elemzésén alapuló akusztikai mérések eredményét mutatja. Míg alacsonyabb f0-értéknél a 3. ábrán közölt adatok egybe-

eshetnek az artikulációs konfigurációra utaló első három rezonancia frekvenciájával, magasabb alapfrekvenciákon a spektrális alul-mintavételezés miatt inkább már csak az adott rezonancia sávzsélességébe tartozó legintenzívebb felharmonikus értékére utalnak.



3. ábra: A beszélve és énekelve ejtett magánhangzók spektrumában mérhető első három spektrális energiamaximum frekvenciaértéke (F1, F2, F3)

Akusztikai adataink szerint azon esetekben, amikor az  $f_0$  emelése során az énekes hangjának alapfrekvenciája megközelítette a magánhangzó a beszédben mérhető F1 formánsának értékét (500 Hz-en), az F1 a megemelt (magas)  $f_0$  értékére vagy a fölé hangolódott. Erre abból következtethetünk, hogy az  $f_0$  és F1 az adott alapfrekvenciától kezdve nem volt elkülöníthető a magánhangzó spektrumában, miközben a spektrum legintenzívebb akusztikai összetevője az első harmonikus, azaz az  $f_0$  volt (az F1 feltehetőleg tehát ezt a komponenset erősítette).

### Következtetések

Percepciós vizsgálatunk eredményeit összegezve megállapítható, hogy kísérletünk nem igazolta az a szakirodalomban olvasható (alapvetően spekulatív) feltételezést, mely szerint az  $f_0$  emelésével (az F1 :  $f_0$  hangolásának következtében) egyre magasabb F1 a megvalósult magánhangzók percepciós nyitódását okozná – ezzel éppen ellentétes tendenciát találtunk. Legáltalánosabban a percepciós adatok az  $f_0$  emelésével fokozatosan egyre bizonytalanabb magánhangzó-minőségekre utalnak.

Akusztikai méréseinkben a kisugárzott hang spektrumának energiamaximumait detektálva az F1 :  $f_0$  hangolás közvetett bizonyítékait láttuk.

## A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a mássalhangzó képzésmódjának hatása

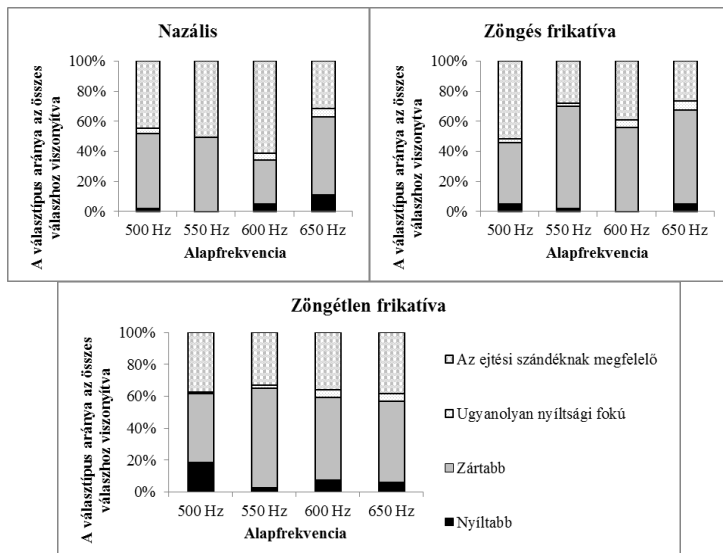
Második vizsgálatunkban azt a feltételezést teszteltük, hogy a magánhangzók zárabbként azonosítása lehetett-e a nazális hangkörnyezet hatásának eredménye, a magánhangzó nazalizálódása ugyanis okozhat zártabb észleletet (vö. Krakow et al. 1987). Ehhez a nazális kontextust zöngés és zöngétlen obstruensekkel vetettük össze.

### Anyag, módszer és kísérleti személyek

A korábban is vizsgált magánhangzókat itt  $mVn$ ,  $zsVs$  és  $sVs$  kontextusban rögzítettük a beszédben és 500, 550, 600, 650 Hz alappfrekvencián ugyanazon énekes ejtésében, és teszteltük egy, az előzőekhez hasonló percepció tesztnben 15 adatközlővel (10 nő, 5 férfi, 21–29 évesek, átlagos életkoruk 23 év).

### Eredmények

A vizsgálat legfontosabb eredményeit a 4. ábra összesíti. Eszerint minden középső vagy alsó nyelvállású magánhangzó esetében az ejtési szándéknál zártabb felismerés volt jellemző minden kontextusban. Ráadásul, bár a statisztikai összevetés szerint nem szignifikáns módon, de a nazális kontextusban jellemzően valamivel magasabb volt az ejtési szándék szerinti azonosítások aránya.



4. ábra: A második és harmadik nyíltsági fokú /ε e: ø:/ hangzókra az egyes alappfrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltsági foka szerint

### Következtetések

Eredményeink alapján megállapíthatjuk, hogy a nazális kontextusban tapasztalt észlelési tendenciák nem térnek el a más obstruens környezetben ejtve hallott énekelt magánhangzókra az ejtési szándék szerinti azonosításától. Az adatok megerősí-



tik azt a feltételezésünket, hogy a produkció és percepció egyszerű, lineáris összefüggéseit sugalló szakirodalmi feltételezés téves, tehát a nyíltabban ejtett hangzókat nem (feltétlenül) nyíltabbként azonosítják a hallgatók.

### **A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: azonosítás mássalhangzó-környezetben, izolált ejtésben, valamint a zöngéindítás mint akusztikai kulcs hiánya esetén**

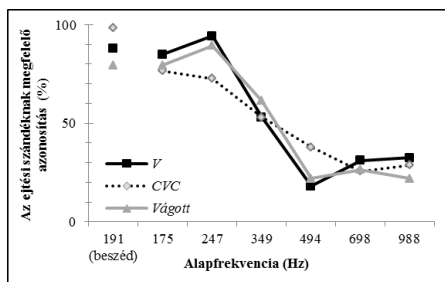
Harmadik kísérletünkben egy korábbi vizsgálat (Smith–Scott 1980) bizonyítottnak vélt felvetését, a mássalhangzó-környezetnek mint koartikulációs formánsátmeneteknek a feltételezett pozitív hatását teszteltük az azonosításra (az idézett tanulmánnyal szemben kiegyenlített elemszámú feltételekkel). A zöngéindítás szerepét korábbi pszichoakusztikai eredményekre alapoztuk, melyek szerint a hangszín felismerésében kiemelt szerepet játszik a hangindítás (vö. pl. Berger 1964).

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

A vizsgálatban az /a: i: u:/ vokálisokat rögzítettük egy szoprán énekesnő (életkor: 27 év) beszéd- és énekprodukciónak (utóbbi esetben 175 és 988 Hz között a zenei *h* és *f* hangokon, 3 oktávon, összesen 6 alapprofrekvencián) izolációban („V” kondíció) és a *bVb* hangkörnyezetben („CVC” kondíció). A zöngéindítás nélküli anyagot („Vágott” kondíció) egy hangeditáló program segítségével hoztunk létre. A teszthangsorokat egy, az előbbiekhöz hasonló percepció tesztnek vetettük alá 22 adatközlő részvételével (21 és 35 év közöttiek, átlagosan 29 évesek).

#### **Eredmények**

A vizsgálat legfontosabb eredményét az 5. ábrán mutatjuk be. Az ábrán azt látjuk (és a statisztikai elemzések ezt megerősítették), hogy a magánhangzóknak az ejtési szándék szerinti azonosítására kizárólag az *f0* hat számottevően (ANOVA:  $F(1) = 33,33$ ,  $p < 0,001$ ). Az azonosítási arányok és az *f0* összefüggését ugyanis közepesen erős negatív korreláció jellemzi ( $r = -0,631$ ,  $p < 0,001$ ).



5. ábra: A három vizsgált magánhangzó (/a: i: u:/) azonosításának összesített százalékos aránya az alaphang függvényében a három kondícióban

#### **Következtetések**

Eredményeink nem igazolták sem a koartikulációs formánsátmenetek, sem pedig a zöngéindítás pozitív hatását az énekelt magánhangzók az ejtési szándék szerinti azo-

nosítására. Az eredmények összhangban vannak azzal a feltételezésünkkel, hogy a magánhangzók akusztikai szerkezete megváltozik az f0 emelésével, így az ejtési szándék szerinti minőséget nem kódolhatják akusztikai kulcsok.

## **A mássalhangzó-környezet hatása az énekelt magánhangzók észlelésére: a magánhangzók azonosítása értelmes és értelmetlen szavakban**

Negyedik kísérletünkben azt a feltételezésünket vizsgáltuk meg, mely szerint a kontextus mint a magánhangzók értelmes szavakba ágyazottságának az ejtési szándék szerinti azonosítást segítő pozitív hatása (melyre Smith és Scott 1980 tanulmányából következtethetünk) eltűnik akkor, ha a szóba jöhető értelmes és értelmetlen szavak mint kontextusok száma hasonlóan magas (de nem feltétlenül azonos). A kísérlet másik kérdése, hogy az énekes hallgatóság vajon képes-e saját produkciós gyakorlata miatt magasabb számban az ejtési szándék szerint azonosítani a magas alaphangfrekvencián énekelt magánhangzókat, mint a „laikus” hallgatók.

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

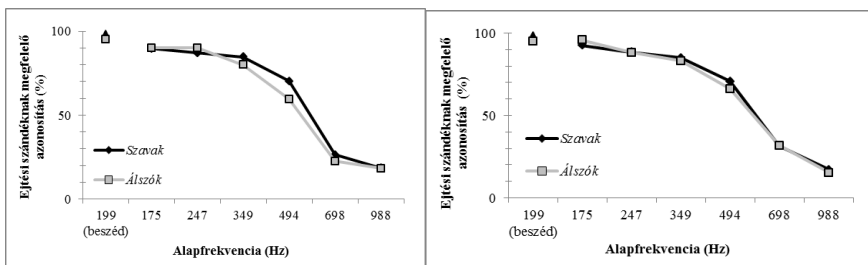
A hipotézisek vizsgálatára egy két ülésből álló percepciós tesztet terveztünk, melyet 20 „énekes” (26 és 45 év közöttiek, átlagos életkoruk 31 év) és 20 „nem énekes” (26 és 52 év közöttiek, átlagos életkoruk 29 év) adatközlővel vettünk fel. A teszthez az /v a: ε i: ø: y:/ magánhangzókat rögzítettük értelmes (*szVr*, pl. *szár*) és értelmetlen (*dVr*, pl. *dár*) szavakban (rendre „Szavak” és „Álszók” kondíció), az előző vizsgálat énekesével az ott megjelölt alaphangfrekvenciákon. A teszt első felében mindkét csoport a „Szavakat”, a második felében pedig az „Álszókat” hallotta. A válaszadáskor mindkét tesztrészben mind a 9 magyar magánhangzóból választhattak a tesztelők (tehát adhattak nem legitim választ), de kivetítettük az elfogadott válaszokat, valamint a kontextust is. Így a „Szavak” esetében valamivel nagyobb lehetett a „helyes találat” esélye, mint az „Álszóknál” (rendre 13 és 11%).

### **Eredmények**

A vizsgálat legfontosabb eredményét a 6. ábrán közöljük. Látható (és a statisztikai elemzések is alátámasztják), hogy sem a két kondícióban, sem pedig a két adatközlőcsoportban kapott adatok nem térnek el egymástól, ugyanis az ejtési szándék szerinti azonosítások aránya minden kondícióban és csoportban csökken az f0 emelésével.

### **Következtetések**

Vizsgálatunk bizonyítékkal szolgált arra, hogy noha az értelmes szavak a rendelkezésre álló kisebb számú válaszlehetőség miatt segíthetik a magas alaphangfrekvencián énekelt magánhangzóknak az ejtési szándéknak megfelelő azonosítását (vö. Smith–Scott 1980), a pozitív hatás eltűnik, ha az értelmes szavak hasonlóan nagyszámú válaszlehetőséget engednek meg az azonosításban, mint az értelmetlen szavak, ill. az izoláltan ejtett magánhangzók. Emellett azt is találtuk, hogy produkciós gyakorlottságuk ellenére az énekesek nem élveznek percepciós előnyt az énekelt magánhangzók az ejtési szándéknak megfelelő azonosításában.



6. ábra: Az egyes alapfrekvenciákon bemutatott összes stimulus ejtési szándék szerinti azonosítási aránya a „nem énekes” (balra) és „énekes” (jobbra) csoportban

## Az énekelt magánhangzók akusztikai szerkezete néhány módszertani újítás tükrében, valamint az akusztikai és percepció jellemzők összefüggései

Negyedik vizsgálatunkban egy alternatív felvételi és mérés módszertani eljárás újszerű kombinációjával igyekeztünk megállapítani az énekelt magánhangzók minőségét kialakító üregi rezonanciák értékét, valamint a rögzített anyagon percepció tesztek is végeztünk azért, hogy következtetéseket fogalmazhassunk meg a produkció és percepció összefüggéseiről. Percepció adatainkat (eddig egyedülálló módon) Traummüller (1981) az észleleti nyíltságot az F1 és f0 tonotópiás távolságából levezető modellje segítségével igyekeztünk értelmezni. Eszerint az  $f_0 = 370$  Hz kiemelt jelentőségű, ugyanis e fölött a beszédben pszichoakusztikai okokból csak első és negyedik nyíltsági fokú magánhangzók észlelhetők. Ugyanakkor a modelltől az F1 : f0 hangolás miatt megváltozó prominenciaviszonyok percepció következményei nem jósolhatók.

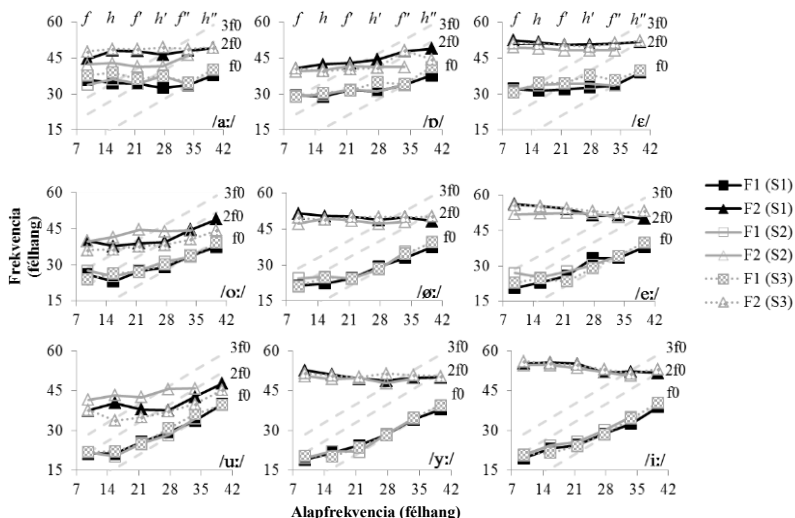
### Anyag, módszer és kísérleti személyek

A vizsgálatban a 9 magyar magánhangzót elemeztük az előzőekben említett 6 énekelt alapfrekvencián és a beszédben, három profi szoprán (28, 28 és 45 évesek) ejtésében. Az akusztikai elemzéseket két, egymást kiegészítő technikával végeztük. Egyfelől manuális inverz szűréssel (DeCap, Svante Granqvist, é. n.) elemeztük a természetes fonációval ejtett hangzókat, melyhez a hangfelvétel közben regisztrált elektroglossogramot (EGG) is használtunk (vö. Hertegård–Gauffin 1993). Másfelől pedig a spektrális alul-mintavételezettség kiküszöbölésére a fonáció mellett egy külső gerjesztésű jellel (elektromos gége) is besugároztuk a toldalékcsovet, majd az így létrejött hullámformát is rögzítettük, és akusztikailag elemeztük (vö. Sundberg 1975). A természetes fonációval ejtett magánhangzókat a korábbiakhoz hasonló percepció tesztnak vetettük alá, melyben 21 felnőtt (26 és 45 közöttiek, átlagos életkoruk 29 év) vett részt.

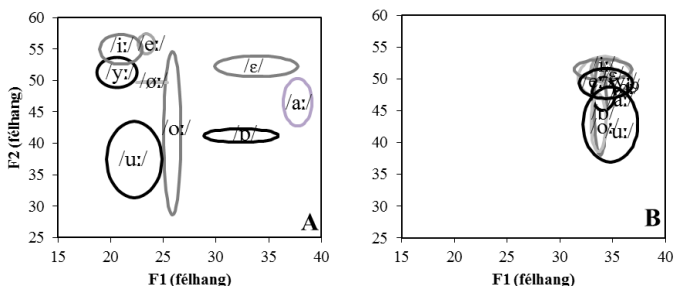
### Eredmények

Akusztikai eredményeink szerint az F1 : f0 hangolás az egyes magánhangzók nyíltsági fokának, illetve F1-ének függvényében jelent meg az f0 emelésével (zárt, első nyíltsági fokú, felső nyelvállásúak /i: y: u:/: f' zenei hang vagy 349 Hz; nyíltabb, második nyíltsági fokú, középső nyelvállásúak /e: ø: o:/: f'' zenei hang vagy 698 Hz;

legnyíltabb, harmadik és negyedik nyíltsági fokúak /ɛ ɒ a/: h'' zenei hang vagy 988 Hz, 7. ábra). Az F1 mellett az F2 f0-függő változásait is láttuk, melyek a magánhangzóknak a magasabb alapfrekvenciákon centralizálódó képzéshelyére utalnak. Ez az F1 emeléséhez szükséges állkapocsnyitásszög-növelés (fiziológiai okokból) törvényszerű velejárója. A változások következtében centralizálódó magánhangzóteret látjuk a 8. ábrán.



7. ábra: A három énekes által ejtett hangzók F1 és F2 formánsfrekvenciái mint üregi rezonanciák ( $y$  tengely) az alapfrekvencia ( $x$  tengely) függvényében. (A világosszürke szaggatott vonalak az alapfrekvenciát és annak első két felharmonikusát jelölik:  $f_0$ ,  $2f_0$ ,  $3f_0$ )

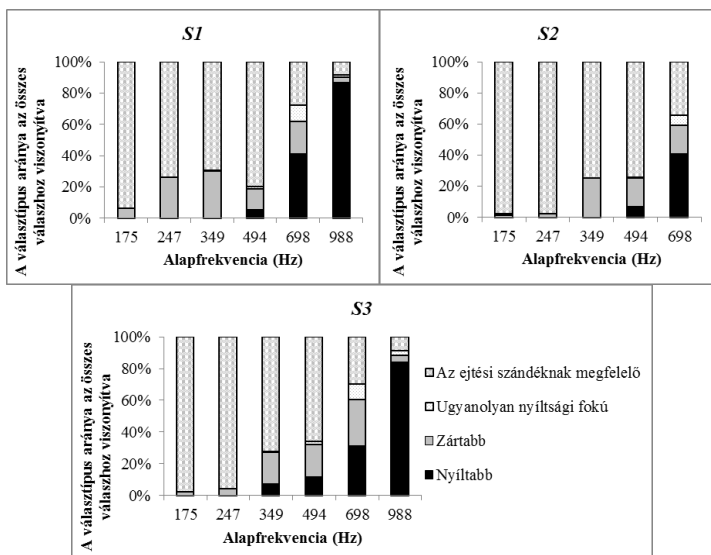


8. ábra: A három énekes által énekelt magánhangzók megvalósulásának eloszlása a (pszi-cho)akusztikai magánhangzóterben a beszédben (A panel) és az  $f''$  zenei alaphangon,  $f_0 = 698$  Hz (B panel). (Az ellipszisek középpontja: az F1 és F2 átlaga; sugara: kétszeres szórás)

Percepciósi adataink szerint az  $f_0$  emelésével a várakozásoknak megfelelően egyre csökken a magánhangzóknak az ejtési szándék szerinti azonosítása<sup>5</sup>. Ezen belül

<sup>5</sup> A három énekes által énekelt magánhangzók azonosítása nem tért el számottevően.

370 Hz alapprofrekvencia alatt a legjellemzőbb a magánhangzóknak az ejtési szándéknál zártabbként azonosítása (a korábbi eredményeinkkel összhangban) egészen 494 Hz-ig ( $h'$  zenei hang), 988 Hz-en ( $f''$  zenei hang) pedig közel 100%-ban a hangzók /v a:/-ként azonosítását láttuk. A kettő között megnövekedett a második és harmadik nyíltságú fokú hangzóválaszok aránya, sok esetben az ejtési szándéknál nyíltabb azonosításokként (vö. 9. ábra). A beszéddel szemben az éneklésre nem bizonyult igaznak az, hogy abban 370 Hz fölött csak első és negyedik nyíltságú fokú hangzók volnának észlelhetők. Az ily módon az  $f_0$  emelése mellett is megmaradó hangszínbeli változatoság feltehetőleg az  $F1 : f_0$  hangolás következménye, emiatt ugyanis a beszédben nem jellemző észleleti prominenciaviszonyok alakulhatnak ki az énekelt magánhangzóknak. Az  $f''$  hangon megnövekedett /v a:/ válaszok aránya is ugyanezen okra, egészen pontosan az  $f_0$  és  $2f_0$  összetevők erősítettségére vezethető vissza (ez ugyanis a beszédben megjelenő /v a:/-ra emlékeztető észleleti prominenciaviszonyokat eredményez), nem pedig a produkció és percepció lineáris összefüggéseire.



9. ábra: A második és harmadik nyíltságú fokú /ε e; ø/ hangzókra az egyes alapprofrekvenciákon kapott válaszok megoszlása a magánhangzók nyíltságú foka szerint

### Következtetések

Korábbi alternatív módszerek újszerű kombinációjával sikeresen demonstráltuk az  $F1 : f_0$  hangolás a nyíltságtól (ill.  $F1$ -től) függő megjelenését, és a veláris-palatáris dimenzió redukcióját az  $f_0$  emelésével az összes magyar magánhangzóban az éneklésben, 3 énekes ejtésében. Percepció adataink Traummüller 1981-es modelljének kereteiben értelmezve jól magyarázhatóak voltak, és megerősítették, hogy az éneklésben az  $f_0$  emelését kísérő produkciós nyitódást (a korábbi javaslatokkal ellentétben) a percepcióban alapvetően nem nyitódás kíséri, a produkció és percepció viszonya ennél jóval összetettebb.

## Összegzés és következtetések

**H1:** Első hipotézisünket kísérleteink igazolták, ugyanis azt találtuk, hogy a szoprán énekesek a megemelt  $f_0$  értékére emelik a magánhangzók kialakításában részt vevő első üregi rezonancia ( $F_1$ ) értékét akkor, ha az alapfrekvencia megközelíti az első formáns beszédbeli megjelenésének frekvenciáját az összes magyar magánhangzó-minőség esetében. Az  $F_1 : f_0$  hangolás tehát a magánhangzó nyíltságának függvényében jelentkezik. A kísérletekhez sikerrel dolgoztunk ki és teszteltünk egy, korábban bemutatott eszközöket ötvöző hangfelvételi és mérési eljárást.

**H2:** A második hipotézis csak részben igazolódott. Egyfelől igazolódott a magánhangzók ejtési szándék szerinti azonosítási arányának csökkenése az alapfrekvencia emelésével. Eredményeink szerint ugyanis az alapfrekvencia emelésével változó akusztikai szerkezetű magánhangzók minősége percepciósan egyre bizonytalanabbá vált, e változás pedig fokozatos (nem pedig kategorikus) természetű volt. Másfelől azonban kísérleteink nem igazolták azt a feltevést, hogy a szoprán alapfrekvencia-tartomány magasabb régióiban egyértelműen a hangzók zártabbként azonosítása volna jellemző. Eredményeink szerint 370 Hz alatt (nagyjából a zenei  $f\#'$  hang) az elvárásoknak (és Traummüllernek a perceptuális nyíltságot az  $F_1$  és az  $f_0$  távolságából levezető 1981-es modellje alapján bejósoltaknak) megfelelően az egyes énekelt magánhangzókat a nem az ejtési szándék szerinti azonosítások esetén alapvetően zártabbként azonosították a hallgatók. Ugyanakkor 370 Hz ( $f\#'$ ) fölött megnőtt a válaszok szórása, 700 Hz alapfrekvencia ( $f''$ ) fölött pedig nagyon megszorodott a nyíltabb hangzók, különösen az / $v$  a:/ válaszok száma. Ezek a tendenciák azonban nagy valószínűséggel nem egyszerűen a hangzók produkciós nyíltságának egyenes következményei, hanem az  $F_1$  és  $f_0$  megnövekedett tonotópiás távolságának, valamint az  $F_1 : f_0$  hangolásnak az eredményei.

**H3:** Percepció vizsgálataink – a kérdést vizsgáló egyetlen korábbi tanulmány (Smith–Scott 1980) eredményeivel ellentétben – azt bizonyították, hogy sem a másalhangzó-környezet mint a koartikulációs formánsátmenetek révén dinamikus akusztikai kulcsokat biztosító kontextus, sem pedig az izolált ejtésű magánhangzók zöngéindulása nem fejt ki pozitív hatást a magas alapfrekvencián énekelt magánhangzók az ejtési szándék szerinti azonosítására. Eredményeink alapján levonhattuk azt a következtetést, hogy az énekelt magánhangzók az ejtési szándék szerinti azonosítását kizárólag statisztikai hatások befolyásolják (ill. a magasabb beszédmegértési folyamatok aktivációja). Ezek a kontextus (a tágabb szövegkörnyezetbe, és mondatba ágyazottság), valamint a fonológia (a fonotaktikai valószínűségek hatásai és a fonológiai szomszédsági hatás, azaz a fonológiai szomszédsági hatás, idegen szóval *phonological neighborhood density*). Ezt támasztja alá az az eredményünk is, mely szerint az énekesek produkciós gyakorlottságuk ellenére semmilyenfajta percepció előnyt nem látszanak élvezni a magas alapfrekvencián énekelt magánhangzók az ejtési szándék szerinti azonosításában.

## A dolgozat tézisei

1. Igazoltuk azt a magas alapfrekvenciákon éneklő szoprán énekesekkel kapcsolatban feltételezett stratégiát a magyar magánhangzókra, mely szerint akkor, ha az alapfrekvencia emelése során az alapfrekvencia magasabb lenne, mint az első formáns, az énekes az első formáns frekvenciáját az alapfrekvencia értékére emeli. Demonstráltuk az összes magyar magánhangzó produkciós nyitódását az alapfrekvencia emelésével, valamint a magánhangzók redukcióját és konvergálását az /a:/ pozíciója felé a magánhangzótérben.

2. Sikeresen teszteltük a manuális inverz szűrés módszerét elektroglottográfia, valamint elektromos géggel kiegészítve a magas alapfrekvenciájú szopránéneklésben realizálódó magánhangzók formánsainak a fonáció alapfrekvenciájától független mérésében.

3. Megcáfoltuk azt a szakirodalomban honoló (de nem jól kontrollált kísérletekre alapozott) feltételezést, mely szerint a magánhangzók az alapfrekvencia emelésével fokozódó produkciós nyitódását a hangzómegvalósulások fokozatos percepciós nyitódása kísérné. Eredményeink szerint a produkció és percepció összefüggése e kérdésben (is) jóval komplexebb. Az alapfrekvencia emelésével alapvetően a magánhangzók percepciós minőségének elbizonytalanodását tapasztaltuk.

3.a A produkciós adatoknak megfelelő, de az ejtési szándéknál nyíltabb (/v a:/ hangzóként történt) magánhangzó-azonosításokat csak abban az esetben találtunk, amikor az alapfrekvencia kellően magas volt ahhoz, hogy az  $f_0$  és  $2f_0$  egymáshoz viszonyított távolsága, valamint a két összetevő erősítetttsége is hasonlóvá tegye az egyes magánhangzószerű hangzókat az /a:/ beszédbeni prominenciamintázatához. Ez az  $F1 : f_0$  hangolás miatt csak akkor valósul meg egyértelműen, amikor az alapfrekvencia emelésével változó frekvenciaértékű  $F2$  formáns eléri (és erősíteni kezdi) a  $2f_0$  összetevőt (itt  $f''$  zenei hang, azaz kb. 988 Hz).

3.b Az 1000 Hz alatti alapfrekvenciákon a produkció és percepció az egyszerű lineáris összefüggésnél sokkal komplexebbnek mutatkozik: 5-600 Hz-ig még alapvetően az egyes „eredeti” minőségük szerint nyíltabbként és zártabbként is azonosítható hangzók az ejtési szándéknál zártabbként azonosítása dominál, az ejtési szándéknál nyíltabb azonosítások száma csak e fölött a frekvenciasáv fölött kezd el növekedni. Feltehetjük, hogy a nyíltabbként azonosításhoz e produkciósan nyíltabb és  $F1 : f_0$  formánshangolt hangzók esetében szükséges a spektrális alul-mintavételezettség fokozódása, és az  $F2$  és  $2f_0$  egybeesése (azaz a  $2f_0$  erősítése), tehát a percepciós nyitódás nem a produkciós nyitódás egyenes következménye.

4. Valószínűnek látszik, hogy éppen a magánhangzók akusztikai szerkezetét megváltoztató  $F1 : f_0$  hangolás következtében lehetséges a második és harmadik nyíltsági fokú magánhangzók megjelenése a percepcióban (bár nem feltétlenül az ejtési szándéknak megfelelően) – a formánshangolási tendencia nélkül elsősorban első és negyedik nyíltsági fokú hangzók megjelenését jósolja a szakirodalom.

5. Megállapítottuk, hogy a mássalhangzó-környezet a dinamikus akusztikus kulcsok (azaz a koartikulációs formánsátmenetek révén) a beszéddel ellentétben a magas alapfrekvenciájú éneklésben nem segíti a magánhangzók az ejtési szándék szerinti

azonosítását. Ugyanakkor a mássalhangzó-környezet mint értelmes szavakba ágyazottság pozitív hatást fejthet ki akkor, ha az adott szó valószínűsége magas, illetve a szónak kevés fonológiai szomszédja van (tehát csak kis számú olyan szósomszédal rendelkezik, melytől csak a kérdéses magánhangzó különbözteti meg).

6. Megállapítottuk, hogy az énekes és nem énekes (tehát énektechnikailag képzetlen) hallgatóság gyakorlatilag ugyanolyan tendenciák szerint azonosítja a magas alapprofekvencián énekelt magánhangzókat, és az éneklésben szerzett gyakorlat nem jelent percepció előnyt a magánhangzóknak az ejtési szándék szerinti felismerésében.

7. Megállapítható, hogy az énekelt magánhangzók akusztikai szerkezete az alapprofekvencia emelésével egyre jobban megváltozik (eltér a beszédbeni megvalósulásától), mely produkciós tendencia pusztán a beszédpercepció szintjén nem kompenzálható – a kompenzációhoz a magasabb szintű beszédmegértési folyamatok aktivációja szükséges.

### Az összefoglaláshoz felhasznált irodalom

- Benolken, Martha S. – Swanson, Charles E. 1990. The effect of pitch-related changes on the perception of sung vowels. *Journal of the Acoustical Society of America* 87(4). 1781–1785.
- Berger, Kenneth W. 1964. Some factors in the recognition of timbre. *Journal of the Acoustical Society of America* 26(10). 1888–1891.
- Boersma, Paul – Weenink, David 2009. *Praat: Doing phonetics by computer (Verison 5.3)*. <http://www.praat.org>. (A letöltés ideje 2014. január 12.)
- 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.
- de Cheveigné, Alain – Kawahara, Hideki 1999. Missing-data model of vowel identification. *Journal of the Acoustical Society of America* 105(6). 3497–3508.
- Fant, Gunnar 1960. *Acoustic Theory of Speech Production*. Mouton & Co., The Hague, Netherlands.
- Garnier, Maëva – Henrich, Natalie – Smith, John – Wolfe, Joe 2010. Vocal tract adjustments in the high soprano range. *Journal of the Acoustical Society of America* 127(6). 3771–3780.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Hertegård, S. and J. Gauffin 1993. Voice source – vocal tract interaction during high-pitched female singing. In: *Proceedings of the Stockholm Music Acoustics Conference 1993 (SMAC 1993)*. Stockholm, Royal Swedish Academy of Music. 177–182.
- Hollien, Harry – Mendes-Schwartz, Ana P. – Nielsen, Kenneth 2000. Perceptual confusions of high-pitched sung vowels. *Journal of Voice* 14(2). 287–298.
- Johansson, Christine – Sundberg, Johan – Wilbrand, Hermann 1982. X-ray study of articulation and formant frequencies in two female singers. In: Askenfell, Anders – Felicetti, S. – Jansson, Erik – Sundberg, Johan (szerk.) *Proceedings of the Stockholm Music Acoustics Conference 1983 (SMAC 1983)*, Stockholm, Royal Swedish Academy of Music. 203–218.
- 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.
- Krakow, Rena A. – Beddor, Patrice S. – Goldstein, Louis M. – Fowler, Carol A. 1987. Coarticulatory influences on the perceived height of nasal vowels. *Status Report on Speech Research*, Haskins Laboratories. 31–53.



- 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.
- Smith, Lloyd – Scott, Brian L 1980. Increasing the intelligibility of sung vowels. *Journal of the Acoustical Society of America* 67(5). 1795–1797.
- Sundberg, Johan 1975. Formant technique in a professional female singer. *Acta Acustica united with Acustica* 32(2). 89–96.
- Sundberg, Johan 1979. Perception of singing. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 20(1). 1–48.
- Sundberg, Johan 1987. *The science of the singing voice*. Northern Illinois University Press, DeKalb Illinois.
- Sundberg, Johan –Skoog, Jörgen 1995. Jaw opening, vowel and pitch. *Speech Transmission Laboratory Quarterly Progress and Status Report (STL-QPSR)* 36(2–3). 43–50.
- Svante Granqvist. é. n. *DeCap*. tolván.com. (A letöltés ideje: 2013. április 26.)
- Traumüller, Hartmut 1981. Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America* 69(5). 1465–1475.

### A szerzőnek a témában megjelent publikációinak listája

- Deme Andrea 2011. Az énekelt magánhangzók fonetikai elemzése. In: Parapatics Andrea (szerk.) *Félúton 6: A hatodik Félúton konferencia (2010) kiadványa*. 25–42.
- Deme Andrea 2011. Az énekelt magánhangzók észlelése réshangkörnyezetben. In: Váradi T (szerk.) *V. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. MTA Nyelvtudományi Intézet, Budapest. 16–28.
- Deme Andrea 2012. Az énekelt magánhangzók fonetikai elemzése. In: Parapatics Andrea (szerk.) *Doktoranduszok a nyelvtudomány útjain: 6. Félúton konferencia*. ELTE – Eötvös Kiadó (Talentum), Budapest. 33–46.
- Deme, Andrea 2013. On the Hungarian sung vowels. *The Phonetician* 105-106(1–2). 73-87.
- Deme Andrea 2014a. A mássalhangzó-környezet és a zöngéindulás hatása az énekelt magánhangzók észlelésére. *Beszédkutatás* 22. MTA Nyelvtudományi Intézet, Budapest. 155–171.
- 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 2014c. Formant strategies of professional female singers at high fundamental frequencies. In: Fuchs, Susanne –Grice, Martine –Hermes, Anne –Lancia, Leonardo – Mücke, Doris (szerk) *Proceedings of the 10th International Seminar on Speech Production Cologne (ISSP)*, Köln. 90–93.
- Deme Andrea (megjelenőben). Speech perception at its best: Extracting linguistic information from phonetically underspecified input. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS) 2015*.