

A beszéd alaphangmagasságának mérése spontán beszédből és izolált [ə] hangokból, férfi és női beszélőknél

GOCSÁL ÁKOS

Bevezetés

A beszéd számos akusztikai paramétere közül az egyik leggyakrabban mért adat az átlagos alaphangmagasság. Ezt a paramétert nem elsősorban a nyelvi értelemben vett jelentésmegkülönböztető szerepe miatt vizsgáljuk, hanem azért, mert – Laver (2003) szóhasználatával élve – a beszéd extranyelvi rétegében a beszélő személy valamely tulajdonságaival áll összefüggésben.

Ilyen tulajdonság például a beszélő személy nemi hovatartozása. Hétköznapi tapasztalatunkat, hogy a férfiak mélyebb, a nők magasabb alaphangon beszélnek, több mérés is megerősíti. Szende (1995, idézi Gósy 2004) hazai és külföldi alaphangkutatások eredményeit foglalja össze. Ezek alapján megállapítható, hogy a magyar anyanyelvű férfiak átlagos alaphangmagassága 80 és 130 Hz közé esik, a nőké pedig 160 és 260 Hz közé. Ugyanez az adat német férfiaknál 103,6 és 146 Hz, nőknél pedig 185 és 247 Hz. Az átlagértékeket tekintve spanyol férfiaknál 130 Hz, nőknél pedig 230 Hz adódott (Trittin – de Santos y Lleó 1995). Portugál beszélőknél 109,2 Hz-et, illetve 186,6 Hz-et mértek (Guimaraes–Abberton 2005).

Az átlagos alaphangmagasságot felhasználjuk a beszélő nemének azonosítása során is. A férfiak és a nők hangja azonban nemcsak az átlagos alaphang magasságában különbözik, hanem többek között a magánhangzók formánsszerkezete is eltér. Gelfer és Mikos (2005) arra a kérdésre kereste a választ, hogy az alaphangmagasságot vagy a magánhangzók formánsszerkezetét használjuk-e inkább, amikor a beszélő nemét a hangja alapján határozzuk meg. Szintetizált magánhangzók lejátszása során azt találták, hogy a hallgatók az esetek 84,2%-ában férfinak, illetve 73,8%-ában nőnek vélték a beszélőt, ha a 120, illetve 240 Hz-es alaphang mellett férfiakra, illetve nőkre jellemző formánsstruktúrát alkalmaztak. Amikor a 240 Hz-es alaphanghoz férfi formánsokat párosítottak, a hallgatók az esetek 19,3%-ában jelezték csak, hogy a beszélő férfi, míg ha a 120 Hz-es alaphang mellett nőkre jellemző formánsokat alkalmaztak, akkor az esetek 20,2%-ában gondolták a hallgatók, hogy a beszélő nő. Ezek az eredmények arra utalnak, hogy a beszélő nemének megállapításakor elsősorban az alaphangjára támaszkodunk.

Beszédtudomány. Az anyanyelv-elsajátítástól a zöngelkedési időig: 1–16.

Az átlagos alaphangmagasság nemcsak a beszélő nemének megállapításához ad kulcsot, de a hallgató és a beszélő közötti rokonszenvi viszonyok kialakulását is nagymértékben befolyásolhatja. A női beszélő magasabb átlagos alaphangja vonzóbb nő benyomását kelti a férfi hallgatókban (Gocsál–Huszár 2003; Collins–Missing 2003), de ez csak egy bizonyos hangmagasságig igaz. Borkowska és Pawlowski (2011) ugyanis kimutatta, hogy a kb. 280 Hz fölötti alaphang esetében a férfi válaszadók már kevésbé vonzóknak vélték a beszélő nőket. A szerzők véleménye szerint ennek oka az, hogy a beszélők gyerekesnek, éretlennek tündek. Egy másik vizsgálat során pedig női hallgatók számára játszottak le férfiak által kiejtett magánhangzókat a kutatók (Vukovic et al. 2010). A magánhangzók alaphangját előzőleg megemelték, illetve csökkentették, így ugyanannak a hangnak egy férfiasabb és egy nőiesebb jellegű változatát is előállították. A válaszadó nőknek el kellett dönteniük, hogy a magánhangzó párok közül melyik alapján tartanak vonzóbbnak a beszélőt. A vonzóság mértékét négy fokozat egyikének megjelölésével jelezheték. Az eredmények azt mutatták, hogy a mélyebb hangú beszélőt szignifikánsan nagyobb valószínűséggel választanak rövid illetve hosszú távú párkapcsolathoz partnernek, mint a magasabb hangút.

A rokonszenvi-vonzalmi kapcsolatok kutatása mellett az utóbbi időben megjelentek olyan publikációk is, amelyek a beszélő alaphangmagassága és a neki tulajdonított dominancia összefüggéseit keresik. Borkowska és Pawlowski (2011) kutatása azt mutatta ki, hogy a mélyebb alaphangú női beszélőket dominánsabbnak gondolják a hallgatók, mint a magasabb hangúakat. Kiderült az is, hogy a női hallgatók a női beszélők hangjából „kihallott” dominanciajegyekre jóval érzékenyebbek, mint a férfi hallgatók. A már említett Vukovic és munkatársai (2010) pedig a férfi beszélőkre vonatkozóan mutatták ki, hogy a mélyebb hang alapján a női hallgatók nagyobb mértékű fizikai dominanciát feltételeztek.

Számos kutatás foglalkozik azzal, hogy az életkor előrehaladtával hogyan változik a hangunk, illetve milyen akusztikai fonetikai paraméterek alapján becsüljük meg a beszélő személy életkorát, ha csak a hangját halljuk. Az ilyen jelenségeknél is központi jelentősége van a beszéd átlagos alaphangmagasságának. Harnsberger és munkatársai (2008) fiatal (21–29 éves) és idős (74–88 éves) férfiak hangját vizsgálták. A két csoportba tartozó kísérleti személyek F_0 -értékei szignifikánsan különböztek. A fiatalok alaphangja átlagosan 107 Hz, az időseké pedig 144 Hz volt. Életkorbecslési kísérletük során felvett, illetve manipulált beszédmintákat használtak. Ez utóbbi esetben külön-külön megváltoztatták az alaphangmagasságot, a beszédtempót, illetve mindkettőt. A hallgatók az eredetileg felvett, módosítatlan beszédmintához képest szignifikánsan fiatalabbnak vélték a beszélő idős embert, amikor a felgyorsított, illetve a felgyorsított és mélyített beszédmintát hallották. Önmagában azonban az alaphang frekvenciájának 20%-os csökkentése nem eredményezte fiatalabb beszélő benyomását.

Különböző korosztályokhoz tartozó női beszélők által kiejtett izolált magánhangzók alaphangját vizsgálta Ferrand (2002). A fiatal (21–34 éves) és középkorú (40–63

éves) beszélőknél az átlagos alaphangmagasság nem különbözött szignifikánsan (209,68 és 204,49 Hz), míg ezektől szignifikánsan eltért az idős (70–90 éves) beszélőknél kapott 175,23 Hz-es érték.

Reubold és munkatársai (2010) öt olyan angol anyanyelvű beszélő alaphangját vizsgálták, akiktől több évtizedre visszamenően rendelkezésre álltak hangfelvételek. Az életkor előrehaladtával az alaphangmagasság lineáris csökkenése mutatkozott, a férfi beszélőnél azonban 87 éves koráig volt megfigyelhető a mélyülés. Ezt követően hirtelen emelkedni kezdett a hangmagassága, a korábbi 100 Hz körüli értékről néhány év alatt közel 140 Hz-re emelkedett. Középkorú beszélők esetében azonban hat év alatt még nem történik számottevő változás a beszéd akusztikai szerkezetében (Gósy 1996).

Bár az alaphangmagasság és a beszélő neme és életkora között összefüggést találgunk, az alaphang és a beszélő más biológiai tulajdonságai – testméretei – között alig vagy nem található kapcsolat. A szakirodalomban gyakran idézett Van Dommelen és Moxness (1995) szerzőpáros megállapítása szerint az átlagos alaphangmagasság nem korrelál a beszélő magasságával és testsúlyával. Ugyanakkor a hallgató pontosabban tudja a beszélő testalkatát megbecsülni, ha mélyebb az alaphangja. Ez különösen a férfi beszélők esetében jellemző (Gósy 2001).

Léteznek olyan, a beszélő szervezetében végbemenő biológiai változások, amelyekről feltételezhető, hogy hatással vannak a beszédképzésre. Hadman és munkatársai (2009) azt vizsgálták, hogy a terhességnek milyen hatása lehet a beszéd akusztikai sajátosságaira. Kutatásuk során a várandós nők hangját szülés előtt, illetve szülés után is megvizsgálták, továbbá kontrollcsoportként olyan nők hangmintáit használták, akik nem voltak terhesek. Az átlagos alaphangmagasságot két módszerrel is megmérték. Egyrészt, hosszan kitartott [a:] hangból mérték (F_0), másrészt egytől tízig számoltatták a kísérleti személyeket (habitual pitch). A három adatsor (kontrollcsoport, szülés előtti és szülés utáni alaphangmagasság) sehol sem mutatott szignifikáns különbséget. Ugyanakkor figyelemreméltó, hogy a kitartott [a:] hangból mért alaphangmagasság $210,1 \pm 18,1$ Hz volt a szülés előtti állapotban, szülés után pedig $218,2 \pm 25,4$ Hz. A szerzők valószínűsítik, hogy a hormonrendszer változása, illetve a hirtelen súly- és folyadékvesztés okozza ezt a különbséget.

Az átlagos alaphangmagasságot számos betegség esetében is mérik a kutatók. Különös jelentősége lehet egyes beszédszervi elváltozásokban szenvedő betegek kezelése során, ilyen például a laryngectomizált betegek rehabilitációja. A gégeeltávolítást követően többféle módja van a beszédrehabilitációnak: nyelőcsőbeszéd, hangprotézis, beszélő sipoly és külső rezgékeltők alkalmazása (Balázs 1997). Most (2000) mérései szerint szignifikánsan mélyebb a nyelőcsőbeszédet, illetve tracheoözofageális beszédet alkalmazók átlagos alaphangja az egészséges beszélőkből álló kontrollcsoportéhoz képest. Gósy (2002) ugyanakkor 185 Hz-es alaphangértéket mért nyelőcsővel beszélőknél, ami az előzőekkel összehasonlítva lényegesen magasabb érték.

A betegségek mellett az átlagos alaphangmagasság változása különböző fiziológiai hatások eredményeként is megfigyelhető. Ilyen például az alaphang mélyülését

eredményező dohányzás (Sorensen 1982), az alkoholfogyasztás következtében esetenként előforduló alaphang-emelkedés (Gyarmathy 2007) vagy a drogfogyasztás, amely egy kísérletben mélyebb alaphangot okozott (Helfrich et al. 1984).

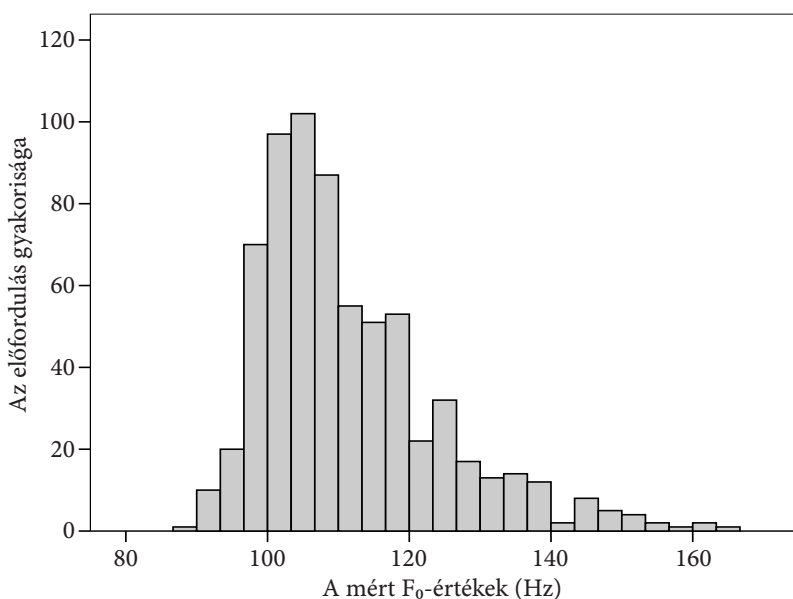
A beszélő személy különböző biológiai tulajdonságai, állapotai mellett a társalgás kontextusa is hatással lehet a beszéd átlagos alaphangmagasságára. Zarick és munkatársai (2006) hat különböző, szimulált beszédhelyzetben vizsgálták női beszélők átlagos alaphangmagasságát (nyilvános beszéd, barátához, főnökhöz, beosztotthoz, házastársához/szülőhöz intézett beszéd, hangdiagnosztika). Az egyes beszédhelyzeteknél mért frekvenciaértékek összehasonlításakor azt találták, hogy a főnökhöz szignifikánsan mélyebb ($F_0 = 195,42$ Hz, $SD = 19,45$ Hz), a beosztottakhoz szignifikánsan magasabb ($F_0 = 231,07$ Hz, $SD = 17,86$ Hz) alaphangon beszélnek a kísérleti személyek. A többi esetben 203–206 Hz adódott. Egy másik kutatás pedig azt vizsgálta (Zarick et al. 2000), hogy hét különböző feladat során (számolás 1-től 10-ig, felolvasás, 10 másodperces spontán beszéd, hosszan ejtett magánhangzó, számolás 1-től 3-ig, az [i] hangzó kitarított ejtése, hümmögés, *aha* ejtése) eltérő alaphangot használnak-e a beszélők. A méréseket a kutatók felnőtt nőknél, felnőtt férfiaknál, illetve 5–10 év közötti gyerekeknél végezték el. A felnőtt férfiaknál és a gyerekeknél az eredmények nem mutatták ki azt, hogy a feladat jellege szignifikánsan befolyásolná az átlagos alaphangmagasságot. A női csoportnál azonban az eredmények szignifikáns különbséget mutattak.

Bár a fentiekben távolról sem törekedhettünk a teljességre, megállapíthatjuk, hogy az átlagos alaphangmagasság mérése rendkívül széles körben elterjedt kutatási módszer, amelyet nem csak a fonetika, de számos más tudományterület is alkalmaz.

Az átlagos alaphangmagasság mérésével kapcsolatban ugyanakkor több módszertani kérdés is megfogalmazódik. Traunmüller és Eriksson (1995) arra hívja fel a figyelmet, hogy az F_0 eloszlása gyakran nem haranggörbe alakú, azaz nem mutat normál eloszlást (vö. még: Horii 1975; Rose 1991; Edlund–Heldner 2007). Az általában mért átlagos F_0 , illetve a szóródás mértékét kifejező standard deviáció azonban normál eloszlású adathalmazt feltételez. Problémát jelenthet tehát, ha normál eloszlást feltételező számításokat végzünk olyan adathalmazon, amelynek nem normál az eloszlása. Az alaphangmagasság-eloszlás értelmezését az 1. ábra illusztrálja.

Az 1. ábra a következőképpen értelmezendő. Egy 23 éves férfi beszélő spontán beszédéből kb. 1 perces részletet választottunk ki. A Praat 5.2 (Boersma–Weenink 2011) program által mért F_0 -értékeket átmenttük az SPSS 13 programba, majd megjelenítettük, hogy az egyes F_0 -értékek hány alkalommal fordultak elő a kb. 1 perces beszédmintában. A grafikon a vízszintes tengelyen 5 Hz-es lépésközzel ábrázolja a mért alaphangértékeket. Így az állapítható meg, hogy a 101–105 Hz közötti frekvenciaértékek valamivel több, mint 150 alkalommal fordultak elő, míg pl. a 111–115 Hz-es tartományban kb. 90 alkalommal mért a szoftver F_0 -értéket.

Az eloszlási grafikon megerősíti Traunmüller és Eriksson (1995) megállapítását, ugyanis szemmel láthatóan nem haranggörbére illeszkedő az eloszlás, hanem balra ferdülő, aszimmetrikus felépítésű. Emiatt azt várhatjuk, hogy a számtani



1. ábra. Az alaphangmagasság-értékek eloszlási grafikonja egy 23 éves férfi beszélő esetében

középértékként kiszámolt átlagos alaphangmagasság értéke nem esik egybe a grafikon maximumhelyével – móduszával –, azaz azzal az F₀-értékkel, amely a leggyakrabban fordul elő a hangmintában. Kérdésként merül fel tehát, hogy az általában használatos számtani középérték statisztikailag is eltér-e az aszimmetrikus eloszlásra kevésbé érzékeny másik két középértéktől, a módusztól és a mediántól, ha több beszélőt vizsgálunk. Hipotézisként azt állítjuk, hogy eltérést találunk, de az eltérés nem szignifikáns.

Második kutatási kérdésünk egy további kutatás-módszertani problémára irányul. Egyes kísérletekben a kutatók az átlagos alaphangmagasságot különböző hosszúságú beszédminták alapján határozzák meg (pl. Lindh 2006; Reubold et al. 2010), míg más kísérletekben izoláltan ejtett magánhangzókat használnak (pl. Collins–Missing 2003; Gelfer–Mikos 2005). Felmerül a kérdés, hogy a kétféle módszerrel mért alaphangértékek megegyeznek-e. Hipotézisünk szerint a kétféle módszer alapján mért F₀-értékek eltérnek, de a különbségük nem szignifikáns.

Itt ismertető kutatásunkban ezt a két problémát kívánjuk saját adatok elemzésével megvizsgálni.

Kísérleti személyek, anyag, módszer

A hipotézisek vizsgálatához 40 beszélőtől származó spontánbeszéd-részletet, illetve [ə]-bemondást használtunk fel. A kísérleti személyek nemi megoszlása: 20 férfi, 20 nő. Életkoruk: 19–27 év. Mindannyian a Pécsi Tudományegyetem

hallgatói. Jellemzőjük, hogy mindannyian másodgenerációs értelmiségiek, magyar anyanyelvűek, a sztenderdnek tekinthető köznyelvi nyelvváltozatot beszéltek. Egyiküknek sem volt semmilyen hangképzéssel, beszédképzéssel kapcsolatos problémája, betegsége.

A kísérleti személyek kötetlen formában saját pályaválasztásukról beszéltek, azaz a kutatáshoz spontán beszédet, ezen belül interjúhelyzetben monológot rögzítettünk (vö. Markó 2006). A beszédminták felvétele csendes helyiségben, nyugodt körülmények között történt. Zavaró háttérzaj nem volt. A monológokból minden esetben egy 50–80 s hosszúságú szakaszt használtunk fel a vizsgálathoz. A monológokat követően minden egyes kísérleti személytől felvettünk egy-egy kb. 3–4 másodpercig kitartott [ə] hangot.

A beszédminták rögzítéséhez Fujitsu Siemens Lifebook C típusú laptopot alkalmaztunk. A hangfelvételekhez AKG D 60 típusú mikrofont használtuk. A hangmintákat digitális formában, a számítógép vincseszterén, .wav formátumban rögzítettük, 16 bites, 20 kHz-es mintavételezéssel. Ehhez, illetve a későbbiekben az akusztikai paraméterek méréséhez a már említett Praat 5.2 programot használtuk. Többféle algoritmus ismeretes, amely az alaphang mérését szolgálja. A Praat esetében az autokorrelációs algoritmust állítottuk be. Minden mérési paramétert a program által felkínált alapértéken tartottunk, kivéve a mérés tartományát. Ezt – tekintettel az alábbiakban ismertető hibákra – a 20–500 Hz-es tartományra állítottuk be.

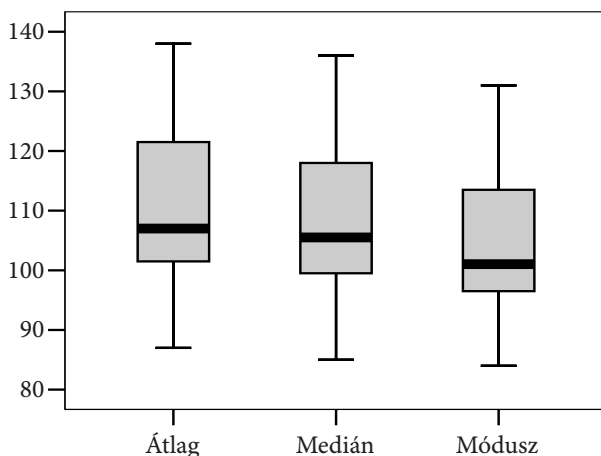
A Praat programból az egyes beszédmintákban mért F_0 -sorozatokat átemeltük az SPSS programba. A statisztikai számításokat megelőzően adattisztítást kellett végezni. Ennek az egyik oka az volt, hogy a Praat néhány esetben réshangokhoz, levegővételhez, illetve egyéb zajokhoz – tévesen – F_0 -értékeket rendelt. Bár nem nagy számban jelentkeztek ezek a téves F_0 -értékek, az adatállományból törölni kellett őket, mivel a statisztikai számításokat megzavarták volna. Hasonlóképpen törölni kellett az érdes zöngéhez rendelt, a normálisnál lényegesen mélyebb F_0 -értékeket is. Beke (2008) is megemlíti, hogy ezek az alaphangértékek nem tartoznak a beszélő hangterjedelméhez. Hasonló módon járt el Lindh (2006), illetve tett javaslatot erre a módszerre Schötz (2004). Az így előkészített adatállományok már alkalmasak voltak a statisztikai számításokra. A számítások során minden esetben 95%-os konfidenciaintervallumot alkalmaztunk.

Az 1. kísérlettel arra a kérdésre kerestük a választ, hogy a spontánbeszéd-mintákból mért alaphangok három középértéke, tehát a számtani közepe (átlaga), módusza és mediánja között mutatkozik-e szignifikáns eltérés. A 2. kísérlettel azt vizsgáljuk, hogy az izoláltan ejtett [ə] hang alaphangja ugyanazon beszélők spontán beszédéből mért alaphang-középértékekhez hogyan viszonyul. Ennek megállapításához kiszámítjuk az egyes beszélőknél kapott [ə] hangok alaphangja, illetve az egyes középértékek közötti korrelációs együtthatókat, másrészt pedig párosított t -próbákkal megállapítjuk, hogy szignifikáns különbség mutatkozik-e az [ə] hangok, illetve a középértékek között.

Eredmények

1. kísérlet

A kutatásban szereplő férfi beszélőkre vonatkozóan a kapott adatokat a 2. ábra illusztrálja.



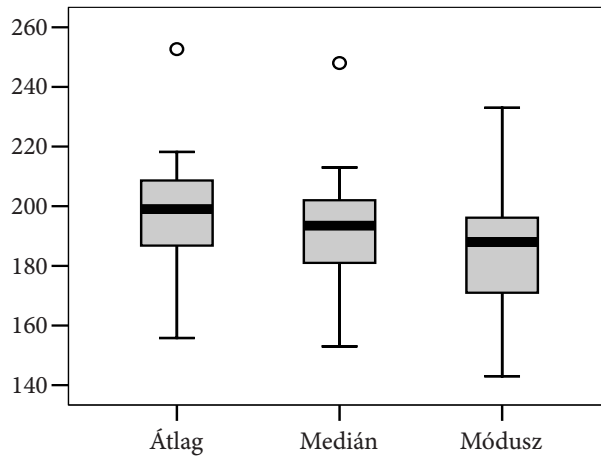
2. ábra. A férfi beszélőknél kapott középértékadatok

A 2. ábra az alábbiak szerint értelmezendő. A vízszintes tengelyen az egyes középértéktípusok találhatók. Mindegyikhez egy-egy függőleges alakzat tartozik. Ezek alsó és felső határoló vonala a 20 beszélőnél tapasztalt legmélyebb, illetve legmagasabb mért átlagot, mediánt, illetve móduzt jelzi. Az átlag két szélsőértéke: 87 és 138 Hz, a mediáné 85 és 136 Hz, a móduzé pedig 84 és 131 Hz. A vízszintes vastag vonal mindhárom esetben az adott adathalmaz mediánját (középső értékét) jelzi, ami az átlagnál 107, a mediánnál 105, a móduzsnál pedig 101 Hz volt. A függőlegesen, téglalappal ábrázolt tartomány az adathalmaz középső két kvartilisét mutatja, azaz tíz válaszadó adatai estek ebbe a tartományba. A téglalap alatt, illetve fölött pedig a határoló vonalig bezáróan az alsó és a felső kvartilis található, ami a jelen esetben a legmélyebb, illetve legmagasabb alaphangon beszélő öt-öt férfit jelenti.

Az ábrából és a számadatokból különbségek olvashatók ki. Azt, hogy az eltérés szignifikáns-e, az összetartozó mintás varianciaanalízis (Repeated Measures ANOVA) segítségével határozzuk meg. Ez a számítási módszer akkor alkalmazható, ha ugyanazon a mintán különböző méréseket végzünk, és ezek eredményeit hasonlítjuk össze. A szórássegyezés megállapításához szfericitástesztet kell végeznünk (Test of Mauchly's sphericity). Ennek eredménye $W(2) = 0,194$; $p < 0,001$, emiatt az összetartozó mintás varianciaanalízis Greenhouse–Geisser-féle változatát vesszük figyelembe. Ennek

eredménye a következő: $F(1,107; 21,039) = 78,196$; $p < 0,001$. Ez az adat arra utal, hogy a három középérték-számítás által kapott értékek szignifikánsan különböznek egymástól. Az átlagos alaphangmagasságok átlagértéke 111,2 Hz, a mediánoké 108,9 Hz, a móduszoké pedig 105,2 Hz. A páronkénti összehasonlítás $p < 0,001$ szinten szignifikáns különbséget mutat minden esetben, azaz a három középérték-számítás szignifikánsan eltérő eredményt ad.

Ugyanezt a számítást a női beszélők esetében is elvégeztük. Az adatokat a 3. ábra összesíti.

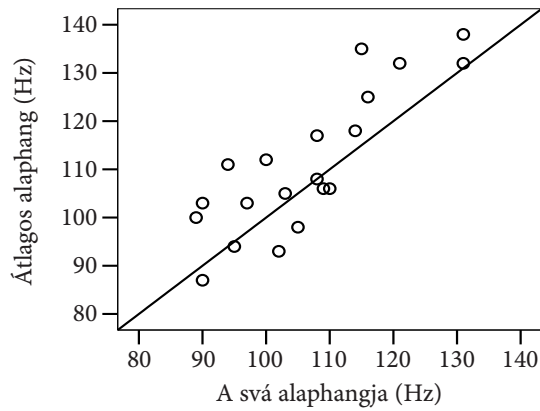


3. ábra. A női beszélőknél kapott középértékadatok

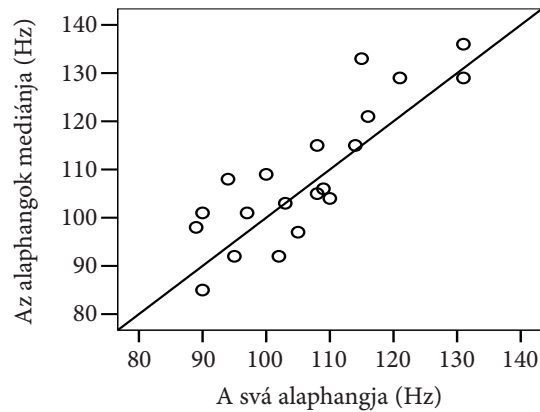
A grafikonon ábrázolt minimum- és maximumértékek a következők: az átlagértéknél 156 és 252 Hz, a mediánnál 153 és 248 Hz, a módusznál pedig 143 és 233 Hz. A három középérték mediánja pedig az alábbiak szerint alakult: az átlagoké 199 Hz, a mediánoké 194 Hz, a móduszoké pedig 188 Hz. Az összetartozó mintás varianciaanalízis eredménye a következő. A sfericitásra szignifikáns érték adódott, $W(2) = 0,492$, $p < 0,001$. Emiatt a Greenhouse–Geisser-féle korrekciót vesszük figyelembe, ennek eredménye: $F(1,327; 25,205) = 104,107$; $p < 0,001$. Ezek az adatok arra utalnak, hogy a grafikonról leolvasható különbségek szignifikánsak. Az átlagos alaphangok átlaga 198,5 Hz, a mediánoké 193,2 Hz, a móduszoké pedig 186,6 Hz. Az adatsorok páronkénti összehasonlítása minden esetben szignifikáns ($p < 0,001$ szinten) különbséget mutat.

2. kísérlet

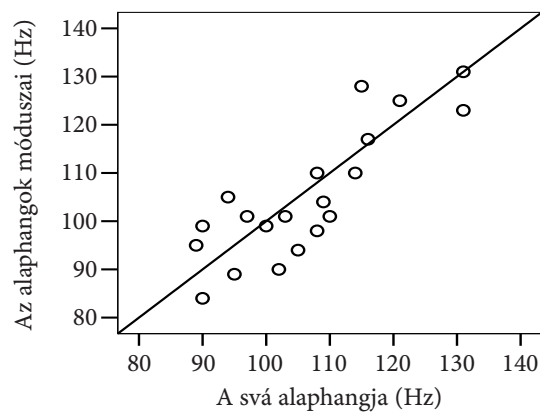
A 4., 5. és 6. ábrákon látható, hogy a férfi beszélőknél az [ə] hangok ugyanazon beszélők esetében hogyan függenek össze a spontán beszédből mért alaphang-középértékekkel.



4. ábra. Az [ə] hang és az átlagos alaphangmagasság kapcsolata a férfi beszélőknél



5. ábra. Az [ə] hang és a mediánok kapcsolata a férfi beszélőknél



6. ábra. Az [ə] hang és a móduszok kapcsolata a férfi beszélőknél

A páros t -próba segítségével megállapítható, hogy szignifikáns különbség mutatkozik-e a grafikonok vízszintes és függőleges tengelyein ábrázolt alaphangértékek között. A t -próbák számítási eredményeit az 1. táblázat foglalja össze (minden esetben $n=20$, $df=19$).

1. táblázat. Az alaphangértékek összehasonlítása t -próbával, férfi beszélőknél

Pár	t	p
[ə] vs. F_0 -átlag	2,645	0,016
[ə] vs. F_0 -medián	1,493	0,152
[ə] vs. F_0 -módusz	-0,724	0,478

Az 1. táblázatból leolvasható, hogy a férfiak esetében a spontán beszédből mért átlagos alaphangmagasság-értékek (átlaguk: 111,2 Hz) szignifikánsan különböztek az izolált [ə] hangok alaphangértékeitől (átlaguk: 106,4 Hz). A [ə] hangok alaphangját összehasonlítva a spontán beszédből mért mediánokkal (átlaguk: 109,0 Hz) és móduszokkal (átlaguk: 105,2 Hz) nem adódott ilyen különbség.

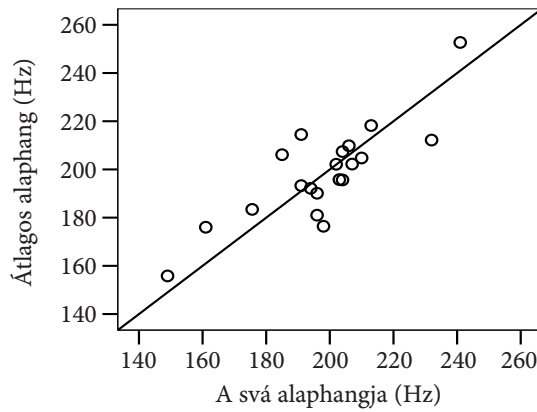
A női beszélőknél a [ə] hangok és a spontán beszédből mért alaphang-középtértékek közötti összefüggéseket a 7., 8. és 9. ábrák szemléltetik.

A korrelációs számítás a következő eredményt adta. Az [ə]-átlagérték páros esetében $r=0,823$; $p < 0,001$. Az [ə]-medián esetében $r=0,836$; $p < 0,001$, az [ə]-módusz páros esetében pedig $r=0,851$; $p < 0,001$. De ez az eredmény – hasonlóan a férfiaknál kapott adatokhoz – nem meglepő, előre feltételezhető, hogy azok a női beszélők, akik mélyebb alaphangon ejtik az [ə] hangot, a spontán beszédükben is mélyebb középtértékkel beszélnek. Ugyanakkor érdemes itt is megvizsgálni, hogy az adatpontok az $y=x$ tengelyhez viszonyítva hol helyezkednek el. A legfeltűnőbb aszimmetria a 9. ábrán az [ə] hangok és a spontán beszéd alaphangértékeinek móduszainál figyelhető meg. Látható, hogy a pontok jelentős része az egyenes alatti felsíkban helyezkedik el. Ez azt jelenti, hogy ugyanazon beszélők magasabb alaphanggal ejtették ki az [ə] hangot, mint amilyen a spontán beszédük alaphangjának módusza volt. A 8. ábrán már kisebb mértékű az adatpontok eloszlása közötti különbség a két felsíkban, a 7. ábrán pedig az látható, hogy csaknem azonos mennyiségű pont helyezkedik el az egyenes alatt és fölött.

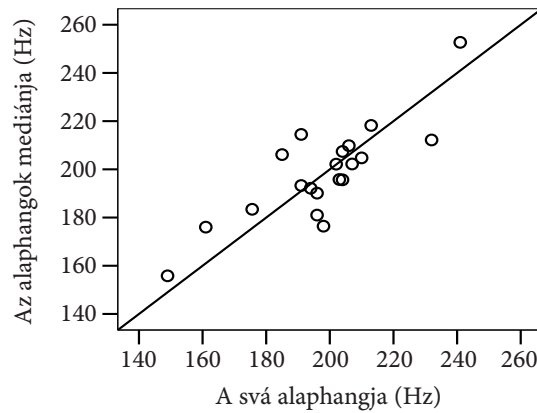
A páros t -próbák eredményeit a 2. táblázat foglalja össze (minden esetben $n=20$, $df=19$).

2. táblázat. Az alaphangértékek összehasonlítása t -próbával, női beszélőknél

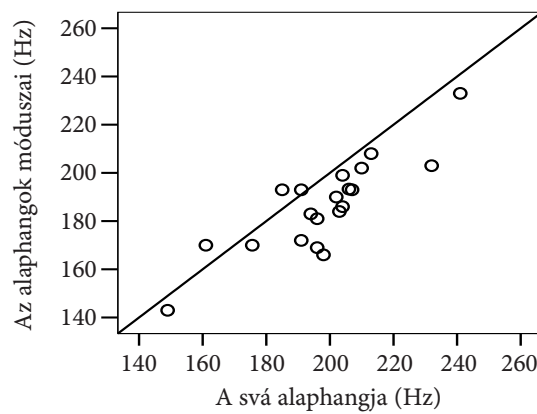
Pár	t	p
[ə] vs. F_0 -átlag	-0,206	0,839
[ə] vs. F_0 -medián	-1,830	0,083
[ə] vs. F_0 -módusz	-4,621	<0,001



7. ábra. Az [ə] hang és az átlagos alaphangmagasság kapcsolata a női beszélőknél



8. ábra. Az [ə] hang és a mediánok kapcsolata a női beszélőknél



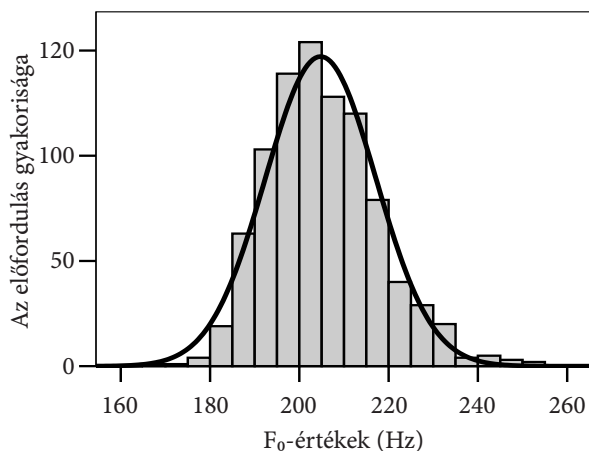
9. ábra. Az [ə] hang és a móduszok kapcsolata a női beszélőknél

A 2. táblázatból leolvasható, hogy az [ə] hangok alaphangmagassága (átlagértékük: 197,0 Hz) nem különbözött szignifikánsan az átlagos alaphangmagasságok átlagértékétől (198,4 Hz), és az alaphangértékek mediánjainak átlagértékeitől (193,2 Hz). Ugyanakkor nagyon jelentős, szignifikáns különbség mutatkozott az alaphangértékek móduszaival való összehasonlítás során (átlagértékük: 186,6 Hz).

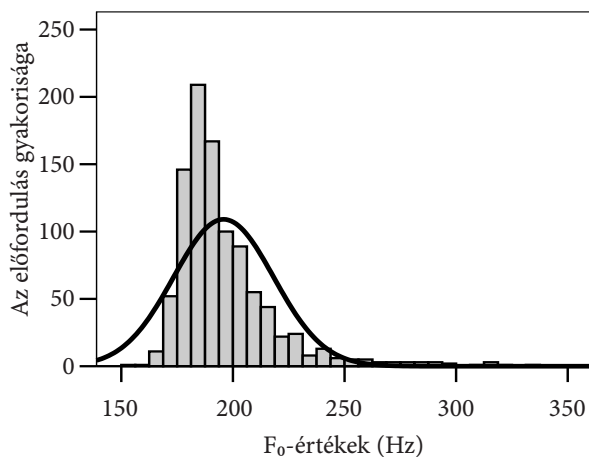
Következtetések

Az első kísérlethez kapcsolódó hipotézisünk szerint a különböző módszerekkel kiszámított középértékek eltérnek, de az eltérés nem szignifikáns. Hipotézisünk anynyiban igazolódott, hogy valóban eltérnek az említett értékek, de az eltérés szignifikáns mindkét nemnél. Mindkét esetben a legmélyebb a módusz, ennél magasabb a medián, a legmagasabb pedig az átlagos alaphang. Ennek az eredménynek a jelentőségét két esetre vonatkozóan célszerű értelmezni. Amennyiben az egyes beszélők alaphangmagassága közötti sorrend a lényeges, a három középérték közötti szignifikáns eltérés ellenére nem valószínű, hogy más eredményt adna a számítás, ha az átlagos alaphangmagasságot, a mediánt vagy a móduszt használjuk. Ilyen jellegű vizsgálat például az, amikor az alaphangmagasság és az életkor korrelációját vizsgáljuk. Ha például megnézzük a férfi beszélők átlagos alaphangmagassága, illetve módusza szerinti sorrendjét, a Spearman-féle korrelációs együttható igen erős összefüggést mutat ($\rho=0,937$; $p < 0,001$). Azaz, ha egy kutatás során egy grafikon vízszintes tengelyén a beszélő életkorát, függőleges tengelyén pedig az alaphangmagasságát ábrázoljuk, a beszélők függőleges tengely menti sorrendje nem fog szignifikánsan különbözni, akár az átlagos alaphangmagasságot, akár a móduszt használjuk. Így arra számíthatunk, hogy akármelyik középértéket választjuk, a számítás az életkorral való összefüggésre ugyanazt vagy közel ugyanazt az eredményt fogja adni.

Ugyanakkor abban az esetben, ha jelentősége van az egyes középértékek mögötti precíz akusztikai tartalomnak, akkor megfontolandó az átlagos alaphangmagasság helyett például a módusz használata. A számtani középérték és a mellette közzölt SD érték ugyanis minden esetben normál eloszlást feltételez. A 10. ábrán látható az egyik kísérleti személy alaphang-eloszlási görbéje. Megfigyelhető, hogy az alaphangértékek eloszlása jól illeszkedik a haranggörbére. Az ő esetében az átlagos alaphangmagasság 204,0 Hz, az alaphang módusza pedig 202,0 Hz volt. Kísérletünkben azonban csak néhány ilyen beszélőt találtunk, lényegesen gyakoribb volt a 11. ábrán láthatóhoz hasonló eloszlás. Ennél a beszélőnél az átlagos alaphangmagasság 195,4 Hz, a módusz pedig 184,0 Hz, tehát 11,4 Hz a különbség. Látható, hogy az egyes frekvenciaértékek előfordulási gyakoriságát jelző oszlopok aszimmetrikus elrendeződést mutatnak, és nem illeszkednek jól a haranggörbére. Ennek ellenére könnyen kiszámolható az átlagos alaphangmagasság és a hozzá kapcsolódó standard deviáció is, azonban ezek használata esetén hallgatólagosan feltételezzük az



10. ábra. Haranggörbére illeszkedő, szimmetrikus alaphang-eloszlási görbe egy női beszélőnél



11. ábra. Haranggörbére nem illeszkedő, aszimmetrikus alaphang-eloszlási görbe egy női beszélőnél

ábrán látható haranggörbe szerinti eloszlást – ami azonban nem felel meg a tényleges alaphangeloszlásnak. Előadódhat olyan kutatási helyzet, hogy egy percepciósi kísérlethez alkalmazott két beszélő hangjának azonos az átlagos alaphangmagassága és a szóródása, azonban csak az egyikük eloszlási grafikonja illeszkedik a haranggörbére, a másikuk esetében erősen aszimmetrikus az eloszlás. Emiatt a hallgatók akusztikai élménye a két esetben nagymértékben különbözni fog, annak ellenére, hogy az alaphangjukhoz kiszámított és az adatfeldolgozás során felhasznált két átlagérték és szóródás megegyezik.

Ha az F_0 -értékek eloszlása haranggörbére illeszkedik, akkor a három középérték egybeesik. Az, hogy kutatásunkban szignifikánsan eltért a három középérték, arra utal, hogy a vizsgált 40 beszélőre általában az aszimmetrikus eloszlás a jellemző. Minél nagyobb mértékű az alaphangmagasság eloszlási grafikonjának az aszimmetriája, annál kevésbé jellemzi jól a haranggörbe, az átlag és a standard deviáció az eloszlást. A probléma megoldásaként mindenképpen megfontolásra érdemesnek tartjuk a módusz és a medián mint középértékek használatát, a standard deviáció mint szóródási mutató helyett pedig az aszimmetriát leíró statisztikai adatok, pl. a csúcsosság és a ferdeség alkalmazását.

Második kísérletünkben arra kerestük a választ, hogy az izoláltan ejtett [ə] hangok alaphangja milyen kapcsolatban áll az ugyanazon beszélők spontán beszédéből mért alaphang-középértékekkel. A férfiak esetében az [ə] hangok alaphangja nem tért el szignifikánsan a spontán beszéd alaphangjának móduszától és mediánjától, csak az átlagértékétől. Ez azt jelenti, hogy a kísérletben szereplő férfi beszélők az [ə] kiejtésekor a spontán beszédükben általában is leggyakrabban előforduló F_0 -értékeket reprodukálták. A női beszélőknél viszont megfigyelhető, hogy az [ə] hangok alaphangja leginkább a spontán beszéd átlagos alaphangjához állt közel, és szignifikánsan magasabb volt, mint a spontán beszédben leggyakrabban előforduló F_0 -érték. Ez a különbség arra világít rá, hogy még az izolált hangzók ejtésében is különbség van a férfiak és a nők beszédstratégiájában. A nők magasabb ejtésű [ə] hangja hátterében számos ok lehet. Elképzelhető például az erőteljesebb felhívó hatás szándéka, az izolált ejtés során koncentráltabb figyelem, de a kedvezőbb, nőiesebb benyomás kialakítása iránti mögöttes motiváció is oka lehet.

Mindkét probléma további kutatását indokoltnak tartjuk. A különféle alaphang-eloszlások ugyanúgy egyedi sajátosságai lehetnek egy-egy beszélőnek, mint például a hangszínezete. Célszerűnek látjuk továbbá korábbi, alaphangmérést tartalmazó kísérletek megismétlését az alaphangeloszlás aszimmetriájának figyelembevételével, így az eddigieknél árnyaltabb képet kaphatunk az alaphang viselkedéséről és percepciójáról. Az izolált hangzók esetében pedig további kérdés, hogy az izolált hangzókkal végzett kísérletek eredményei milyen szempontból általánosíthatók, milyen körülmények között, hogyan helyettesíthetik, illetve reprezentálhatják a spontán beszéddel végzett kísérletek eredményeit, különös tekintettel a különböző nemű beszélőkre és hallgatókra.

Irodalom

- Balázs Boglárka 1997. Hangterápia nyaki műtétek után. *Beszédgyógyítás* 97/2–3. 7–19.
 Beke András 2008. A felolvasás és a spontán beszéd alaphangszerkezeteinek vizsgálata. *Beszéd-kutatás* 2008. 93–107.
 Boersma, Paul – Weenink, David 2011. Praat: Doing phonetics by computer. Version 5.2. <http://www.fon.hum.uva.nl/praat/>

- Borkowska, Barbara – Pawlowski, Boguslaw 2011. Female voice frequency in the context of dominance and attractiveness. *Animal Behavior* 82. 55–59.
- Collins, Sarah A. – Missing, Caroline 2003. Vocal and visual attractiveness are related in women. *Animal Behavior* 65. 997–1004.
- van Dommelen, Wim A – Moxness, Bente H. 1995. Acoustic parameters in speaker height and weight identification: sex-specific behaviour. *Language and Speech* 38. 267–287.
- Edlund, Jens – Heldner, Mattias 2007. Underpinning /nailon/: Automatic estimation of pitch range and speaker relative pitch. In Müller, Christian (ed): *Speaker classification II*. Springer-Verlag, Berlin–Heidelberg, 229–242.
- Ferrand, Carole T. 2002. Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice* 16. 480–487.
- Gelfer, Marylou Pausewang – Mikos, Victoria A. 2005. The relative contributions of speaking fundamental frequency and formant frequencies to gender identification based on isolated vowels. *Journal of Voice* 19. 544–554.
- Gocsál Ákos – Huszár Ágnes 2003. Csábító hangok. *Beszédkutatás* 2003. 9–18.
- Gósy Mária 1996. A beszéd akusztikai szerkezetének állandóságáról. In Terts István (szerk.): *Nyelv, nyelvész, társadalom. Emlékkönyv Szépe György 65. születésnapjára barátaitól, kollégáitól, tanítványaitól II*. Keraban Könyvkiadó – JPTE, Pécs, 66–75.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125. 137–148.
- Gósy Mária 2002. Beszédképzés zöngé nélkül. *Beszédkutatás* 2008. 18–37.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Guimaraes, Isabel – Abberton, Evelyn 2005. Fundamental frequency in speakers of Portuguese for different voice samples. *Journal of Voice* 19. 592–606.
- Gyarmathy Dorottya 2007. Az alkohol hatása a spontán beszédprodukcóra. *Beszédkutatás* 2007. 108–120.
- Hadman, Abdul-Latif – Mahfoud, Lorice – Sibai, Abla – Seoud, Muheiddine 2009. Effect of pregnancy on the speaking voice. *Journal of Voice* 23. 490–493.
- Harnsberger, James D. – Shrivastav, Rahul – Brown, William S. Jr. – Rothman, Howard – Hollien, Harry 2008. Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice* 22. 58–62.
- Horii, Yoshiyuki 1975. Some statistical characteristics of voice fundamental frequency. *Journal of Speech and Hearing Research* 18. 192–201.
- Laver, John 2003. Three semiotic layers of spoken communication. *Journal of Phonetics* 31. 413–415.
- Lindh, Jonas 2006. Preliminary F₀ statistics and forensic phonetics. In: *Proceedings of IAFPA 2006*. http://www.ling.gu.se/konferenser/iafpa2006/Abstracts/Lindh__IAFPA2006.pdf (A letöltés ideje: 2010. május 1.)
- Markó Alexandra 2006. A spontán beszéd – monologikus és dialogikus szövegek. *Acta Academiae Paedagogicae Agriensis* XXXIII. 66–78.
- Most, Tova – Tobin, Yishai – Mimran, Ravit Cohen 2000. Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production. *Journal of Communication Disorders* 33. 165–181.
- Reubold, Ulrich – Harrington, Jonathan – Kleber, Felicitas 2010. Vocal aging effects on F₀ and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52. 638–651.
- Rose, Phil 1991. How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication* 10. 229–247.
- Schötz, Susanne 2004. The role of F₀ and duration in perception of female and male speaker age. In: *Proceedings of Speech Prosody 2004*. Nara, Japan, 23–26 March 2004. 379–382.

- Sorensen, David 1982. Cigarette smoking and voice fundamental frequency. *Journal of Communication Disorders* 15. 135–144.
- Traunmüller, Hartmut – Eriksson, Anders 1995. The frequency range of the voice fundamental in the speech of male and female adults. http://www2.ling.su.se/staff/hartmut/fo_m&f.pdf (A letöltés ideje: 2011. október 26.)
- Trittin, Pamela Jean – de Santos y Lleó, Andrés 1995. Voice quality analysis of male and female Spanish speakers. *Speech Communication* 16. 359–368.
- Vukovic, Jovana – Jones, Benedict C. – Feinberg, David R. – DeBruine, Lisa M. – Smith, Finlay G. – Welling, Lisa L. M. – Little, Anthony C. 2010. Variation in perceptions of physical dominance and trustworthiness predicts individual differences in the effect of relationship context on women's preferences for masculine pitch in men's voices. *British Journal of Psychology* 102. 37–48.
- Zarick, Richard I. – Gentry, Mollie A. – Smith-Olinde, Laura – Gregg, Brent A. 2006. The effect of speaking context on elicitation of habitual pitch. *Journal of Voice* 20. 545–554.
- Zarick, Richard I. – Skaggs, Sheri D. – Montague, James C. 2000. The effect of task determination of habitual pitch. *Journal of Voice* 14. 484–489.