

GRÁCZI TEKLA ETELKA<sup>1,2</sup> – KREPSZ VALÉRIA<sup>1</sup> – MARKÓ ALEXANDRA<sup>2,3</sup> – HUSZÁR ANNA<sup>1</sup> – SZÁRAZ BETTINA<sup>1</sup>

<sup>1</sup>Nyelvtudományi Intézet, <sup>2</sup>MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport, <sup>3</sup>ELTE BTK  
Alkalmazott Nyelvészeti és Fonetikai Tanszék  
graczi.tekla.etelka@nytud.hu, krepsz.valeria@nytud.hu, marko.alexandra@btk.elte.hu,  
huszar.anna@nytud.hu, szaraz.bettina@nytud.hu

Gráci Tekla Etelka – Krepsz Valéria – Markó Alexandra – Huszár Anna – Száraz Bettina: Az f<sub>0</sub>-  
jellemzők felolvasásban és spontán beszédben  
Alkalmazott Nyelvtudomány, XIX. évfolyam, 2019/2. szám  
doi:<http://dx.doi.org/10.18460/ANY.2019.2.003>

## Az f<sub>0</sub>-jellemzők felolvasásban és spontán beszédben

A large number of studies investigated the differences in f<sub>0</sub> characteristics between reading aloud (RA) and spontaneous speech (SpS) in various languages. Their basic assumption was that the different production strategies lead to difference in the prosodic features, however, their results were not consistent as to which speech style was realized with higher mean f<sub>0</sub> or a larger f<sub>0</sub> range. Hungarian data have been only analyzed on small numbers of speakers. Therefore our goals are: (i) to provide a large sample (82 subjects) based comparison of the f<sub>0</sub> characteristics of RA and SpS in Hungarian, and (ii) to analyze the individual differences behind general tendencies.

Mean f<sub>0</sub> and pitch range (of the interpausal units) were higher in RA, while the f<sub>0</sub> range in SpS. The interspeaker differences played an important role in the mean f<sub>0</sub> results: no speech style characteristic difference was found in women, while this was apparent in men.

### Bevezetés

A felolvasás és a spontán beszéd alaphfrekvencia-jellemzőit több nyelvben is összehasonlították korábban, ám az eredmények gyakran ellentmondók az egyes tanulmányok között (Daly-Zue, 1992 és Skarnitzl-Vaňková, 2017). A magyar nyelv esetében Beke (2008) és Markó és Bóna (2010) végzett vizsgálatot, 5, illetve 14 fővel. A jelen tanulmányban az alaphfrekvencia átlagát, szórását, a hangterjedelmet és a hangközt elemeztük. Hangterjedelem alatt a teljes beszédidő alatti, míg hangköz alatt az egyes beszédszakaszok (két szomszédos szünet közötti beszéddel kitöltött egység) legalacsonyabb és legmagasabb f<sub>0</sub>-értékének eltérését értjük. A tanulmány 82 magyar kísérleti személy beszédmintáját elemzi annak érdekében, hogy a két beszédstílus egyes alaphfrekvencia-jellemzőit a beszélők közötti variabilitás tekintetbe vételével tárja fel magyar beszédben.

A beszéd alaphfrekvencia-jellemzőit számos tényező befolyásolja, amelyeket nem lehet figyelmen kívül hagyni a beszédstílusok vizsgálata során. Többek között a beszélők neme, életkora (pl. Markó–Bóna, 2010), a nyelv és a nyelvjárás (de Silva et al., 2003) is meghatározhatja azt.

A felolvasás és a spontán beszéd két olyan beszédstílus, amelyeket a beszéd tervezési és kivitelezési sajátosságai mentén egy képzeletbeli skála két ellentétes pontjának lehet tekinteni (Eskenazi, 1993), ha a további tényezőket (például címzett, beszélői attitűd) állandónak tekintjük. Korábbi kutatások alátámasztották (Tøndering, 2011, Obin, 2011), hogy a prozódiai tervezés különbözik a beszédstílusok között, ami azonban a szupragmentális jellemzők alakulására is hatással van. Az egyes kutatások azonban eltérő eredményeket mutattak. Felmerül a kérdés, hogy esetleg az egyes nyelvek közötti eltérés vagy más ok (is) áll ennek hátterében.

Daly és Zue (1992) az ún. Voyager corpusban szereplő 44 férfi és 45 nő adatközlő mintáján azt találták, hogy az átlagos alapfrekvencia ( $f_0$ ) magasabb volt a spontán beszédben, mint a felolvasásban, ugyanakkor a szórásban és a hangköz értékeiben nem találtak szignifikáns különbséget a két beszédstílus között. Swerts és kollégái (1996) magasabb átlagos  $f_0$ -értéket találtak felolvasásban, mint spontán beszédben egy svéd férfi megnyilatkozásainak vizsgálata alapján. Skarnitzl és Vaňková (2017) pedig 100 cseh férfi beszédmintái alapján azt találták, hogy az átlagos  $f_0$ , valamint a szórás magasabb a felolvasásban, mint a spontán beszédben. Abu-Al-Makarem és Petrosino (2007) egy táblázatot mutatott be, amely összefoglalja a korábbi vonatkozó szakirodalmak eredményeit. Összesen 11 vizsgálatot tartalmaz, amelyek összehasonlították a felolvasás és a spontán beszéd sajátosságait az alapfrekvencia átlaga és szórása mentén. A 11 vizsgálat közül hat magasabb  $f_0$ -átlagot talált a felolvasásban, mint a spontán beszédben. A többi idézett tanulmány 5 Hz-nél kisebb különbséget adatolt a két beszédstílus  $f_0$ -átlaga között (az átlagos  $f_0$  értéke négyben a felolvasás esetében volt nagyobb, egy esetben pedig a spontán beszédben). De Silva és munkatársai (2003) három nyelvben hasonlították össze a spontán beszédben és a felolvasásban mért  $f_0$ -értékeket. Ők kizárólag a szonor beszédhangokban végeztek méréseket a többi itt bemutatott tanulmánytól eltérően (akik ilyen megkötést nem tettek). Az eredmények a következő tendenciaszerű különbségeket mutatták (szignifikáns eltérés sehol nem igazolódott a vizsgált szempontok szerint): hét holland beszélő közül hatnál magasabb volt az átlagos  $f_0$  a spontán beszédben, míg egynél azonos értékeket találtak a két beszédstílusban. A hat finn adatközlő közül háromnál nem volt szignifikáns különbség az átlagos  $f_0$ -ban a két beszédstílus között, míg a másik adatközlő esetében a felolvasásban, kettőnél pedig a spontán beszédben volt magasabb az  $f_0$  értéke, egy beszélőnél nem volt kimutatható különbség. Tanulmányaikban az öt orosz beszélő közül kettőnél magasabb volt az átlagos  $f_0$  felolvasásban, a másikban kettőben az spontán beszédben, és egynek nem volt szignifikáns különbsége a két beszédstílus között. A tendenciák az olaszul beszélők körében is különböztek (Caldognetto et al., 1997). Abu-Al-Makarem és

Petrosino (2007) 15 férfi beszédét elemezve magasabb értéket talált a felolvasásban, mint a spontán beszédben arab és angol beszédben. Az eltérés ugyan alacsony az átlagok között (1 Hz, ill. 4 Hz), de az ismételt méréses ANOVA alapján szignifikáns. A két nyelv között nem volt szignifikáns eltérés. Hudson és Holbrook (1982) 100 férfi és 100 nő beszélő mintáit vizsgálták. Magasabb átlagos f<sub>0</sub>-értéket találtak a felolvasásban, ugyanakkor a két beszédstílus közötti átlagos f<sub>0</sub>-különbség csupán 2-4 Hz volt.

A hangterjedelem esetében is eltérő eredmények születtek. A spontán beszédben találták szélesebbnek a hangterjedelmet Caldognetto és kollégái (1997: 3 olasz beszélő), Hudson és Holbrook (1982: 100 férfi és 100 nő, angol beszéd), míg Skarnitzl és Vaňková a felolvasásban (2017: 100 cseh férfi). Tønderin (2011) a hangközt 15 fiatal és idős dán kísérleti személy beszédében vizsgálta, és azt találta, hogy a fiatalok a felolvasásban beszélnek szélesebb f<sub>0</sub>-tartományban, míg az idősek a spontán beszédben.

A felolvasás és a spontán beszéd közötti alaphfrekvencia-különbségeket magyar beszédben is vizsgálták. Markó és Bóna (2010) a felolvasás és a spontán beszéd f<sub>0</sub>-jellemzőit hét fiatal és hét idős női beszélőben elemezte. A tanulmány 1. ábráján (Markó–Bóna, 2010: 46) az látható, hogy a f<sub>0</sub> mediánértéke alacsonyabb volt három fiatal és két idős adatközlőnél spontán beszédben, egy-egy fő esetében pedig magasabb, mint felolvasásban. Eredményeik összességében azt mutatták, hogy az f<sub>0</sub>-tartomány (hangterjedelem) nagyobb volt a felolvasásban a fiatal beszélők esetében, de a beszélők közötti variabilitás nagyobb volt az idősebbek körében. A hangköz tekintetében azt találták, hogy az mindkét életkori csoportban magasabb volt a felolvasásban, mint spontán beszédben. A korcsoportok között nem volt szignifikáns különbség ezekben az értékekben a felolvasásban, míg spontán beszédben igen. Egy másik tanulmány (Beke, 2008) szerint két férfi és három női beszélő mintájában a két beszédstílus között a hangterjedelemben nem volt szignifikáns eltérés, míg a leggyakrabban használt tartományok, melyek az adatközlő f<sub>0</sub>-adatainak legalább 70%-át teszik ki (Beke, 2008: 95–96), a felolvasásban nagyobb volt a nők esetében.

Amint az előző adatokból kitűnik, az alaphfrekvencia-jellemzők beszédstílusfüggő különbségei nagy eltéréseket mutathatnak az egyes tanulmányok között. Számos tanulmány viszonylag kevés adatközlőt vizsgált, valamint különféle nyelveket elemeztek, és az olvasott szövegek és a spontán beszéd rögzítésének módja is a legtöbb esetben szükségszerűen különbözött a vizsgálatok során. Meg kell azonban jegyezni, hogy sok nemzetközi tanulmány a „Rainbow passage”-t (Fairbanks 1960) vagy annak egy részét használta, és hogy a két magyar tanulmány, Beke (2008), valamint Markó és Bóna (2010), ugyanazon adatbázis ugyanazon beszédfeladatain

végezte el vizsgálatát. Továbbá, míg egyes szerzők részletesen bemutatták módszereiket (pl. Mennen et al., 2007), mások, különösen azok, amelyek automatikus méréseket alkalmaztak, nem voltak részletesen kifejtettek (például Beke, 2008). Ezért számos esetben nem ismert például, hogy milyen tartományban mérték alapfrekvenciát, milyen további beállításokat alkalmaztak, vagy hogy alkalmaztak-e valamilyen korrekciós módszert például az irregulárisan képzett részek és az esetleges oktávugrás torzításának csökkentésére, és ha igen, milyen.

Felmerülhet a kérdés – főként azokat a forrásokat elemezve, amelyek esetében bemutatták a beszélők között eltérő eredményeket is, hogy nem csak a nyelv/kultúra közötti, de a beszélők közötti eltérések is jelentősek. A nemzetközi irodalom ellentmondásos eredményei és a magyar nyelvre található kevés adatközlővel készült tanulmányok miatt a jelen tanulmány célja, hogy megvizsgálja a felolvasás és a spontán beszéd alapfrekvencia-jellemzőinek különbségét nagyszámú magyar beszélő beszédében, hogy felfedje a magyar beszédben a két stílus közötti eltérések általános tendenciáit a beszélők közötti eltérések figyelembevételével.

Az eddigi irodalomban tehát alapvetően ellentmondó eredmények születtek a két beszédstílus  $f_0$ -jellemzőire. Mindezek fényében azt a hipotézist állítottuk fel, hogy i) a felolvasott és a spontán beszéd között nem található nagymértékű eltérés az átlagos  $f_0$ -értékben, illetve hogy ii) mind az  $f_0$ -átlag, mind a hangköz és a hangterjedelem a két beszédstílus közötti eltérése jelentős egyéni eltéréseket mutat.

### **Kísérleti személyek, anyag és módszer**

40 férfit és 42 nőt választottunk ki a BEA magyar beszédatadtbázisból (Neuberger et al., 2014). A kísérleti személyek magyar anyanyelvűek voltak, és nem volt hallási vagy beszédzavaruk. Életkoruk 20 és 45 év között volt, átlagosan 28,1 évesek voltak (SD 6,6 év).

Az adatbázis két beszédstílusát elemeztük. Spontán beszédnek a felvételek „interjú” részét választottuk. Ebben a beszélők az életükről, munkájukról, hobbijukról beszélnek. Ez a felvételi egység kvázimonologikus, a felvételvezető akkor szólal meg, ha további kérdéseket kell feltenni a beszélő ösztönzésére. Felolvasásként pedig a felvételek olvasott szövegét használtuk. Ebben a feladatban egy ismeretterjesztő szöveget olvasnak fel a beszélők.

Az adatbázisban rendelkezésre állt a felvételek beszédszakaszszintű lejegyzése (Gyarmathy et al., 2014). Egy beszédszakasz két bármilyen hosszú és típusú (pl. kitöltött vagy néma) szomszédos szünet közötti beszédegység. A zöngétlen zárszakasszal kezdődő szakaszok esetében 30 ms-ot számítottak zárszakasznak. Nem végeztünk mérést a bármilyen jellegű információs értelemben vett zajt (például nevetést, átfedő beszédet, háttércsatorna-jelzéseket) tartalmazó részleteken, csak

azon időpontokban végeztünk mérést, ahol kizárólag az adatközlő beszéde jelent meg az akusztikumban.

A beszédanyagot a Praat szoftver (Boersma–Weenink, 2018) használatával elemeztük. Az alapprofrekvencia-értékeket automatizálva mértük a Pitch (cc) paranccsal. A várható  $f_0$ -tartományt a két nem esetében eltérően állítottuk be: 75–300 Hz-re a férfi és 100–500 Hz-re a női beszélők esetében; a további paraméterek esetében a sztenderd  $f_0$ -mérési beállításokat alkalmaztuk (time step = automatikus, max. number of candidates = 15, silence threshold = 0,03, voicing threshold = 0,45, octave cost = 0.01, octave-jump cost = 0,35, voiced/unvoiced cost = 0,14). Az adatokat 50 ms-os ablakkal és 25 ms-os csúsztatással nyertük ki. Az irreguláris zöngével képzett beszédrészletek esetében gyakori, hogy nem valós értéket mér a szoftver. Ezen értékek és az esetleges oktávugrások kiküszöbölése céljából az alábbi szűrést alkalmaztuk: A két beszédstílusban mért adatokat összevontuk, és kiszámítottuk beszélőnként az  $f_0$ -értékek interkvartilis tartományát (IQR). Azon értékek, amelyek a mediántól nagyobb mértékben tértek el, mint az IQR 1,5-szerese, kizártuk az elemzésből. Ez a módszer a kiugró értékeket távolítja el. A tanulmányban az  $f_0$ , a hangterjedelem és a hangköz értékeit összehasonlítottuk a beszédstílusok és a nemek között. Az alapprofrekvencia értékét Hz-ben, a hangterjedelmet és a hangközöt félhangokban adtuk meg. A félhangértékeket a hqmisc csomaggal számoltuk ki úgy, hogy a legkisebb  $f_0$ -értéket adtuk meg alapnak, és a legmagasabb értéknek az ettől félhangban mért eltérését számoltuk ki (Harrel–Dupont, 2017; az R programban (R Core Team, 2018)).

Az adatokat lineáris kevert modellel (LMM) elemeztük az lme4 csomag (Bates et al., 2015) segítségével az R programban (R Core Team, 2018). A függő változók az  $f_0$ , a hangterjedelem, a hangköz voltak, a független változók pedig a beszédstílus és a nem. Ezeket egyesével, majd együtt, az interakciójukat is megengedve adtuk a modellhez. Az egyes változókra készítettünk intercept és random slope modelleket egyesével, illetve a két faktort tartalmazó modellben a kettőre együtt is az  $f_0$  és a hangköz elemzése során. A hangterjedelem elemzésekor a beszélőnkénti egy adat miatt csak random intercept modelleket készítettünk. A beszélőt minden esetben random hatásként adtuk meg. A modelleket az ANOVA segítségével hasonlítottuk össze az lmerTest csomag segítségével (Kuznetsova et al., 2017). A kapott modellek közül minden esetben a legalacsonyabb AIC-számút (Akaike információs kritérium; Akaike, 1974) tartottuk meg. Mivel a lineáris kevert modellek nem számítanak  $p$ -értékeket (Lawrence, 2013), ezeket Satterthwaite közelítéssel számoltuk ki az lmerTest csomagban (Kuznetsova et al., 2017) az ANOVA módszerével. Minden statisztikai tesztben 95%-os konfidenciaintervallumot használtunk.

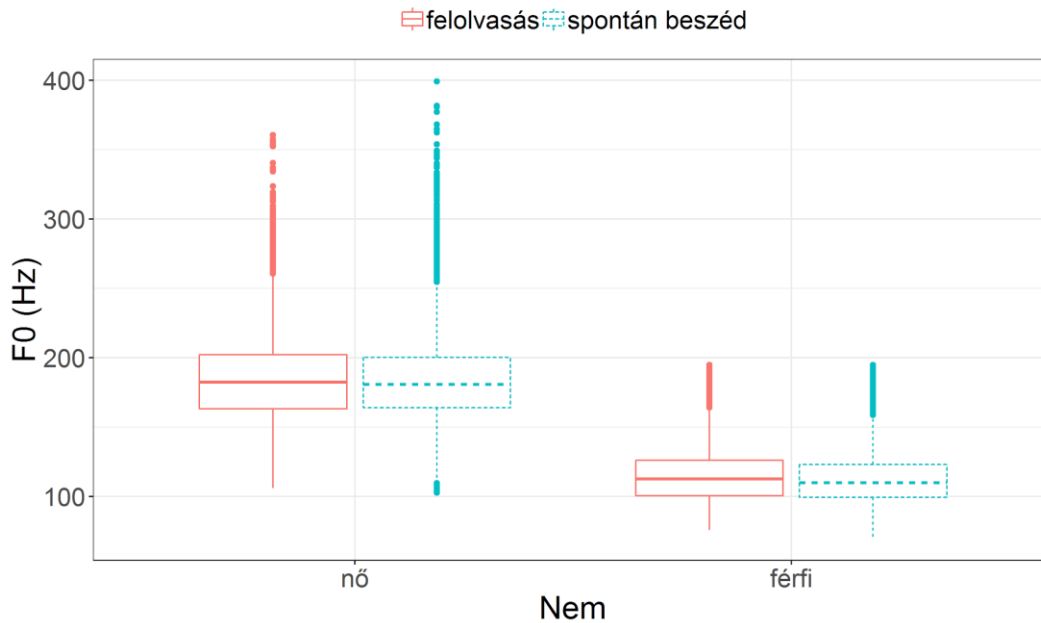
## Eredmények

Az 1. táblázat nemenként mutatja az  $f_0$ -ra kapott átlagértékeket a két beszédstílusban. Az  $f_0$  átlagértékei között a nők esetében kevesebb, mint 1 Hz, a férfiak esetében majdnem 3 Hz az eltérés a két beszédstílus között. Mindkét esetben a felolvasásban mértünk magasabb értéket. Ez a fentebb ismertetett irodalmi adatokhoz hasonló kismértékű eltérés, melyet általában szignifikánsnak találtak. A statisztikai próba alapján a beszédstílus a jelen vizsgálatban is szignifikánsan meghatározó szerepet játszott az  $f_0$ -értékek alakulásában:  $F(1; 81,984) = 8,237$ ;  $p = 0,005$ . A nemek között is szignifikáns eltérést kaptunk, ahogyan várható volt ( $F(1; 72,483) = 386,9262$ ;  $p < 0,001$ ). A két faktor interakciója nem volt szignifikáns. A legjobb leíró modellnek a nemet és a beszédstílust is faktorként, illetve a mindkettőre random görbét tartalmazó modell bizonyult. A legalacsonyabb  $f_0$ -érték a felolvasás esetében alacsonyabb volt a férfiak és a nők esetében is (5,3 Hz-cel, 3,3 Hz-cel), a legmagasabb  $f_0$ -érték pedig a nők esetében mutatott csak eltérést, a spontán beszédben volt magasabb, több mint 30 Hz-cel. A hangterjedelem és a hangköz átlagos értékei között csak néhány tizednyi eltérést mértünk a két beszédstílus között mindkét nem esetében. A hangköz kivételével minden változó értéke a nők esetében volt valamivel magasabb.

Az 1. ábra az  $f_0$ -értékeket mutatja be nemenként és beszédstílusonként. Az egyes nemek esetében közel hasonlóak a dobozdiagramok, ahogyan az átlag- és szórás-értékek alapján várható, csoportszinten nincs jelentősnek látszó eltérés a beszédstílusok között.

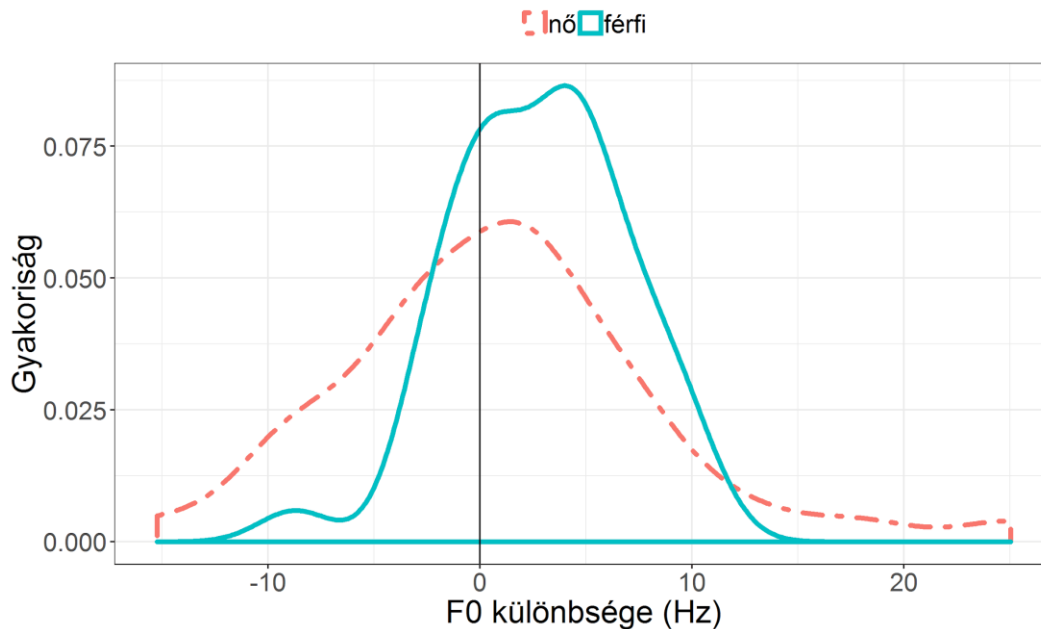
1. táblázat. Az alaphfrekvencia-jellemzők alakulása a beszédstílus és a nemek függvényében (fh. = félhang,  $\bar{a} \pm SD$  = átlag és SD).

Függő változók	Spontán beszéd		Hangos olvasás	
	férfi	nő	férfi	nő
$f_0$ (Hz, $\bar{a} \pm SD$ )	110,7 $\pm$ 12,3	182,8 $\pm$ 19,6	113,6 $\pm$ 12,7	183,8 $\pm$ 19,2
$f_{0\min}$ – $f_{0\max}$ (Hz)	70,3–194,6	102,7–398,9	75,6–194,6	106,0–360,3
hangterjedelem (fh., $\bar{a} \pm SD$ )	10,38 $\pm$ 1,89	10,63 $\pm$ 3,28	10,21 $\pm$ 1,98	10,46 $\pm$ 3,07
hangköz (fh, $\bar{a} \pm SD$ )	7,08 $\pm$ 2,55	7,04 $\pm$ 3,39	7,64 $\pm$ 2,24	7,32 $\pm$ 2,75



1. ábra. Az alapfrekvencia-értékek (Hz) tartománya a felolvasásban és a spontán beszédben a nemek függvényében.

A két beszédstílus között talált kis átlagos eltérés szignifikáns volta mögött a beszélők közötti variabilitást feltételeztük okként. Ezért kiszámítottuk beszélőnként a két beszédstílusra kapott átlagos  $f_0$ -értéket, és a felolvasás értékéből kivontuk a spontán beszédre kapott adatot (2. ábra). A pozitív érték azt jelenti, hogy a felolvasásban mértünk magasabb átlagos  $f_0$ -értéket az adott beszélő esetében, míg a negatív azt mutatja, hogy ez az érték a spontán beszédben volt magasabb. (Pontosan 0 Hz különbség nem volt, de ahhoz közeli természetesen nagy számban.) 23 nő, azaz 54,76% esetében mértünk a felolvasásban (átlag:  $6,07 \pm 5,90$  Hz), 19, azaz 45,24% esetében a spontán beszédben magasabb  $f_0$ -értéket (átlag:  $-5,16 \pm 3,94$  Hz). A férfiak esetében jelentősen több férfi, 31 fő, azaz 77,50% esetében a felolvasásban volt magasabb az átlagos alapfrekvencia értéke ( $4,37 \pm 3,01$  Hz), és csak 9 fő, azaz 22,50% esetében a spontán beszédben (átlag:  $-2,43 \pm 2,56$  Hz). Ez utóbbi esetében az eltérés is jóval alacsonyabb a férfiaknál.



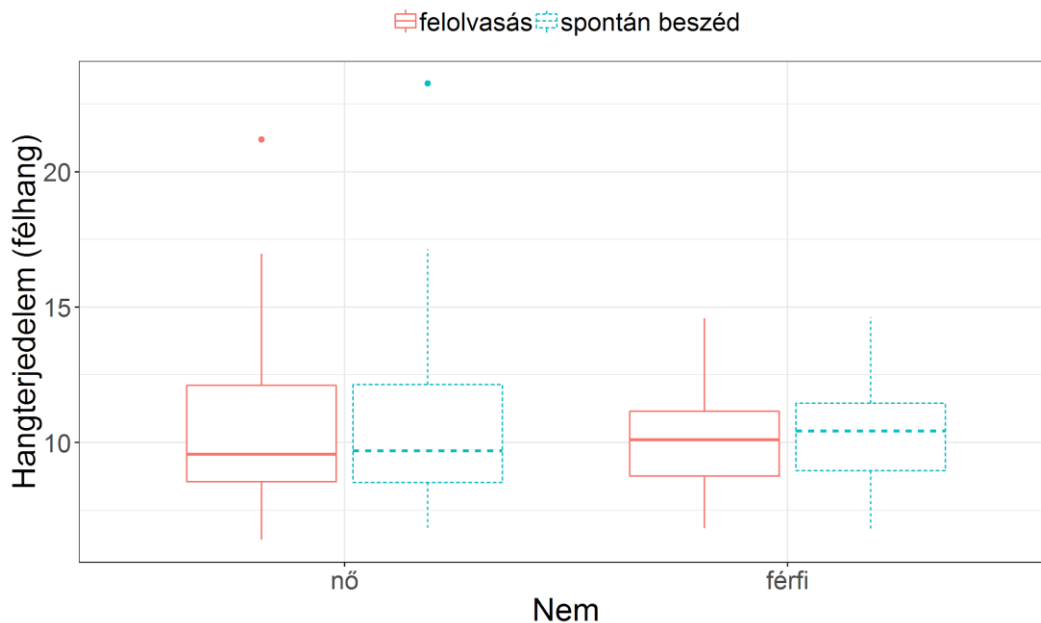
2. ábra. A beszélőnként számított átlagos alaphfrekvencia-érték különbsége (Hz) a felolvasás és a spontán beszéd között. (A pozitív értékek azt jelentik, hogy az átlagos f0 a felolvasásban volt magasabb, míg a negatívak azt, hogy a spontán beszédben.)

A hangterjedelem esetében (3. ábra) ugyancsak nagyon minimális eltérést tapasztalhatunk a csoportátlagok alapján a két beszédstílus között (1. táblázat), de ez az eltérés szignifikánsnak bizonyult. A legalacsonyabb AIC-számokat az a statisztikai modell adta, amelyben a nem és a beszédstílus is szerepelt faktorként, és mindkét faktorra tartalmazott random görbét. A nemek között nem, de a két beszédstílus között szignifikáns eltérést kaptunk:  $F(1, 82) = 7,087, p = 0,009$ . A két faktor interakciója nem bizonyult statisztikailag szignifikánsnak. A hangterjedelem esetében is elemeztük, hogy az egyes beszélők esetében melyik beszédstílusban kaptunk magasabb átlagértéket. A 4. ábra ábrázolja az egyes beszélők esetében kapott hangterjedelem-különbséget. A pozitív számok ismét azt jelentik, hogy a felolvasásban, a negatívak pedig azt, hogy a spontán beszédben mért érték volt magasabb. (Pontosan 0 félhang eltérés ismét nem volt, de ehhez nagyon közel eső különbségértékek előfordultak gyakran.) Az ábrán látható, hogy a negatív hangterjedelem-értékek mindkét nem esetében gyakoribbak, mint a pozitív értékek. Ez arra utal, hogy a beszélők többségére jellemző, hogy a spontán beszéd esetén nagyobb hangterjedelemben beszél, mint a felolvasás esetén. A férfiak 75,00%-a (30 fő), a nők 76,19%-a (32 fő) nagyobb átlagos hangterjedelemben beszélt a spontán beszéd esetében (férfiak átlaga:  $-0,39 \pm 0,43$  félhang, nők átlaga:  $-0,29 \pm 0,40$  félhang). A férfiak 25,00%-a (10 fő) és a nők 23,81%-a (10 fő) beszélt nagyobb

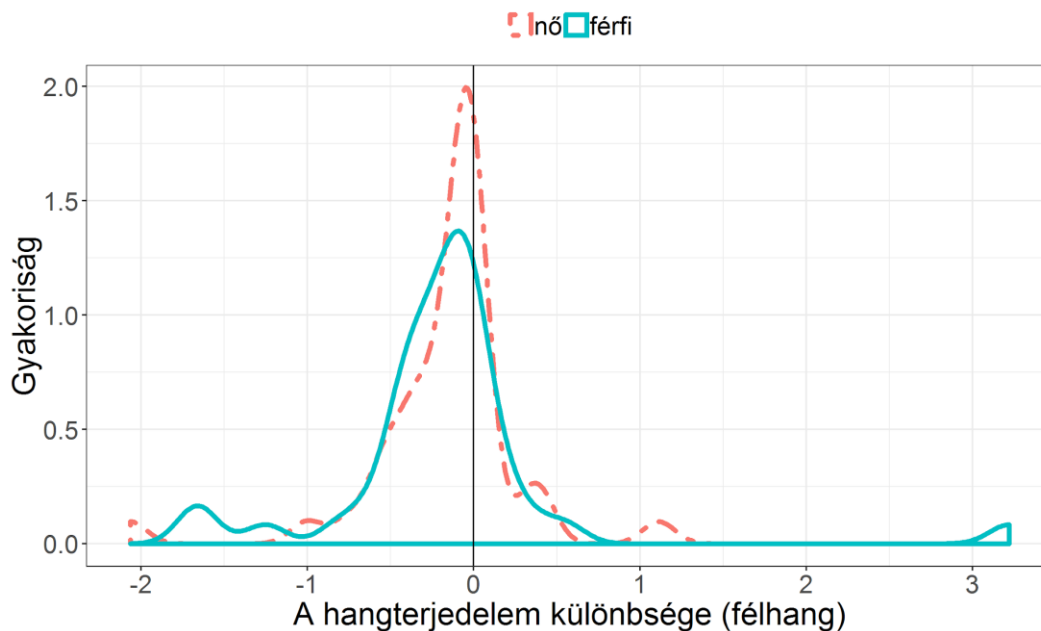


hangterjedelemmel a felolvasás esetében (férfiak átlaga:  $0,46 \pm 0,98$  félhang, nők átlaga:  $0,24 \pm 0,35$  félhang).

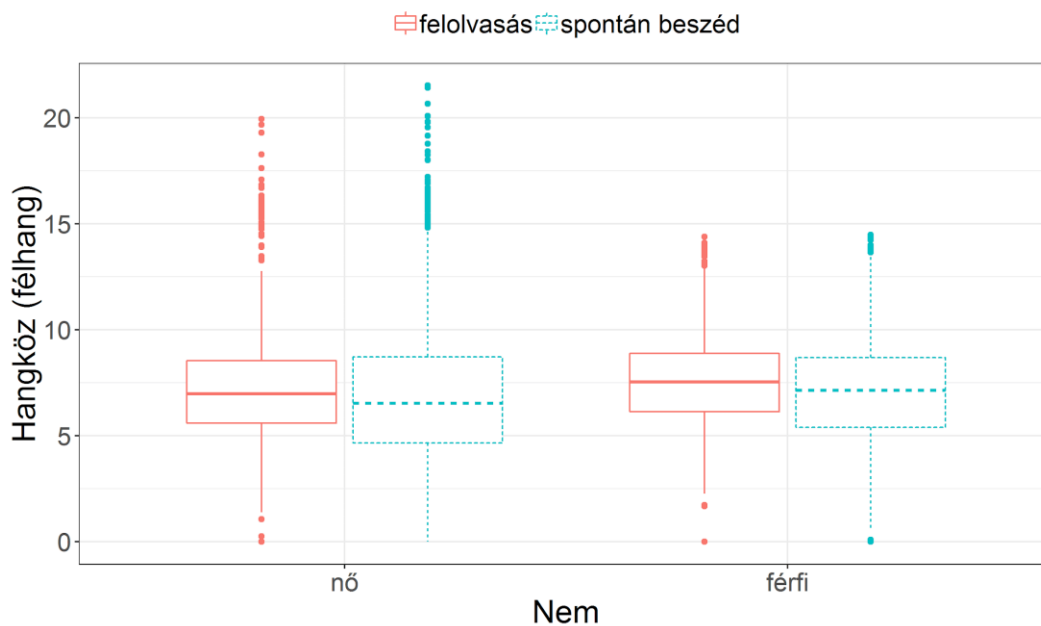
A nemenként számított átlagos hangközértékek nem mutatnak nagymértékű eltérést a két beszédstílus között (5. ábra), a statisztikai elemzés azonban szignifikáns különbséget mutatott. Az adatokat legjobban leíró modellnek ismét az a modell bizonyult, melyben mind a beszédstílus, mind a nem szerepel faktorként, és mindkettőre random görbét illesztettünk. A két beszédstílus között tehát a hangközértékek szignifikánsan eltértek ( $F(1, 73,286) = 20,885; p < 0,001$ ). A nemek között nem volt szignifikáns különbség, és a két faktor interakciója sem bizonyult szignifikáns hatásnak. Az 5. ábráról leolvasható, hogy a kiugró értékek a nők esetében jellemzőbbek beszédstílustól függetlenül.



3. ábra. A hangterjedelem-értékek (félhang) tartománya a felolvasásban és a spontán beszédben a nemek függvényében.

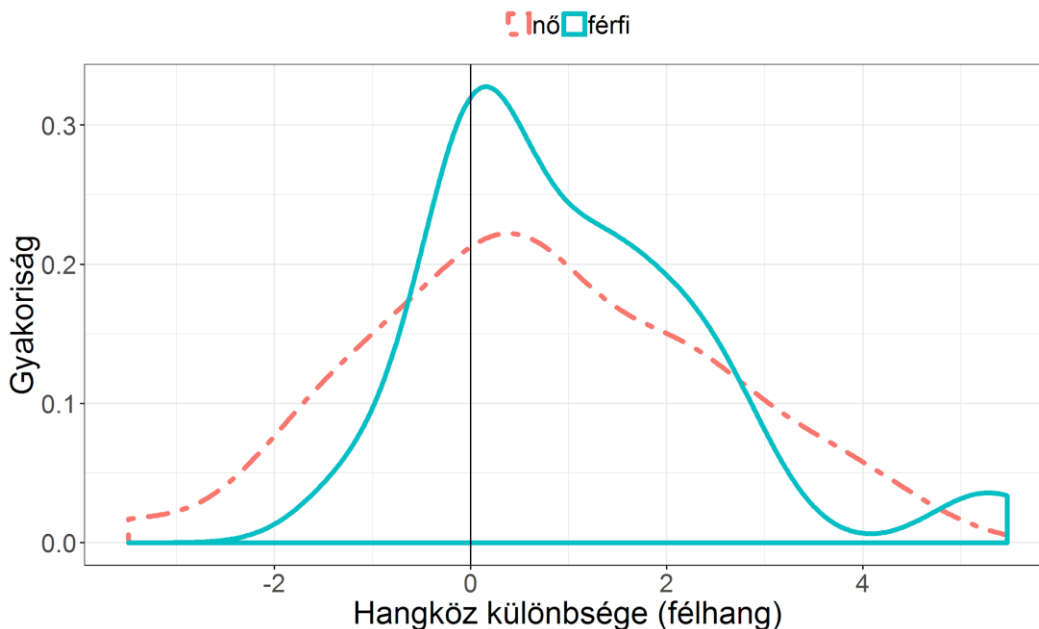


4. ábra. Az egyes beszélők hangterjedelmének különbsége (félhang) a felolvasás és a spontán beszéd között. (A pozitív értékek azt jelentik, hogy a felolvasásban, a negatív értékek azt, hogy a spontán beszédben volt tágabb a hangterjedelem.)



5. ábra. A hangközértékek (félhang) tartománya a felolvasásban és a spontán beszédben a nemek függvényében.

Ismét elemeztük, hogy az egyes beszélők esetében melyik beszédstílusban magasabbak az átlagos hangközértékek (6. ábra). A pozitív értékek azt jelentik, hogy az adott beszélő esetében a felolvasásban volt nagyobb az átlagos hangköz, míg a negatívak azt, hogy a spontán beszédben. A férfiak 77,50%-a (31 fő, átlag:  $1,46 \pm 1,34$  félhang), a nőknek pedig a 69,05%-a (29 fő; átlag:  $1,59 \pm 1,26$  félhang) a felolvasásban használt nagyobb hangközöket. A férfiak 22,50%-a (9 fő, átlag:  $-1,45 \pm 0,45$  félhang), a nőknek pedig a 30,95%-a (13 fő, átlag:  $-1,27 \pm 0,85$  félhang) beszélt nagyobb hangközzel a spontán beszéd esetében.

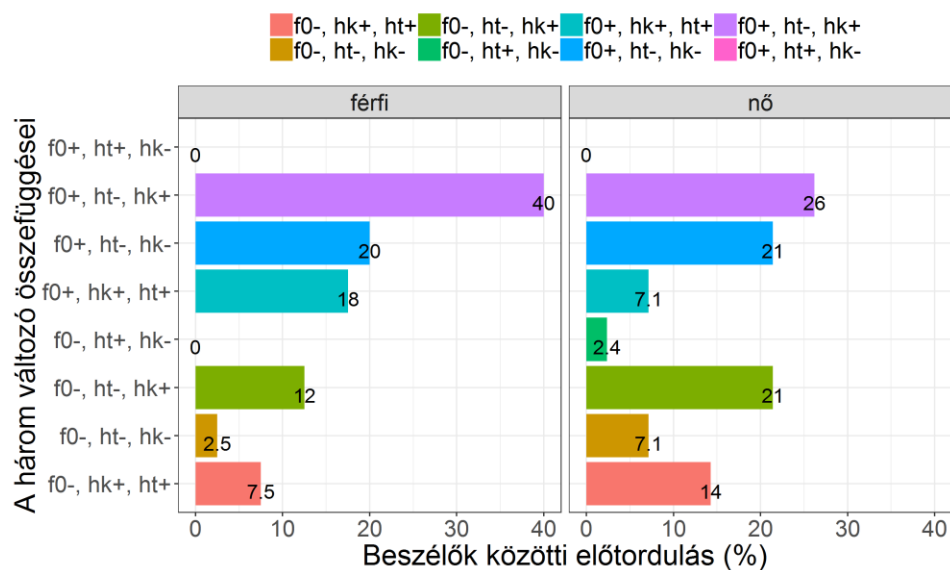


6. ábra. A beszélőnként számított átlagos hangközérték különbsége (félhang) a felolvasás és a spontán beszéd között (A pozitív értékek azt jelentik, hogy a felolvasásban volt tágabb az átlagos hangköz, a negatívak azt, hogy a spontán beszédben.)

A három változóra kapott különbségek eloszlását figyelembe véve (2., 4., 6. ábra) elmondhatjuk, hogy a hangterjedelem esetében kisebb a beszélők közötti variabilitás, míg a másik két változónál laposabb a görbe meredeksége, azaz változatosabbak a beszélőnkénti adatok.

Felmerülhet a kérdés, hogy ha nagyjából a beszélők háromnegyedére igaz, hogy a hangterjedelem a spontán beszédben, az átlagos hangköz pedig a felolvasásban tágabb, illetve a férfiak esetében a spontán beszédben magasabb az átlagos  $f_0$ -érték, akkor van-e összefüggés olyan értelemben az adatok között, hogy akikre jellemző, hogy nem igazodnak az egyik változóban talált jellemző tendenciához, esetleg a többiben sem. A 7. ábrán az adatok olyan csoportosításban láthatóak, hogy az adott

változó két beszédstílus közötti eltérése milyen tendenciát mutat (pozitív: a felolvasásban magasabb, negatív: a spontán beszédben magasabb az (átlag)érték), és ezek együttjárása milyen gyakori a két nemi csoportban. A leggyakrabban mindkét csoportban az fordult elő, hogy az f0-átlag és a hangközátlag a felolvasásban, a hangterjedelem a spontán beszédben volt magasabb, azaz azok a beszélők voltak a legtöbben, akik mindhárom tendenciát követték: a nőknél 26% és a férfiaknál 40%. A nők esetében ugyancsak gyakori (21%), hogy az átlagos f0 a spontán beszédben magasabb, míg a másik két változó a fő tendenciát követi. Ez a két eredmény várható volt, hiszen a két nemben ezek voltak a domináns eredmények. Mindkét csoportban gyakori azonban (20-21%), hogy ugyan az f0-átlag és a hangterjedelem követi a fő tendenciát, de az átlagos hangközérték a spontán beszédben magasabb. Egyik nemben sem jellemző, hogy a hangterjedelem és a hangköz a csoportban kimutatott tendenciától eltér (ht+, hk-, f0+/-). A két nemben az így felállított kategóriák előfordulási aránya összefügg (Pearson-korreláció:  $r^2 = 0,652$ ), azaz amelyik gyakoribb az egyik csoportban, az általában a másik csoportban is gyakoribb. Összességében mindkét nemben a három fő tendenciával egyező mintázat a leggyakoribb, ahogyan az várható volt, de gyakori, hogy azoktól a hangközben térnek csak el. A többi esetben a leglényegesebb eltérés, hogy a férfiak esetében ritka, hogy az f0 átlaga ne a felolvasásban legyen magasabb, míg a nőknél ez gyakori volt, ezért a másik két tényező mentén csoportosítva és az f0-különbséget nem figyelembe véve ugyanazt a mintázatot kaptuk.



7. ábra. A három változó (hk = hangköz, ht = hangterjedelem) különbségei által mutatott tendenciák összefüggései a beszélők közötti előfordulás arányában (%).  
(+: az olvasásban magasabb az (átlag)érték, -: a spontán beszédben magasabb az (átlag)érték)

## **Következtetések**

A tanulmányban az alapfrekvencia-jellemzőket elemeztük 82 anyanyelvi magyar beszélő (40 férfi és 42 nő) spontán beszédében és felolvasásában. A nemzetközi irodalomban eltérő eredményeket kaptak az egyes tanulmányokban abban a vonatkozásban, hogy a két beszédstílus között milyen eltérés áll fenn az  $f_0$ -értékek és a hangterjedelem tekintetében. Felmerül a kérdés, hogy a magyar nyelvben hogyan alakulnak a két beszédstílus  $f_0$ -jellemzői. A korábbi, magyar beszéden végzett kutatásokban kis számú adatközlő szerepelt. A jelen kutatás célja az volt, hogy az alapfrekvencia-jellemzőket nagyszámú magyar anyanyelvű adatközlő beszédében vesse össze felolvasás és a spontán beszéd között, hogy a két beszédstílus közötti eltéréseket leírja, illetve hogy a beszélők közötti variabilitást figyelembe vegye.

Mind az alapfrekvencia-, mind a hangterjedelem- és a hangközértékek eltértek a két beszédstílus között. Az alapfrekvencia és a hangköz esetében a felolvasásban, a hangterjedelem esetében pedig a spontán beszédben kaptunk magasabb értékeket mindkét nemre vont átlag esetében. Ezek a különbségek habár nagyon kicsik, mégis szignifikánsak. Ennek háttérében a beszélők közötti variabilitást feltételeztük. Azaz azt a kérdést tettük fel, hogy az egyes beszélőket külön vizsgálva is jellemző-e az átlagok alapján mért tendencia, és az eltérések maguk ilyen alacsonyak, avagy az adatközlők között jelentősen eltérő-e, hogy mely beszédstílusban mérhető magasabb érték, és emiatt kiegyenlítődik a csoportátlag eredménye.

Az alapfrekvencia-különbségek esetében jelentős eltérést tapasztaltunk a két nem eredményei között. Míg a nőknél lényegében a fele csoport az egyik, a másik fele a másik beszédstílusban realizált magasabb átlagos  $f_0$ -t, a férfiak esetében túlnyomórészt, közel 80%-ban a felolvasásban mértünk magasabb átlagos alapfrekvencia-értékeket. Ez azt jelenti, hogy a férfiak eredményei egységesebbnek mondhatók, az adatközlők ötöde kivételével, míg a nők esetében nem találtunk jellemző tendenciát. Az is elmondható, hogy a nők esetében mindkét esetben 5-6 Hz körüli az eltérés a két beszédstílus között, a férfiaknál a spontán beszédben mért magasabb átlagos  $f_0$  esetében jóval alacsonyabb a különbség, mint a másik csoportjukban.

A hangterjedelem és a hangköz esetében mindkét nemnél találtunk jellemző mintázatot, mindkét esetben azonos irányba mutatót. A hangterjedelemben átlagos intervalluma mindkét nem esetében a beszélők háromnegyedénél a spontán beszédben volt tágabb, míg negyedük pedig ellentétes mintázatot mutatott. A hangköz pedig a felolvasásban volt tágabb a férfiak több mint háromnegyedének és a nők közel 70%-ának esetében. A hangterjedelem esetében kisebbek az eltérések a

két beszédstílus között, a hangköz esetében valamivel nagyobbak, és ez utóbbinál a variabilitás, azaz a beszélők közötti változatosság is nagyobb.

A két beszédstílus közötti eltérés mértéke a hangterjedelem esetében kevésbé volt változatos a beszélők között, míg az f0 és a hangköz esetében a különbségértékek nagyobb változatosságot mutattak.

Összegezve tehát elmondhatjuk, hogy a két beszédstílus összevetésénél nem elegendő az összes beszélő átlagának vett különbségét megadni, a beszélők közötti eltérések jelentősek, akár egy-egy esetben nem is mutatnak jellemző tendenciát (vö. a nők f0-adatai). Mindennek azért van jelentősége, mert a legtöbb tanulmány abból indul ki, hogy a két beszédstílus produkciós folyamatainak eltérésére alapozva különbséget feltételez, mely a két beszédstílus viszonyában leírható. Ugyanakkor ez az eltérés nem feltétlenül jár azonos következményekkel a beszélők esetében, hanem további hatások is érvényesülnek. Az eredmények alapján tehát önmagában a beszédmód nem meghatározó az itt vizsgált alaphangjellemzők tekintetében. Úgy véljük, hogy ennél nagyobb jelentősége lehet a témának, az ahhoz való érzelmi viszonyulásnak, a spontán beszéd esetében a helyzethez és a beszédpartnerhez való viszonynak és így tovább. Hozzá kell mindehhez tennünk továbbá, hogy – a korábbi szakirodalmi forrásokhoz hasonlóan – a mért paraméterek globális elemzése nincs tekintettel olyan lokális jelenségekre, amelyek adódhatnak a beszédtervezési eltérésekből, például a felolvasásban az alkalmazható intonációs sémákból, a spontán beszédben az esetlegesen változatosabb, vagy témafüggően akár egységesebb dallammodulációkból, hiszen ezek a teljes anyagban ki is egyenlíthetik egymást.

Mindezek az eredmények nem mondanak ellent ugyanakkor annak az állításnak, amely szerint az észlelésben jól elkülönülnek a két vizsgált beszédmód formai jellemzői, azaz a hallgatók nagy arányban helyesen tudják megítélni, hogy spontán beszédet vagy felolvasást hallanak-e – vö. pl. Bóna (2011) tanulmányát, ahol a felismerési arányi mindkét szövegtípus esetében 80% fölött volt magyar nyelvű anyagokon, hasonlóan a korábbi nemzetközi vizsgálatok eredményeihez (pl. Levin et al., 1982). Laan (1997) holland nyelvű beszédmintákon végzett vizsgálatában azt találta, hogy a felolvasás a spontán beszédhez képest lassabb artikulációs tempóval, nagyobb frekvenciavariációval, több ereszkedő dallamkontúrral, alacsonyabb shimmerértékekkel és kevesebb magánhangzó-redukcióval jellemezhető. Természetesen ezek az eredmények a nyelvi különbségek (pl. a magyarban nincs fonológiai magánhangzó-redukció, a szórend és az intonáció, illetve ezek kapcsolata is jellegzetesen eltérő a két nyelvben) miatt nem vonatkoztathatók automatikusan a magyarra, arra azonban rávilágítanak, hogy a lokális jellemzőkben érdemes keresni azokat az eltéréseket, akár az alaphangmagasság sajátosságain belül is, amelyek alapján felismerjük és megkülönböztetjük egymástól a két beszédmódot.

A jelen vizsgálat fő hozadéka véleményünk szerint az, hogy tisztázta a korábbi magyar nyelvű szakirodalomban a részben a kis adatközlőszámból eredő ellentmondásokat, és felhívta a figyelmet az egyéni variancia jelentőségére, amely nemcsak a beszélők között, hanem a beszélőn belül is megnyilvánul. Ezek alapot szolgáltatnak mind további alapkutatásokhoz ezen a téren, mind pedig alkalmazott kutatásokhoz. Az ilyen jellegű vizsgálatok eredményei ugyanis olyan gyakorlati területeken is felhasználhatók, amelyekben különös jelentőséggel bír az egyéni belüli és az egyének közötti variancia figyelembevétele. Ilyen például az igazságügyi fonetika, ahol gyakran fordul elő, hogy az inkriminált beszédminta spontán szöveget tartalmaz, míg a gyanúsítottal felveendő beszédminta csak utánmondásra (ismétlésre) vagy felolvasásra épülhet, és a két- (vagy három-) féle beszéd típus összevetésében jelentőséggel bírnak a hasonló kutatások tanulságai.

## Irodalom

- Abu-Al-Makarem, A. & Petrosino, L.** (2007) Reading and Spontaneous Speaking Fundamental Frequency of Young Arabic Men for Arabic and English Languages: a Comparative Study. *Perceptual and Motor Skills* 105. pp. 572–580.
- Akaike, H.** (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19/6. pp. 716–723.
- Bates, D., Mächler, M., Bolker, B. & Walker, S.** (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, pp. 1–48.
- Beke A.** (2008) A felolvasás és a spontán beszéd alaphangszerkezeteinek vizsgálata. *Beszédkutatás* 2008. 93–107.
- Boersma, P. & Weenink, D.** (2018) *Praat: Doing phonetics by computer. (computer program version 6.0.19)*. www.praat.org.
- Bóna J.** (2011) A különböző beszédstílusok az akusztikai-fonetikai és a percepció vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* 11/ 1-2. sz. 39–48.
- Caldogetto, E. M., Zmarich, C. & Ferrero, F.** (1997) A Comparative Acoustic Study of Spontaneous and Read Italian Speech. *Proceedings of Eurospeech*. 770–782.
- Daly, N. A. & Zue, V. W.** (1992) Statistical and linguistic analyses of F0 in read and spontaneous speech. *Proceedings of the International Conference on Spoken Language Processing* 1. 763–766.
- de Silva, V., Iivonen, A., Bondarko, L. V. & and Pols, L. C. W.** (2003) Common and Language Dependent Phonetic Differences between Read and Spontaneous Speech in Russian, Finnish and Dutch. *Proceedings 15th International of Phonetic Sciences* 3. 2977–2980.
- Eskenazi, M.** (1993) Trends in Speaking Styles Research. *Proc EUROSPEECH'93* 1. 501–509.
- Fairbanks, G.** (1960) Voice and articulation drillbook. New York: Harper & Row. pp. 124–139.
- Gyarmathy D., Neuberger T., Gráci T. E.** (2014) Lejegyzési útmutató a BEA Spontánbeszéd-adatbázis háromszintű annotálásához. *Alkalmazott Nyelvtudomány* 14., pp. 35–44.
- Harrel, F. E. & Dupont, C.** (2017) *Hqmisc: Harrell Miscellaneous. R package version 3.16-0*.
- Hudson, A. I. & Holbrook, A.** (1982) Fundamental Frequency Characteristics of Young Black Adults: Spontaneous Speaking and Oral Reading. *Journal of Speech and Hearing Research* 25. 25–28.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B.** (2017) LmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82. pp.1–26.
- Laan, G. P. M.** (1997) The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication* 22. 43–65.

- Lawrence, M. A.** (2013) *Ez: Easy Analysis and Visualization of Factorial Experiments*. R Package Version 4.
- Levin, H., Schaffer, C.A. & Snow, C.** (1982) The prosodic and paralinguistic features of reading and telling stories. *Language and Speech* 25/1. 43–54.
- Markó, A. & Bóna, J.** (2010) Fundamental frequency patterns: The factors of age and speech type. *Proceedings of the Workshop Sociophonetics, at the crossroads of speech variation, processing and communication*. 357–366.
- Mennen, I., Schaeffler, F. & Docherty, G.** (2007) Pitching it differently: A comparison of the pitch ranges of German and English speakers. *16th International Congress of Phonetic Sciences*. 1769–1972.
- Neuberger T., Gyarmathy D., Grácsi T. E., Horváth V., Gósy M. & Beke A.** (2014) Development of a large spontaneous speech database of agglutinative Hungarian language. In: Sojka, P., Horák, A., Kopeček, A. & Pala, K (eds.) *Proceedings of TSD*. New York: Springer. 424–431.
- Obin, N.** (2011) *MeLos: Analysis and Modelling of Speech Prosody and Speaking Style*. *Signal and Image processing*. Doctoral Thesis. Université Pierre et Marie Curie. Paris.
- R Core Team** (2018) *R: A Language and Environment of Statistical Computing*. R Foundation for Computing, Vienna.
- Skarnitzl, R. & Vaňková, J.** (2017) Fundamental Frequency statistics for Male Speakers of Common Czech. *Acta Universitatis Carolinae Philologica* 3, pp. 7–17.
- Swerts, T., Strangert, E. & Heldner, M.** (1996) F0 Declination in readaloud and spontaneous speech. *Proceeding of Fourth International Conference on Spoken Language* 96. 1501–1504.
- Tønderin, J.** (2011) Preplanning of Intonation in Spontaneous versus Read Aloud Speech: Evidence from Danish. *Proceedings of the XVIIth International Congress of Phonetics Sciences* 4. 2010–2013.

A kutatást a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal FK-128814 számú pályázata támogatta.