

BUZA ÁKOS¹ – DEME ANDREA²

¹Pázmány Péter Katolikus Egyetem

²Eötvös Loránd Tudományegyetem

akos.bz@gmail.com

<https://orcid.org/0000-0002-6834-3769>

deme.andrea@btk.elte.hu

<https://orcid.org/0000-0001-5395-8309>

Buza Ákos – Deme Andrea: Magyar magánhangzók akusztikai szerkezetének sajátosságai
spontán beszédben és felolvasásban

Alkalmazott Nyelvtudomány, XXIII. évfolyam, 2023/1. szám, 97–119.

doi:<http://dx.doi.org/10.18460/ANY.2023.1.006>

Magyar magánhangzók akusztikai szerkezetének sajátosságai spontán beszédben és felolvasásban¹

The present study focuses on possible acoustic differences between vowels in read and spontaneous speech. We addressed whether vowels are positioned more to the hyperarticulation end of the hypo-hyper speech continuum (established by Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W.J., Marchal, A. (eds) *Speech Production and Speech Modelling*. NATO ASI Series, 55. Dordrecht Dordrecht: Springer, 403–439.), that is, they are more peripheral in read than spontaneous speech, as suggested by previous literature.

We analysed three vowels of the Hungarian language, /a:/, /i:/, and /u/, which are positioned farthest from each other in the acoustics and articulatory vowel spaces, and thus cover the most salient contrasts between Hungarian vowels. We measured duration, fundamental frequency, first and second formant frequencies, and compared them in th

e two speech styles. We recorded speech samples from 10 Hungarian females in the two speech modes; target vowels were produced in CVCVCVCVC structured Hungarian words embedded in sentences. First, the spontaneous speech was recorded using a guided elicitation technique that applied pictures and word lists to prompt speakers to produce sentences precisely as intended by the authors. Second, participants read the same sentences they produced in the first section of the experiment.

Results showed that as opposed to expectations, /a:/ was longer, while /i:/ had a higher fundamental frequency in spontaneous speech than in read speech. Furthermore, /i:/ realisations also showed a higher first formant frequency, that is, more open/centralised production in read speech than in spontaneous speech. Lastly, second formants did not differ in the two speech styles. Our findings countered expectations and did not corroborate the hypothesis that read speech is positioned more to the hyperspeech end of the hypo-hyperspeech continuum (i.e., articulated more precisely) than read speech. What is more, our results did not show a clear difference between the two styles at all.

Keywords: spontaneous speech, read speech, vowels, formant, duration

1 Bevezetés

A beszéd hangzó (akusztikus) tulajdonságai a beszédmódok szerint eltérők lehetnek. A jelen tanulmányban egy olyan kísérletet ismertetünk, amelynek célja az, hogy új adatokkal szolgáljunk arra nézve, hogy miként befolyásolja az ún. beszédstílus a magyar beszéd hangjainak megvalósulását: a magyar nyelv három, egymástól akusztikailag és artikulációsan is legtávolabb eső (azaz legjobban

¹ A kutatást a TKA–DAAD 177375 sz. projektje támogatta.

elkülönülő) magánhangzóját, az /a:/, /i:/ és /u/ hangokat vizsgáljuk meg felolvasásban és spontán beszédben.

1.1. Beszédstílusok: a spontán beszéd és a felolvasás

A spontán beszédet és felolvasást magába foglaló kategória/fogalom elnevezése problematikus; a hazai szakirodalom számos névvel illeti. *Beszédstílus*nak hívja többek között Markó (2009), Bóna (2011), Deme (2012) és Váradi (2010). *Beszéd*típusként hivatkozik rá többek között Beke (2008) és Krepsz (2016). *Beszéd*módként nevezi meg például Wacha (1974) alapján Markó (2006). Ugyanerre ráadásul az angol szakirodalom is számos, akár a fentiekől eltérő gyakorisággal megjelenő látszó elnevezést használ. De Ruiter (2015), valamint Mehta és Cutler (1988) például *speech mode*-nak, Wagner, Trouvan és Zimmerer (2015), valamint Huber-Darling (2011) *speech style*-nak hívja ezt a kategóriát. Látható tehát, hogy nincs világos megegyezés a spontán beszédet és felolvasást átfogóan leíró kategória megnevezését illetően. A magyar szakirodalmi hagyományokat követve az egyszerűbb követhetőség érdekében mi ezt a fogalmat a *beszédstílus* terminussal írjuk le.

Wacha (1974) a beszéd hangzó sajátosságai alapján négy fő stíluskategóriát különböztet meg, melyeket a meghatározások szerint a tervezettség mértéke mentén is elkülönít. Ezek egyike a spontán beszéd. Spontán beszédnek a szöveg- és hangzás azon együttesét tartja, amely „a gondolatok születésekor, vagy a tudatban előre megfogalmazott gondolatok elmondásakor szólal meg” (209). E mellett Wacha (1974) azt is írja, hogy ebben a beszédstílusban a gondolkodás és szöveg megalkotása, illetve a mondatok megkonstruálása és kiejtése egyszerre zajlik le, szinkron tevékenység. Ezzel szemben áll – egy képzeletbeli tengely másik végén – a felolvasás: a korábban, a beszélő vagy más személy által nyelvileg megfogalmazott és leírt gondolatok megszólaltatása. Felolvasás során Wacha (1974) szerint a beszélő igyekszik minél pontosabban átadni az előtte lévő nyelvi anyagot vagy kódot. Ezzel együtt a felolvasó önkéntelenül a hangjával is kifejezi, hogy nem az ő alkotásáról van szó, így az előre megalkotott szöveg meghangosításának hangzása alapvetően eltér a spontán beszédétől.

A beszédtudományban mára elterjedt vélekedés szerint a beszédstílusok hangzásbeli eltérései elsősorban a beszédtervezés különbségeire vezethetők vissza, éppen ezért a spontán beszéd (illetve a felolvasás és a kettő összevetésének) vizsgálatában az artikulációt megelőző folyamatoknak is nagy jelentőségük van. Markó (2015) a spontán beszédet Levelt (1989, 1999) beszédprodukciós modellje nyomán a következőképpen írja le. A beszélő (mind a munkamemóriájának, mind a hosszútávú memóriájának aktivitása révén) létrehozza a preverbális üzenetet, amelynek tervezése egyfelől a kommunikációs célok és a kimondani szándékolt információk kiválasztásával jár (ezt nevezzük makrotervezésnek), másfelől pedig azzal, hogy a mikrotervezésben előállt információdarabokhoz propozíciós sémákat rendelünk (ezt nevezzük

mikrotervezésnek). A konceptuális struktúrákat az ún. formulátor alakítja nyelvi jelekké, a grammatikai kódolás a felszíni szerkezetet hozza létre, a kódot rendezi struktúrákba, a fonológiai kódolás pedig az artikulációs tervet alakítja ki. Ez teszi lehetővé a beszéd fizikai megvalósítását: az így előállt „program” képezi az ún. artikulátor bemenetét, amely a beszéd meghangosításáért felel (Markó, 2015).

A fentiek alapján feltehető, hogy a felolvasás a beszédtervezés szintjén alapvetően eltér a spontán beszédétől, hiszen felolvasás során biztosan nem a spontán beszédben tapasztalható módon jön létre a preverbális üzenet (ha egyáltalán beszélhetünk ilyenről), ugyanis a mondandó tartalma adott, mi több már nyelvi formába is van öntve. Ugyanezen okból a felolvasásban a spontán beszédben tapasztalhatótól minden bizonnyal eltérő (és feltehetően korlátozottabb működésű) a makro- és mikrotervezés (ismét csak: ha egyáltalán beszélhetünk ezeknek a bevonódásáról). Végeredményben az feltételezhető, hogy a beszélő a felolvasás során – a spontán beszéd létrehozásával szemben – minden bizonnyal főként (vagy kizárólag) a meghangosítandó szöveg artikulációs tervének, azaz az artikuláció előkészítésének kivitelezésére fókuszál.

A képet némiképpen bonyolítja, hogy a felolvasás és a spontán beszéd között több további beszédstílus is elképzelhető lehet. Wacha (1974) két ilyen különít el, a reprodukív és fél-reprodukív beszédet, és ezeket alapvetően két, egymástól elkülöníthetetlenül összevegyített szempont mentén írja le. A szempontok egyike a szöveg előre tervezettségére, a másik pedig a hangzás természetességére vonatkozik. Wacha (1974) elgondolásai szerint a reprodukív beszéd alapvetően csak hangzásában „természetes” meghangosítása az előre megalkotott, a beszélő részéről a tervezési folyamatok aktivitását nem igénybe vevő szövegeknek. A félreprodukív beszéd pedig hol a tervezés és kivitelezés párhuzamosságát mutató spontán beszéd, hol pedig csupán az előzőekhez hasonló „meghangosítás”, azaz lényegében természetes beszédnek ható, de teljes mértékben előre tervezett megszólalás.

Krepsz (2016) szintén több beszédstílust különböztet meg egymástól. A spontán beszédet Gósy (1998) alapján úgy határozza meg, hogy abban a beszélő az adott pillanatban válogatja ki a közlésre szánt gondolatokat, és tervezi meg a nyelvi formákat, így ebben a beszédtervezés és kivitelezés párhuzamosan zajlik. Egy további kategóriaként jelenik meg a félspontán beszéd, amely – Wacha (1974) *fél-reprodukív beszéd* fogalmától nagymértékben eltérően – Krepsz (2016) meghatározásában annyiban különbözik a spontán beszédétől, hogy a beszélő (tartalmilag) átgondolja a mondanivalóját, de a nyelvi formát továbbra is az adott beszédhelyzetben rendeli csak hozzá. Krepsz (2016) harmadik kategóriája a reprodukív beszéd, amelyben a beszélő egy előre elkészített szöveget tanul meg, majd mond el fejből. A negyedik kategória a felolvasás, amelynek során a beszélő egy előre rögzített, megtervezett, grammatikailag és szókincsében elkészített szöveget hangosít meg.

Látható tehát, hogy a kategóriákat kijelölő szempontok vegyülnek az egyes irodalmi forrásokban, illetve a felolvasás és spontán beszéd között feltételezhető köztes kategóriák meghatározása sem egybehangzó. Ezeknek a kérdéseknek az alaposabb körüljárása, tisztázása azonban túlmutat a jelen tanulmány céljain és lehetőségein. A jelen tanulmányban Krepsz (2016) meghatározásaiból indulunk ki, és azt a megközelítést választjuk, hogy kizárólag a tervezettség és meghangosítás idői viszonyai alapján különböztetjük meg a beszédstílusokat egymástól. Ez alapján elkülönítünk spontán beszédet, felolvasást, illetve félspontán beszédet. Spontán beszédként nevezzük meg azt, amikor a beszéd tartalmi és nyelvi tervezése, illetve a beszéd kivitelezése lényegében egyidejűleg zajlik. Felolvasásként értelmezzük azt, ha a beszélőnek egy teljes mértékben előre megformált szöveget kell meghangosítania, tehát a tervezés és kivitelezés időben a legtávolabb áll egymástól. Félspontán beszédnek pedig azt tartjuk, amikor a tervezés folyamatai csak részben fednek át a kivitelezéssel, mert a nyelvi formák kialakítása, a nyelvi tervezés aktuálisan zajlik, míg a megszólalás tartalmi elemei már részben adóttak.

A vizsgálatunk az itt felsorolt beszédstílusok közül a felolvasásra és a félspontán beszédre fókuszál. A fentiekkel összefüggésben ugyanakkor azt feltételezzük, hogy a beszédben lezajló, az alulartikuláltságot jellemző folyamatok a spontán beszédben hasonlóak a félspontán beszédben tapasztalhatóakhoz, csak azok mértéke különbözik úgy, hogy a spontán beszédben tapasztaljuk a legnagyobb mértékű alulartikuláltságot, illetve redukciót (lásd lentebb). Ilyen formában tehát a félspontán beszéd vizsgálatában találtakból vélekedésünk szerint közvetve a spontán beszédre nézve is levonhatunk következtetéseket.

1.2. A H&H elmélet

Lindblom (1990) a beszédprodukciónak vizsgálata során megalkotta a H&H elméletet (a két *h* a *hyperspeech*, azaz túlartikulált beszéd, valamint *hypospeech*, azaz alulartikulált beszéd fogalmára utal). Elmélete szerint a beszélők megváltoztatják az akusztikai megformáltság mértékét, „tisztaságát” (vö. *clear speech*, pl. Lindblom, 1990: 428) a hallgató információigényeihez igazodva. A beszélő túlartikulál akkor, amikor a hallgatónak maximális akusztikai megformáltságra van szüksége, de csökkenti az artikulációs pontosságot (és az arra tett erőfeszítések mértékét) akkor, amikor a hallgató helyettesíteni tudja az akusztikumban ebből előálló hiányokat más információkkal. Az alulartikuláció folyamatát és a megformáltság mértékét a lexikai megkülönböztető képesség irányítja azáltal, hogy a beszélő túlzott artikulációs gazdaságosságát korlátozza: az alulartikulációnak ugyanis gátat szab az, hogy a megformáltság túlzott mértékű csökkentése potenciális értelmezhetetlenséget okozna az egyes jelalakok vagy jelek (fogalmak) összemosódása révén. Lindblom (1990) modellje szerint a

beszéd létrehozásában ugyanis kiemelt jelentőségű az, hogy a hallgató képes legyen megkülönböztetni a célfogalmat a szókincs egyéb elemeitől.

A fentieket úgy is összefoglalhatjuk, hogy Lindblom (1990) szerint a beszédprodukción belül a megszólalás módja hallgatóvezérelt folyamat, melyet két egymással ellentétes vezérelv működtet. A gazdaságosságra törekvés értelmében a beszélő igyekszik minél kevesebb energiát elhasználni a beszédprodukción során (Lindblom, 1990). (Ilyen, a gazdaságosságot elősegítő tényező lehet például a beszédhangok koartikulációja, azaz a következő hangok létrehozásához szükséges artikulációs mozzanatok anticipációja, vö. pl. Kassai, 2001; valamint a beszédhangoknak a koartikuláció következtében előálló redukciója, vö. a magánhangzók koartikulációs ellenállásának kérdését pl. Deme et al. 2019b; 2021; 2022a; 2022b; előkészületben.) Az érthetőségre törekvés ezzel szemben növeli vagy csökkenti a közlés által hordozott információ hozzáférhetőségét a hallgató számára, melynek kulcsa a „kontraszt”, azaz a fentebb említett megkülönböztetés biztosítéka a beszédjelben. Kontraszton – éppen a megkülönböztetés biztosítékaként való értelmezés miatt – a legtöbb kutatásban a jelalakok és jelentések megkülönböztetéséért felelős nyelvi kontrasztot értik a kutatók, mint például a magánhangzók minőségét, a mássalhangzók zöngességét vagy a nyelvi hosszúságot. Az elmélet ugyanakkor teret ad tágabb értelmezéseknek is (Deme et al., 2019a).

A beszélő megszólalásai az elmélet értelmében két pólus között ingadoznak, a *hyperspeech* (túlargikulált) és *hypospeech* (alulargikulált) beszéd között. Ezt az ingadozást alapvetően meghatározza a hallgató hozzáférése a beszélő által közölni kívánt információhoz, továbbá a kommunikációs körülmények is (Lindblom 1990). Túlargikulációt számos tényező kiválthat. Ezek lehetnek kognitív, a beszélő vagy a hallgató személy mentális/érzelmi állapotából fakadó, illetve a kommunikációs csatornában rejlő tényezők is: túlargikulált stílusnak vagy regiszternek szokás tartani például a környezeti vagy háttérzajban produkált beszédet, az idősekhez vagy hallássérültekhez szóló beszédet, valamint a dajkanyelvet (Deme et al., 2019a).

Mind a spontán beszéd és felolvasás (összevető) elemzése, mind pedig a hiper- és alulargikulált megszólalásmódok elemzése egyre gyarapodó szakirodalommal bír, ezekből a következőkben adunk áttekintést, de csak a jelen vizsgálat szempontjából releváns kérdésekben, a beszéd ún. szegmentális vagyis beszédhangszintű vetületére fókuszálva. (A prozodiáról vö. pl. Lieberman et al., 1984; Howell–Kadi-Hanifi, 1991; Blauuw, 1994; Markó, 2012; Beke, 2008). Emellett pedig már ezen a ponton is kiemeljük azt a tényt, hogy a fentiek alapján egyes feltételezések szerint a spontán beszéd és felolvasás által képzett, valamint a túl- és alulargikulált beszédstílusok mentén kirajzolódó dimenziók között párhuzamosság van, mégpedig úgy, hogy a spontán beszéd az alulargikulátnak nevezett beszédhez áll közelebb, míg a felolvasás inkább a túlargikulált beszédstílusok csoportjába tartozik (vö. Markó et al., 2009). Úgy látszik, a

szakirodalom egy részében – expliciten vagy impliciten – ez a feltételezés adja a beszédstílusok összehasonlításának apropóját, és ez az a feltevés, amit a jelen tanulmányban bemutatott kísérlet is vizsgálni kíván.

1.3. Korábbi eredmények a beszédstílusok hangzásbeli jellemzőinek kutatásából

A spontán beszéd és a felolvasás fentebb vázolt különbözőségeit felismerve a szakirodalomban az elmúlt évtizedekben jelentősen megnövekedett az ezeknek a beszédstílusoknak a hangzásbeli sajátosságait összehasonlító vizsgálatok száma.

Ladefoged és munkatársai (1976) hét különböző beszédstílusban vizsgálták magánhangzók megvalósulását a teljesen informális beszélgetéstől a listaszerűen felolvasott szósortokig, kilenc felnőtt amerikai angol anyanyelvű férfi beszélő ejtésében. A vizsgálat meghatározott /b/V vagy /b/V/b/ szerkezetű valódi szavakban előforduló magánhangzókra fókuszált, melyeket a spontán beszédműfajokban, illetve a felolvasásban is többféleképpen elicitáltak, és az egyes módokon elicitált szavakat külön-külön beszédstílusú megvalósulásoknak tekintették. Megjegyzendő, hogy a beszédstílus fogalmát a szerzők inkább a formális-informális tengely, nem pedig a tervezettség dimenziójának tekintették – erre a megfontolásra már utaltunk –, így az elicitációs módok és az elemzés is ennek fényében értelmezendők. Az egyik fajta spontán beszédstílust úgy elicitálták a szerzők, hogy az adott szóhoz (pl. *bee* 'méh') kapcsolódó beszédtemát kezdeményeztek, és a célszó felé terelgették a résztvevő beszélgetőtársakat (pl. rákérdeztek, hogy meg tud-e nevezni egy, a bögölyhöz hasonló állatot). Egy második fajtát rímelési feladattal hívták elő (pl. *Mi rímel arra, hogy toy? 'játék' – Boy 'fiú'*). Egy harmadik beszédstílust képezték az elemzésben azok a megvalósulások, melyek úgy keletkeztek, hogy a kísérletvezetők megkérték a résztvevőket arra, hogy szabadon alkossanak mondatokat a kézhez kapott célszavakkal. Bár a szerzők nem teszik ezt ilyen formában nyílttá, a korábban ismertett megfontolásaink alapján azt mondhatjuk, hogy sorrendben a negyediktől kezdődnek a tervezettebb (egyértelműen nem spontánnak tekinthető) stílusok. A negyedik stílusban a résztvevők a kísérletvezető által felolvasott mondatokat ismételték el. Az ötödik kondícióban egy összefüggő rövid szövegben olvasták fel a beszélők a célszavakat. A hatodik stílusban egyetlen, azonos szerkezetű hordozómondatban olvasták fel a résztvevők az összes célszót (*The word is: ____; 'a szó a következő: ____'*). Végül a célszavakat listaszerűen is meghangosították a résztvevők. Az elemzés az első négy toldalékcsovi rezonanciára, azaz formánsra fókuszált (ezekből az első kettő, az F_1 és F_2 értékeit artikulációsán a vízszintes és függőleges nyelvhelyzet, illetve az ajakműködés alakítja), és azt állapította meg, hogy a beszédhangok megvalósítása nem függ a beszédstílustól, a szerzők ugyanis nem találtak szisztematikus, és a beszélőkön belül vagy a beszélők között általánosítható különbségeket a beszédstílusok között.

DiCanio és munkatársai (2015) egy Mexikóban beszélt veszélyeztetett nyelv beszélőivel végeztek vizsgálatot, itt azonban a spontán beszéd összevetése nem egy felolvasási, hanem egy utánmondási feladattal történt. Erre a megszólalási formára, hasonlóan a felolvasáshoz, kisebb mértékű tervezettség jellemző, de ez az előhívási feladat révén (ami minden bizonnyal nagyobb/más típusú kognitív terhelést jelent, mint a felolvasás) nem tekinthető azonosnak azzal. A kutatás azt mutatta, hogy a spontán beszédben rövidebbek a (hangkörnyezetükben, ill. prozódiaiilag nem kontrollált előfordulású) magánhangzók, illetve jobban is hasonulnak minőségükben (azaz jobban redukálódnak), mint az elicitált beszédben (szavak utánmondásában). A spontán beszédben a fentiek mellett nagyobb szóródást is találtak a kutatók a magánhangzók ejtésében. Mindezek ellenére a kutatók arra következtettek, hogy a beszédstílusok hatása nem jelentős olyan értelemben, hogy a spontán beszélgetésből származó adatok lényegében ugyanazt a mintázatot rajzolták ki a magánhangzók ejtését illetően, mint az elicitált anyagból származók. (A vizsgálat célja elsősorban az volt, hogy a kutatók felmérjék, ugyanolyan eredményeket kapnak-e akkor, ha spontán hanganyagokat elemeznek a laboratóriumi beszéd, azaz felolvasás/utánmondás helyett, ez ugyanis a korlátozott hozzáférhetőség miatt kardinális kérdés a veszélyeztetett nyelvek kutatásában.)

Docherty és munkatársai (2019) az elől képzett /i/, /ɛ/ és /æ/ magánhangzók megvalósulását elemezték ausztrál angol anyanyelvű beszélők spontán beszélgetéseiben az első két formáns mentén. A vizsgált magánhangzók az anyagban változatos (lényegében kontrollálatlan) fonetikai helyzetekben (pl. hangkörnyezetekben, illetve prozódiai helyzetekben) jelentek meg. Az elemzésből származó adatokat korábbi, listaszerűen ejtett szavak vizsgálatából származókhöz hasonlítva a kutatók – egyebek mellett – arra következtettek, hogy a spontán beszélgetésben ejtett magánhangzók megvalósulásai a két paraméter mentén számolt átlagértékekben redukáltabbnak tűntek a spontán beszélgetésben, illetve ugyanitt nagyobb változatossággal is jelentkeztek. A kutatás körülményeiből ugyanakkor természetesen következik, hogy a két összevetett adathalmaz esetében sem a vizsgált anyag, sem a beszélők, sem az elemzések módszertana nem egyezett vagy illeszkedett egymáshoz tökéletesen, így a következtetések inkább iránymutatóként kezelhetők további kutatási irányok tervezésében.

Wang (2019) szavak és egy szöveg felolvasását, valamint egy spontán beszélgetést (interjút), pontosabban az ezekben előforduló magánhangzókat hasonlította össze 20 kínai anyanyelvű nő angol nyelvű beszédprodukciónak. A 11 vizsgált magánhangzó a három kondícióban nem kontrollált (és sejtetően nem azonos) környezetekben fordult elő; az elemzésbe a vokális-időtartamokat és az első két formáns frekvenciaértékét vonták be a szerzők. Az eredmények szerint a szólista felolvasásakor hosszabb beszédhangokat ejtettek a beszélők, mint a másik két megszólalási formában, míg azok egymástól nem különböztek. A

magánhangzók elhelyezkedése az akusztikai magánhangzótérben, illetve az ezekből kirajzolódó magánhangzótér azonban nem tért el a beszédstílusok mentén, azaz a magánhangzók nem különültek el jobban vagy kevésbé az egyik beszédstílusban, mint a másik(ak)ban – bár az egyes magánhangzók szóródása eltért egymástól a három beszédstílusban együttesen tekintve az adatokat. A magyar nyelvre vonatkozóan három olyan vizsgálatról számolhatunk be, amelyek a spontán beszéd és a felolvasás közötti különbségek feltárásához közvetlenebbül kapcsolódnak.

Grácsi és Horváth (2010) tanulmányukban a magánhangzók realizációját vizsgálták spontán beszédben. Egy 9848 magánhangzóból álló korpusz vizsgálata során megállapították a vizsgált vokálisok első és második formánsának értékeit, majd összehasonlították ezeket a nyelv vízszintes helyzetével összefüggő elölség, a nyelv függőleges helyzetével összefüggő nyelvállásfok/nyíltság, illetve az ajakműködés és a nyelvi hosszúság fonológiai kategóriái alapján. A nyelvállásfok/nyíltság esetében az F_1 , az elölség és az ajakműködés esetében pedig az F_2 értékét elemezték a szerzők (ezek között találni ugyanis szorosabb összefüggéseket, vö. Stevens, 1998; Deme, 2016). Az eredmények szerint a formánsok frekvenciaértékei nagyban átfedtek az egyes kategóriák között, azaz a magánhangzó-megvalósulások minden egyes releváns csoportja között (pl. az elől és hátul képzett vagy a kerekítéses és kerekítetlen magánhangzók között). Az adatok alapján a szerzők megállapították, hogy a spontán beszédben a hátul képzett magánhangzók gyakran realizálódtak (akusztikai értelemben) a vártnál előrébb (azaz magasabb F_2 -értékkel), míg a hátul képzett magánhangzók F_2 -je gyakran a palatálisok tartományában realizálódott, ami előrébb képzettséget sejtet. A fentebb említett átfedések kisebb mértékűnek látszottak ugyanakkor a fonológiailag hosszú magánhangzók esetében. A nyíltság szerint a szerzők azt is feljegyezték, hogy még a legszélsőségesebb felső és legalsó nyelvállású kategóriák is mutattak átfedést.

Markó, Grácsi és Bóna (2009) a zöngésségi hasonulás megvalósulását vizsgálták spontán beszédben és felolvasásban. A szerzők itt expliciten azt feltételezték, hogy az eltérő tervezési folyamatok miatt a hasonulások realizációiban is eltérést találni a beszédstílusok között, és ilyen módon a spontán beszéd a Lindblom (1990) H&H elméletében alulartikulálnak nevezett beszédhez áll közelebb, míg a felolvasás inkább a túlartikulált beszédstílusok csoportjába tartozik. Az eredmények alapján úgy tűnt, hogy olvasás során az értelmi tagolás megvalósítására tett erőfeszítések több esetben is felülírták a zöngésségi hasonulás szabályait (azaz megfigyelhető volt, hogy nem egyszer nem játszódott le a hasonulás ebben a beszédstílusban, és „betűejtettek” a beszélők), míg spontán beszédben gyakran szünet jelent meg a vizsgált mássalhangzók között, ami ugyancsak gátolta a hasonulást. A spontán beszédben a szünetek megjelenését a beszédtervezésre vezették vissza a szerzők (amihez a szünetek időt biztosítanak). A fentiek alapján Markó és munkatársainak (2009) kiinduló hipotézisét, mely

szerint a spontán beszéd alulartikuláltabb, mint a felolvasás, a vizsgálat eredményei alátámasztani látszottak. A szerzők arra következtettek, hogy a spontán beszédben a spontán beszéd egyes típusai az alulartikulált beszéd különböző (és adott esetben különböző fokban alulartikulált) megjelenései (Markó et al., 2009).

Végezetül Deme, Grácsi, Horváth és Markó (2011) a spontán beszédben és felolvasásban megvalósuló magánhangzókat hasonlították össze akusztikai szerkezetük szerint. A szerzők kutatásukban arra a kérdésre keresték a választ, hogy találhatók-e különbségek a spontán beszédbeli és felolvasásbeli magánhangzó-realizációk között, továbbá arra, hogy (Markó et al., 2009 következtetéseinek értelmében) a túl- és alulartikuláltság kontinuumuma valóban megfeleltethető-e a spontaneitás/tervezettség kontinuumának. Ilyen értelemben ez a tanulmány a jelen kutatás közvetlen előzménye. Deme és munkatársai (2011) vizsgálták a vokálisok formánsstruktúráját és időtartamát, és elemezték a beszédhangejtésben tapasztalható eltéréseket a két beszédstílusban, valamint az egyes, fonológiailag elkülöníthető magánhangzócsoportok közötti kontrasztokat, azaz a magánhangzókat elkülöníthetőségét automatikus osztályozó segítségével. A szerzők a magánhangzó formánsstruktúrájában eltérést fedeztek fel az /v, a:, ε, e:, i, i:/ és /y:/ hangok esetében a két beszédstílus között. A hangok fizikai időtartama túlnyomó többségben tükrözte fonológiai hosszúságukat, és (elsősorban a rövid magánhangzókat esetében) a felolvasásban hosszabb megvalósulásokat találtak a szerzők, mint a spontán beszédben. A függőleges és vízszintes nyelvhelyzet, illetve az ajakműködés fonológiai kategóriái szerint azonban az automatikus osztályozó segítségével ugyanolyan mértékű eltéréseket találtak mindkét beszédstílusban. Ez pedig azt jelenti, hogy annak ellenére, hogy a magánhangzókat eltérni látszottak a spontán beszédben és a felolvasásban, az egyes kontrasztok nagyon hasonlóan alakultak ezekben, azaz a megfelelő akusztikus sajátosságok ugyanolyan mértékben látszottak elkülöníteni a kontraszt páriáit egymástól. Ennélfogva a szerzők végül arra következtettek, hogy a magánhangzó-realizációk lényegében nem tértek el a két beszédstílusban, már ami a köztük lévő kontrasztot illeti (Lindblom, 1990 elméletére utalva). Így Deme és munkatársainak (2011) eredményei végeredményben arra utalnak – Markó és munkatársainak (2009) következtetéseivel ellentétben –, hogy a spontán beszéd és a felolvasás szembenállása nem feleltethető meg a túl-, illetve alulartikuláltság dimenziójának.

1.4 A jelen kutatási célja és kérdése

A jelen tanulmányban három magyar magánhangzónak, az /i:/-nek, /u/-nak és az /a:/-nak hasonlítjuk össze három meghatározó akusztikai tulajdonságát félspontán beszédben és felolvasásban: az első és második formánsok frekvenciáját (F_1 , F_2), valamint az időtartamot. Azt a kérdést igyekszünk megválaszolni, hogy akusztikailag valóban jobban megformáltak-e a magánhangzókat a felolvasásban, mint a félspontán beszédben (és közvetve a spontán beszédben). A kutatásunk

kiindulópontját Markó és munkatársainak (2009), valamint Deme és munkatársainak (2011) eredményei képezik, melyek a feltett kérdésre látszólag ellentmondó válaszokat adnak.

A két beszédstílus vizsgálatának egyik szempontja a vizsgált magánhangzók időtartama volt. Lindblom H&H elmélete alapján arra számítottunk, hogy a felolvasásban mért időtartamok az alaposabb megformáltság és az ehhez szükséges nagyobb időtartam miatt átlagosan nagyobbak, mint a félszpontán beszédben. E mellett a H&H elmélet alapján, illetve a korábbi vizsgálatok eredményei alapján azt feltételeztük, hogy a felolvasásban túlartikuláltabbak a magánhangzók, mint a félszpontán beszédben, ami szélsőségesebb (periferikusabb) pozíciókat jelent az F_1 és F_2 mentén felvett akusztikai magánhangzótérben.

2 Módszerek

A kísérletben 10 magyar anyanyelvű, 20 és 21 év közötti női beszélő vett részt. Saját bevallásuk szerint mindannyian éphallók, ép beszédűek és egészségesek voltak a felvételek készítésekor.

A kísérletben három magyar magánhangzót vizsgáltunk, melyek a magyar magánhangzók rendszerében egymástól a legtávolabb helyezkednek el artikulációs és akusztikai tekintetben is: a közepen képzett, legalsó nyelvállású, illabiális, nyílt és hosszú /a:/-t, a hátul képzett, felső nyelvállású, labiális, zárt és rövid /u/-t, valamint a hátul képzett, felső nyelvállású, illabiális, zárt és hosszú /i:/-t. A magánhangzók kiválasztásának legfőbb szempontja a képzésükben tapasztalható különbség volt. Mivel ezek az egymástól legmesszebb elhelyezkedő vokálisok a magyarban, jól reprezentálják a magánhangzótér kiterjedését, illetve megmutatják a maximális minőségbeli kontrasztot a magyar magánhangzórendszerben. Látható ugyanakkor, hogy a vizsgált magánhangzók fonológiai hosszúsága nem egyezik. Ennek az az oka, hogy a lentebb részletezett, jól kontrollált kontextusban a hosszú-rövid párok itt megadott tagjaival volt lehetséges egymástól nem jelentősen eltérő gyakoriságú és nem túl ritka előfordulású szóalakokat találni.

A célmagánhangzókat CVCVCVCVC szerkezetű magyar szavak második, tehát hangsúlytalan szótagjában, nem szóvégi helyzetben vizsgáltuk. A célhangok tehát elméletileg semmilyen szempontból nem szerepeltek „erősített” vagy kiemelt helyzetben – hangsúlyos szótagban, illetve a frázis végén ugyanis tudható, hogy a magánhangzók hosszabbak, illetve minőségükben is eltérők lehetnek a hangsúlytalan, illetve nem frázisvégi helyzetűekhez képest (vö. Kohári, 2018; Markó et al., 2018; Deme et al., 2019b; 2021; megjelenőben). A vizsgált magánhangzókat a környező mássalhangzók szempontjából is kontrolláltuk: minden magánhangzót minden esetben két, azonos képzési helyű, alveoláris mássalhangzó (/n t s r/ vagy /l/) zárt közre (szisztematikusan, tehát pl. egy felolvasott *szít* /si:t/ szótaghoz mindig tartozott egy megfelelő *szít* /si:t/ vagy *rit*

/ri:s/ vagy *nír* /ni:r/ stb. szótag a félspontán beszédben) azért, mert ismeretes, hogy a képzési hely befolyásolja a szomszédos magánhangzó megvalósítását. Az így előállított célszavakra az 1. táblázatban láthatók példák.

A célmagánhangzókat tartalmazó szavakat mind felolvasásban, mind félspontán beszédben megvizsgáltuk. A félspontán beszéd irányított előhívását a következő feladattal végeztük. A kísérletben résztvevők azt a feladatot kapták, hogy egy diasoron plusz „(+)” jelekkel összekapcsolt szótöveket összeolvasva alkossanak mondatokat – a mondatok létrehozását a kontextust megteremtő, továbbá a mondat jelentésének érzékeltetését szolgáló képekkel segítettük (1. ábra). Ezzel a módszerrel sikerült úgy előhívni a résztvevőkkel az adott magánhangzó vizsgálatát szolgáló szavakat, hogy bár a végeredményül előálló mondat minden beszélő esetében azonos volt, a kísérleti személyek maguk alkották meg a szöveget.

1. táblázat. Példák az /i:/, /a:/ és /u/ hang vizsgálatához használt példamondatokra a két beszédstílus és a mássalhangzó-kontextus tükrében (a célszavakat félkövérrel, a célmagánhangzót pedig aláhúzással jelöltük; a félspontán beszéd esetében a kivetített szöveget, illetve az elicitált mondatot is megadtuk)

Célhang	Beszéd-stílus	Mássalhangzó-kontextus	Példamondatok, illetve a képernyőre vetített szövegek
/i:/	félspontán	/r/_/t/	A kivetített szöveg: <i>A pincér(többesszám) + egy + asztal + terít</i> Az elicitált mondat: <i>A pincérek egy asztalt terítenek.</i>
	felolvasás	/r/_/t/	<i>Ponyvát terítenek az autóra, hogy megvédjék.</i>
/a:/	félspontán	/n/_/l/	A kivetített szöveg: <i>A nő + megnéz + kínálat(tárgyeset)</i> Az elicitált mondat: <i>A nő megnézi a kínálatot.</i>
	felolvasás	/t/_/r/	<i>A határozat értelmében már a miénk a lakás</i>
/u/	félspontán	/l/_/t/	A kivetített szöveg: <i>Nyár + mindig + hosszú + délután(többesszám)</i> Az elicitált mondat: <i>Nyáron mindig hosszúak a délutánok.</i>
	felolvasás	/n/_/l/	<i>A riporter megkérdezte a tanulókat is.</i>

A második beszéd feladatban teljes, grammatikus magyar mondatokat olvastak fel a kísérleti személyek (1. táblázat, 2. ábra). Ezekben a mondatokban azonos alakban jelentek meg a vizsgált szavak, mint a kísérlet első felében a félspontán módon megalkotott mondatokban. A levetített mondatokhoz itt is képeket rendeltünk a könnyebb értelmezhetőség érdekében.

A félspontán beszéd felvétele minden résztvevő esetében megelőzte a felolvasást, hiszen a vizsgált szavak megegyeztek a kísérlet két részében. Így a célszavakkal a vizsgált formájukban a félspontán beszédben találkoztak először a résztvevők, ami biztosította a körülményekhez képest legnagyobb fokban tervezetlen megszólalást.

Az összes ingert kétszeri ismétlésben rögzítettük minden kísérleti személlyel. Így kísérleti személyenként (3 magánhangzó × 3 példaszó és mondat × 2 ismétlés × 2 beszédstílus =) 36 tokenet, összesen pedig (36 token × 10 kísérleti személy =) 360 realizációt vizsgáltunk a kísérletben.

A felvételek az ELTE Fonetikai tanszékén, csendesített szobában készültek. Az akusztikai jelet az MTA–ELTE Lendület Lingvális Artikuláció Kutatócsoport eszközeivel készítettük: egy fejre rögzített omnidirekcionális kondenzátoros mikrofonnal és külső hangkártyával 44,1 kHz-en és 16 biten digitalizálva.

A célmagánhangzókat az anyagban a Praat programban (Boersma–Weenink 2019) kézzel címkéztük. A magánhangzók elejét és végét a második formáns megjelenése és eltűnése szerint jelöltük. Manuálisan mértük a célhangok első és második formánsát a magánhangzó időbeli középpontjában. Szükség esetén (a pontos mérés érdekében) a méréshez a beállításokat beszélőnként, de akár tokenenként is változtattuk. Az időtartamot automatikusan nyertük ki a címkesor alapján. A formánsokat a hangszín számszerűsítésére, az időtartamot pedig az ejtés időtartamának vizsgálatára használtuk.

1. ábra. A félspontán beszéd elicitálására használt bemutatások egyike, melyben az /i:/ célhang a /ri:t/ hangsorban szerepel



A pincér(többesszám) + asztal + terít

2. ábra. A felolvasásra használt bemutatások egyike, melyben az /i:/ célhang a /ri:t/ hangsorban szerepel



Ponyvát terítenek az autóra, hogy megvédjék.

A statisztikai elemzések az R programban (R Core Team, 2018) készültek az lme4 csomag (Bates et al., 2015) segítségével. A p - és F -értékeket Satterthwaite-approximáció segítségével nyertük, amely az lmerTest csomagban (Kuznetsova et al. 2017) elérhető. A függő változókra (időtartam, F_1 , F_2) külön-külön illesztettünk modellt (összesen tehát hármat), melyekben minden esetben a BESZÉDHANG (/a:/, /i:/ vagy /u/) és BESZÉDSTÍLUS (félspontán beszéd vagy felolvasás) változók hatását teszteltük (interakcióban is). A modellekben random hatásként (interceptként/metszéspontként, illetve meredekségként) felvettük a beszélőt, valamint random meredekségként a két tesztelt faktort, de csak akkor, ha ezek szignifikánsan növeltek a modell prediktív erejét (az anova függvény és az AIC információs kritérium alapján). A páros összevetéseket (Tukey-féle post hoc teszt) az lsmeans csomagban elérhető függvényekkel végeztük (Lenth, 2016).

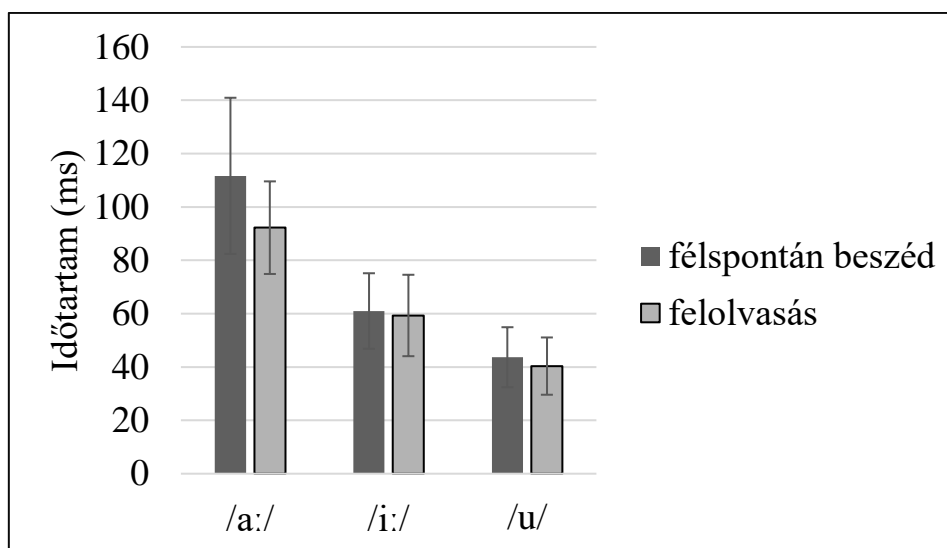
3. Eredmények

3.1. A vizsgált magánhangzók összesített adatai a beszédstílusok függvényében

Legelőször a három magánhangzó tulajdonságait együttesen elemeztük a két beszédstílus összehasonlításában.

A két beszédstílus vizsgálatának egyik szempontja a vizsgált magánhangzók időtartama volt. Lindblom H&H elmélete alapján arra számítottunk, hogy a felolvasásban mért időtartamok átlagosan nagyobbak, mint a félspontán beszédben. Az általunk mért időtartamok átlaga alapján valóban megállapítható volt egy közel 10 ms-os, tehát csekély mértékű különbség a két beszédstílus között, de meglepő módon a félspontán beszéd magánhangzói bizonyultak hosszabbnak ($\text{időtartam}_{\text{félspontán beszéd}} = 74 \pm 35$ ms; $\text{időtartam}_{\text{felolvasás}} = 65 \pm 26$ ms). A statisztikai elemzés szerint az időtartamra a BESZÉDHANG és BESZÉDSTÍLUS változók interakcióban hatottak ($F(2, 348) = 10,24$; $p < 0,05$), ami a post hoc elemzés szerint arra vezethető vissza, hogy a két beszédstílusban az /a:/ megvalósulásai eltértek egymástól ($p < 0,05$) ennek a paraméternek a mentén ($\text{időtartam}_{\text{/a:/, félspontán beszéd}} = 112$ ms; $\text{időtartam}_{\text{/a:/, felolvasás}} = 92$ ms; ami mintegy 20 ms eltérés) (3. ábra). A másik két magánhangzó esetében azonban lényegében nem találtunk különbséget (az /i:/ esetében 1 ms, az /u/ esetében 3 ms eltérést mértünk). A két beszédstílus között mért időtartambeli különbség tehát az /a:/ hangok megvalósulásai között mérhető eltérésnek volt köszönhető.

3. ábra. A három magánhangzó időtartama a beszédstílus szerint (átlag és szórás)



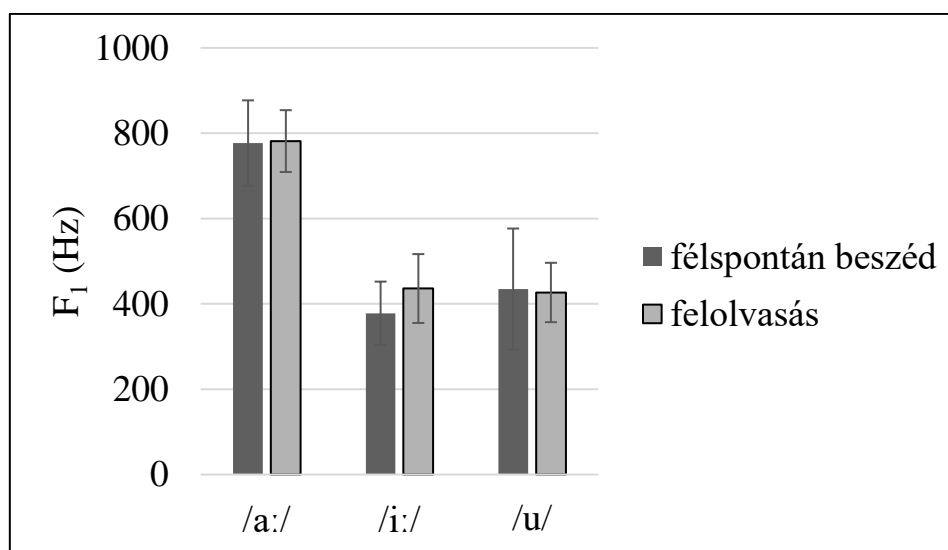
Ez után megvizsgáltuk a három magánhangzó első két formánsának frekvenciaértékét is. Lindblom H&H elmélete (Lindblom, 1990), illetve a korábbi vizsgálatok eredményei alapján lehetségesnek tartottuk, hogy a felolvasásban túlartikuláltabbak a magánhangzók, mint a félspontán beszédben. Ez az akusztikai szerkezetet illetően azt jelentené, hogy az /i:/ és /u/ esetében az F_1 értékei magasabbak, az /a:/ esetében pedig alacsonyabbak a felolvasásban, mint a félspontán beszédben, hiszen – rendre – az /i:/ és /u/ zártabb, az /a:/ pedig nyíltabb ebben a beszédstílusban.

Az összesített adatokban nem figyeltünk meg számottevő eltérést a két beszédstílus első formánsainak értékei között: az eltérés mindösszesen 11 Hz volt (F_1 félspontán beszéd = 539 ± 209 Hz; F_1 felolvasás = 550 ± 182 Hz). Ennek az egyik oka éppen az lehet, hogy a zárt és nyílt magánhangzók esetében várható ellentétes tendenciák kiátlagolódtak az összesített adatokban. A statisztikai elemzés szerint azonban nem ez a helyzet. Hasonlóan a korábbiakhoz az F_1 függő változót a BESZÉDHANG és BESZÉDSTÍLUS szignifikáns interakcióban befolyásolta úgy, hogy mindkét főhatás ugyancsak szignifikánsnak mutatkozott ($F(2, 347) = 5,30$; $p < 0,05$). Az előbbi főhatás megjelenése várt eredmény, és azt mutatja, hogy a magánhangzók eltérő minőségűek (hangszínűek). Utóbbi pedig arra utalna, hogy a beszédhangok eltérnek a beszédstílus mentén. Azonban az, hogy nemcsak ezek a főhatások, hanem az interakciós hatás is szignifikáns, valójában azt jelenti, hogy az egyik változó (BESZÉDSTÍLUS) a másik (BESZÉDHANG) szintjeinek mentén eltérően hatott, tehát nem minden magánhangzó esetében tért el a beszédhangok minősége a két beszédstílusban – ami egybevág(hatna) a várakozásokkal. A post hoc elemzések ugyanakkor azt mutatták, hogy a korábbiakhoz képest itt is csak az /i:/ megvalósulásaiban találni eltérést ($p < 0,05$) ($F_{1/i:/, \text{félspontán beszéd}} = 378$ Hz; $F_{1/i:/, \text{felolvasás}} = 436$ Hz; tehát az eltérés átlagosan 58 Hz), a másik két beszédhangban nem (az eltérés az /a:/ esetében átlagosan 4 Hz, az /u/ esetében

8 Hz volt a beszédstílusok között). Ráadásul az /i:/-ben is a feltételezésünkkel ellentétes tendenciát találtunk: a felolvasásban nem periferikusabban, azaz zártabban, hanem nyíltabban/alacsonyabb nyelvállással (magasabb F_1 -gyel) valósultak meg az /i:/-k, mint a félspontán beszédben (4. ábra).

Az F_2 összesített adataiban csakúgy, mint az F_1 esetében, nem találtunk jelentős különbségeket a beszédstílusok között (F_2 félspontán beszéd = 1847 ± 536 Hz; F_2 felolvasás = 1829 ± 547 Hz), az eltérés mindösszesen 18 Hz volt. Itt arra számíthattunk, hogy az /i:/ előrébb, azaz magasabb F_2 -vel az /u/ pedig hátrébb, alacsonyabb F_2 -vel valósul meg akkor, ha túlartikulált (azaz periferikusabb), amely tendenciák szintén kiátlagolódnak. A statisztikai elemzés – a BESZÉDHANG tényező bevonásával – azonban itt sem erősítette meg a várakozásokat: azt találtuk, hogy a BESZÉDSTÍLUSnak nem volt szignifikáns hatása az adatokra, pusztán a BESZÉDHANG-nak ($F(2, 347) = 5,30$; $p < 0,05$) (ami a beszédhangok eltérő minőségéből következő, fentebb az F_1 kapcsán is említett trivialisitás) (5. ábra). A fentiek szerint talált – a statisztikai elemzés alapján tehát lényegében nem releváns – különbségek a következőképpen alakultak: /i:/ 112 Hz (a félspontán beszédben mértünk magasabb értéket); /a:/: 22 Hz (a felolvasásban mértünk magasabb értéket); /u/: 8 Hz (a felolvasásban mértünk magasabb értéket).

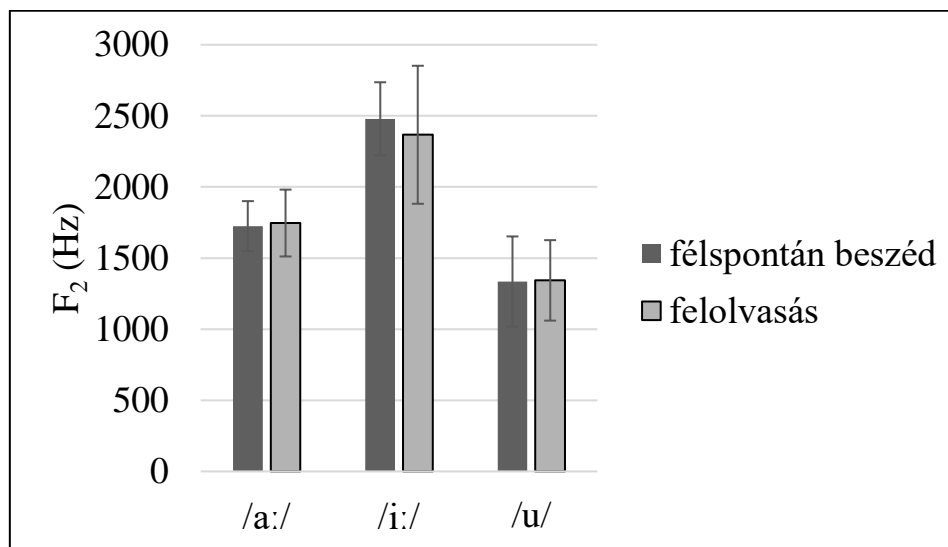
4. ábra. A három magánhangzó első formánsa a beszédstílus szerint (átlag és szórás)



A kutatás legfőbb kérdése az, hogy akusztikailag jobban megformáltak, túlartikuláltak-e a felolvasásban lévő beszédhangok, mint a félspontán beszédben. Ez az első és második formánsok átlagának együttes ábrázolásával jól érzékeltethető, ugyanis ez az ún. magánhangzótér az a tartomány, melyben az /a:/, /i:/ és /u/ szélsőséges pozíciókat elfoglalva mintegy körülrajzolja a beszélők által kihasznált akusztikai teret, és ahol a túlartikuláltság periferikusabb, szélsőségebb pozíciókat jelent. Ráadásul akkor, ha az ábrán a szórást is

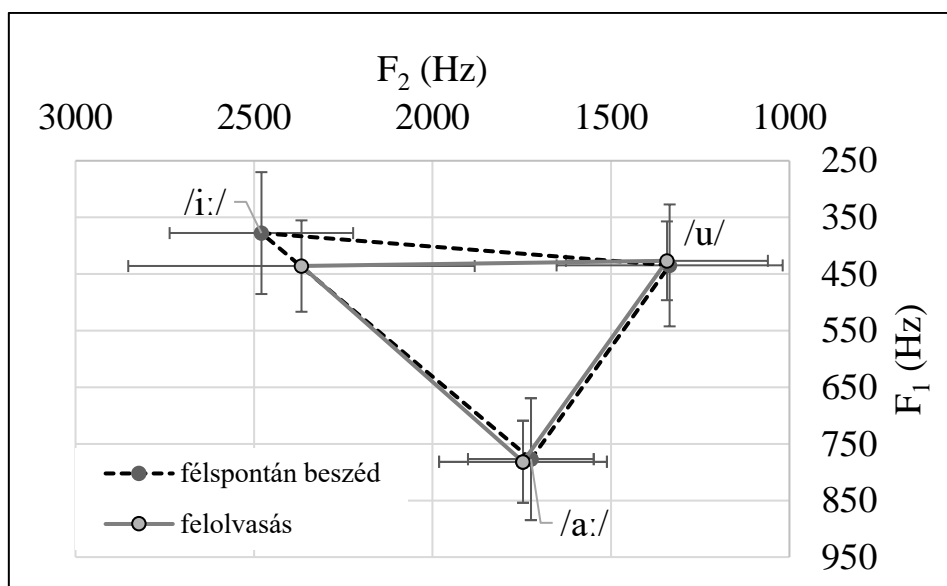
szerepeltetjük, az egyes beszédhang-kategóriák változatossága és elkülönülése is jól látható. Az adatokat ilyen módon ábrázoltuk a 6. ábrán.

5. ábra. A három magánhangzó második formánsa a beszédstílus szerint (átlag és szórás)



A 6. ábráról leolvasható, hogy az /a:/ és /u/ megvalósulásai a két beszédstílusban közel azonos átlagértékekkel és erős átfedésben valósultak meg a kategóriákon belül. Ezzel ellentétben azonban az /i:/-ben nagyobb különbségek figyelhetők meg a kategórián belül, de a beszédstílusok között: ahogyan azt már említettük, itt az akusztikai nyíltságban a várakozással ellentétben a félszpontán beszédben találtunk szélsőségesebb, túlartikuláltabb, azaz zártabb (inkább az ábra teteje felé elhelyezkedő) megvalósulásokat. Ezzel szemben az akusztikai palatalizáltságban – elől képzettségben – nem számottevőek az egyébként a várakozásokkal egyező tendenciák, tehát az, hogy a grafikon bal oldala felé helyezkedik el a felolvasásban mért adatok átlaga a félszpontán beszédhez képest.

6. ábra. A három magánhangzó az akusztikai magánhangzótérben (átlag és szórás)

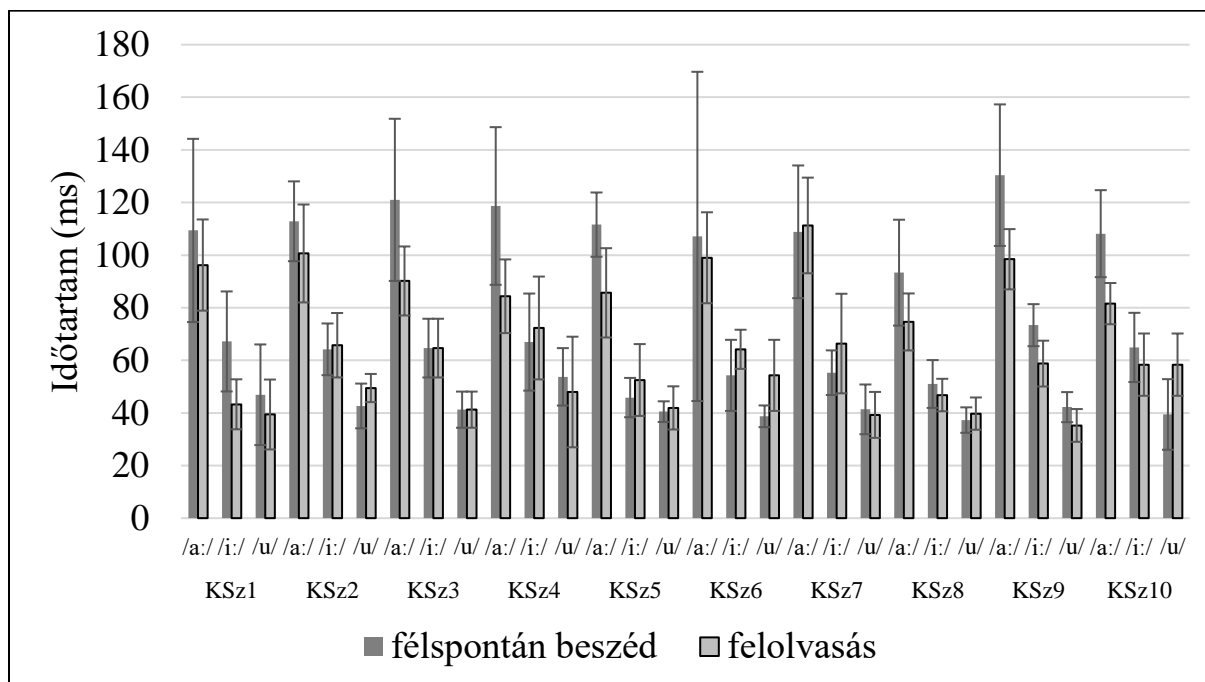


3.2. A vizsgált magánhangzók adatai kísérleti személyenként

Az adatelemzés második szakaszában az egyes kísérleti személyekre lebontva is megvizsgáltuk a kinyert adatokat azért, hogy megállapítsuk mennyiben voltak egységesek a beszélői stratégiák a kétféle beszédstílus közti különbségeket illetően (7–9. ábra).

Az időtartam adataiban (7. ábra) csak az /a:/-ban találtunk eltérést úgy, hogy a félspontán beszédben jelentek meg hosszabb magánhangzók. Érdekes látni, hogy bár tíz kísérleti személyből kilenc esetében ezt a tendenciát figyeltük meg, egy esetben a várakozásokkal egyező, bár valójában csekély mértékű különbséget mutató tendenciák is megjelentek. A 7. kísérleti személy esetében ugyanis hosszabb átlagos időtartamot mértünk felolvasásban (111 ms), mint félspontán beszédben (109 ms). Bár nagyjátlagban nem mutatott jelentős eltérést, egyes beszélők esetében az /i:/ is jelentősen hosszabban valósult meg az egyik beszédstílusban, mint a másikban: az ejtés tartama hosszabb volt a félspontán beszédben, mint felolvasásban az 1. kísérleti személy (67 ms vs. 43 ms; az eltérés: 24 ms), illetve a 9. kísérleti személy (73 ms vs. 59 ms; az eltérés: 14 ms) esetében. Hosszabbnak mutatkozott továbbá az /u/ a felolvasásban, mint a félspontán beszédben például a 6. és 10. beszélő ejtésében (a különbség rendre 39 és 54 ms volt).

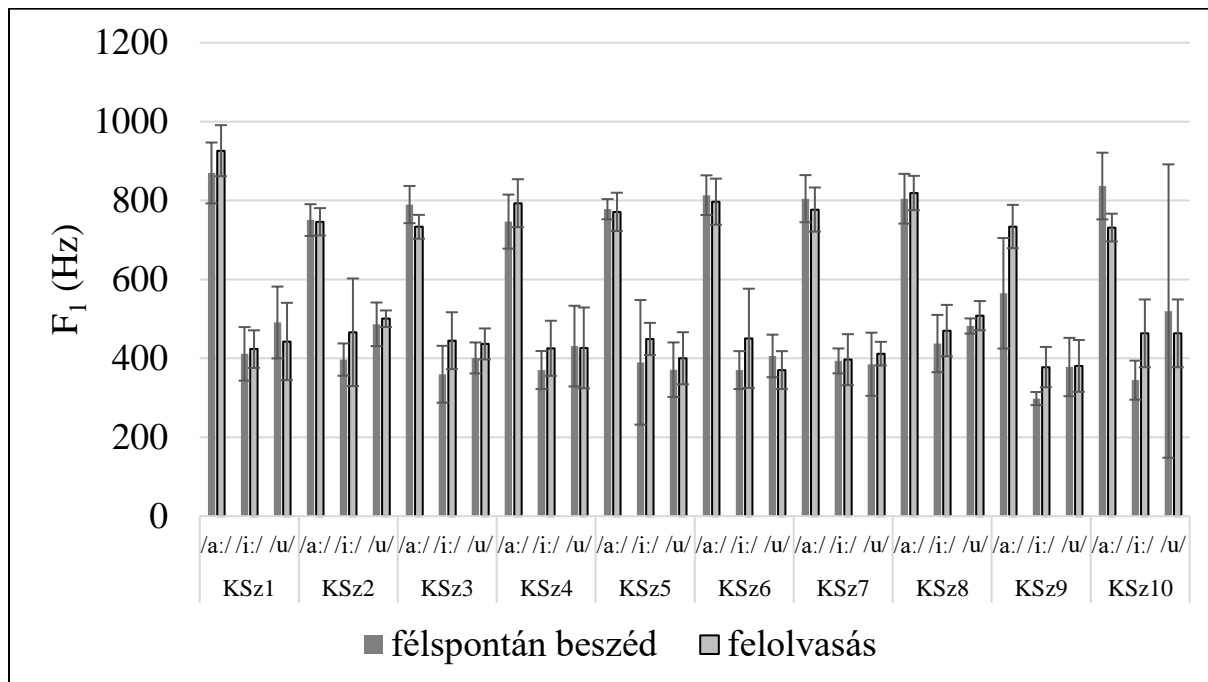
7. ábra. Az egyes kísérleti személyek magánhangzóinak időtartama a két beszédstílusban (átlag és szórás)



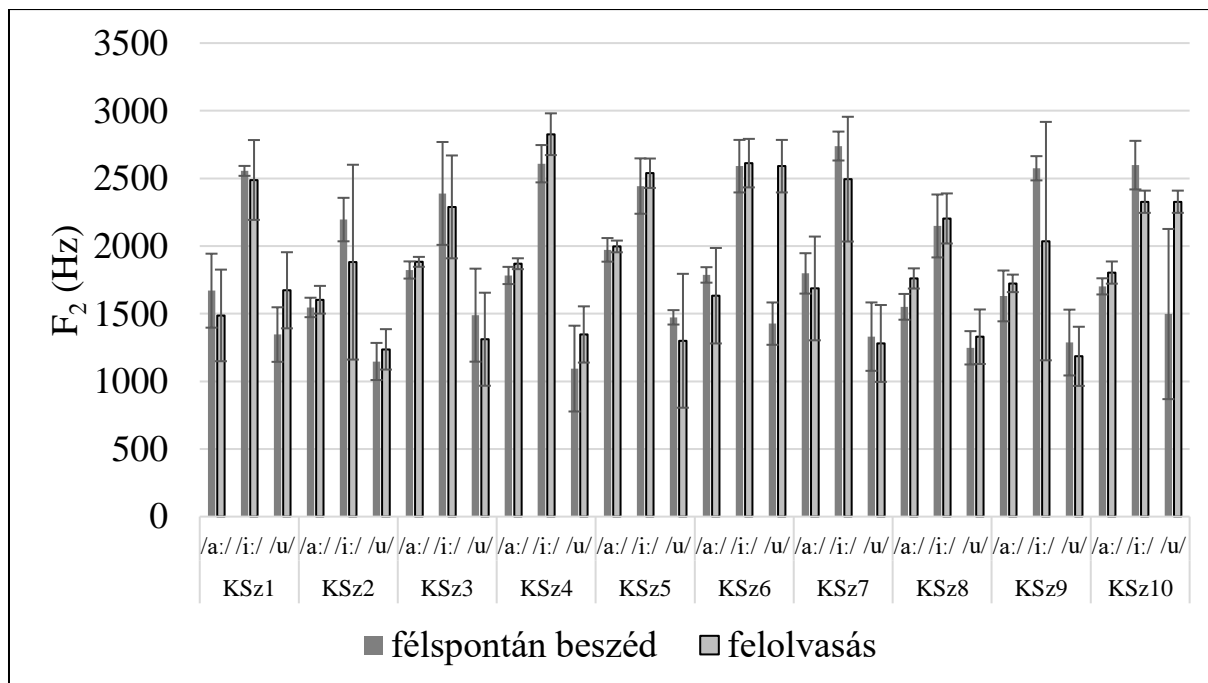
Az F_1 esetében az összesítésben csak az /i:/-ben találtunk eltérést: a felolvasásban magasabb volt ez az érték, azaz akusztikailag nyíltabban valósultak meg a hangzók. Az egyéni adatok (8. ábra) hozzávetőlegesen ezt a képet tükrözték, bár kiemelhető, hogy például a 9. és 10. beszélő az /a:/-k esetében is nagyobb eltéréseket produkált úgy, hogy míg a 9. a felolvasásban akusztikailag nyíltabb (F_1 félspontán beszéd = 565 vs. F_1 felolvasás = 764 Hz), a 10. a felolvasásban akusztikailag zártabb (F_1 félspontán beszéd = 837 Hz vs. F_1 felolvasás = 731 Hz) /a:/-kat ejtett, mint a másik beszédstílusban.

Végül a második formáns értékeit is megvizsgáltuk személyekre bontva (9. ábra). Az összesítés itt azt mutatta, hogy a BESZÉDSTÍLUS nem befolyásolja az adatok alakulását. Ehhez képest az egyéni trendekben látni eltéréseket, különösen az 1., 2., 6., 9. és 10. beszélő esetében. Az 1., 6. és 10. beszélő /u/-inak akusztikai megvalósítása felolvasásban palatálisabb volt, azaz a nem várt módon centralizáltabb (F_2 félspontán beszéd, KSz1 = 1346 Hz; F_2 felolvasás, KSz1 = 1673 Hz; F_2 félspontán beszéd, KSz6 = 1427 Hz; F_2 felolvasás, KSz6 = 2590 Hz; F_2 félspontán beszéd, KSz10 = 1497 Hz; F_2 felolvasás, KSz10 = 2328 Hz), a 2. és 9. beszélő esetében az /i:/ felolvasásban ugyancsak centralizáltabbnak (hátrébb képzettnek) bizonyult (F_2 félspontán beszéd, KSz2 = 2196 Hz; F_2 felolvasás, KSz2 = 1881 Hz; F_2 félspontán beszéd, KSz9 = 2575 Hz; F_2 felolvasás, KSz9 = 2036 Hz).

8. ábra. Az egyes kísérleti személyek magánhangzóinak első formánsai a két beszédstílusban (átlag és szórás)



9. ábra. Az egyes kísérleti személyek magánhangzóinak második formánsai a két beszédstílusban (átlag és szórás)



4. Következtetések

A tanulmány a félspontán beszéd és felolvasás közötti különbségeket vizsgálta három magánhangzó, az /a:/, /i:/ és /u/ időtartama, illetve első és második formánsának frekvenciaértéke alapján.

Időtartamát tekintve csak az /a:/ esetében láttunk eltérést a beszédstílusok között úgy, hogy ezek a magánhangzók a félspontán beszédben hosszabban valósultak meg, mint felolvasásban. Ez ellentmondhat a várakozásoknak, hiszen a túlartikuláció, azaz a periférikusabb ejtés kivitelezése hosszabb tartamot igényelne. Az eredmény alapján feltételezhető, hogy félspontán beszédben valamilyen mértékben lassabb lehetett a beszélők artikulációs tempója (vö. pl. Cucchiari et al., 2002; Váradi, 2010; de pl. Bóna, 2011 ennek a feltevésnek ellentmondó adatokkal is szolgál), ami következhetett a feladat jellegéből, hiszen a kísérlet komplex tervezési feladat elé állította az adatközlőket, és a nagyobb mentális erőfeszítés, ami lassíthatta az artikulációs tempót.

Az F_1 frekvenciaértéke ugyancsak kizárólag az /i:/ esetében mutatott eltérést a beszédstílusok között, ráadásul úgy, hogy a felolvasásban magasabb értékeket kaptunk. Ez azt jelenti, hogy az /i:/ a várakozásokkal ellentétes módon a félspontán beszédben valósult meg periférikusabban, tehát ebben a beszédstílusban állt közelebb a túlartikulált stílushoz, míg a felolvasásban nyíltabb volt az ejtése ennél.

Végezetül az F_2 frekvenciaértéke lényegében egyezett a két beszédstílusban.

Az eredmények alapján a kiinduló kérdésre (az anyag és az adatok korlátait figyelembe véve) az a választ adható, hogy ha van is a túl- és alulartikuláltság, illetve a tervezettség dimenzió között megfelelés, az nem egyértelmű, és nem lehet feltétlennek tekinteni, hiszen a magyar nyelv három szélsőséges magánhangzó-minőségét vizsgálva nem volt tapasztalható ilyen jellegű összefüggés. Az említett korlátok miatt azonban a kérdés további vizsgálat tárgya lehet: megfontolandó, miként tágítható a tervezettség dimenzióinak a hangzásban (és kifejezetten a beszéd szegmentális vetületében) megjelenő sajátosságainak vizsgálata. Zárásképpen ezekről, az eredmények alakulásához feltehetően hozzájáruló fentebb említett korlátokról, illetve a további lehetőségekről ejtünk szót.

Az itt félspontán beszédként vizsgált irányított megszólalási mód bizonyos tekintetben biztos, hogy nem meríti ki a spontán beszéd fogalmát, hiszen a jelen vizsgálatban a nyelvi formák részben adottak voltak a beszélők számára. Ez pedig okozhatta azt, hogy ezekre a már adott (szó)elemekre a beszélők nagyobb figyelmet fordítottak a meghangosításban annál, mint amit a teljesen tervezetlen megszólalásban várhatnánk. Másként megfogalmazva feltehető, hogy az itt félspontán beszédként vizsgált beszédstílusban a tervezés a mentális erőforrásokat kisebb mértékben vonta el, mint akkor tette volna, ha a lexémák (azaz a szavak hangzói formáinak reprezentációi) nem, csak a lemmák (azaz a szójelentések hangalak nélküli reprezentációi) adottak a közlésben (Warren, 2021), és a

hangalak előhívása is a feladat része. Ez valószínűleg ahhoz a beszédstílushoz is közelebb áll, amit a szakirodalomban spontán beszédként hivatkoznak, bár megjegyzendő, hogy ez a kérdés nem teljesen egyértelmű, hiszen általában kevésbé specifikáltak az egyes beszédstílus-kategóriák a tervezés egyes pontos részfolyamatait és azok aktivitását illetően.

Az itt felvetett probléma részben kiküszöbölhető, ha olyan módszereket használunk az elicitálásban, ahol a nyelvi forma nem, csak a jelentés jelenik meg vizuálisan, például képi formában. A fő ok, ami miatt ugyanakkor a jelen vizsgálatban nem ezt az eljárást választottuk, az, hogy ez a módszer nagy teret enged az asszociációknak, ennél fogva könnyen előfordulhat, hogy a résztvevők nem az elicitálni kívánt lexikai elemet aktiválják és ejtik ki. Az itt használt, félspontánnak tekintett beszéd használata mellett ezen túlmenően azért is döntöttünk, mert ennek segítségével jól kontrollált fonetikai környezetben tudtuk vizsgálni a kérdéses hangokat, így számos tényezőt kiküszöböltünk az ejtési változatosságból, amik másként elfedhették volna a beszédstílus hatását. Lehetséges lehet azonban a jövőben olyan módszerek kidolgozása, melyek kizárólag képi ingerrel támogatott előhíváson alapulnak.

Úgy véljük, hogy a kísérletünk eredményei a fentebb említett korlátok ellenére is rámutatnak arra, hogy a közlések hangzásának az ejtésbeli megformáltság (és az erre fordított erőfeszítés) mentén feltételezhető skálája nem tekinthető azonosnak a beszéd tervezettségével, illetve a tervezés és meghangosítás egyidejűsége mentén felállítható dimenzióval. Ez az eredmény pedig hozzájárul a beszédhangzás kialakítását befolyásoló tényezőkről alkotott ismereteink mélyebb megértéséhez.

Irodalom

- Bates, D., Mächler, M. Bolker, B. & Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Beke András** (2008). A felolvasás és a spontán beszéd alaphangszerkezetének vizsgálata. *Beszédkutató*, 2008, 93–107.
- Blaauw, E.** (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication*, 14, 359–375. doi: 10.1016/0167-6393(94)90028-0
- Boersma, P. & Weenink, D.** (2019). Praat: doing phonetics by computer [Computer program]. Version 5.4. Letöltés: <http://www.praat.org/>
- Bóna Judit** (2011). A különböző beszédstílusok az akusztikai-fonetikai és a percepció vizsgálatok tükrében. *Alkalmazott Nyelvtudomány*, 11, 39–48.
- Cucchiari, C., Strik, H. & Boves, L.** (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, 111, 2862–2973. doi: 10.1121/1.1471894
- Deme, A., Bartók, Márton, Grácsi, T. E., Csapó, T. G. & Markó A.** (előkészületben). The effect of pitch-accent on the acoustic and articulatory variability of vowels: a real-word EMA study.
- Deme Andrea, Bartók Márton, Csapó Tamás Gábor, Grácsi Tekla Etelka, Juhász Kornélia & Markó Alexandra** (2022a). A magánhangzók centralizációja és produkciós homogenitása az előrefelé és hátrafelé ható magánhangzók közti koartikulációban – artikulációs és akusztikai adatok. *Általános Nyelvészeti Tanulmányok*, 34, 15–50. Letöltés: <https://mersz.hu/altalanos-nyelveszeti-tanulmanyok-xxxiv//>

- Deme Andrea, Bartók Márton, Grácsi Tekla Etelka, Csapó Tamás Gábor, Juhász Kornélia & Markó Alexandra** (2022b). A koartikulációs ellenállás és agresszió hangsúlyos helyzetű magánhangzókbán: artikulációs és akusztikai adatok. *Nyelvtudományi Közlemények*, 118, 257–288. doi: 10.15776/NyK.2022.118.9
- Deme Andrea – Grácsi Tekla Etelka – Horváth Viktória – Markó Alexandra** 2011. *Magánhangzó-realizációk az olvasásban és a spontán beszédben* (előadás). Elhangzott: *Beszédkutatás* 2011.
- Deme Andrea, Kohári Anna, Reichel Uwe D., Szalontai Ádám & Mády Katalin** (2019a). A magánhangzós hosszúsági fonológiai kontraszt a dajkanyelvben a csecsemő életkorának függvényében. *Beszédkutatás*, 27, 221–242. doi: 10.15775/BeszKut.2019.221-242
- Deme Andrea, Bartók Márton, Grácsi Tekla Etelka, Csapó Tamás Gábor & Markó Alexandra** (2019b). A mondathangsúly hatása a magánhangzók megvalósulásának változatosságára. *Nyelvtudományi Közlemények*, 115, 199–232. doi: 10.15776/NyK/2019.115.7
- Deme, A., Bartók, M., Csapó, T. G., Grácsi, T. E. & Markó, A** (2021). Acoustic and articulatory vowel variation as quality shift and increased variance in anticipatory and carryover vowel-to-vowel coarticulation. In M. Tiede., D. H. Whalen, V. Gracco (szerk.), *Proceedings of the 12th International Seminar on Speech Production* (32–35). New Haven (CT), Amerikai Egyesült Államok: Haskins Press.
- Deme Andrea** (2016). *Magánhangzók ejtése és észlelése a szopránéneklésben*. Budapest: ELTE Eötvös Kiadó.
- de Ruiter, Laura E.** (2015). Information status marking in spontaneous vs. read speech in story-telling tasks – Evidence from intonation analysis using GToBI. *Journal of Phonetics*, 48, 29–44. doi: 10.1016/j.wocn.2014.10.008
- DiCanio, Ch., Nam, H., Amith, J. D., Castillo García, R. & Whalen, D. H.** (2015). Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics*, 48, 45–59. doi: 10.1016/j.wocn.2014.10.003.
- Docherty, G., Gonzalez, S. Mitchell, N. & Foulkes, P.** (2019). An acoustic analysis of short front vowel realizations in the conversational style of young English speakers from Western Australia. In Calhoun, S., Escudero, P., Tabain, M. & Warren, P. (eds), *Proceedings of the 19th International Congress of Phonetic Sciences (1759–1783)*. Melbourne, Australia.
- Gósy Mária** (1998). A beszédtervezés és a beszédkivitelezés paradoxona. *Magyar Nyelvőr*, 122, 3–15.
- Grácsi Tekla Etelka és Horváth Viktória** (2010). A magánhangzók realizációja spontán beszédben. *Beszédkutatás*, 2010, 5–16.
- Howell, P. & Kadi-Hanifi, K.** (1991). Comparison of prosodic properties between read and spontaneous speech material. *Speech Communication*, 10, 163–169. doi: 10.1016/0167-6393(91)90039-V
- Huber, J. E. & Darling M.** (2011). Effect of Parkinson’s disease on the production of structured and unstructured speaking tasks: respiratory psychologic and linguistic considerations. *Journal of Speech, Language, and Hearing Research*, 54, 33–46. doi: 10.1044/1092-4388(2010/09-0184)
- Kohári Anna** (2018). *Időzítési mintázatok a magyar beszédben*. Budapest: ELTE Eötvös Kiadó.
- Krepsz Valéria** (2016). Fonetikai hasonlóságok és különbözőségek a beszéd típusokban. *Fonetikai olvasókönyv*. Budapest: ELTE Fonetikai Tanszék. 178–182.
- Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B.** (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26.
- Ladefoged, P., Kameny, I., and Brackenridge, W.** (1976). Acoustic effects of styles of speech. *Journal of the Acoustical Society of America*, 59, 228–231. doi: 10.1121/1.380856
- Lenth, R. V.** (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, 69, 1–33.
- Levelt, W. J. M.** (1989). *Speaking: From intention to articulation*. A Bradford Book. Cambridge–London: MIT Press.
- Levelt, W. J. M.** (1999). Producing spoken language: a blueprint of the speaker. In Brown, C. M. & Hagoort, P. (eds.), *The neurocognition of language* (83–122). Oxford: Oxford University Press,
- Lieberman, P., Katz, W., Jongman, A., Zimmerman, R. & Miller, M.** (1984). Measures of the sentence intonation of read and spontaneous speech in American English. *The Journal of the Acoustical Society of America*, 77. doi: 10.1121/1.391883

- Lindblom, B.** (1990). Explaining phonetic variation: a sketch of the H&H theory. In Hardcastle, W.J., Marchal, A. (eds), *Speech Production and Speech Modelling* (403–439). NATO ASI Series, 55. Dordrecht Dordrecht: Springer. doi: 10.1007/978-94-009-2037-8_16
- Markó Alexandra** (2006). A spontán beszéd – monologikus és dialogikus szövegek. *Acta Academiae Pedagogicae Agriensis. Nova Series Tom. XXXIII. Sectio Linguistica Hungarica.* 66–78.
- Markó Alexandra** (2012). A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In Markó Alexandra (szerk.), *Beszédtudomány: Az anyanyelv-sajátítástól a zöngelkedési időig* (277–303). Budapest: MTA Nyelvtudományi Intézet – ELTE BTK.
- Markó Alexandra** (2015). *A spontán beszéd prozódiai szerkezete.* Budapest: Akadémia Kiadó.
- Markó Alexandra, Bartók Márton, Grácsi Tekla Etelka, Deme Andrea & Csapó Tamás Gábor** (2018). Mondathangsúlyos és hangsúlytalan helyzetű magánhangzók néhány artikulációs és akusztikai jellemzője a magyarban. *Beszédkutatás*, 26, 85–109.
- Mehta G. & Cutler A.** (1988). Detection of target phonemes in spontaneous and read speech. *Language and Speech*, 31, 135–156.
- R Core Team** (2018). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. Letöltés: <https://www.R-project.org/>
- Stevens, K. N.** (1998). *Acoustic phonetics.* Cambridge, Massachusetts–London: The MIT Press.
- Váradí Viola** (2010). A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutatás*, 2010, 100–110.
- Wacha Imre** (1974). Az elhangzó beszéd főbb akusztikus stíluskategóriáiról. *Általános Nyelvészeti tanulmányok*, X., 203–216.
- Wagner, P., Trouvain, J. & Zimmerer, F.** (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, 48, 1–12. doi: 10.1016/j.wocn.2014.11.001
- Wang, Q.** (2019). Stylistic variation in vowel production: evidence from China English. In Calhoun, S. Escudero, P., Tabain, M. & Warren, P. (eds), *Proceedings of the 19th International Congress of Phonetic Sciences* (1089–1093). Melbourne, Australia.
- Warren, P.** (2012). *Introducing Psycholinguistics.* Cambridge: Cambridge University Press.