

A karakterábrázolás alapfrekvencia-jellemzői

Huszár Anna

ELTE BTK Nyelvtudományi Doktori Iskola, MTA Nyelvtudományi Intézet
huszar.anna@nytud.hu

Kivonat: A meseolvasás egy speciális területe a beszédnek. A mesét olvasók beszédük változatosságával ragadják meg, illetve tartják fenn a hallgatóság figyelmét, az egyes karakterek jellegzetességeit a hangjukkal jelenítik meg. Ebből adódóan a meseolvasás szupraszegmentális szerkezete eltérhet más beszédműfajokétól (Fackrell et al. 2000; Theune et al. 2006). A mesén belül az egyes karakterek, illetve bizonyos karaktertípusok megjelenítésének módja is különbözik (Doukhan et al. 2011). A jelen kutatásban sztereotíp tulajdonságokkal rendelkező karakterek ábrázolását vizsgáltam. Elemeztem az átlagos alapfrekvenciát, a hangterjedelmet és a zöngeminőséget. Az eredmények szerint szignifikánsan eltért az átlagos alapfrekvencia és hangterjedelem a karakterek mentén. A irreguláris részek aránya is tendenciaszerűen eltérést mutatott a karakterek mentén.

1 Bevezetés

A beszédprodukción számos tényező befolyásolhatja, többek között a beszélő személye, a hallgatóság, a beszédtema, a beszédhelyzet és a közlés célja (Gósy 2004). A beszélő életkora, neme, szociológiai háttere, személyisége, aktuális érzelmi és fizikai állapota mind hatással lehetnek a beszédprodukcóra (Balázs–Bóna 2016; Bóna 2016; Gocsál 1998; Gocsál–Huszár 2003; Gósy 2004; Gósy–Krepsz 2015; Gósy–Krepsz 2018; Gyarmathy 2007; Németh 2013). A beszélő személyén kívül a hallgatóság is meghatározó lehet, a résztvevők száma, illetve az, hogy ismerik-e egymást, illetve milyen viszonyban állnak egymással (pl. barátok, anya-gyermek, tanár-diák stb.). A beszéd típus is meghatározó tényező, más a spontán beszéd, és más a felolvasás prozódiai szerkezete (Bata–Grácsi 2009; Bóna 2011; Jacewicz et al. 2010; Markó 2009; Váradi 2010). Prozódiai vagy szupraszegmentális szerkezet alatt értem a beszéd azon nem szegmentális jellemzőit, melyek az idő, a frekvencia és az intenzitás változásaival írhatók le (Hardcastle–Laver 1997).

Az egyes beszédműfajok szintén eltérhetnek egymástól (pl. hírolvasás, politikai beszéd, reklámszöveg, regény és meseolvasás) több paraméter mentén is: beszédtempó, artikulációs tempó, szünetezés, beszédszakaszok (azaz a szünetről szünetig tartó egységek) hossza, alapfrekvencia (átlagos alapfrekvencia és hangterjedelem) és formánszerkezet (Abe–Mizuno 1994; Eskenazi 1993; Hirschberg 2000; Roekhaut et al. 2010).

Huszár Anna: A karakterábrázolás alapfrekvencia-jellemzői. In Váradi Tamás (sorozatszerkesztő), Ludányi Zsófia, Grácsi Tekla Etelka (szerkesztő): *Doktoranduszok tanulmányai az alkalmazott nyelvészet köréből 2019. XIII. Alkalmazott Nyelvészeti Doktoranduszkonferencia*. Budapest: MTA Nyelvtudományi Intézet. 2019. 59–71. DOI: 10.18135/Alknyelvdok.2019.13.5

A meseolvasás egy speciális területe a beszédnek, prozódiai szerkezete eltér más beszédműfajokétól (Fackrell et al. 2000; Theune et al. 2006). A mesét olvasók beszédük változatosságával tartják fenn a hallgatóság figyelmét, az egyes karakterek jellegzetességeit a hangjukkal jelenítik meg.

Többen vizsgálták a mese prozódiai szerkezetét gépi felolvasás létrehozásának céljából. Ahhoz, hogy minél tökéletesebb gépi felolvasóalgoritmust hozzanak létre, szükséges minél pontosabban ismerni a mese szupraszegmentális szerkezetét, több akusztikai elemzés született a témában (Fackrell 2000; Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015). A meseolvasás temporális szerkezete jelentősen eltér más beszédműfajokétól, lassabb beszéd- és artikulációs tempó, hosszabb szünetek, magasabb szünetarány és hosszabb beszédszakaszok jellemzik (Fackrell et al. 2000; Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015). A meseolvasás alapfrekvencia-jellemzői is különböznek a többi beszédműfajokétól, magasabb átlagos alapfrekvencia és nagyobb hangterjedelem, nagyobb fokú hangsúlyozás és az intenzitás nagyobb fokú variabilitása jellemzi (Jokisch et al. 2005; Roekhaut et al. 2010; Theune et al. 2006; Verma et al. 2015).

Magyar kutatások is vizsgálták a mese prozódiai szerkezetét (Imre 2005; Olaszky 2005). Egy kutatásban Olaszky (2005) a hírolvasással, novella és reklámok felolvasásával vetette össze a mesemondás prozódiját. Az eredmények szerint a mesemondás időszerkezete a magyar nyelvben is eltér más beszédműfajokétól: a mesemondásra lassabb beszéd- és artikulációs tempó jellemző a hírolvasáshoz képest. A mesemondásra általában rövidebb, jól tagolható mondatok, több mondatbelei szünet jellemző, melyek időtartama hosszabb, mint más beszédműfajokban. A mesemondás frekvenciaszerkezete szintén eltér más beszédműfajokétól, általában magasabb és variábilisabb alapfrekvencia és nagyobb hangterjedelem jellemzi.

A mesemondás egyik jellemzője, hogy a mesemondó nemcsak narrátor, hanem sok esetben ő jeleníti meg beszédével az összes karaktert. A mesemondó prozódiai eszközöket használ annak érdekében, hogy az egyes karaktereket elkülönítse egymástól, illetve kifejezően ábrázolja azokat. Adell és munkatársai (2005) többek között a meseolvasás szupraszegmentális szerkezetét vetették össze más beszédműfajokkal (érzelmes beszéd és parlamenti beszéd). Többek között mérték alapfrekvencia-jellemzőket: az alapfrekvencia átlagát, szórását, minimumát, maximumát, a hangterjedelmet és a dallammenetet elemezték. Az intenzitásnak szintén nézték az átlagát, szórását, minimumát és maximumát is. Az artikulációs tempót és a szünetek időtartamát is elemezték. A mese különböző részeinek szupraszegmentális szerkezetét is vizsgálták, és megállapították, hogy a narratív részek, a leírások és az egyes karakterek megszólalásai prozódiailag elkülönülnek egymástól.

Egy másik kutatás (Montaño et al. 2013) szintén a meseolvasás narratív, leíró és dialogikus részeinek prozódiai szerkezetét vizsgálta. A narrátorhoz tartozó beszédegységeket különböző részekre bontva elemezték: semleges, leíró, karakter megszólalása utáni és többféle érzelmi töltetű rész. Az átlagos alapfrekvencia és annak szórásának, az átlagos intenzitás és az artikulációs tempó alakulását vizsgálták. Az említett kategóriák között főleg az alapfrekvenciában (átlag és szórás tekintetében is) és az intenzitás-ban találtak eltéréseket, az artikulációs tempóban a kategóriák többsége között nem volt jelentős különbség. A karakterek megszólalása utáni részekre alacsonyabb alapfrekvencia és intenzitás volt jellemző. A leíró részekre magasabb alapfrekvencia volt jellemző (átlag és szórás tekintetében is), intenzitásában azonban nem különböztek jelentősen a semleges stílusú részeketől. A különböző érzelmi töltettel rendelkező részek prozódiai szerkezetét is elemezték, és lassabb artikulációs tempót mérték a semleges részekhez viszonyítva. Továbbá magasabb átlagos intenzitást és alapfrekvenciát mérték (illetve

az alapfrekvencia nagyobb szórását) a következő érzelmek esetében: düh, öröm, meglepetés és félelem. A szomorúságra alacsonyabb átlagos intenzitás és alapfrekvencia, illetve az alapfrekvencia kisebb szórása volt jellemző.

Doukhan és munkatársai (2011) szintén a meseolvasás szupraszegmentális sajátosságait vizsgálták. Tizenkét mese felolvasásában elemezték az artikulációs tempót, a szünetek százalékos időtartamarányát, az átlagos alapfrekvenciát, az átlagos intenzitást, a szótagok közti átlagos alapfrekvencia- és intenzitáskülönbségeket és a hangterjedelmet. A narrátor szakaszait a következő részekre bontva vizsgálták: cím, bevezető szöveg, a kiváltó esemény bemutatása, jelenetek, refrén, utószó. Eredményeik szerint a cím prozódiai sajátosságai jelentősen elkülönültek a narrátor többi szakaszának sajátosságaitól: a címnél mérték átlagosan a legmagasabb intenzitást, a legnagyobb szótagok közti alapfrekvencia-különbséget, a legnagyobb hangterjedelmet, és a leglassabb artikulációs tempót (a szünetarány is a cím esetében volt a legalacsonyabb, azonban a címek általában lényegesen rövid szövegek). A bevezetésben általában nagy volt a szótagok közötti alapfrekvencia-különbség, ami a hallgatóság figyelmének megragadásával is magyarázható. A refrénekre alacsony szünetidőtartam-arány volt jellemző, illetve alacsony volt a szótagok közti alapfrekvencia-különbség is a többi narrátori részhez képest. Az utószóra szintén alacsony szótagok közti alapfrekvencia-különbséget mérték, alacsony alapfrekvenciát és intenzitást, kis hangterjedelmet és intenzitásváltozást, illetve lassú artikulációs tempót. A narrátor szakaszain kívül a karakterek megjelenítését is vizsgálták. A megjelenített karaktert neme és életkora (gyerekszereplő, felnőtt vagy idős) alapján csoportosították. A meseolvasó beszédének alapfrekvenciája és intenzitása magasabb volt, amikor karaktereket jelenített meg, mint amikor a narrátor részeit (ez alól a nagy medve kivétel, ennél a karakternél mérték a legalacsonyabb alapfrekvenciát). A gyerekszereplők esetében általában magasabb és egyenletesebb alapfrekvenciát figyeltek meg; a fiúknál magasabb, a lányoknál alacsonyabb intenzitást. A karakterek az alapján is elkülöníthetőek voltak prozódiai szempontból, hogy fő- vagy mellékszereplők voltak. A gyerek főszereplőkre általában jellemző volt a magasabb átlagos alapfrekvencia és a kisebb szótagok közti alapfrekvencia-különbség. A felnőtt mellékszereplők megjelenítése a narrátoréhoz volt hasonló prozódiailag. Az idősebb szereplők esetében nagyobb volt a szótagok közti alapfrekvencia-különbség. Nem elhanyagolható azonban a megjelenítendő karakterek életkora sem, a szupraszegmentális szerkezetre a karakterek életkora és a szerep típusa együttesen lehetett hatással.

A fenti kutatásokból látható, hogy a meseolvasás prozódiai szerkezete több tényező mentén is különbözik a többi beszédműfajétól. A mese egyes részei (pl. narrátor és a különféle karakterek megjelenítése) több prozódiai jegyben is eltérhetnek egymástól. A fenti kutatások főleg adott mesék felolvasásának prozódiai elemzését tűzték ki célul. A narrátor szerepén kívül az egyes karakterek megjelenítését is vizsgálták, a karakterek megjelenése egy mesében azonban nem egyenletes, vannak karakterek (pl. főszereplők), akik gyakran szóhoz jutnak a mese során, vannak azonban olyan karakterek is (pl. mellékszereplők), akik csak ritkán kapnak szót, ez nehezíti az összehasonlíthatóságukat. Ezen túl az egyes karakterek szövege eltérő, tehát ez is nehezíti a megjelenítésük vizsgálatát. A prozódia befolyásolhatja a mondatok hossza, szótagszáma, összetettsége, hogy tartalmazznak-e alá- vagy mellérendeléseket, így az egyes karakterek közti különbségek egy része adódhat az eltérő mondat szerkezetből is. Az összehasonlítás érdekében célszerű ugyanazon szövegen vizsgálni az eltéréseket.

A jelen kutatás célja sztereotíp tulajdonságokkal rendelkező karakterek ábrázolásának vizsgálata adott szövegen. A karakterek megválasztásánál törekedtem arra, hogy

minél több tulajdonság mentén eltérjenek egymástól, a következő karakterek megjelenítését vizsgáltam: IZGÁGA HÖRCSÖG, LUSTA LAJHÁR, KECSSES ŐZ és RAVASZ RÓKA. Az elemzés során arra a kérdésre keresem a választ, hogy mennyiben térnek el az egyes alapfrekvencia-jellemzők a karakterek mentén, illetve a semleges felolvasáshoz viszonyítva. Hipotéziseim szerint a karakterek mentén eltérés mutatkozik (i) az átlagos alapfrekvenciában, (ii) a hangterjedelemben és (iii) az irregularitás megjelenésében is.

Az IZGÁGA HÖRCSÖGHÖZ társítható az izgatottság és a vidámság. Az izgatottságra az alapfrekvencia nagyobb mértékű variabilitása (Trouvain–Barry 2000), a vidámságra pedig magasabb és variábilisabb alapfrekvencia jellemző (Scherer et al. 1991). Tehát az IZGATOTT HÖRCSÖG esetében magasabb alapfrekvenciát és nagyobb hangterjedelmet várok a kontroll kondícióhoz viszonyítva.

A LUSTA LAJHÁRHOZ társítható a szomorúság, a csalódottság vagy akár az unottság. A szomorúságot jellemezheti alacsonyabb és kevésbé variábilis alapfrekvencia, kisebb hangterjedelem és nagyobb fokú irregularitás (Burkhardt–Sendlmeier 2000; Paeschke 2004; Scherer et al. 1991). Az unottságra jellemző lehet alacsonyabb alapfrekvencia, nagyobb hangterjedelem és nagyobb fokú irregularitás (Burkhardt–Sendlmeier 2000; Paeschke 2004; Scherer et al. 1991). A csalódottságra kisebb hangterjedelem jellemző (Maekawa 1998). Tehát a LUSTA LAJHÁR esetében alacsonyabb alapfrekvenciát és nagyobb fokú irregularitást várok a kontroll kondícióhoz viszonyítva.

A KECSSES ŐZHÖZ társítható a nőieség. A magasabb alapfrekvenciájú hangot nőiesebbnek ítélik meg (Feinberg et al. 2008). A KECSSES ŐZ esetében tehát magasabb alapfrekvenciát várok a kontrollkondícióhoz képest.

A RAVASZ RÓKAHOZ társítható az irónia. Az iróniára jellemző a magasabb és variábilisabb alapfrekvencia (Bryant–Fox 2005). A RAVASZ RÓKA esetében tehát magasabb alapfrekvenciát és nagyobb hangterjedelmet várok.

2 Módszertan

Adatközlőim középiskolás dráma szakos diákok voltak (nyolc tizenegyedik osztályos lány). Feladatuk egy adott szöveg felolvasása volt, melyet előre megkaptak. A felolvasandó szöveg a kísérletre készült, öt kisebb egységből áll. Összesen 40 mondatot tartalmaz, melyek között vannak rövidebb (kb. 5-6 szótagos) és hosszabb (kb. 9-11 szótagos) mondatok is. A szöveg felolvasása körülbelül egy percet vett igénybe. A szöveget először semleges stílusban kellett felolvasniuk (kontroll kondíció), majd a négy különböző sztereotíp tulajdonsággal rendelkező állatfigurákat (RAVASZ RÓKA, LUSTA LAJHÁR, KECSSES ŐZ és IZGÁGA HÖRCSÖG) megjelenítve. Az adatközlők rajz-illusztrációkat (1. ábra) kaptak segítségül a karakterek megformálásához, arra azonban nem kaptak semmilyen instrukciót, hogy az adott karaktereket hogyan jelenítsék meg.

A felvételeket csendesített szobában, fejmikrofonnal rögzítettem Audacity (Audacity Team 2018) programmal. A hanganyagot Praat szoftverrel (Boersma–Weenink 2018) szakasz- és mondatszinten annotáltam. Az esetleges megakadásokat az elemzés során nem vettem figyelembe. Szegmentáltam az irregularis részeket a rezgéskép, a hangszínekép és az auditív információk alapján (Dilley et al. 1996; Markó 2013). Az alapfrekvenciát a mondatok nem irregularis részein 5 ms-onként, 75–600 Hz között, egy Praat szkript segítségével mértük. Az átlagos alapfrekvenciát mondatokra számoltam, Hz-ben kifejezve. A hangterjedelmet félhangban kifejezve adtam meg mondatokra az alábbi képletet használva:

$$(1) 12 * \log_2 \frac{f_{0max}}{f_{0min}}$$



1. ábra. A megjelenítendő karakterek rajz-illusztrációi (RAVASZ RÓKA, LUSTA LAJHÁR, KECSES ŐZ és IZGÁGA HÖRCSÖG)

Ahol $f_{0_{max}}$ az adott mondatban mért maximális alapfrekvencia, $f_{0_{min}}$ pedig az adott mondatban mért minimális alapfrekvencia.

Az irreguláris részek arányát százalékosan adtam meg. Egyrészt kiszámoltam az irreguláris részek időtartamának arányát, azaz megadtam, hogy a mondatok teljes idejének hány százaléka az irreguláris szakaszok összydőtartama (az időtartamértékeket szintén Praat szkripttel nyertem ki). Másrészt az irreguláris szótagok arányát is megadtam, tehát azt, hogy a teljes szöveg szótagszámának hány százalékát teszik ki az irreguláris szótagok.

Az adatokon statisztikai elemzéseket végeztem, kevert modelleket állítottam fel (Linear Mixed Model, LMM), az SPSS 20.0 szoftverben számoltam. A kondíció mint független változó hatását vizsgáltam a következő függő változókra: átlagos alapfrekvencia és hangterjedelem. A beszélőt és a mondatot mindkét esetben random faktorként kezeltem. A csoportok közötti összehasonlításhoz a Tukey-féle post hoc tesztet használtam.

A kondíciónak az irregularitásra gyakorolt hatását a Friedman-féle összetartozó mintás varianciaanalízissel vizsgáltam, melyben a függő változó az irreguláris részek időtartam-, illetve szótagaránya volt, a független változó pedig a kondíció. A csoportok közötti összehasonlítást a Wilcoxon Signed Rank tesztel végeztem.

3 Eredmények

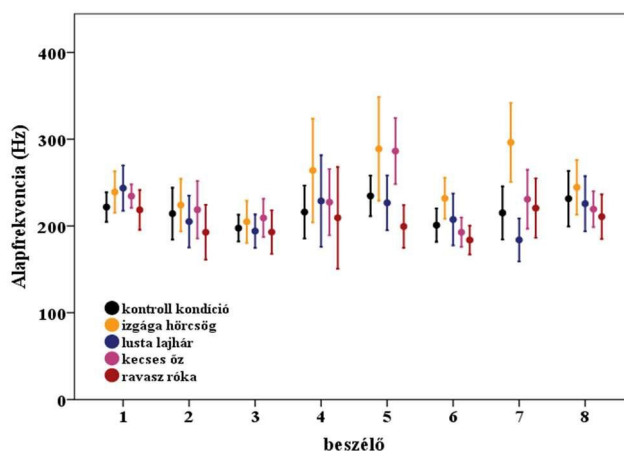
3.1 Átlagos alapfrekvencia

Az alapfrekvencia értékeit a 2. ábra mutatja Hz-ben kifejezve, az ábrán az átlag- és szórásértékek láthatók. A grafikon az egyes kondíciók (kontroll kondíció és a négy karakter) közti eltéréseket szemlélteti, beszélőnkénti bontásban.

Az átlagos alapfrekvencia minden beszélő esetén magasabb az IZGÁGA HÖRCSÖG megvalósításakor, mint a kontroll kondícióban. Egy adatközlő kivételével mindenki átlagosan alacsonyabb alapfrekvencián valósította meg a RAVASZ RÓKA karakterét a kontrollhoz viszonyítva. A KECSES ŐZ esetében a beszélők többségénél (nyolcból hat beszélő) átlagosan magasabb alapfrekvencia volt jellemző a kontroll kondícióhoz viszonyítva. A LUSTA LAJHÁR ábrázolása volt a legváltozatosabb, az adatközlők közül öten alacsonyabb, hárman magasabb átlagos alapfrekvencián valósították meg ezt a karaktert.

A grafikonról továbbá leolvasható, hogy az adatközlők eltérő mértékben változtattak az alapfrekvencián. Voltak beszélők, akiknél az átlagos alapfrekvencia karakterenként csak kis mértékben tért el (pl. első és harmadik beszélő), azonban náluk a szórás is kisebb volt, tehát a karakterek közti változtatások lehetnek következetesek. Néhány adatközlő azonban jelentősen nagyobb mértékben változtatott az alapfrekvencián az egyes karakterek esetében (pl. negyedik, ötödik és hetedik beszélő), az ő esetükben a szórás is nagyobb volt.

Az 1. táblázat az átlagos alapfrekvenciát és annak szórását mutatja kondíciókra bontva.



2. ábra. Az alapfrekvencia alakulása a kondíció függvényében, beszélőnkénti bontásban (átlag- és szórásértékek)

kondíció	f_0 -átlag (Hz)	szórás (Hz)
kontroll	216,4	28,1
izgága hörcsög	249,6	49,4
lusta lajhár	213,5	37,0
kecses őz	227,5	38,3
ravasz róka	203,3	34,3

1. táblázat. Az átlagos alapfrekvencia és szórása kondícióként

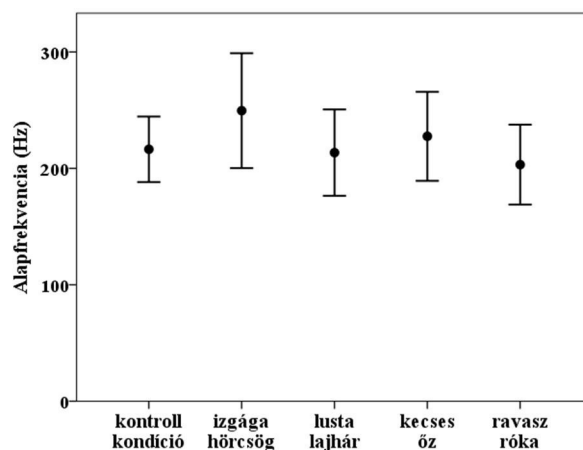
A táblázatról leolvasható, hogy az alapfrekvencia szórása a kontroll kondícióban volt a legkisebb és az IZGÁGA HÖRCSÖG esetében a legnagyobb.

A 3. ábra az alapfrekvencia alakulását szemlélteti a kondíció függvényében. (Az ábrán átlag- és szórásértékek láthatóak.)

A 3. ábrán látható, hogy az IZGÁGA HÖRCSÖG és a KECSES ŐZ karakterét általában magasabb, a LUSTA LAJHÁR és a RAVASZ RÓKA karakterét általában alacsonyabb alapfrekvencián valószínűsítették meg a beszélők. A kontrolltól való legnagyobb mértékű eltérés az IZGÁGA HÖRCSÖG, míg a legkisebb mértékű eltérés a LUSTA LAJHÁR esetében volt tapasztalható. Általában megállapítható, hogy az adatközlők nagyobb mértékben változtatták az alapfrekvenciájukat felfelé, mint lefelé a kontrollhoz képest.

A statisztikai modell (LMM) alátámasztotta, hogy a kondíció hatással van az alapfrekvenciára ($\chi^2(1) = 324,32; p < 0,05$). A Tukey-féle post hoc teszt alapján elmondható, hogy az alapfrekvencia a LUSTA LAJHÁR kivételével mindegyik karakter esetében szignifikánsan eltért a kontrolltól (IZGÁGA HÖRCSÖG: $Z = 12,259; p < 0,001$; LUSTA

LAJHÁR: $Z = -0,894$; $p = 0,815$; KECSSES ŐZ: $Z = 3,291$; $p = 0,004$ és RAVASZ RÓKA: $Z = -3,298$; $p = 0,004$).



3. ábra. Az alapfrekvencia alakulása a kondíció függvényében (átlag- és szórásértékek)

3.2 Hangterjedelem

A hangterjedelem beszélőnkénti alakulását a karakterek mentén a 4. ábra mutatja félhangban kifejezve. Az ábrán az átlag- és szórásértékek láthatók.

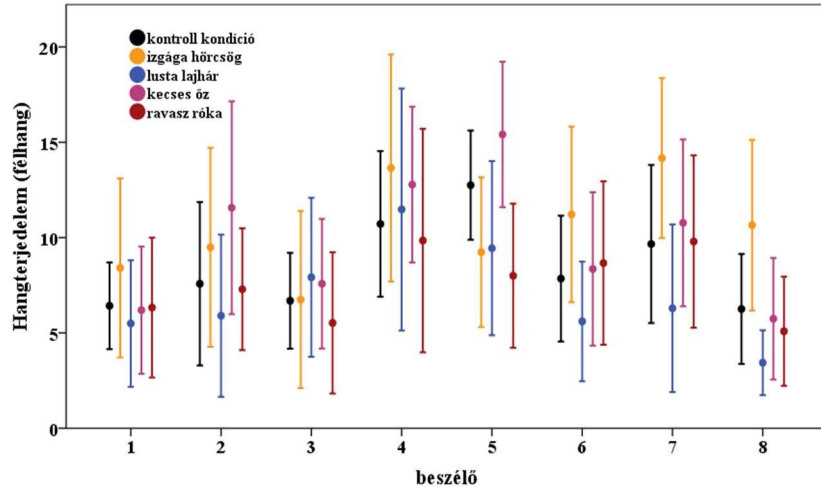
Egy adatközlő kivételével mindenki átlagosan nagyobb hangterjedelemben valósította meg az IZGÁGA HÖRCSÖGöt a kontroll kondícióhoz képest. Két beszélő kivételével a LUSTA LAJHÁRT és a RAVASZ RÓKÁT átlagosan kisebb, a KECSSES ŐZet nagyobb hangterjedelemben valósították meg a kontrollhoz képest.

Beszélőnként eltérő mértékű változtatások figyelhetők meg, vannak azonban hasonló stratégiák: az első, a hatodik, a hetedik és a nyolcadik beszélőnél egyaránt megállapítható, hogy IZGÁGA HÖRCSÖG hangterjedelme jelentősen nagyobb volt a többi karakterhez viszonyítva, a LUSTA LAJHÁRÉ a legalacsonyabb, míg a KECSSES ŐZ és RAVASZ RÓKA nem sokban tért el a kontroll kondíciótól.

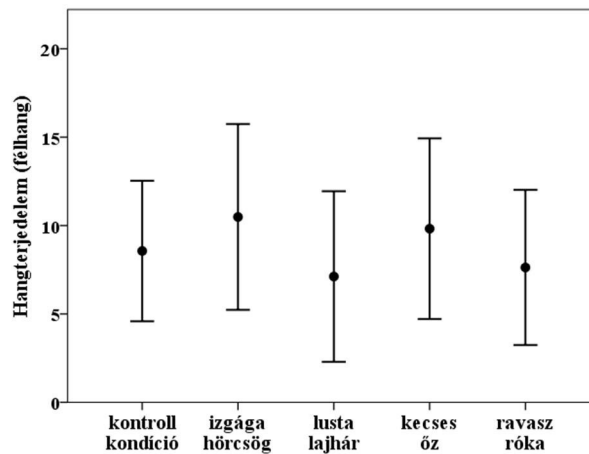
Az 5. ábra a hangterjedelmet a kondíció függvényében mutatja, beszélőnkénti bontás nélkül.

Az 5. ábra alapján megfigyelhető, hogy azon karaktereket, melyeket a kontroll kondícióhoz képest átlagosan magasabb alapfrekvencián valósítottak meg, általában nagyobb hangterjedelemmel is, illetve, melyeket átlagosan alacsonyabb alapfrekvencián valósítottak meg, azok esetében a hangterjedelem is kisebb volt.

A statisztikai modell (LMM) megerősítette, hogy a kondíció hatással van a hangterjedelemben ($\chi^2(4) = 157,66$; $p < 0,05$). A Tukey-féle post hoc teszt alapján elmondható, hogy a hangterjedelem mindegyik karakter esetében szignifikánsan eltért a kontrolltól (IZGÁGA HÖRCSÖG: $Z = 5,976$; $p < 0,001$; LUSTA LAJHÁR: $Z = -4,407$; $p < 0,001$; KECSSES ŐZ: $Z = 3,711$; $p < 0,001$ és RAVASZ RÓKA: $Z = -2,556$; $p = 0,0459$).



4. ábra. A hangterjedelem alakulása a kondíció függvényében, beszélőnkénti bontásban (átlag- és szórásértékek)



5. ábra. A hangterjedelem alakulása a kondíció függvényében (átlag- és szórásértékek)

3.3 Zöngeminőség – irregularitás

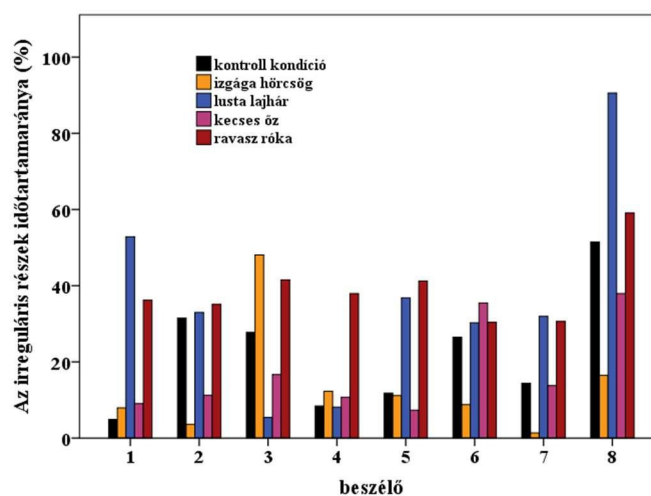
A 6. ábra az irregularis részek időtartamarányát mutatja a kondíció függvényében, beszélőkre bontva, százalékban kifejezve.

Az irregularis részek időaránya minden beszélő esetében magasabb volt a RAVASZ RÓKA megvalósításakor, mint a kontroll kondícióban. Egy adatközlő kivételével a LUSTA LAJHÁR esetében is magasabb volt az irregularis részek aránya a kontrollhoz képest. A KECSÉS ŐZ és az IZGÁGA HÖRCSÖG esetében a beszélők többségénél alacsonyabb volt az irregularis részek időaránya. Megfigyelhető azonban, hogy nagyok voltak az egyéni eltérések, melyek már abból is adódhattak, hogy a kontroll kondícióban is jelentősen eltért az irregularitás aránya az egyes beszélőknél. Az első, a negyedik és az

ötödik adatközlőre kevésbé, míg a második, a harmadik, a hatodik és a nyolcadik adatközlőre nagyobb mértékben volt jellemző az irregularitás.

A 7. ábra az irregularis részek időtartamarányát mutatja a kondíció függvényében, beszélőnkénti bontás nélkül, százalékban kifejezve.

A 7. ábra alapján az mondható el, hogy a LUSTA LAJHÁR megvalósítása tért el a legnagyobb mértékben a többi kondíciótól. Az 6. ábrával összevetve (ahol látszik a beszélőnkénti bontás is) azonban látható, hogy a nyolcadik beszélő magas értékei jelentősen befolyásolták az arányokat. Megfigyelhető továbbá, hogy azon karakterek esetében, melyeket a kontrollhoz képest alacsonyabb alapfrekvencián valósítottak meg, általában nagyobb volt az irregularis részek aránya is.



6. ábra. Az irregularis részek időtartamarányának alakulása a kondíció függvényében, beszélőnkénti bontásban

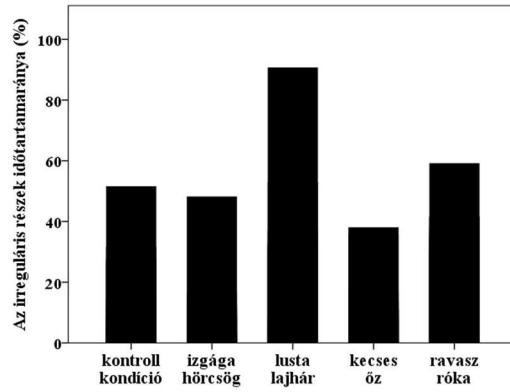
A statisztikai modell alátámasztotta, hogy a kondíció hatással van az irregularis részek időtartamarányára ($\chi^2(4) = 10,9; p = 0,028$). A Wilcoxon Signed Rank teszt alapján elmondható, hogy csak egy karakter, a RAVASZ RÓKA esetében tért el szignifikánsan az irregularis részek időaránya a kontrolltól (IZGÁGA HÖRCSÖG: $Z = -0,98; p = 0,327$; LUSTA LAJHÁR: $Z = -1,68; p = 0,093$; KECSSES ŐZ: $Z = -1,12; p = 0,263$ és RAVASZ RÓKA: $Z = -2,521; p = 0,012$).

Az irregularis szótagok arányát is elemeztem, melyek alakulását a 8. ábra szemlélteti. Felül az irregularis szótagok aránya a kondíció függvényében, alul ugyanez beszélőkre bontva látható. Mindkét grafikon az irregularis szótagok százalékban kifejezett értékét mutatja az összes szótagszámhoz viszonyítva.

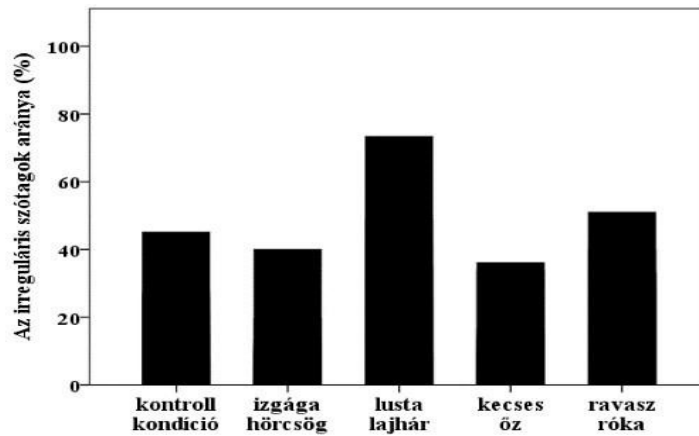
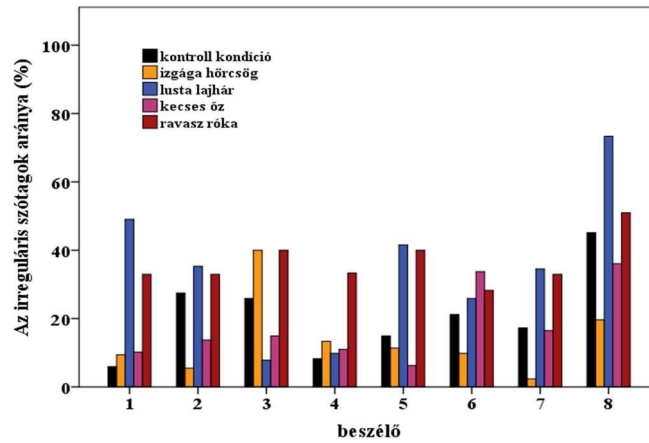
A 8. ábra grafikonjait összevetve a 6. és a 7. ábrával megállapítható, hogy nem volt jelentős különbség az időtartam- és a szótagok aránya között. A tendenciák hasonlóak voltak a két különböző ábrázolási módban.

A statisztikai modell alátámasztotta, hogy az irregularis szótagok arányára is hatással van a kondíció ($\chi^2(4) = 11,899; p = 0,018$). A Wilcoxon Signed Rank teszt alapján pedig elmondható, hogy csak egy karakter, a RAVASZ RÓKA esetében tért el szignifikánsan az irregularis szótagok aránya a kontrolltól (IZGÁGA HÖRCSÖG: $Z = -1,12; p =$

0,263; LUSTA LAJHÁR: $Z = -1,82$; $p = 0,069$; KECSSES ÖZ: $Z = -0,84$; $p = 0,401$ és RAVASZ RÓKA: $Z = -2,521$; $p = 0,012$).



7. ábra. Az irreguláris részek időtartamaránya a kondíció függvényében



8. ábra. Az irreguláris szótágok arányának alakulása a kondíció és a beszélő függvényében

4 Következtetések

A jelen kutatásban a karakterábrázolás alaphfrekvencia-jellemzőinek alakulását vizsgáltam. Arra a kérdésre kerestem a választ, hogy elkülönülnek-e a karakterek mentén a következő paraméterek: átlagos alaphfrekvencia, hangterjedelem és irregularitás.

Első hipotézisem az volt, hogy az átlagos alaphfrekvencia eltérést fog mutatni a karakterek mentén. Ezt a hipotézisemet alátámasztották az adatok, egy karakter kivételével (LUSTA LAJHÁR) valóban szignifikánsan eltértek az értékek a kontroll kondícióban mérttől. Az IZGÁGA HÖRCSÖG és a KECSSES ŐZ esetében az elvárásokkal egyezően a kontrollnál magasabb alaphfrekvenciát mértem. A RAVASZ RÓKA esetében a kontrollnál alacsonyabb átlagos alaphfrekvencia volt megfigyelhető a beszélők többségénél, ez ellentétben van a megfogalmazott elvárással.

A második hipotézisem a hangterjedelemre vonatkozott. Ez a hipotézisem is megerősítést nyert, a kondíció valóban hatással volt a hangterjedelemre, minden karakter esetén szignifikánsan eltértek az értékek a kontroll kondícióban mérttől. Az IZGÁGA HÖRCSÖG esetében a várttal megegyezően nagyobb hangterjedelmet mértem a kontrollhoz képest, míg a RAVASZ RÓKA esetében a várttal ellenkezően kisebbet. A KECSSES ŐZ esetében nagyobb, a LUSTA LAJHÁR esetében pedig kisebb hangterjedelmet mértem a kontrollhoz képest.

A harmadik hipotézisem az volt, hogy a kondíció hatással lesz az irregularitás mértékére, tehát az irreguláris részek aránya eltérést fog mutatni a karakterek mentén. A hipotézis csak részben nyert megerősítést az eredmények által: csupán egy karakter esetében (RAVASZ RÓKA) tértek el szignifikánsan az értékek a kontrolltól. Tendenciaszerűen azonban a többi karakter is különbözött egymástól az irregularitás tekintetében: a RAVASZ RÓKA és a LUSTA LAJHÁR esetében (a várttal egyezően) a kontrollhoz képest gyakoribb, az IZGÁGA HÖRCSÖG és a KECSSES ŐZ esetében ritkább volt az irreguláris szakaszok előfordulása. A kutatás eredményei hozzájárulnak a karakterábrázolás alaphfrekvencia-jellemzőinek részletesebb megismeréséhez.

Köszönetnyilvánítás

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

Irodalom

- Abe, M. – Mizuno, H. 1994. Speaking styles conversion by changing prosodic parameters and formant frequencies. *ICSLP 94*. Yokohama, Japan: 1455–1458.
- Adell, J. – Bonafonte, A. – Escudero, D. 2005. Analysis of prosodic features towards modelling of emotional and pragmatic attributes of speech. *Procesamiento del lenguaje natural*, no. 35: 277–283.
- Audacity Team 2018. Elérhető: https://www.audacityteam.org/?fbclid=IwAR1X0rCeE0RuaW3T1ONRRhc2ic3F2TyNb2t5hGy7PwO2jtGy18nBUbBz_-g
- Balázs B. – Bóna J. 2016. Életkori sajátosságok a beszédképzésben és a beszédfeldolgozásban. In: Bóna J. (szerk.) *Fonetikai olvasókönyv*. Budapest: ELTE BTK Fonetikai Tanszék. 7–19.

- Bata S. – Grácz T. E. 2009. Hatással van-e a beszédpartner életkora a beszélő beszédének szupraszegmentális jellegzetességeire. In: Keszler B. – Tátrai Sz. (szerk.) *Diskurzus a grammatikában, grammatika a diskurzusban*. Budapest: Tinta kiadó. 74–83.
- Boersma, P. – Weenink, D. 2018. *Praat: doing phonetics by computer*. Elérhető: http://www.fon.hum.uva.nl/praat/download_win.html (letöltés ideje: 2018. 12. 01.)
- Bóna J. 2011. A különböző beszédstílusok az akusztikai-fonetikai és percepció vizsgálatok tükrében. *Alkalmazott Nyelvtudomány 11*: 39–48.
- Bóna J. 2016. *Női beszéd - férfi beszéd a fonetikai és a pszicholingvisztikai vizsgálatok tükrében*. Nyelvtudományi értekezések 166. Budapest: Akadémiai Kiadó.
- Bryant, G. A. – Fox, J. E. 2005. Tree Is there an ironic tone of voice? *Language and Speech 48*(3): 257–277.
- Burkhardt, F. – Sendmeier, W. F. 2000. Verification of acoustical correlates of emotional speech using formant-synthesis. In: *Proceedings ISCA Workshop on Speech and Emotion*, 5–7. September 2000, Newcastle, Northern Ireland: 151–156.
- Dilley, L. – Shattuck-Hufnagel, S. – Ostendorf, M. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics 24*(4): 423–444.
- Doukhan, D. – Rilliard, A. – Rosset, S. – Adda-Decker, M. – d’Alessandro, C. 2011. Prosodic analysis of a corpus of tales. *InterSpeech 2011*: 3129–3132.
- Eskenazi, M. 1993. Trends in Speaking Styles Research. *Proceedings Eurospeech*. Berlin: 501–509.
- Fackrell, J. – Vereecken, H. – Buhmann, J. – Martens, J-P. – Van Coile, B. 2000. Prosodic variation with text type. *Proceedings of 6th International Conference on Spoken Language Processing (ICSLP 2000)*, Vol.3: 231–234.
- Feinberg, D. R. – DeBruine, L. M. – Jones, B. C. – Perrett, D. I. 2008. The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception 37*: 615–623.
- Gocsál Á. 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutatás 1998*: 122–135.
- Gocsál Á. – Huszár Á. 2003. Csábító hangok. *Beszédkutatás 2003*: 9–18.
- Gósy M. 2004. *Fonetika, a beszéd tudománya*. Budapest: Osiris Kiadó.
- Gósy M. – Krepesz V. 2015. Magánhangzók temporális jellemzői az idő múlásának függvényében. *Beszédkutatás 2015*: 53–65.
- Gósy, M. – Krepesz, V. 2018. Phrase-final Lengthening of Phonemically Short and Long Vowels in Hungarian Spontaneous Speech across Ages. In: Gósy M. – Grácz, T. E. (szerk.) *Challenges in analysis and processing of spontaneous speech*. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences (RIL HAS). 99–126.
- Gyarmathy D. 2007. Az alkohol hatása a spontán beszédprodukcóra. *Beszédkutatás 2007*: 108–121.
- Hardcastle, W. J. – Laver, J. (eds.) 1997. *The handbook of phonetic sciences*. Blackwell Publishers, Oxford.
- Hirschberg, J. 2000. A Corpus-based Approach to the Study of Speaking Style. In: Horne, M. (szerk.) *Prosody: Theory and Experiment*. The Netherlands, Kluwer. 335–350.
- Imre A. 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr 129*(4): 510–520.
- Jacewicz, E. – Fox, R. A. – Wei, L. 2010. Between-speaker and within-speaker variation in speech tempo of American English. *The Journal of the Acoustical Society of America 128*: 839.
- Jokisch, O. – Kruschke, H. – Hoffmann, R. 2005. Prosodic reading style simulation for Text-to-Speech synthesis. In: Tao, J. – Tan, T. – Picard, R. W. (szerk.) *Affective Computing and Intelligent Interaction, ser. Lecture Notes in Computer Science*. Springer Berlin Heidelberg, vol. 3784: 426–432.
- Maekawa, K. 1998. Phonetic and phonological characteristics of paralinguistic information in spoken Japanese. *Proceedings International Conference Spoken Language Processing*, vol. 2: 635–638.
- Markó A. 2009. Stigmatizált hanglejtésforma a spontán beszédben. *Beszédkutatás 2009*: 88–106.
- Markó A. 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. Budapest: ELTE Eötvös Kiadó.
- Montaño, R. – Alas, F. – Ferrer, J. 2013. Prosodic analysis of storytelling discourse modes and narrative situations oriented to text-to-speech synthesis. *8th ISCA Workshop on Speech Synthesis*, Barcelona, Spain, August 31: 191–196.
- Németh M. 2013. A tanári beszéd és személyiség. *Anyanyelv-pedagógia 2013/1*. Elérhető: <http://anyanyelv-pedagogia.hu/cikkek.php?id=434> (letöltés ideje: 2019. január 3.)
- Olaszy G. 2005. Prozódiái szerkezetek jellemzése a hírfelolvasásban, a mesemondásban, a novella és a reklámok felolvasásában. *Beszédkutatás 2005*: 21–50.

- Paeschke, A. 2004. Global trend of fundamental frequency in emotional speech. *Proceedings of Speech Prosody 2004*, Nara: 671–674.
- Roekhaut, S. – Goldman, J. – Simon, A. 2010. A Model for Varying Speaking Style in TTS systems. In: *5th International Conference on Speech Prosody*, Chicago: IL.
- Scherer, K. R. – Banse, R. – Wallbott, H. G. – Goldbeck, T. 1991. Vocal cues in emotion encoding and decoding. *Motivation and Emotion* 15: 123–148.
- Theune, M. – Meijs, K. – Heylen, D. – Ordelman, R. 2006. Generating expressive speech for storytelling applications. *Speech, and Language Processing* 14(4): 1137–1144.
- Trouvain, J. – Barry, W. J. 2000. The prosody of excitement in horse race commentaries. In: *Proceedings ISCA Workshop on Speech and Emotion*. 5–7. September, Newcastle, Northern Ireland: 86–91.
- Váradi V. 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszédkutató 2010*: 100–109.
- Verma, R. – Sarkar, P. – Rao, K. 2015. Conversion of neutral speech to storytelling style speech. *Advances in Pattern Recognition (ICAPR)*: 1–6.