

## SUMMARY

### **Speech rate and articulation rate in the speech of children and adolescents depending on speech tasks and units of measurement**

*Bóna, Judit – Váradi, Viola*

The study examines speech rate and articulation rate in 6-, 9-, 13-, and 17-year-old speakers in three types of tasks (interview, storytelling based on a series of pictures, and summarizing the content of a heard text). The tempo values are presented and compared in several ways: sound/s, syllable/s, and word/minute. Results show that although age has an effect on speech and articulation rate, its effect is not present in all speech tasks and for all tempo values. The effect of speech tasks on tempo values was significant from the age of 13 on. However, it is also important which unit the tempo values are given in: they can lead to quite different results and therefore different conclusions. While we measured significant differences between the data obtained in one unit, there were no statistical differences across age groups or speech tasks in the other units. Our data enrich the knowledge related to first language acquisition with new results on the one hand, and draw attention to an important methodological issue, namely the importance of units of measurement, on the other hand.

**Keywords:** speech rate, articulation rate, speech task, units of measurement, children

### **A beszélőn belüli variabilitás alakulása a beszéd típus, a szövegbeli helyzet és az életkor mentén**

#### **1. Bevezetés**

Az öregedő nyugati társadalmakban a kor észlelése és a kapcsolódó sztereotípiák megvitatása kulcsfontosságú tényezők, amelyek befolyásolják a nemzedékek közötti sikeres kommunikációt. A beszédben megjelenő időszedés (legyen az bármely fiatal vagy idős korban) folyamatával kapcsolatos érdeklődés megjelenése egészen a '60-as évekre vezethető vissza, amikor is Ptacek és Sander (1966) kísérleti eredményei szerint a lehallgatók 78%-ban helyesen meg tudják mondani, hogy a lejátszott magánhangzókat idős vagy fiatal beszélők produkálták-e. Az olvasás alapján történő megítéltetés esetében ez az arány 99%-os volt. Egy másik kutatás hasonló eredményekről számol be a produkció és a percepció oldalát egyidejűleg figyelembe véve: az artikulációs tempó az életkor előrehaladtával lassul, és a lassabb tempójú beszélőket idősebbnek ítélik meg, ez a tendencia azonban csak a felolvasásban volt kimutatható, a spontán beszédben nem találtak ilyen irányú összefüggést a tempó, a valódi és a becsült életkor között (Brückl–Sendlmeier 2003).

A humánlehallgatás jó teljesítménye számos kérdést vet fel arra vonatkozóan, hogy vajon hogyan (és) mely paraméterek mentén képesek meghatározni a beszélők

életkorát, mely akusztikai információkat használják fel a lehallgatók a beszélők kronológiai életkorának a megállapításához, és a becsült életkor milyen mértékben tér el a biológiai koruktól fiatal- és időskorban.

Kutatások szerint a prozódiai jellemzők (és azok változása) központi szerepet játszanak a kor megítélésében (Winker 2007). Az akusztikus jegyek alapján történő értékelést főként a következő fogalmakhoz kapcsolódóan írták le: kapcsolatot találtak az alapfrekvencia, a formánsértékek, a hangszalagok rezgésének modulációja, az intonáció, valamint a spektrális energiaeloszlás és az életkor között (Brückl 2007).

A legtöbbet vizsgált paraméter a közlés tempója. A tempó nemcsak az életkortól, hanem számos más tényezőtől függ, egy szövegen belül is változó a temporalitás. Az artikulációs tempó kisebb egységeken belüli alakulását vizsgálták egy másik kutatásban (Dankovičová 1999), amelyben kimutatták, hogy az intonációs frázisokon belül (tehát olyan beszédrészen, amelyek hanglejtésük és hangsúlyozásuk alapján egy egységet alkotnak) lassulás figyelhető meg az artikulációs tempóban. Kohári (2016) a felolvasást vizsgálva azt találta, hogy tagmondaton belül jellemzőbb a lassulás, de nagyobb egységre, így például teljes szövegre kiegyenlítődés jellemző. Gósy és munkatársai (2011) a spontán beszéd temporális alakulását vizsgálta a tematikus egységek mentén (új tematikus egység alatt azt értik, amikor az interjúvezető témát vált). Eredményeik szerint a tematikus egységek rövidülnek a narratívák végéhez közeledve, a tempóban azonban nem találtak tendenciaszerű változást.

Az időzítést a beszéd típus is meghatározza: több kutatásban alátámasztották, hogy a felolvasás artikulációs tempója gyorsabb, mint a spontán beszédé (Raming 1983; Duchin–Mysak 1987; Olasz 2006; Bóna 2014). Vannak azonban ennek ellentmondó eredmények is. Jacewicz és munkatársai (2010) vizsgálatában a spontán beszéd bizonyult gyorsabbnak, míg például Kohári (2016) nem talált számottevő különbséget a két beszéd típusban. A tempón kívül a szünetezés is eltérő lehet, a spontán beszédben magasabb szünetarányt, gyakoribb szüneteket és ebből adódóan rövidebb beszédszakaszokat mértek az olvasáshoz képest (vö. pl. Walker 1988; Váradi 2010; Bóna 2013). A beszédritmusban szintén találtak eltérést a két beszéd típus között: Kohári (2016) eredményei szerint a magánhangzós és a mássalhangzós szakaszok időtartamának variabilitása nagyobb a spontán beszédben, mint a felolvasásban. Ezek a kutatások több korosztály beszédének sajátosságait foglalják össze, és jutnak arra a következtetésre, hogy minden korosztályban különbség mutatkozik a beszéd típusok temporális jellemzőit illetően.

Az artikuláció sebessége tekintetében több tanulmány is kimutatta, hogy az idősek lassabban beszélnek, mint a fiatalabbak (Oyer–Deal 1985; Smith et al. 1987; Brown et al. 1989; Zellner–Keller 2006; Yuan et al. 2006). Smith és társai (1987) 10 fiatal és 10 idős beszélő normál és gyorsított tempóban történő szó- és mondatvisszmondása alapján azt találta, hogy az idősek 20-25%-kal lassabb produkcióra voltak képesek, mint a fiatalok. A lassulást számos tényezővel magyarázzák, beleértve a hosszabb szótagidőtartamot és a hosszabb szünetet is a fiatalabb beszélők produkciójához viszonyítva. A növekvő számú és időtartamú szünet kapcsolódik az időskort jellemző fiziológiai változásokhoz, amelyekkel mind a produkció megváltozását, mind a kognitív eltéréseket magyarázzák a fiatalabb beszélőkhöz viszonyít-

va, például a diskurzustervezésben (Linville 2001). Emellett az idősebb beszélőknél a nagyobb számú szünet megjelenése csökkentheti a beszédszakaszok időtartamát (beszédszakasz alatt értjük azon beszédegységeket, amelyeket két néma vagy kitöltött szünet határol), amelyet a tüdőszövet csökkent rugalmassága magyaráz (Linville 1996). A lassabb artikulációs tempó adódhat a változó gégefunkciókból, a szájizmok és az ajkak csökkenő aktivitásából is (Linville 2001).

Más tanulmányok eredményei azonban mindennek ellentmondó eredményekre jutottak: néhány vizsgálat nem talál összefüggést az életkor növekedése és az artikulációs tempó csökkenése között (pl. Hoit et al. 1994 – 80 fő vizsgálata alapján, akik 20–30, 40–50, 60–70 és 80+ évesek voltak, Linville 2000 – 80 fő alapján, akik közül negyvenen 19 és 24 év közötti beszélők és negyvenen 62 és 79 év közötti beszélők voltak). A LangAge korpusz alapján, ahol 48 férfi és női beszélő (70 év feletti) hanganyagát elemezték, a beszédszakasz időtartama negatívan korrelált a beszélők korával, miközben a szünetek és az életkor között nem volt szignifikáns összefüggés (Gesternberg 2011). Más kutatások szerint az idősebbek artikulációs tempójára nagyobb fokú variabilitás jellemző, mint a fiatalokéra, így nem vonható le egységes következtetés a beszélők produkciójára vonatkoztatva csupán azok életkora alapján (Linville 2001).

A korábbi kutatások keresztmetszeti módszertant alkalmaztak, azaz különböző életkorú csoportok eredményeinek összevetése alapján vizsgálták az életkor hatását. Ezzel szemben a longitudinális módszertannal dolgozó kutatások száma alacsony, és ezek legtöbbször csupán kevés adatközlő bevonásával készültek, és rövid időtartamot ölelnek fel (vö. pl. Brückl 2007). Ennek oka elsősorban a módszertani nehézségekben keresendő: az adatgyűjtés, az azonos felvételi körülmények megteremtése összetett módszertani feladat, ahogyan az adatközlők hosszabb idő után történő felkutatása is. Emellett azonban számos előnnyel is rendelkezik ez az eljárás: például segítségével kiszűrhető a kohorteffektus; a keresztmetszeti vizsgálatok esetében felmerülő, az eltérő életpályák sajátosságaiból adódó különbségek nehezíthetik az összevetést, így például az eltérő biológiai hajlam vagy a különböző kórtörténetek. Ezen eljárás segítségével könnyebb kezelni azt a heterogenitást, amelynek hatását kimutatták a prozódiai sajátosságok területén: az öregedéskutatásban az egyik legmeghatározóbb eredmény az extrém egyéni variáció (Ringel–Chodzko–Zajko 1987).

A temporális sajátosságok és az életkor közti összefüggést általában gyermekkorban vagy patológiás esetekben vizsgálják longitudinális módszertan segítségével, és csupán néhány esetben kutatták korábban egészséges felnőtt beszélők esetében. Gesternberg és munkatársai az életkorhoz kapcsolódó lassulás fogalmát vizsgálták idős (70–80 éves) felnőtt német és francia beszélők alapján, az artikulációs tempó különböző prozodikus paramétereinek, a beszédszakaszok időtartamának és a szünetek közötti összetett kölcsönhatásnak az összefüggésében közel 10 éves különbséggel. Az eredmények elsődleges eredménye az, hogy nagyon jelentős egyéni eltérések rajzolódtak ki az életkor előrehaladtával minden vizsgált tényezőben, így az artikulációs tempóban, a beszédszakaszok időtartamában és a beszédszakaszonkénti szótagszámokban is; egyénenként eltérő módon egyaránt kimutatható volt növekedés és csökkenés. Emellett részben nyelvspecifikus sajátosságokat igazoltak: a franciák esetében a tem-

pó inkább növekedett, míg a németeknél inkább lassult. Az eredmények magyarázata részben csatlakozik a nemzetközi trendekéhez: a temporális sajátosságokat magyarázza a kardiovaszkuláris rendszer változása, ami együtt jár a rövidebb légzési idővel és a kisebb tüdőkapacitással, noha fiziológiás változások mellett az egyéni életvitel (pl. dohányzás, hangterhelés) is nagyon jelentős hatással bír.

Shum (2008) szakdolgozatában érintőlegesen kitért II. Erzsébet király artikulációs tempójának változására a Christmas Broadcast alapján. Noha a leírás nem szolgál részletes leírással a vizsgálat módszertanáról és eredményeiről, a közölt ábra alapján leolvasható, hogy a királynő artikulációs tempója fokozatosan nőtt 20 és kb. 55 éves kora között, majd ezt követően tendenciális csökkenés volt megfigyelhető a 80-as éveinek végéig.

Egy másik kutatás (Quené 2013) Beatrix királynő beszédének temporális sajátosságait vizsgálta longitudinális módszertannal az 1980 és 2012 között rögzített „Troonrede”-ei („trónbeszédei”) alapján (a királynő életkora ebben az időintervallumban 42 és 74 év közé esett, összesen 9 felvételt elemeztek). Az eredmények kevéssé támasztották alá az életkorral összefüggő lassulást: noha a tempóértékek kismértékben csökkennek az első néhány évtizedben, az utóbbi időszakban ezzel ellentétesen gyorsulás volt megfigyelhető. Tempóváltozás a beszédeken belül is megfigyelhető volt: míg a korábbi felvételeket (1980 és 1996 között) a lassabb-gyorsabb-lassabb mintázat, a későbbi felvételeket (1996 és 2012 között) fokozatos gyorsulás jellemezte.

A korábbi kutatások legtöbbször az idősebb korosztály beszélőit érintették, egészséges, fiatal beszélőknél továbbra is kevés információval rendelkezünk az életkor előrehaladtával bekövetkező változásokról. A jelen kutatásban azt vizsgáljuk, hogy mintegy 10 év elteltével (az első felvétel esetén 19 és 40 év közöttiek az adatközlők, a 2. felvétel esetén életkoruk 29 és 50 év közé esik) miként változnak a beszéd egyes időzítési sajátosságai két különböző beszéd típusban (spontán beszédben és olvasott szövegben). Hipotéziseink szerint különbségek mutatkoznak meg a) az artikulációs tempóban és b) a szünetezésben (szünetarány és gyakoriság) egészséges felnőtt férfiak közlései alapján. 1. az egyes beszéd típusok szerint: az olvasott szöveg esetében i) gyorsabb artikulációs tempót és ii) nagyobb arányú és gyakoriságú szüneteket várunk. 2. A felvételek között eltelt idő szerint i) lassulást várunk a tempóban 10 év elteltével, valamint ii) nagyobb arányú, gyakoribb szüneteket. Illetve 3. hipotézisként az egyes felvételeken belül szintén i) lassulást és ii) gyakoribb szüneteket feltételeztünk.

## 2. Kísérleti személyek, anyag és módszer

A jelen kutatásban 13 egészséges, fiatal felnőtt férfi beszélő felvételét elemeztük. A korábbi felvételeket (a továbbiakban 1. felvétel) a BEA spontánbeszéd-adatbázisból választottuk ki (a rögzítés körülményei és a felvételek jellemzőiről: Neuberger et al. 2014), az adatközlők életkora 19 és 40 év közé esett (átlag: 27 év, SD: 5 év). A mintegy 10 évvel későbbi felvételeket (a továbbiakban 2. felvétel) egy longitudinális adatbázisból választottuk ki (a rögzítés körülményei és a felvételek jellemzőiről:

Grácsi et al. 2020), az adatközlők életkora 29 és 50 év közé esett (átlag: 38 év, SD: 5 év). Az adatközlők magyar anyanyelvűek voltak, és nem volt hallási vagy beszédzavaruk.

Mindkét adatbázisból két beszéd típus felvételeit elemeztük: spontán beszédet és olvasott szöveget. A spontán beszéd részben a kísérletvezető az adatközlőt a munkájáról, a hobbijáról, a családjáról és egyéb szabadidős tevékenységeiről kérdezte. Ez a felvételi egység kvázimonologikus, mivel a kísérletvezető nagyrészt olyan céllal szólal meg, hogy kérdésekkel, hozzászólásokkal segítse az adatközlő folyamatos beszédét. Az olvasott szövegrészben az adatközlő feladata egy ismeretterjesztő szöveg felolvasása volt. Az adatközlőnek lehetősége volt a szöveget átolvasni a hangos felolvasás előtt.

A felvételeket beszéd szakasz szinten manuálisan jegyeztük le a Praat szoftver segítségével (Boersma–Weenink 2019). Beszéd szakasz alatt azt a beszédegységet értjük, amelyet két bármilyen szünet – néma vagy kitöltött – határol. A zöngétlen zárhanggal kezdődő beszéd szakaszok esetében minden esetben 30 ms-ot számoltunk zárszakasznak. A beszéd anyagot szintén a Praat szoftverrel elemeztük (Boersma–Weenink 2019) a következő módon: az artikulációs tempót egy erre a célra létrehozott szkripttel számítottuk ki az egyes beszéd szakaszokra, szótag/s-ban kifejezve. A szünetezési sajátosságokat a szünetek időtartamarányának és gyakoriságának vizsgálatával elemeztük, ezeket a mérőszámokat szintén egy Praat szkript segítségével nyertük ki. Az egyes felvételeken belüli esetleges temporális mintázatok elemzésére a felvételeket egy Praat szkripttel öt egyenlő időtartamú részre bontottuk (a továbbiakban 20%-onként felosztott részek), majd a szkript segítségével hozzárendeltük az adott részekhez az azokban lévő beszéd szakaszok artikulációs tempóit. Az adott részek időzítési sajátosságainak pontosabb összevetése céljából az artikulációstempó értékeket normalizáltuk (z-transzformáció:  $Z = X - \mu / \sigma$ , ahol jelen esetben  $Z$  a normalizált artikulációs tempó,  $X$  az aktuális artikulációs tempó,  $\mu$  a mintában mért átlagos artikulációs tempó,  $\sigma$  pedig a mintában mért artikulációs tempó szórása) felvételeként (azaz beszélőnként + beszéd típusonkénti bontásban). Az időzítés lokális vizsgálatára egy másik felosztási módszert is alkalmaztunk: a beszélő teljes beszéd idejét felosztottuk félperces szakaszokra, majd mindegyik félperces szakaszra számoltunk egy artikulációstempó-értéket (egy Praat szkript segítségével).

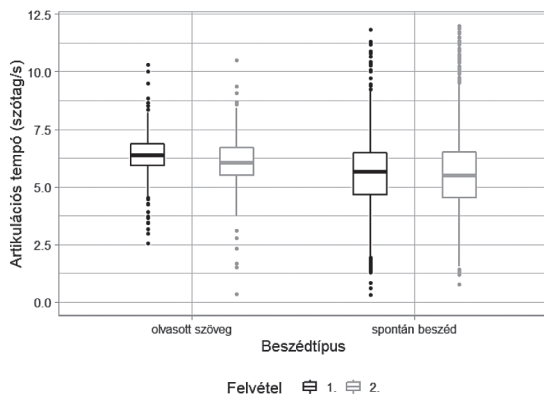
A szünetek időtartamarányát a beszélő teljes beszéd idejéhez viszonyítottuk ( $t_{\text{SIL}} / (t_{\text{SIL}} + t_{\text{SZAKASZ}})$ ), ahol  $t_{\text{SIL}}$  a szünetek időtartamának összege,  $t_{\text{SZAKASZ}}$  pedig a beszéd szakaszok időtartamának összege). A szünetek gyakoriságát a teljes beszéd időre vetítve elemeztük ( $n_{\text{SIL}} / (t_{\text{SIL}} + t_{\text{SZAKASZ}})$ ), ahol  $n_{\text{SIL}}$  a szünetek száma,  $t_{\text{SIL}}$  a szünetek időtartamának összege,  $t_{\text{SZAKASZ}}$  pedig a beszéd szakaszok időtartamának összege). (Azért tartottuk fontosnak a szünetarány mellett a gyakoriságot is vizsgálni, mert azonos szünetarány esetében is lehet eltérő a gyakoriság, például kevesebb, de hosszabb szünetek vagy több de rövidebb szünetek esetén a szünetarány hasonló lehet, a gyakoriság azonban a második esetben magasabb.) A szünetarány és a szünetek gyakoriságának alakulását a felvételek 20%-os bontásában is megnéztük, szintén egy Praat szkript segítségével nyertük ki az egyes, 20%-os részekhez tartozó szünetjellemzőket (tehát nem a teljes időtartamra vetített értékeket elemeztük, hanem a 20%-os részek beszéd idejére vetítve határoztuk meg az értékeket).

Az artikulációs tempó és a szünetek alakulását lineáris kevert modellekkel elemeztük az R programban (R Core Team 2018) az lme4 csomaggal (Bates et al. 2015), a  $p$ -értékeket Satterthwaite-approximáció segítségével nyertük ki (lmerTest csomag, ANOVA-függvény; Kuznetsova et al. 2017). Fix hatásokként (intercept) a felvétel időpontját (1. felvétel [BEA], 2. felvétel [Longi]), a beszédtypust [spontán beszéd vagy olvasott szöveg]) és a 20%-onként felosztott részeket adtuk meg, random hatásként pedig a beszélőket, a függő változók pedig az artikulációs tempó, a szünetarány és a szünetgyakoriság voltak. Minden paraméterre készítettünk egy random intercept és egy random slope modellt is (a beszélővel mint random faktorról, minden változóra), és a két modellt összehasonlítottuk (az lmerTest csomagban elérhető ANOVA-függvénnyel; Kuznetsova et al. 2017). A modellek között nem volt szignifikáns különbség, így a következőkben az alacsonyabb AIC-számot (Akaike 1973) adó random slope értékeit ismertetjük az eredményekben. A 20%-os bontásban szignifikáns különbség esetén a csoportokat Tukey post hoc teszttel hasonlítottuk össze.

### 3. Eredmények

#### 3.1. Az artikulációs tempó alakulása

Az artikulációs tempó értékeinek alakulását elsőként globálisan mutatjuk be. Az 1. ábra a beszédtypusok és a felvételek időpontja szerinti artikulációs tempó értékeit ábrázolja.



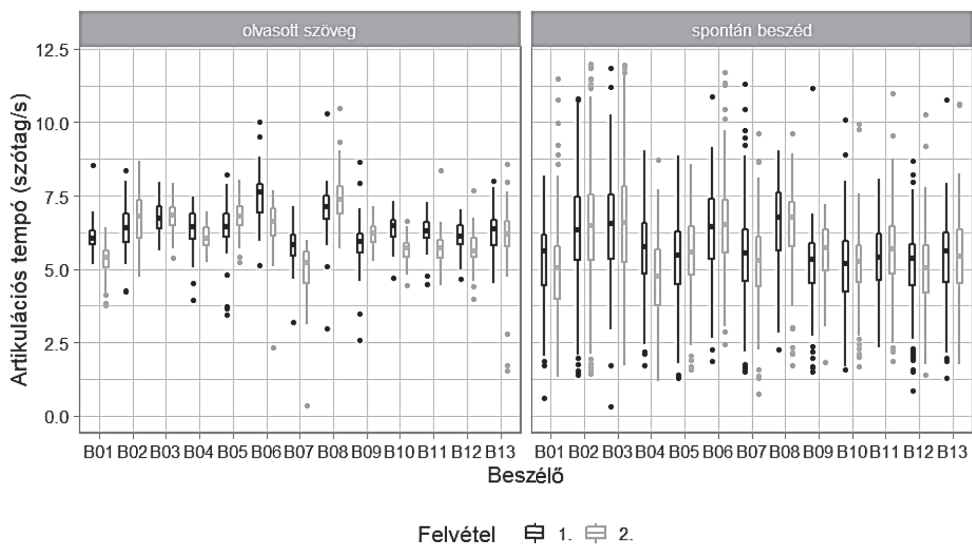
**1. ábra.** Az artikulációs tempó alakulása a beszédtypus és a felvétel időpontja szerint (medián és interkvartilis tartomány)

Különbség mutatkozott a két beszédtypus között az artikulációs tempóban: felolvasás esetén gyorsabban beszéltek az adatközlők (átlag: 6,24 szótag/s), mint a spontán beszéd esetében (átlag: 5,60 szótag/s). Ez a különbség a mediánértékekben is megmutatkozott (1. ábra). További eltérés volt megfigyelhető a két beszédtypusra jellemző varianciában, ahogy a grafikonokon is látható, hogy az interkvartilis tarto-

mányok a spontán beszédben nagyobbak voltak, mint az olvasott szöveg esetében. A varianciák közti különbségre utal az is, hogy a szórás csaknem 1 szótag/s-mal nagyobb a spontán beszéd esetén (SD: 1,83 szótag/s), mint az olvasott szövegben (SD: 0,97 szótag/s).

Az 1. ábra alapján elmondható továbbá, hogy a felvételek ideje (10 éves eltérés) alapján nincs olyan mértékű különbség az artikulációs tempók között, mint amit a beszéd típus szerint tapasztaltunk. A spontán beszéd esetén szinte teljes volt az átfedés az interkvartilis tartományok között (1. ábra), az átlagok között is kevesebb, mint 0,05 szótag/s különbség volt adatható (átlag<sub>1</sub> = 5,59 szótag/s; SD<sub>1</sub> = 0,68 szótag/s; átlag<sub>2</sub> = 5,61 szótag/s; SD<sub>2</sub> = 1,91 szótag/s). Az olvasott szöveg esetében valamivel nagyobb különbséget tapasztaltunk a két felvétel értékei között, de az interkvartilis tartományok több mint fele ebben az esetben átfedés mutatkozott, illetve az átlagok közti különbség nem haladta meg a 0,3 szótag/s-ot, tehát elhanyagolhatónak mondható (átlag<sub>1</sub> = 6,39 szótag/s; SD<sub>1</sub> = 0,91 szótag/s; átlag<sub>2</sub> = 6,1 szótag/s; SD<sub>2</sub> = 1,01 szótag/s).

A 2. ábra a beszélőnkénti bontással árnyaltabb képet ad az artikulációs tempo alakulásáról.



**2. ábra.** Az artikulációstempó-értékek a felvétel időpontja és a beszélők bontásában interjúban és olvasott szövegben (medián és interkvartilis tartomány)

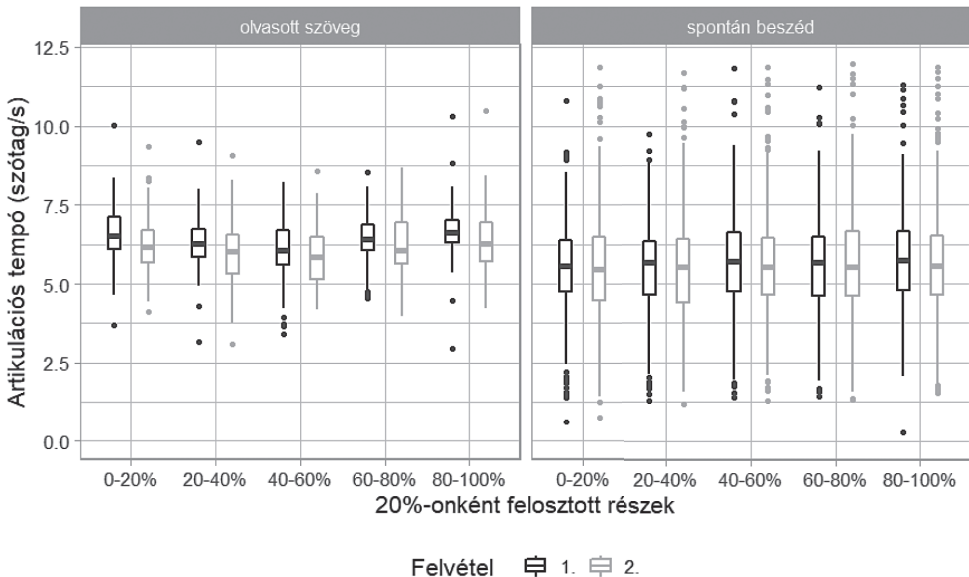
A beszélőnkénti bontás rámutat arra, hogy a variancia nemcsak csoportszinten volt kisebb az olvasott szöveg esetében, hanem beszélőkön belül is minden esetben jelentősen kisebb volt a spontán beszédhez viszonyítva. A beszélők többségénél legalább kétszerese volt a spontán beszédben mért interkvartilis tartomány az olvasott szövegben mértnek. A beszélők többségénél (10 beszélő) megmutatkozott a fentebb említett tendencia, hogy a felolvasás tempója gyorsabb volt a spontán beszédéhez



képest. Három beszélő esetén a két beszéd típus átlagai közti különbség nem haladta meg a 0,2 szótag/s-ot, a különbség esetükben elhanyagolhatónak tekinthető.

A felvételek ideje alapján vannak beszélők, akiknél különböztek az artikulációs tempó értékei, mindkét irányban tapasztaltunk azonban eltéréseket, nem mutattak egységes tendenciát: az artikulációs tempó mediánjait összevetve megállapítható, hogy a spontán beszédben 5 beszélő értékei voltak alacsonyabbak, és 5 beszélő értékei voltak magasabbak a 2. felvételen az elsőhöz viszonyítva. Az 0,1 szótag/s-nál kisebb eltéréseket elhanyagolhatónak tekintettük, ez 3 beszélőt érintett. Az olvasott szöveg esetében 8 beszélő mediánértékei voltak alacsonyabbak, és 4 beszélő értékei voltak magasabbak a 2. felvételen az elsőhöz viszonyítva, illetve egy beszélőnél volt elhanyagolható a különbség a két felvétel között. A beszéd típusok közötti eltérés szignifikáns volt:  $F(1, 6025) = 23,858$ ;  $p < 0,001$ .

Ahogy a 3. ábrán is látható, nem mutatkozott jelentős különbség az artikulációs tempóban attól függően, hogy az adott beszéd szakasz a közlés mely pontján helyezkedett el. A spontán beszéd esetében az interkvartilis tartományok nagyarányú átfedése is utal arra, hogy elhanyagolható az egyes részek közti tempókülönbség. Ezt támasztja alá az is, hogy az artikulációs tempó átlagértékei közti eltérések sem haladják meg a 0,3 szótag/s-os értéket. Az első felvétel spontán beszédének átlaga a 20%-os bontás szerinti 5 szakaszban 5,79 szótag/s (SD: 2,21 szótag/s), a másodikban pedig 5,80 szótag/s (SD: 2,13 szótag/s) volt, a különbség a legnagyobb és a legkisebb érték között pedig az első és a második esetben mindössze 0,36 szótag/s.

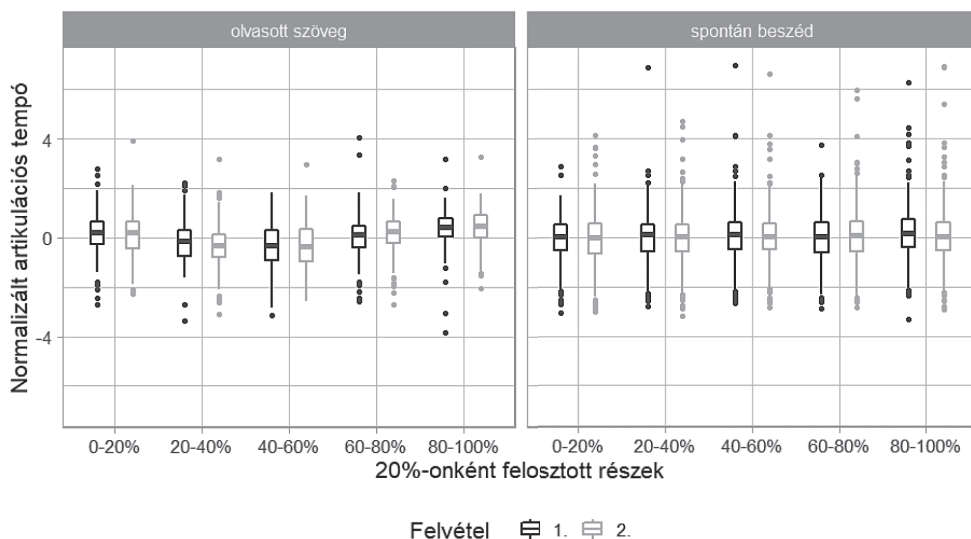


**3. ábra.** Az artikulációs tempó alakulása 20%-os bontásban (medián és interkvartilis tartomány)



Az olvasásban ugyanezek értékek eltérése már valamivel nagyobb volt, mint a spontán beszédben: mind az első, mind a második felolvasásában az első és az utolsó 20%-os szakasz artikulációs tempója bizonyult a legnagyobbak (a legkisebb pedig a középső, 40–60%-os szakaszé). A különbség a legnagyobb és legkisebb érték között az első esetben 0,42 szótag/s, a második esetben 0,65 szótag/s volt. A statisztikai elemzés szerint a szakaszok száma szerint nem volt szignifikáns különbség az artikulációs tempóban a 20%-os bontás tekintetében.

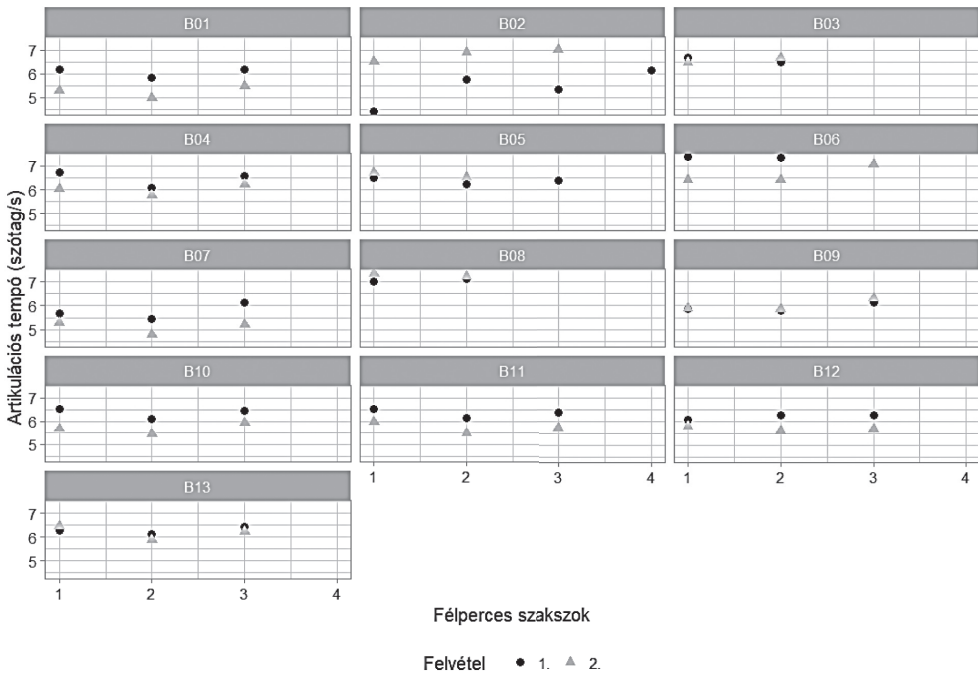
A pontosabb összevetéshez normalizáltuk  $z$ -transzformációval az artikulációs-tempó-értékeket. Az eredmények alátámasztották a korábban bemutatott tendenciákat, azaz, hogy 1) a spontán beszédben mért értékek elhanyagolható különbséget mutattak az egyes szakaszok között (az első felvétel esetében a legnagyobb és a legkisebb különbsége 0,06; a második esetében 0,05 volt), 2) az olvasásban a legnagyobb értékeket a felvétel elején és végén, a legkisebb értékeket a felvételek közepén mértük. Azonosítható tehát egy olyan tendencia, miszerint a felolvasás elején gyorsabban olvasnak az adatközlők, majd fokozatosan lassulnak, és a vége felé újra gyorsabban olvasnak (ez az 1. és 2. felvétel esetén is látható). A statisztikai elemzés azonban nem mutatott szignifikáns különbséget az egyes részek értékei között.



4. ábra. A normalizált artikulációs tempó alakulása 20%-os bontásban

A szövegek 20%-os bontása lehetővé teszi, hogy minden közlést azonos arányok szerint vessük össze, ám a felvételek közti időtartam-különbségeket elfedi, tehát ezen számítási móddal nem mutatkoznak meg a rövidebb és hosszabb felvételek között kirajzolódó esetleges különbségek. Ezen módszertani nehézség feloldásához a következőkben a közléseket 30 másodperces szakaszokra bontottuk fel (5. ábra). Az eredmények a következő tendenciákat tükrözik az olvasás vizsgálatá alapján:

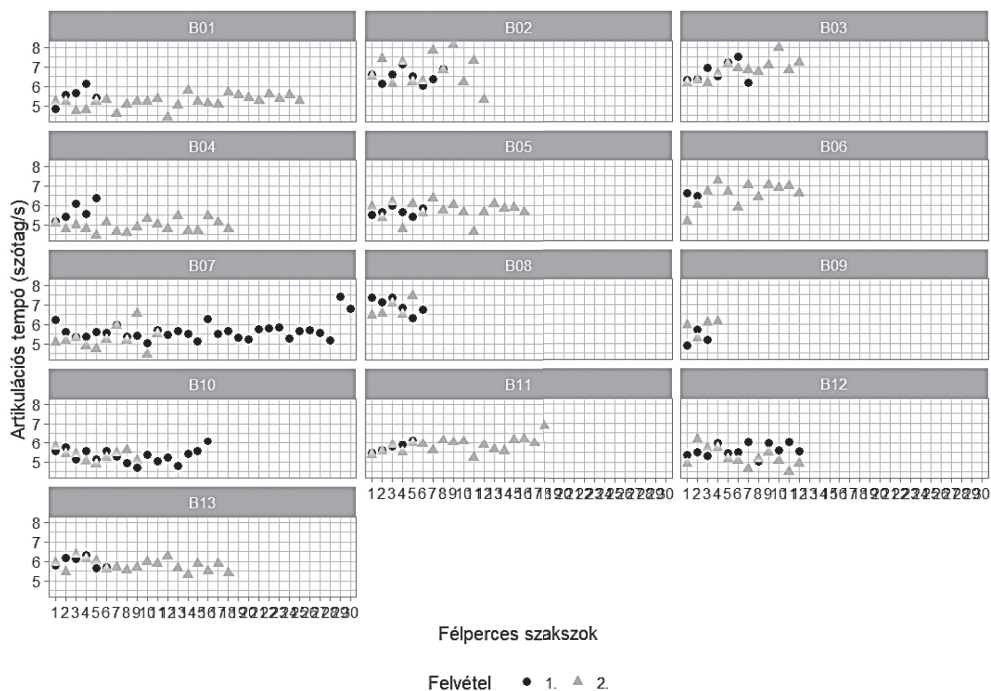
a legtöbb beszélő esetében vagy az első, vagy a második felvétel gyorsabb, ám ez a különbség szisztematikus, azaz minden félperces szakasz esetében alátámasztható, és nem csupán az átlagértékek különbsége tükrözi a két felvétel eltéréseit. Emellett látható, hogy az olvasási sebesség változása – noha egyénileg eltér, hogy melyik beszélő melyik szakaszon gyorsít vagy lassít – mindkét felvétel esetén tükrözi az egyéni tendenciákat, azaz a beszélő közel azonos sebességi mintázattal hangsította meg ugyanazt a szöveget 10 év különbséggel. A félperces bontást bemutató ábra (5. ábra) – a korábbiakhoz hasonlóan – jól tükrözi a jelentős individuális különbségeket és az adott beszélők sajátos, egyéni temporális jellemzőit. Az a tendencia, amit a 20%-os bontásban láttunk, hogy az adatközlők a felolvasásuk közepén lassabban olvasnak, a félperces bontásban is nyomon követhető a beszélők csaknem felénél. A többi adatközlőnél általában nem találunk jelentős különbséget az egyes részek tempóértékei között, illetve egy beszélőnél található csak másfajta tendencia (B02).



**5. ábra.** Félperces időintervallumokra számolt átlagos artikulációstempó-értékek az olvasott szövegekben (a számozás az egyes beszélőket jelöli)

A spontán beszéd értékeinek feldolgozása neheztett, ugyanis – ahogyan láthatjuk a szürke és fekete pontok kiterjedéséből és elhelyezkedéséből – jelentősen eltér a közlések hosszúsága az 1. és 2. felvételek között, valamint sokkal nagyobb variációt mutat a temporális mintázatokban is, mint az olvasás esetében (6. ábra). Az artikulációs tempó értékei jelentős átfedést mutattak a 10 év különbséggel rögzített felvé-

telek között a B04-es beszélő kivételével (akinek egyértelmű lassulás mutatható ki a vizsgált intervallum alatt) minden adatközlőnél. A félperces szakaszok között azonban szintén nem találtunk tendenciaszerű eltérést sem, ahogy az a 20%-os bontásban is látható volt.

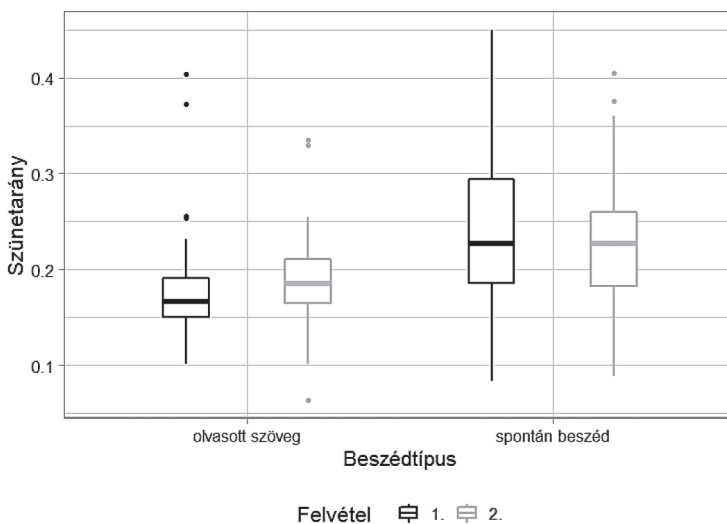


6. ábra. Félperces időintervallumokra számolt artikulációstempó-értékek a spontán beszédben (beszélőkre bontva)

### 3.2. A szünetek alakulása

A közlések temporális mintázatát meghatározza a közlésben tartott szünetek gyakorisága és időtartama. Amennyiben hosszabb és/vagy gyakoribb szünettartás jellemzi a megszólalást, úgy kevésbé tűnik fluensnek az adott közlés, míg rövidebb és/vagy ritkább szünettartással a folyamatosság benyomását kelti a beszélő. (Noha fontos megjegyezni, hogy a fluencia fogalma nem egységes, idegen és anyanyelven eltérő definíciók alkalmazhatók. A jelen esetben mi Lennon [1990] meghatározását alkalmazzuk, aki szerint ez egy, a hallgatóban kialakult benyomás a beszélőről, akinek a tervezési és produkciós folyamatai könnyedén és hatékonyan működnek.) Megvizsgáltuk, hogy hogyan változott az egyes beszélők megnyilatkozása a szünetezési mintázatok szempontjából 10 éves különbséggel az egyes beszéd típusok esetében (7. ábra).

A szünetek időtartamarányának az átlaga a spontán beszédben az 1. felvétel alkalmával 0,24 (SD: 0,06), a 2. felvétel alkalmával pedig 0,23 (SD: 0,06) volt. Az olva-

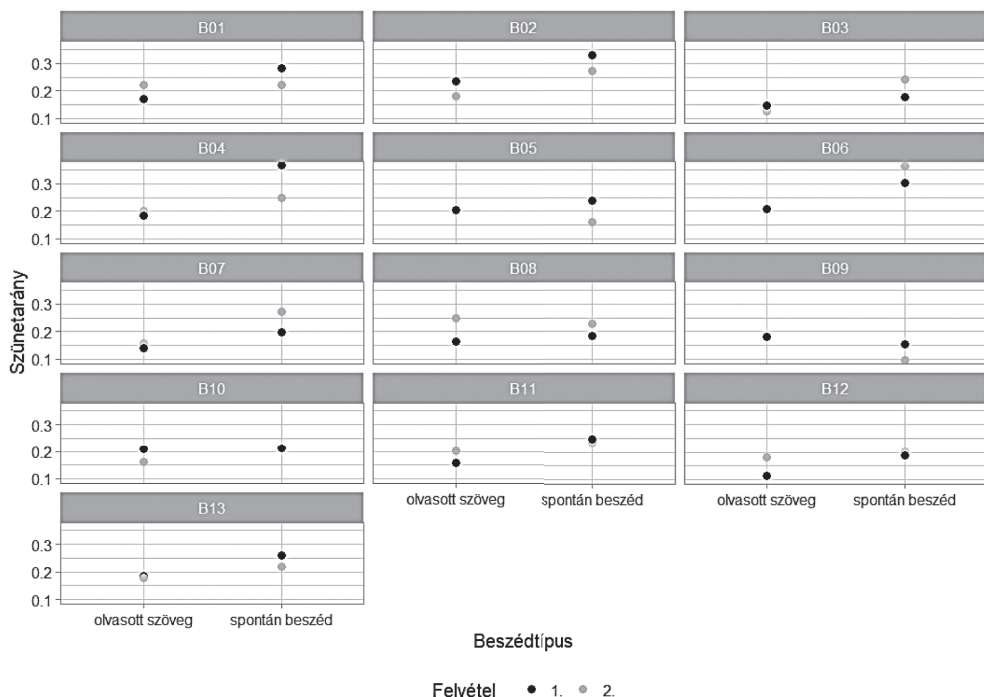


**7. ábra.** A szünetek időtartamaránya a teljes beszédidőre vetítve a beszédtípus és a felvételek szerint (medián és interkvartilis tartomány)

sásában az 1. felvétel alkalmával 0,18 (SD: 0,03), a 2. felvétel alkalmával 0,19 (SD: 0,03) volt. Habár az átlagok alapján elhanyagolhatónak tűnik a különbség a felvételek között, a mediánok és az interkvartilis tartományok az olvasott szöveg esetében valamivel nagyobb mértékben mutatnak eltérést a két felvétel között: 10 év elteltével nagyobb a szünetek aránya, ezt az eltérést azonban okozhatják individuális sajátosságok is.

A beszédtípusok között egyértelműbben mutatkozik különbség a szünetek időtartamarányát tekintve is: a spontán beszédben átlagosan magasabb arányban fordultak elő szünetek. A beszélőnkénti bontás (8. ábra) lehetővé teszi az individuális sajátosságok figyelembevételét is.

Az eredmények alakulása a korábbi vizsgálati szempontokhoz hasonlóan nagyon jelentős egyéni különbségeket mutatott. Az olvasásban 4 beszélőnél 10 évvel később nagyobb volt a szünetek aránya, mint az eredeti felvétel esetében, 3 beszélőnél ellentétes tendencia mutatkozott, azaz kevesebb szünetet produkáltak a második felvétel alkalmával, míg a beszélők közel felénél, 6 beszélőnél nem mutatkozott különbség a két felvétel értékei között. A spontán beszédben csupán 3 beszélőnél nem tapasztaltunk jelentős különbséget az értékekben, 6 esetben csökkent a szünetek aránya, 4 esetben pedig nőtt 10 év elteltét követően. Az eredmények tehát nagyon erős egyéni tendenciákat igazolnak, ahol kevésbé érvényesülnek általánosnak tekinthető jellemzők. Ha a beszédtípusoktól függetlenül elemezzük az arányokat, azt látjuk, hogy a tendencia csupán két beszélőnél volt egyértelmű, a B02-es adatközlő esetében egyértelműen csökkent a szünetek aránya, míg a B08-as adatközlőnél egyértelmű növekedés volt kimutatható függetlenül a beszédtípustól. A többi adatközlőnél vagy vegyes tendencia mutatkozott (pl. B01-es adatközlő: az olvasásban növekedés, a spontán beszédben csökkenés volt megfigyelhető), vagy csupán az egyik beszéd-



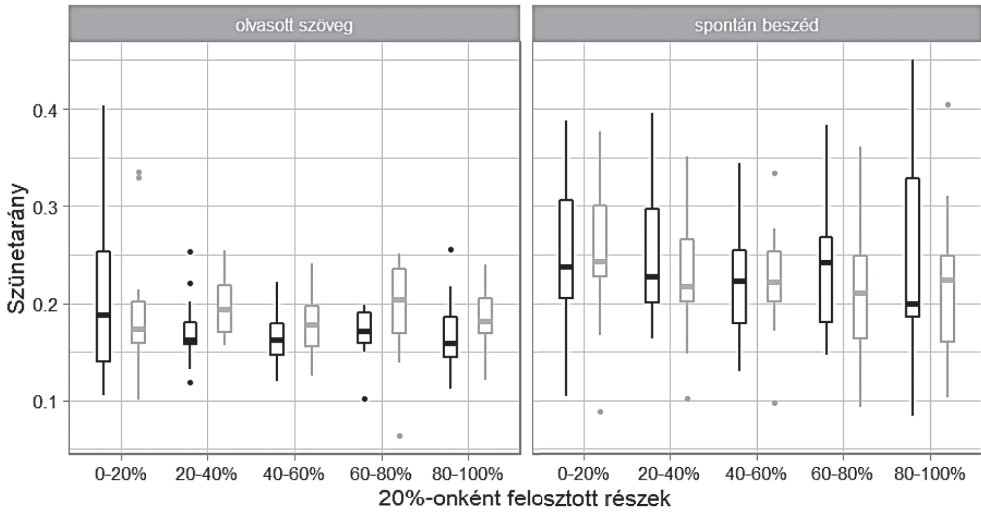
**8. ábra.** A szünetek időtartamaránya a teljes beszédidőre vetítve a két beszéd típusban, a felvétel és a beszélők szerinti bontásban

típusban történt változás. Az adatközlők többségénél (8 beszélő) mindkét felvételen egyértelműen megmutatkozik a fent említett tendencia, miszerint a spontán beszédben nagyobb a szünetek időtartamaránya. A többi beszélőnél is általában az egyik felvételen látható ez a tendencia, a másikon pedig elhanyagolható a különbség a két beszéd típus között. Egyetlen beszélő esetében (B09) tapasztalunk fordított irányú tendenciát a két felvételen, nála az olvasásban mutatkoznak nagyobb arányban a szünetek.

A 9. ábra a szünetarány alakulását mutatja a felvételeken belül, 20%-os bontásban. Nem fedezhető fel tendencia sem a beszéd típus, sem pedig a felvétel ideje szerint az egyes részek között. A 20%-os részek átlagai közti különbség sem a spontán beszédben, sem az olvasott szövegben nem haladja meg a 0,05-öt sem az 1., sem a 2. felvételek esetében, a különbség tehát elhanyagolhatónak tekinthető ( $SD_{olv} = 0,05$ ;  $SD_{sp} = 0,07$ ).

A statisztikai elemzés szerint a szünetarányban a beszéd típus szerint szignifikáns különbség volt az olvasás és a spontán beszéd között ( $F(1, 52) = 61,839$ ,  $p < 0,001$ ), míg a 20%-os szakaszok és a felvételek között nem volt kimutatható statisztikai különbség.

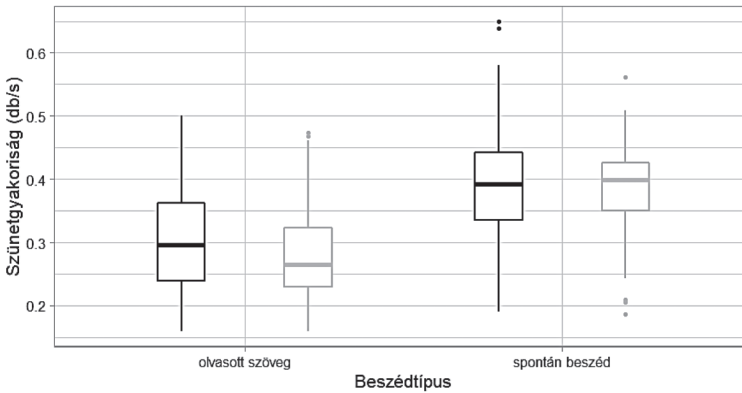
A szünetek gyakorisága az olvasott szövegben az 1. felvétel alkalmával átlagosan 0,3 db/s (SD: 0,06 db/s), a 2. felvétel alkalmával pedig 0,28 db/s (SD: 0,05 db/s)



Felvétel ☐ 1. ☐ 2.

9. ábra. Szünetek aránya 20%-os bontásban (medián és interkvartilis tartomány)

volt. A spontán beszédben az 1. felvétel alkalmával átlagosan 0,4 db/s (SD: 0,07 db/s), a 2. felvétel alkalmával pedig 0,39 db/s (SD: 0,06 db/s) volt (10. ábra).



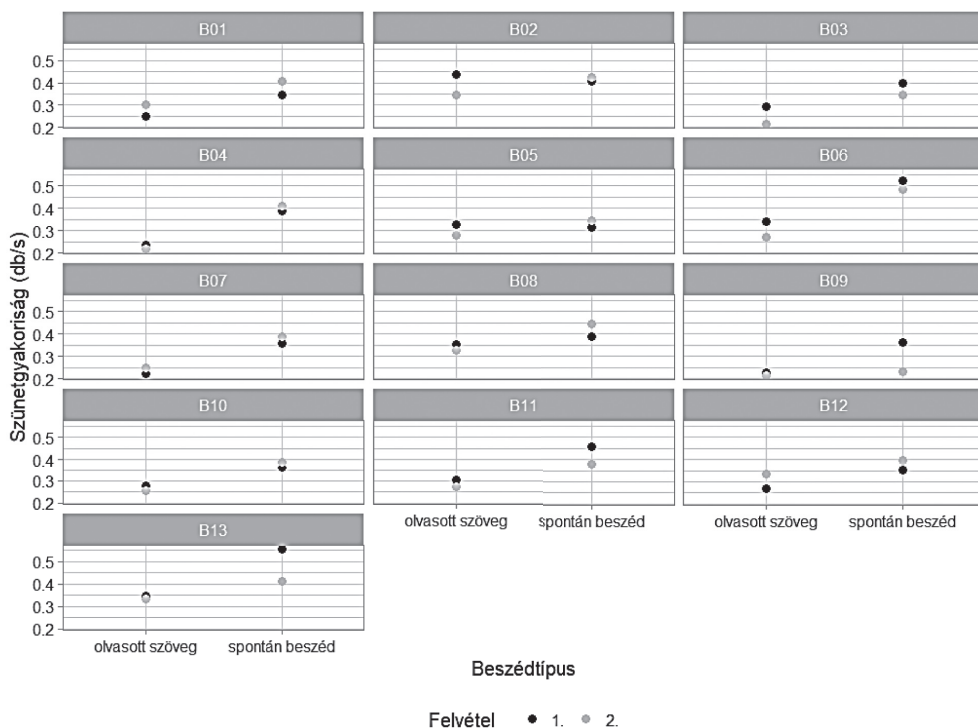
Felvétel ☐ 1. ☐ 2.

10. ábra. A szünetek gyakorisága a beszédtípus és a felvételek szerinti bontásban (medián és interkvartilis tartomány)

A szünetarányhoz hasonlóan a gyakoriságnál is látható, hogy a beszédtípusok között van elsősorban különbség, a spontán beszédben gyakoribbak a szünetek: nagyjából 0,1 db/s-nyi eltérés mérhető a gyakoriságok átlagában, ami a 0,06 db/s-os szórás mellett nem elhanyagolható.

A felvételek ideje között, szintén, ahogy a szünetarányban is láttuk, nincsenek jelentős különbségek, illetve az olvasott szöveg esetén az látszik az ábráról, hogy valamivel kisebb a szünetek gyakorisága. Ha az átlagokat hasonlítjuk össze, akkor azonban nem tapasztalunk jelentős változásokat, bár az értékek ugyanabban az irányban módosulnak (mindkét beszéd típusban a 2. felvételen kisebb a gyakoriságok átlaga), ezek a változások csak 0,01–0,02 db/s-osak, tehát elhanyagolhatóak a 0,06 db/s-os szórások mellett.

A 11. ábra az individuális különbségeket is mutatja, a korábbiakhoz hasonló vegyes tendencia mutatkozott meg. Az olvasásban két esetben a második felvétel alkalmával gyakoribbá váltak a szünetek, mint az első esetben, négy esetben ritkábbak lettek, hét esetben azonban nem mutatkozott jelentős különbség a két felvételi alkalom között. A spontán beszédben három esetben nőtt, öt esetben csökkent a szünetek gyakorisága, öt esetben pedig nem mutatkozott jelentős változás a két felvételi alkalom között. Két beszélőnél beszéd típusától függetlenül nőtt, szintén két esetben csökkent a szünetgyakoriság, kilenc esetben azonban vagy vegyes, vagy elhanyagolható mértékű változás történt 10 év különbséggel. A felvételek időpontja szerint tehát vegyes az adatközlőknél mért szünetgyakoriság alakulása, a beszéd típus szerint azonban jól látható a fent említett tendencia: a beszélők többségénél (10 beszélő) mindkét felvételen egyértelműen gyakoribbak a szünetek a spontán beszéd esetében

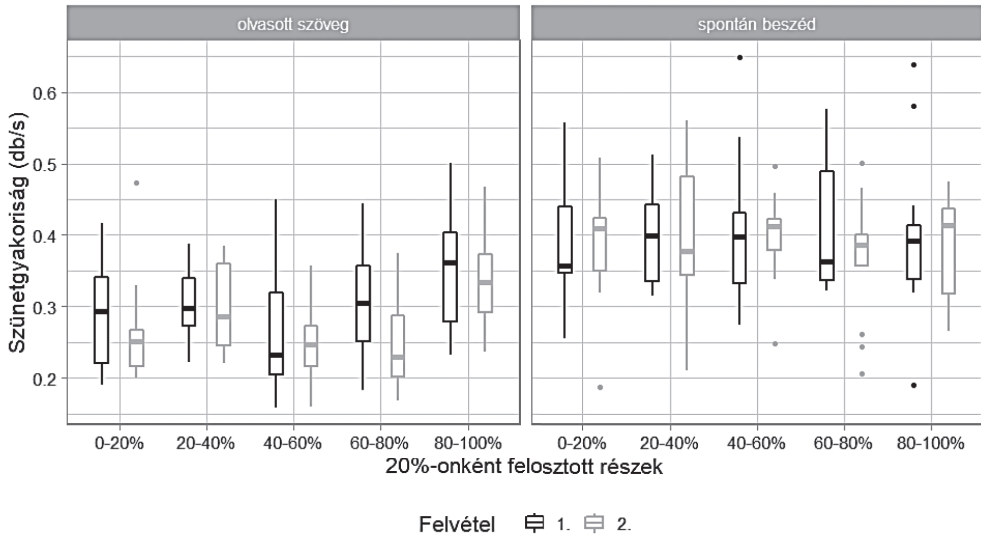


11. ábra. A szünetek gyakorisága a két beszéd típusban, a felvétel és a beszélők szerinti bontásban



az olvasott szöveghez viszonyítva. A maradék három adatközlő esetében is látható a fenti tendencia, de csak az egyik felvételükön, a másik felvételen a szünetgyakoriságok közti különbség nem haladja meg a 0,03 db/s-ot tehát elhanyagolhatónak tekinthető.

A 12. ábra a szünetek gyakoriságát mutatja a felvételeken belül, 20%-os bontásban. Az olvasott szöveg esetén nem mutatkozik tendencia a részek értékeiben egyik felvétel esetében sem, az átlagok közti különbségek nem haladják meg a 0,03 db/s-os értéket, tehát elhanyagolhatóak ( $SD_1 = 0,09$  db/s,  $SD_2 = 0,08$  db/s). Az olvasott szöveg esetében azonban (ahogy az artikulációs tempónál is) látható tendencia: az 1. felvételen az látszik, hogy a szöveg közép részén kevesebb szünetet tartottak az adatközlők, a vége felé pedig többet, ez a tendencia a 2. felvételre is elmondható, bár leginkább az utolsó rész különül el a többtől abban, hogy mindkét felvételen ott a leggyakoribb a szünettartás. Az 1. felvétel esetében a legalacsonyabb és legmagasabb átlagérték közti különbség 0,08 db/s, a 2. felvétel esetén pedig 0,1 db/s ( $SD_1 = 0,08$  db/s,  $SD_2 = 0,07$  db/s).



**12. ábra.** Szünetek gyakorisága 20%-os bontásban (medián és interkvartilis tartomány)

A statisztikai elemzés alátámasztotta, hogy a beszéd típus ( $F(1, 260) = 159,450$ ;  $p < 0,001$ ), a felvételek rögzítésének időpontja ( $F(1, 260) = 4,312$ ;  $p = 0,039$ ) és a 20%-os részek ( $F(4, 260) = 3,629$ ;  $p = 0,007$ ) is hatással vannak a szünetek arányára. A Tukey post hoc szerint az olvasott szövegben az utolsó szakasz (80–100%) tért el minden másik szakasztól, ebben a szakaszban több szünetet tartottak az adatközlők időegységre vetítve (0-20%:  $p = 0,021$ ; 20-40%:  $p = 0,024$ ; 40-60%:  $p = 0,007$ ; 60-80%:  $p = 0,027$ ).

Megvizsgáltuk, hogy kimutatható-e összefüggés, és ha igen, milyen jellegű kapcsolat mutatkozik a szünetarány és a szünetgyakoriság változása között a beszéd típusok között (1. táblázat). A 8. és 11. ábrát összevetve elmondható, hogy beszélőnként hasonló eltérések láthatók a spontán beszéd és olvasás között a szünetarányok és a szünetgyakoriságok tekintetében: amelyek beszéd típusban időarányosan több a szünet, ott általában gyakoribb is, természetesen vannak individuális különbségek a szünettartásban. Az olvasásban mért eredmények alapján elmondható, hogy nem volt egységes a beszélők vizsgált paramétereinek változására vonatkozóan. A legtöbb beszélőnél (4 fő) nem változott sem a szünetek aránya, sem azok gyakorisága. 3 fő esetében, noha a szünetek aránya nem változott, a gyakoriságuk csökkent, azaz hosszabb szüneteket produkáltak a második felvétel alkalmával. 2 beszélőnél nőtt a gyakoriság és az arány is, azaz gyakrabban és közel azonos hosszúságú szüneteket valósítottak meg, a többi adatközlő tendenciái egyéniek voltak, és nagy változatosságot mutattak.

A spontán beszédben, noha a legtöbb beszélőnek (3 fő) csökkent szünetaránya, nem változott azok gyakorisága, tehát rövidebb szüneteket valósítottak meg a második felvételen, mint az első esetében. A többi esetben az egyéni különbségek domináltak.

**1. táblázat.** A szünetarány és a szünetgyakoriság változása az olvasásban és a spontán beszédben (a + jel a vizsgált tényező növekedését [min. 10%], a – jel a [min. 10%] csökkenést mutatja a 2. felvétel esetében mért értékeknek az 1. rögzítéshez képest, a x jel esetében nem volt változás a két tényező között)

Beszélő	Szünetarány		Szünetgyakoriság	
	Olvasás	Spontán beszéd	Olvasás	Spontán beszéd
B01	+	–	+	+
B02	–	–	–	x
B03	x	+	–	–
B04	x	–	x	x
B05	x	–	–	x
B06	x	+	-	-
B07	x	+	x	x
B08	+	+	–	+
B09	x	-	x	–
B10	–	x	x	x
B11	+	x	x	–
B12	+	x	+	+
B13	x	–	x	–

#### 4. Következtetések

Az utóbbi években a beszéddel foglalkozó tudományterületek figyelme a beszédben előforduló extra- és a paralingvisztikus információk vizsgálata felé fordult. Ennek oka elsősorban az, hogy ezek jelentősen befolyásolják a nyelvi információ értékelését, illetve további ösztönzésül szolgálnak a technológiai lehetőségek javításával kapcsolatban, valamint egyre növekszik a társadalmi érdeklődés az életkorra vonatkozóan, és ez szintén sürgetőleg hat az ilyen típusú kutatások kivitelezésére. A jelen vizsgálat ezeknek az igényeknek eleget téve 13 fiatal felnőtt férfi beszélő közlései alapján azt vizsgálta, hogy mintegy 10 év elteltével miként változnak a beszéd egyes időzítési sajátosságai két különböző beszéd típusban, interjúformájú spontán beszédben és olvasott szövegben.

Eredményeink azt mutatták, hogy 1. a spontán beszédre lassabb artikulációs tempó jellemző, 2. a két beszéd típus inkább varianciájában tért el egymástól: a spontán beszédet nagyobb variabilitású artikulációs tempó jellemezte, mint az olvasást. Az 1. hipotézisük első felét tehát megerősítették az adatok. Ennek egyik lehetséges magyarázata lehet az, hogy a spontán beszéd és az olvasás eltérő kognitív terhelést jelent a különböző mértékű előkészítettségéből és ezzel szorosan kapcsolódó, eltérő tervezési és kivitelezési folyamatokból adódóan (vö. összefoglalóan Bóna 2019).

Ismert, hogy az eltérő nehézségű tervezési folyamatokra többek között a beszéd temporális jellemzői is utalnak (Fletcher 2010; Bóna 2013; Markó 2014). A spontán beszéd létrehozása során a beszédtervezés és -kivitelezés egyidejűleg zajlik, a beszélő az adott pillanatban, beszédhelyzetben egyszerre tervezi meg az üzenet tartalmát és nyelvi formáját, a felolvasás során ezzel szemben nincs szükség a magasabb tervezési folyamatok működtetésére, azaz az üzenet és a nyelvi forma megtervezésére, hiszen a felolvasandó nyelvi anyag adott (Váradai 2011; Krepsz 2016). Mindez erősen meghatározó mind az artikulációs tempó nagysága, mind annak variabilitása szempontjából. Az előkészítettség mellett befolyásoló tényező lehet a szövegek hossza is: az olvasás időtartama rövidebb, átlagosan 1-1,5 percet vett igénybe a szöveg meghangosítása, a spontán beszéd az egyéni habitustól, aktuális fizikai és lelki állapottól függően 15 perces is lehetett, és ez szintén nagyobb varianciát idézhetett elő.

A felvételek között eltelt 10 év nem gyakorolt jelentős hatást a közlések tempójára, az 1. és 2. felvétel között nem mutatkozott a beszélőkre jellemző, azonos tendenciájú különbség: azonos számú beszélőnél volt növekvő és csökkenő a tendencia, 3 esetben egyáltalán nem mutatkozott változás, tehát a 2. hipotézisünk első felét nem erősítették meg az adatok. Az eredmények részben alátámasztják, hogy fiatal felnőttkorban kevésbé jelentős változások várhatók még ilyen hosszabb időtartamot, akár egy évtizedet követően is egészséges beszélőknél. Emellett eredményeink egybehangzanak Gesternberg és munkatársainak (2011) eredményeivel, akik idősebb beszélőknél (a vizsgált nyelvektől függetlenül) szintén nem egységes tendenciáról és jelentős individuális sajátosságokról számoltak be. Az idősödés folyamata (függetlenül attól, hogy mely életszakaszban vizsgáljuk) erősen függ az egyéni jellemzőktől, az öröklött tényezőktől, valamint a környezeti sajátosságoktól is.

A spontán beszéd artikulációs tempójának alakulása mindkét felvétel (1. és 2. rögzítés) esetében „stabilnak” mutatkozott, tehát nem volt meghatározó, hogy az adott felvétel mely pontján mértük a tempóértékeket. Ezzel szemben az olvasásban egy kis kitérésű, de jellemzően parabolajellegű görbe volt leírható, azaz az olvasás elején és végén nagyobb, a közepén kisebb tempóértékek voltak jellemzők. Ezek az eredmények ellentmondanak a Beatrix királynőnél tapasztalt kúp alakú elrendeződésnek, nála ugyanis a háromperces szakaszok átlagos tempóértéke fokozatos növekedést, majd körülbelül a közlés felétől csökkenést mutattak. Ennek egyik lehetséges oka, hogy a királynői beszéd lehet begyakorlott, hosszabb felkészülési idővel előadott mű, míg a naiv beszélők esetében rövid felkészülési idő áll rendelkezésre (átolvashatják néhány másodpercben a szöveget), illetve adatközlőink nem professzionális beszélők.

A beszéd típusok közti tendenciákat a normalizált értékek is alátámasztották. A 3. hipotézisünk első felét tehát nem erősítették meg az adatok. Mindez valószínűsíthetően azzal magyarázható, hogy míg az olvasás esetében belátható a szöveg hossza, ilyen módon a beszélő jobban tudja kontrollálni, tervezni a saját tempóját, addig a spontán beszéd helyzetben az esetlegesen felmerülő kérdések, új témák folyamatosan alakítják a közlés sebességét, ami egy „kényelmes” beszédtempóra való beállást enged meg a beszélőnek. Emellett adódhat mindez a szöveg felépítéséből, tagolásából is, noha a tendencia nem volt azonos minden beszélőnél.

A kisebb időegység szerinti bontás jelentős egyéni különbségeket mutatott az olvasás feladatában, amelyek azonban szisztematikus sajátosságokat mutattak a 10 éves különbség összevetése alapján is a temporális sajátosságokban. Ugyanez a tendencia a spontán beszédben nem volt kimutatható, ami elsősorban a korábban felsorolt tervezési és kivitelezési indokokkal magyarázható.

A szünetelési sajátosságok vegyes képet mutattak. Mind a szünetarány, mind a szünetek gyakoriságának tekintetében tapasztalható volt különbség a két beszéd típus között, az 1. hipotézisünk második felét tehát szintén megerősítették az adatok.

A változás mértéke a 10 éves különbségnél a legtöbb beszélő esetében igen kis mértékű volt. A spontán beszéd esetében a felvételek közti különbségek teljesen elhanyagolhatónak bizonyultak, az olvasott szöveg esetén már tendenciaszintű eltérést tapasztaltunk a szünetarány tekintetében: a 2. felvételek esetében nagyobb volt a szünetek aránya az elsőhöz viszonyítva. A szünetek gyakoriságát tekintve szignifikáns különbséget találtunk a két felvétel között: az olvasott szöveg esetében gyakoribbak voltak a szünetek az 1. felvételen, mint a másodikon. A 2. hipotézisünk második felét tehát részben (a szünetek gyakorisága tekintetében) támasztották alá az adatok. A felvételeken belüli változásokban a spontán beszéd esetében nem volt felfedezhető tendencia sem a szünetarányt, sem a szünetek gyakoriságát tekintve. Az olvasott szöveg esetében a szünetarány tekintetében szintén nem mutatkozott jellemző tendencia az egyes részek között, a gyakoriság esetében a felvételek utolsó 20%-a tér el szignifikánsan a többi részhez tartozó értékektől: az utolsó részben nagyobb volt a szünetek gyakorisága. A 3. hipotézisünk második felét tehát szintén részben (a szünetek gyakorisága tekintetében) alátámasztották az adatok.

A temporális jellemzőket összevetve elmondható, hogy az olvasott szövegre nagyobb artikulációs tempó, kisebb szünetarány és kevesebb szünet, ezáltal hosszabb beszédszakaszok jellemzőek a spontán beszédhez viszonyítva. Az olvasott szövegben található – több esetben csak tendenciaszerű – összefüggés a felvételen belüli részek jellemzői között arra enged következtetni, hogy a felolvasások elején és végén gyorsabb artikulációs tempó, a végén pedig gyakoribb szünetek, ebből adódóan hosszabb beszédszakaszok jellemzőek. Ezen temporális sajátosságok adódhatnak a felolvasott szöveg szerkezetéből és tagolhatóságából. Ezen túl számos egyéb tényező is hathat az időzítési sajátosságokra (vö. Krepsz 2016).

Az időződés folyamatának jobb megértése releváns az alkalmazott nyelvészet, az ember és gép kommunikációja, valamint a beszédfelismerés szempontjából, valamint szorosan kapcsolódik a napjainkban egyre nagyobb figyelmet kapó idősekhez irányuló beszédvel („elderspeak”) és az idősebb felnőttek életminőségét segítő és javító eszközök fejlesztésével kapcsolatban.

### SZAKIRODALOM

- Boersma, Paul – Weenink, David 2019. *Praat: Doing phonetics by computer* (computer program version 6.0.19).
- Bóna Judit 2013. A beszédszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszéd kutatás 2013*: 60–75.
- Bóna Judit 2014. Temporal characteristics of speech: The effect of age and speech style. *Journal of the acoustical society of America* 136/2: EL116–EL121. <https://doi.org/10.1121/1.4885482>
- Bóna Judit 2019. A spontán beszéd és a felolvasás temporális jellemzői kisiskolás korban. *Beszéd kutatás 2019*: 272–90.
- Brown, W. S. – Morris, Richard, J. – Michel, John F. 1989. Vocal jitter in young adult and aged female voices. *Journal of Voice* 3/2: 113–9. [https://doi.org/10.1016/S0892-1997\(89\)80137-7](https://doi.org/10.1016/S0892-1997(89)80137-7)
- Brückl, Markus – Sendlmeier, Walter 2003. Ageing female voices: an acoustic and perceptive analysis. *VOQUAL'03*, 163–8.
- Brückl, Markus 2007. *Women's vocal aging: a longitudinal approach*. Interspeech 2007 – 8th Annual Conference of the International Speech Communication Association. Antwerp, 1170–3.
- Dankovičová, Jana 1999. Articulation rate variation within the intonational phrase in Czech and English. In: *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco, 269–72.
- Duchin, Sandra W. – Mysak, Edward D. 1987. Disfluency and rate characteristics of young adult, middle-aged, and older males. *Journal of Communication Disorders* 20: 245–57. [https://doi.org/10.1016/0021-9924\(87\)90022-0](https://doi.org/10.1016/0021-9924(87)90022-0)
- Endres, G. – Bambach, A. – Flösser, M. 1971. Voice Spectrograms as a Function of Age, Voice Disguise and Voice Imitation. *JASA* 49:1842–7. <https://doi.org/10.1121/1.1912589>
- Gerstenberg, Annette 2011. *Generation und Sprachprofile im höheren Lebensalter: Untersuchungen zum Französischen auf der Basis eines Korpus biographischer Interviews*. Klostermann, Frankfurt am Main.
- Gósy Mária – Beke András – Horváth Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszéd kutatás 2011*: 5–31.

- Grácsi Tekla Etelka – Huszár Anna – Krepsz Valéria–Százaz Bettina – Damásdi Nóra – Markó Alexandra 2020. Longitudinális korpusz magyar felnőtt adatközlőkről. In: Berend Gábor–Gosztolya Gábor – Vincze Veronika (szerk.): *XVI. Magyar Számítógépes Nyelvészeti Konferencia*, 103–14.
- Hoit, Jeannette D. – Hixon, Thomas K. – Altman, Morris – Morgan, Wayne J. 1989. Speech breathing in women. *Journal of Speech and Hearing Research* 32: 353–65. <https://doi.org/10.1044/jshr.3202.353>
- Hoit, Jennette De. – Watson, Peter J. – Hixon, Kimberly E. – McMahon, Patrícia – Jacewicz, Eva – 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. <https://doi.org/10.1121/1.3459842>
- Johnson, Cynthia L. 1994. Age and velopharyngeal function during speech production. *Journal of Speech, and Hearing Research* 37/2: 295–302. <https://doi.org/10.1044/jshr.3702.295>
- Kohári Anna 2016. *Időzítési mintázatok a magyar beszédben*. Disszertáció. ELTE BTK, Budapest.
- Lennon, Paul 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning* 40/3: 387–417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- Linville, Sue Ellen 1996. The sound of senescence. *Journal of Voice* 10/2: 190–200. [https://doi.org/10.1016/S0892-1997\(96\)80046-4](https://doi.org/10.1016/S0892-1997(96)80046-4)
- Linville, Sue Ellen 2000. The aging voice. In: Kent, Raymond D. – Ball, Martin J. (eds.): *Voice Quality Measurement*. Singular Thomson Learning, Australia, San Diego.
- Linville, Sue Ellen 2001. *Vocal aging*. Singular Thomson Learning, Australia, San Diego. [https://doi.org/10.1016/S0892-1997\(96\)80046-4](https://doi.org/10.1016/S0892-1997(96)80046-4)
- Neuberger Tilda – Gyarmathy Dorottya – Grácsi Tekla Etelka – Horváth Viktória – Gósy Mária – Beke András 2014. Development of a large spontaneous speech database of agglutinative Hungarian language. In: Sojka, Petr – Horák, Aleš – Kopeček, Ivan – Pala, Karel (eds.): *Proceedings of TSD*. Springer, New York, 424–31. [https://doi.org/10.1007/978-3-319-10816-2\\_51](https://doi.org/10.1007/978-3-319-10816-2_51)
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Akadémiai Kiadó, Budapest.
- Oyer, Herbert – Deal, Leo V. 1985. Temporal aspects of speech and the aging process. *Folia Phoniatrica* 37:109–12. <https://doi.org/10.1159/000265788>
- Ptacek, Paul – Sander, Eric, S. 1966. Age recognition from voice. *Journal of Speech and Hearing Research* 9/2: 273–7. <https://doi.org/10.1044/jshr.0902.273>
- Quené, Hugo 2013. Longitudinal trends in speech tempo: The case of Queen Beatrix. *The Journal of the Acoustical Society of America* 133/6: 452–7. <https://doi.org/10.1121/1.4802892>
- Ramig, Lorraine A. 1983. Effects of physiological aging on speaking and reading rates. *Journal of Communication Disorders* 16: 217–26. [https://doi.org/10.1016/0021-9924\(83\)90035-7](https://doi.org/10.1016/0021-9924(83)90035-7)
- Ringel, Robert L. – Chodzko-Zajko, Wojtek J. 1987. Vocal indices of biological age. *Journal of Voice* 1/1: 31–7. [https://doi.org/10.1016/S0892-1997\(87\)80021-8](https://doi.org/10.1016/S0892-1997(87)80021-8)
- Shum, Stephen 2008. The Language of an Aging Queen. *Linguistics 110 Final Paper*, December 2008. [http://people.csail.mit.edu/sshum/ucb\\_papers/aging\\_queen.pdf](http://people.csail.mit.edu/sshum/ucb_papers/aging_queen.pdf) (Letöltve: 2020. 02. 17.)
- Smith, Bruce L. – Wasowicz Jan – Preston Judy 1987. Temporal characteristics of the speech of normal elderly adults. *Journal of Speech and Hearing Research* 30/4: 522–9. <https://doi.org/10.1044/jshr.3004.522>
- Vaughan, Nancy E. – Furukawa, Izumi – Balasingam, Nirmala – Mortz, Margaret – Fausti, Stephen A. 2002. Time-expanded speech and speech recognition in older adults. *Journal of Rehabilitation Research and Development* 39/5: 559–66.

- Váradi Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszéd kutatás 2010*: 100–109.
- Walker V. G. 1988. Durational characteristics of young adults during speaking and reading tasks. *Folia Phoniatrica et Logopaedica* 40: 12–20. <https://doi.org/10.1159/000265879>
- Winkler, Ralf 2007. Influences of pitch and speech rate on the perception of age from voice. 16th International Congress of Phonetic Sciences. *Proceedings*. Saarbrücken, Germany, 1849–52.
- Yuan, Jiahong–Lieberman, Mark–Cieri, Christopher 2006. Towards an integrated understanding of speaking rate in conversation. INTERSPEECH – ICSLP, 9th International Conference on Spoken Language Processing. *Proceedings*. Pittsburgh (PA), USA, 541–4.
- Zellner-Keller, Brigitte 2006. Ageing and speech prosody. Speech Prosody – 3th International Conference. *Proceedings*. Dresden, Germany.

*Huszár Anna*

tudományos kutató

Nyelvtudományi Kutatóközpont

<https://orcid.org/0000-0002-2951-1918>

*Krepsz Valéria*

tudományos kutató

Nyelvtudományi Kutatóközpont

<https://orcid.org/0000-0002-2099-6285>

**A kutatást az NKFIH-FK128814-es számú pályázat támogatta.**

## SUMMARY

### **The development of variability within the speaker based on type of speech and position in the recordings in a ten-year perspective**

*Huszár, Anna – Krepsz, Valéria*

In recent years, attention given to speech-related disciplines has shifted to analysing extra- and paralinguistic information occurring in speech. This is mainly due to the fact that they significantly influence the evaluation of linguistic information and provide an additional incentive to improve technological opportunities. Moreover, public interest in the factor of age is growing which, in turn, may encourage the conduct of this type of research. Earlier research has mostly analysed older speakers, and, to date, we have little information about the changes taking place in the voices of healthy, young speakers as they get older. Our study examines how certain timing characteristics of speech change in two different types of speech over a period of about 10 years. 13 healthy, young adult male speakers' reading and spontaneous speech were selected from two databases. The development of temporal and pausing patterns was analysed with regard to the type of speech, position, the time of the recording (10-year difference). The paper focuses on timing differences within and between the recordings. The results showed that not only the types of speech but also the differences within and between the recordings affected the pausing strategies. In the case of articulation rate, individual differences dominated.

**Keywords:** temporal patterns, longitudinal data, articulation rate, age, speech type