



KOHÁRI
ANNA

BESZÉD • KUTATÁS • ALKALMAZÁS

IDŐZÍTÉSI
MINTÁZATOK
A MAGYAR
BESZÉDBEN



ELTE
EÖTVÖS
KIADÓ

Kohári Anna

IDŐZÍTÉSI MINTÁZATOK A MAGYAR BESZÉDBEN

Beszéd • Kutatás • Alkalmazás

KOHÁRI ANNA

**IDŐZÍTÉSI MINTÁZATOK
A MAGYAR BESZÉDBEN**

BUDAPEST, 2018





A kötet megjelenését a Magyar Tudományos Akadémia támogatta.

Lektorálták:

Beke András

Bóna Judit

Gocsál Ákos

© Kohári Anna, 2018

ISBN 978-963-312-295-2

ISBN 978-963-312-300-3 (pdf)

ISSN 2064-4442

A borító Vincze Miklós és Vincze Benedek Gergő *Masató* c. festményének felhasználásával készült.



www.eotvoskiado.hu



Felelős kiadó: Hunyady András ügyvezető igazgató

Projektvezető: Sándor Júlia

Kiadói szerkesztő: Brunner Ákos

Borítóterv: Csele Kmotrik Ildikó

Nyomdai előkészítés: Manzana Bt.

Nyomdai kivitelezés: Multiszolg Bt.

Tartalom

Sorozatszerkesztői előszó	7
Előszó.....	9
1. Bevezető.....	11
1.1. A beszéd időbeli szerveződésének lehetséges aspektusai	11
1.2. A szegmentumok időtartamára ható, eddig feltárt tényezők	13
1.2.1. A megnyilatkozás vagy intonációs frázis végi nyúlás	14
1.2.2. A hangsor eleji pozíció	16
1.2.3. A hangsúly és az időtartam kapcsolata	17
1.2.4. A szóhosszúság hatása	19
1.3. A beszédritmus időbeli dimenziója	20
1.3.1. A beszédritmus fogalma	21
1.3.2. Az időzítésen alapuló beszédritmusosztály-hipotézis	21
1.3.3. A beszédritmus-mérőszámok meghatározása.....	23
1.3.4. A beszédritmus-mérőszámok és más tényezők kapcsolata egy-egy nyelven belül.....	30
1.3.5. A magyar nyelven végzett akusztikai beszédritmus-kutatások.....	34
1.4. Az artikulációs tempó és a beszédtempó	35
1.4.1. Tempóingadozások a beszédben.....	36
1.4.2. Az artikulációs és beszédtempóra ható általános tényezők.....	38
1.4.3. A szünetezés	40
1.4.4. Az artikulációs és beszédtempó szélsőséges értékeinek korlátai	41
1.5. A beszédtechnológia és az időszerkezet.....	42
2. A kísérletsorozat kérdésfeltevései, célja, hipotézisei	47
3. Általános módszertan	49
3.1. Kísérleti személyek és anyag	49
3.2. Módszer, szegmentálási nehézségek	51

4. Kísérletsorozatok az időzítési mintázatok vizsgálatára	55
4.1. Az időzítés mint a beszédritmus egyik dimenziója.....	55
4.1.1. Kísérleti személyek, anyag és módszer	56
4.1.2. Eredmények	59
4.1.2.1. Különböző beszéd módok általános időzítési jellemzőinek összehasonlítása a beszédritmus-mérőszámok alapján	59
4.1.2.2. A szünetek és megakadások kiküszöbölésével kapott értékek	63
4.1.2.3. A nemek hatása a beszédritmus-mérőszámok értékeire	82
4.1.2.4. A különböző beszélők szünettartási szokásainak hatása a beszédritmusra	89
4.1.2.5. A magánhangzók és a mássalhangzók időtartamának az aránya és a globális, valamint a lokális artikulációs tempó összefüggései	94
4.1.3. A beszédritmus-mérőszámok eredményei alapján a magyar beszédre levonható általános következtetések	100
4.2. A lassulások és a gyorsulások a magyar beszédben	103
4.2.1. Kísérleti személyek, anyag és módszer	104
4.2.2. Eredmények.....	109
4.2.2.1. Tempóbeli változások általános trendje a szövegfelolvasások tagmondataiban	109
4.2.2.2. Lassulás és gyorsulás a szakasz nagyságtól függően.....	116
4.2.2.3. A lépésstatistikai eredmények különböző alapegységekre számolva	119
4.2.2.4. A lépésstatistikai eltérések beszéd mód szerint.....	126
4.2.2.5. A nemek közti különbségek a lépésstatistikában.....	130
4.2.2.6. Tömbösödés a növekvő és csökkenő lépések alapján	131
4.2.3. A lépésstatistikai eredmények értelmezése és a levonható általános következtetések	134
4.3. A temporális szerkezet és a zöngeminőség kapcsolata a megnyilatkozás végén.....	138
4.3.1. Kísérleti személyek, anyag és módszer	139
4.3.2. Eredmények.....	142
4.3.3. Következtetések.....	149
4.4. Összegzés és kitekintés	151
 Irodalom.....	 155
 Mellékletek	 171
 Temporal Patterns in Hungarian Speech.....	 175

Sorozatszerkesztői előszó

A Beszéd – Kutatás – Alkalmazás sorozat nyolcadik köteteként megjelenő mű egy olyan vizsgálatsorozatot mutat be, amely hiánypótló a magyar beszéddel foglalkozó szakirodalomban. Bár általánosságban a magyar beszéd időzítésével kapcsolatban sok leírás megjelent már, a beszédritmus eddigi megközelítései, az ezzel kapcsolatos korábbi állítások a ritmus sajátosságainak megragadhatatlanságát, a sok változóból adódó bizonytalanságot sugallták, nem véletlenül. A jelenség valóban soktényezős, az adatok gyakran ellentmondásosak, nehéz fogódzókat találni a megfelelő módszertan kialakításához. Kohári Anna azonban vette a bátorságot, hogy erre az ingoványos talajra lépjen, és útját siker koronázta. A nemzetközi szakirodalom széles körének ismeretében új, korábban a (magyar) beszédre még nem alkalmazott módszertanokat használva, különféle metódusokat ötvözve, szorgalmas, aprólékos és szisztematikus elemző munkával jutott el azon eredményekig és megállapításokig, amelyek ebben a kötetben napvilágot látnak, és amelyek a legkorszerűbb ismereteinket foglalják össze a (magyar) beszéd ritmusának vonatkozásában.

A kötet bevezetése tananyagként is használható, mivel áttekinti és értelmezi a vonatkozó tudományos fogalomkészletet és terminológiát, továbbá számot ad a nemzetközi és a magyar kutatási eredményekről a legutóbbi időkig bezárólag. A második fejezettől kezdődően a szerző saját kutatásának lépéseit ismerjük meg, az elemzett anyag, az alkalmazott módszerek és az eredmények részletes, jól illusztrált áttekintését kapja az olvasó. Kohári Anna arra is rámutat, hogy a kapott eredmények mely területeken és milyen módon hasznosulhatnak, valamint kijelöli a további kutatások lehetséges irányait is.

Mindezek alapján a kötet nemcsak a szűkebb, fonetikusokból álló olvasótábor érdeklődésére tarthat számot, hanem olyan területek művelői is építhetnek a benne foglalt ismeretekre, amelyek a beszéd időzítéséhez bármilyen módon kapcsolódnak, a logopédiától a beszédtechnológiáig. Az a kutatómódszertani innováció, amelyre a kötet példát ad, azonban még távolabbi, a beszédhez nem vagy kevésbé kapcsolódó, de az időzítés mintázatait magában rejtő jelenségek leírásában is haszonnal kecsegtet. Így a könyv bátran ajánlható a szélesebb érdeklődő közönség számára is.

Markó Alexandra

A beszéd hangjainak időtartamát számtalan tényező befolyásolja, úgymint a hangzó minősége, a hangzó környezete, a hangsorban elfoglalt helye stb. (vö. MAGDICS 1966; KASSAI 1982; VAN SANTEN 1992; KOVÁCS 2002; GÓSY–BEKE 2010; FLETCHER 2010). Léteznek olyan faktorok is, amelyek több egymás mellett lévő hang időtartamát egyszerre befolyásolják, ilyen például a közlésbeli pozíció, a hangsúlyos helyzet és a szóhosszúság hatása is (VAN SANTEN 1992; WHITE 2002; FLETCHER 2010; TURK–SHATTUCK–HUFNAGEL 2014). A különböző tényezők interakciójának eredményeképpen az egyes beszédhangok időtartama nehezen jósolható meg, de a folyamatos beszédben is megjelennek szabályokkal megragadható időtartambeli struktúrák, amelyeket időzítési mintázatoknak nevezünk.

Jelen kötet célja az, hogy akusztikai mérések segítségével első közelítést adjon a magyar beszédben megjelenő alapvető időzítésbeli mintázatokról. A témával kapcsolatos nemzetközi és hazai szakirodalom áttekintését követően a beszéd időzítésbeli sajátosságait feltáró saját kutatásaimat ismertetem.

A kötet első része a beszédhangok időtartamára ható tényezőket, a beszédritmus időbeli dimenzióját és az artikulációs tempó összefüggéseit tárgyalja, majd egy rövid kitekintést nyújt a beszéd időbeli szerveződésének modellezési lehetőségeiről a beszédtechnológiában. Ezek után külön fejezetben tárgyalom a kötet kérdésfelvetéseit. Saját kutatásaim ismertetése az anyag és a használt módszerek leírásával kezdődik, ezután ismertetem az egyes kérdésekre választ adó kutatások eredményeit. Legvégül összegzem a kutatások eredményeit, és röviden kitérek azok alkalmazhatóságára, hasznosíthatóságára.

Ez a munka nem jöhetett volna létre számos tanárom, mentorom, kollégám, barátom és a családom nélkül, szeretném nekik megköszönni mindazt, amit értem tettek. Nagy hálával tartozom Markó Alexandrának, témavezetőmnek az inspiráló órákért, a lelkesítő és elgondolkodtató konzultációkért, az építő kritikákért, amelyek mindig hasznosnak bizonyultak, és legfőképp köszönöm azt a rengeteg emberi és szakmai támogatást, amit kaptam tőle. Köszönöm Gósy Máriának, hogy megteremtette a kutatás elindításához szükséges alapvető feltételeket, és doktorandusz éveim alatt is segítette munkám létrejöttét. Köszönöm Beke Andrásnak, Bóna Juditnak, Gocsál Ákosnak, Csapó Tamás Gábornak az értekezéssel kapcsolatos tanácsaikat, kritikai észrevételeiket. Köszönet illeti Grácsi Tekla Etelkát a mérés módszertani segítségért, Mády Katalint a konzultációs lehetőségekért, és mert idejüket, energiájukat nem kímélve tanítottak, s tanulhattam tőlük. Köszönettel tartozom Kovácsné Lapu Máriának, a tanár-géniusznak, aki elindított egy úton, amelyen azóta is járok. Hálával tartozom édesanyámnak, édesapámnak, testvéreimnek, családomnak türelmükért, támogatásukért és érdeklődésükért. Köszönet illet mindenkit, aki bármilyen módon hozzászólt, meghallgatta az elmúlt években tartott előadásaimat, véleményt formált cikkeimről, érdeklődött munkám felől. Végül

szeretném megköszönni a megköszönhetlent Vincze Miklósnak, aki a nap bármely órájának bármely percében hajlandó volt meghallgatni a kutatással kapcsolatos gondolataimat, igazi társként kísért végig ezen a hosszú úton. Köszönöm neki és tüneményes kisfiamnak, Benedeknek, hogy időt varázsoltak számomra a kötet befejezéséhez.

A szerző

1. Bevezető

1.1. A beszéd időbeli szerveződésének lehetséges aspektusai

A beszéd időbeli szerveződése vagy más néven időszerkezete alatt azt a struktúrát értjük, amelyben az adott beszédesemény megvalósul (GÓSY 2004). A szakirodalomban hagyományosan két eltérő aspektus különül el egymástól: a szupraszegmentális időzítés (suprasegmental timing) és a szegmentális időzítés (segmental timing). Szupraszegmentális időzítéshez sorolják az olyan jelenségeket, amelyek a beszédhangoknál nagyobb egységek időzítési mintázatához köthetők. Hagományosan ilyen jelenségnek tekintik például az artikulációs és beszédtempót, illetve ezek változásait, valamint a szünettartást és a beszédritmust is. A szegmentális szerkezet időzítésén a beszédhangok időtartamának vizsgálatát értik (GÓSY 2004: 105; HARDCASTLE–LAVER–GIBBON 2010).

Szegmentális aspektusból a beszédhangok időtartamát befolyásoló elsődleges tényező a hangzó minősége. LEHISTE (1970: 18) bevezeti a „benne rejlő időtartam” (intrinsic duration) fogalmát, amelyet pontosan ez a fonetikai minőség, tehát a beszédhang artikulációs módja és helye határoz meg. Ez egy olyan alapvető, elméleti fogalom, amely az időtartam meghatározásakor kizárólag az adott beszédhang artikulációjából eredő hatásokat veszi figyelembe (LEHISTE 1970). Több olyan alapvető szabályszerűséget is kimutattak, hogyan hat a beszédhangok képzési módja (például spiráns-explozíva-likvida), képzéshelye (például elől képzett-hátul képzett) a hangok időtartamára (lásd MADDISON 1997; a magyar nyelvben például MAGDICS 1966; KASSAI 1982; KOVÁCS 2000; OLASZY 2006; GÓSY–BEKE 2010; GRÁCSI 2012). Az egyik ilyen, több nyelvben is kimutatott összefüggés, hogy a felső nyelvvállású magánhangzók mérhetően rövidebbek, mint az alsó nyelvvállású magánhangzók (PETERSON–LEHISTE 1960; VAN SANTEN 1992). Azt azonban hozzá kell tennünk, hogy a természetes beszédképzés során nem ejtünk izoláltan hangzókat. A hangkörnyezeten túl számtalan más tényező is hatással van az időtartamokra, így ehhez a teoretikus fogalomhoz közvetlenül nem társítható az élő-beszédből nyert mért adat. A benne rejlő időtartam fogalmának elkülönítése azért is tekinthető jelentősnek, mert kiindulási alapot ad további szabályszerűségek, tendenciák megfogalmazásához. Az egymást követő beszédhangok időtartamának ehhez hasonló megközelítése megtalálható a beszédszintézis szabályalapú modelljeiben is (például VAN SANTEN 1992; OLASZY 2000). Ezekben a modellekben viszont nemcsak a hangzó minőségét veszik alapul, hanem az adott beszédhang környezetét is figyelembe veszik (ALLEN et al. 1987; OLASZY 2010c). Az ún. specifikus időtartam olyan alapidőtartamot jelent, amely a beszédhang artikulációs konfigurációjából és a közvetlen hangkörnyezetének koartikulációs hatásából fakad. A specifikus

időtartamok meghatározásakor ezért nem az egyes hangok átlagos időtartamát vették figyelembe, hanem hármás hangkapcsolatok középső elemének időtartamát vizsgálták. A pontos érték meghatározásához nem nagy mennyiségű mért adatokat, hanem percepciós ítéleteket használtak fel. Az így kapott értékeket szabályok révén módosították, hogy megkapják az egyes beszédhangok végleges időtartamát a folyamatos beszédben (OLASZY 2000, 2010).

A beszéd időbeli szerveződésének másik lehetséges aspektusa a szupraszegmentális megközelítés. A szupraszegmentális szerkezetet általánosságban az alapján különböztetik el a szegmentális aspektustól, hogy a szupraszegmentális jellemzők – szemben a másik szint tulajdonságaival – nem határozhatók meg egy-egy beszédhang alapján, hanem több beszédhangon keresztül relatív különbségekként érvényesülnek (LEHISTE 1970; NOOTEBOOM 1997: 640; LAVER 1994; MARKÓ 2005; LADEFOGED 2011). Amíg a szegmentális aspektus vizsgálatához az időtartamnak, a frekvenciának, az intenzitásnak egy adott beszédhangra, beszédhangrészletre jellemző abszolút értékeit használják fel általában, addig a szupraszegmentális szerkezet feltárásához inkább ezen tényezők relációinak, változásainak (tempóingadozás, alaphangmagasság-változás) tanulmányozását helyezik előtérbe (HARDCASTLE–LAVER–GIBBON 2010; MARKÓ 2005). Ebből is látszik, hogy a szupraszegmentális és szegmentális aspektus csak teoretikusan különíthető el élesen egymástól, a gyakorlatban nehezen határolható el a szupraszegmentális jellemzők a hangsorban előforduló beszédhangok tulajdonságaitól. Az időbeli szerveződés esetében is nehezen választható szét a szegmentumok időzítése olyan jelenségektől, mint a közlés/megnyilatkozás végén mérhető nyúlás vagy több nyelvben a hangsúlynak az időtartamra gyakorolt hatása. Ezeket a jelenségeket gyakran egy-egy hang időtartamának mérése alapján vizsgálják (például KASSAI 1979; KOVÁCS 2002; WHITE–MÁDY 2008), ugyanakkor előfordul, hogy több egymás melletti hangot vagy a szótagot, szótagokat tekintik a kutatás alapegységének (például CAMBIER-LANGEVELD 1997; SUOMI 2007). Az egyik legnagyobb kérdés a szupraszegmentális időzítés vizsgálatakor, hogy a hangsornak mekkora részét vizsgáljuk egyben, és annak tulajdonságait milyen kisebb egységeken keresztül határozzuk meg.

Szintén az elkülönítés problematikusságát mutatja, hogy egyes szerzők más csoportosításban tárgyalják azokat a tényezőket, amelyek meghatározzák az időzítést a beszédben. Jelen értekezésben alapvetően a TURK és SHATTUCK-HUFNAGEL (2014) munkájában megjelent kategorizálást követjük. A szerzők tanulmányukban három nagy kategóriába sorolják ezeket a jelenségeket. Az első tényező a magánhangzók és mássalhangzók típusát jelenti. Ez lényegében megfeleltethető a szegmentális időzítésnek. Ez a kategorizáció annyiban tér el a hagyományos megközelítésektől, hogy a beszédhangok közvetlen környezetét már máshova sorolják. Külön csoportot alkotnak együttesen a kontextuális faktorok, amelybe beletartozik a környező hangok típusa mellett a prominencijelenség (szóhangsúly, frázishangsúly), a szintaxis és a jósolhatóság hatása az időtartamokra. Hangsúlyozzák, hogy ezen tényezők közül a szintaxisnak és a jósolhatóságnak nincs közvetlenül kimutatható hatása a hangzók időtartamára. A szintaxis esetében azzal érvelnek, hogy bár néhány szintaktikai módosítás hatással van a hangok időszerkezetére, de nem mindegyik. Továbbá ugyanazon szerkezetű

mondatok eltérő időzítéssel rendelkezhetnek: máshol tartanak a beszélők szünetet, máshol jelenik meg a frázis végi nyúlás stb. Sőt, olyan helyeken jelenhetnek meg szünetek, nyúlások, ahol a szintaxis nem jósná előre. A szintaxis helyett ezért a beszédprodukciónban közvetlenül megfigyelhető egységek, az ún. prozódiai szerkezet egységeinek (szótagok, szavak, frázisok, intonációs frázisok, megnyilatkozások stb.) vizsgálatát helyezik előtérbe. Hozzá kell tennünk, hogy a prozódiai és szintaktikai szerkezet szorosan összefügg, de a köztük lévő kapcsolat természete még nem egészen tisztázott (például SAVY–VOGHERA 2010). A következőkben ezen okok miatt a prozódiai egységek és az időzítés kapcsolatának leírására törekszem.

A szintaxishoz hasonlóan problematikusnak találják azt, hogy a jóslhatóság közvetlenül hatna a beszédhangok időszerkezetére. A jóslhatóság egy szó adott szövegkörnyezetben való előfordulásának valószínűségét jelenti. Az előre könnyebben jóslható szavak vagy szótagok rövidebb időtartamban valósulnak meg, mint az előre nehezen jóslhatók (például GREGORY et al. 1999). Felmerül ugyanakkor a kérdés, hogy vajon az időszerkezetre külön-külön ható tényezőnek tekinthető-e a jóslhatóság és az akusztikai feltűnőség (frázis- és szóhangsúly) jelensége. Ha ugyanis valami feltűnőbb, hangsúlyosabb egy mondatban, azt hajlamosabbak vagyunk hosszabban ejteni (*1.2. fejezet*), így a kevésbé fontos vagy hangsúlyos részek időtartama rövidebb. Egyfajta kiegészítő viszony feltételezhető a két jelenség között. TURK és SHATTUCK-HUFNAGEL (2014) a harmadik kategóriába olyan globálisnak tekintett, azaz az egész beszédre jellemző, általános faktorokat sorolnak, mint az artikulációs tempó és a beszédstílus. Ebben a csoportosításban nem szerepel külön a beszédritmus mint időzítésbeli jelenség, aminek lehetséges okai a megfelelő fejezetben (*1.3. fejezet*) kerülnek bemutatásra. Szintén fontos megemlítenünk, hogy a beszélő aktuális érzelmi állapota is hatással lehet a beszéd időbeli tulajdonságaira. Ugyanakkor ezen kutatások elsősorban az artikulációs tempó és a beszédtempó vizsgálatával kapcsolatosak, ezért ezt a *1.4. fejezetben* tárgyaljuk részletesebben.

A következő fejezetekben (*1.2–1.4. fejezetek*) a szupraszegmentális időzítéshez köthető jelenségek és lehetséges megközelítések összefoglalására kerül sor. A következő részben (*1.2. fejezet*) végigvesszük azokat a tényezőket, amelyek közvetlen hatással lehetnek több, egymást követő hang időtartamára, majd a további fejezetekben olyan globálisnak tekinthető időzítésbeli jelenségekkel foglalkozunk, mint az artikulációs tempó és a beszédritmus.

1.2. A szegmentumok időtartamára ható, eddig feltárt tényezők

Ahogy már korábban említettük, az időtartamokkal foglalkozó akusztikai kutatások több olyan tényezőt is feltártak, amelyek szisztematikusan megnyújtanak vagy megrövidítenek több egymást követő hangzót. Ezen faktorok közé tartozik a közlésbeli pozíció, a hangsúlyos

helyzet és a szóhosszúság is (VAN SANTEN 1992; WHITE 2002; FLETCHER 2010; TURK–SHATTUCK-HUFNAGEL 2014). A következőkben ezeknek a tényezőknek az időtartamokra gyakorolt hatását fogjuk részletesebben tárgyalni. Továbbá olyan tényezőkről (szintaxis, szavak jósolhatósága) is szó esik, amelyeknek az egymás melletti beszédhangok időtartamára gyakorolt hatása nem teljesen tisztázott.

1.2.1. A megnyilatkozás vagy intonációs frázis végi nyúlás

A közlésben megjelenő különböző pozíciók és a hangzók időtartama között több összefüggést is sikerült kimutatni. Az intonációs frázis és/vagy megnyilatkozás végén jelentkező nyúlást sokan univerzális jelenségnek tartják, bár a mértéke és a kiterjedése (az, hogy hány beszédhangot érint) nyelvenként változó. FLETCHER (2010) összefoglalásában egyebek között felsorolja az orosz, az angol több változatát, a franciát, az olaszt, a görögöt, a csehet, a németet, a hollandot, a svédet, a finnt, a japánt, az arabot, a hébert és a mandarint mint olyan nyelveket, amelyekben kimutatták ezt a jelenséget (a primer szakirodalmi forrásokat lásd a hivatkozott helyen).

A megnyilatkozás végi nyúlás időtartamokra gyakorolt hatását elsősorban az utolsó szótagon, illetve az utolsó szótag hangjain figyelték meg (például angol: KLATT 1975; francia: TABAIN 2003; holland: CAMBIER-LANGEVELD 1997). De a jelenség nemcsak az utolsó szótagon tapasztalható, hanem az utolsó előtti szótagon is mérhető (például NAKAI et al. 2009), sőt, vannak olyan kutatások, amelyekben a megnyilatkozás végétől számítva a harmadik szótag hangjain is mérték az időtartamok nyúlását (például WHITE 2002; TURK–SHATTUCK-HUFNAGEL 2007).

A megnyilatkozás végi nyújtás kezdetét összefüggésbe hozták a hangsúly elhelyezkedésével. Az egyik hipotézis szerint az utolsó főhangsúlyos szótagon kezdődik, és onnantól kezdődve folyamatos. Több mérési eredmény összhangban van ezzel az állítással (például angol: WHITE 2002; TURK–SHATTUCK-HUFNAGEL 2007; holland: CAMBIER-LANGEVELD 1997). Ugyanakkor német nyelven mért eredmények alapján felmerült, hogy a nyúlás már az utolsó hangsúlyos szótag előtt elkezdődhet (SILVERMAN–PIERREHUMBERT 1990). A rendelkezésünkre álló kevés mért adat alapján nem határozható meg egyértelműen, hogy milyen összefüggés van a hangsúly és a közlés vagy frázis végi nyújtás kezdete között, és hogy ez mennyire tekinthető univerzálisnak.

Nehéz tehát megjósolni, hogy pontosan melyik szótagon és minek a hatására jelentkezik először a lassítás a határhoz közeledve. Azt viszont több vizsgálat is alátámasztotta, hogy a nyúlás mértéke az utolsó szótagon a legnagyobb (angol: KLATT 1975; német: KOHLER 1983; WIGHTMAN et al. 1992). A nyúlás mértéke trendszerűen növekszik a határhoz közeledve (CAMBIER-LANGEVELD 1997; WHITE 2002), ugyanakkor ez a folyamat nem teljesen következetes, ezt igazolják a legújabb mérési eredmények (TURK–SHATTUCK-HUFNAGEL 2007). A hangzó időtartamának változása ugyanis függ a beszédhang minőségétől, képzésmódjától is.

A magánhangzók és a mássalhangzók egyaránt nyúlhatnak a megnyilatkozás végén (például OLLER 1973; HOFHUIS–GUSSENHOVEN–RIETVELD 1995; NAKAI et al. 2009; TURK–SHATTUCK–HUFNAGEL 2007). HOFHUIS, GUSSENHOVEN és RIETVELD (1995) kísérletében a mássalhangzóra végződő szavakban a mássalhangzók többségének (nazális, tremuláns, zárhang) időtartama kevésbé növekedett meg megnyilatkozás végi helyzetben, mint az azt megelőző magánhangzóké. A spiránsok viszont hosszabbak voltak abszolút megnyilatkozás végén, mint a spiránsokat megelőző magánhangzók. Ezt az összefüggést korábban már más vizsgálat is feltárta (OLLER 1973).

Az eddigiekben a megnyilatkozások végére koncentráltunk, a határhoz közeli nyúlás azonban nemcsak ezen a ponton figyelhető meg. Több kutatás igazolta, hogy a különböző prozódiai egységek végéhez közeledve a beszédhangok, szótagok időtartama hosszabb. Nincs egyetértés abban, hogy milyen prozódiai szintek különülnek el az egyes nyelvekben, de mindenhol feltételeznek a megnyilatkozásnál kisebb egységeket (például angol: WIGHTMAN et al. 1992; BYRD–KRIVOKAPIC–LEE 2006; francia: TABAIN 2003; svéd: HORNE–STRANGERT–HELDNER 1995). A kutatási eredmények nem egységesek a tekintetben, hogy a megnyilatkozás végén a nyúlás jelentősebb lenne, mint kisebb egységek határain, például az intonációs frázis végén (lásd FLETCHER 2010). Ezért ezeknek a kérdéseknek a megválaszolásához komplexebb, a prozódiai struktúra és a diskurzusszerkezet közötti kapcsolatot többértően figyelembe vevő elemzésekre volna szükség.

A megnyilatkozás végi nyúlást már több kutatás vizsgálta magyar nyelvű szövegeken is. Mondatfelolvasásokban a magánhangzók hosszabbnak bizonyultak abszolút hangsor végi helyzetben, mint közlés közben (MAGDICS 1966; KASSAI 1979; GÓSY–KREPSZ 2016). A felolvasásokban nemcsak abszolút hangsor végén, hanem CVC felépítésű utolsó szótagban is nyúlnak a magánhangzók, akár rövid, akár hosszú magánhangzóról van szó. Ezzel kapcsolatban WHITE és MÁDY (2008) végzett szisztematikus kutatást. Minimálpármondatokat hoztak létre egy, két és három szótagos, *o-t* és *ó-t* kváziazonos fonetikai kontextusban tartalmazó szavakkal, figyelembe véve a fókuszhangsúly mondatbeli helyzetét is. Eredményeik szerint a megnyilatkozás végi nyúlás a magyarban egyértelműen létezik, annak ellenére, hogy a kvantitás megkülönböztető szereppel bír a magánhangzók körében. Ez a fonológiai különbség a nyúlás megvalósulásában is nyomot hagy: a rövid magánhangzók nyúlása csak a megnyilatkozás utolsó szótagjában volt megfigyelhető, míg a hosszú magánhangzók esetében az utolsó előtti szótagban is időtartamtöbbletet mértek. Ezzel szemben KOVÁCS (2002) vizsgálatai alapján a rövid magánhangzók is kimutatható az időtartambeli növekedés az utolsó előtti szótagban. (Az idézett két kutatás módszere eltérő volt.) A magánhangzók mellett a mássalhangzókat is hosszabbnak találták a felolvasások hangsorvégi pozíciójában (MAGDICS 1966; KASSAI 1979). A hangsor végi lassítást tényként kezelik a gépi felolvasórendszerek modelljeiben is (vö. OLASZY 2006; TÓTH 2013). A magánhangzók és a mássalhangzók időtartama nem ugyanúgy változik a megnyilatkozás végi pozícióban. Zárt szótagok esetében azt találták, hogy a mássalhangzók általában jobban nyúlnak a magánhangzókénál (HOCKEY–FAGYAL 1999), bár ez függ a mássalhangzók minőségétől is (KOVÁCS 2009).

*A megnyilatkozás vagy frázis végi nyúlás megjelenésének
lehetséges magyarázatai*

A jelenség magyarázatára több elképzelés is található a szakirodalomban. LINDBLOM és RAPP (1975) úgy véli, hogy a beszéd soron következő elemei egy hipotetikus „frázispuffer”-ben tárolódnak, és a frázis végi nyúlás lényegében egy általános lassulási tendencia eredménye, amelyet az idéz elő, hogy a pufferből egyre fogynak a beszéd elemei, majd a puffer kiürül. ÖHMAN (1967) ugyanakkor az artikulációs gesztusok renyhülésével magyarázza a jelenséget, és ezt sugallják TABAIN (2003) artikulációs fonetikai kutatásának eredményei is. Más megközelítések (például KLATT 1975; TURK–SHATTUCK–HUFNAGEL 2000) szerint a frázis végi lassítás valójában egy hallgatóorientált stratégia az elemek különféle szintű összetartozásának jelzésére.

1.2.2. A hangsor eleji pozíció

A hangsor végi pozíció mellett a hangsor kezdete is kitüntetett lehet. Több, angol nyelven végzett kutatás mérési eredményei is azt támasztják alá, hogy a megnyilatkozások kezdetén (szünet után) a nazális mássalhangzók rövidebbek hangsúlytalan szótagban, mint a megnyilatkozás közepén (FOURAKIS–MONAHAN 1988; FOUGERON–KEATING 1997). A rendelkezésünkre álló kevés adatból azonban nehéz lenne általános szabályszerűséget megfogalmazni. Felmerül a kérdés például, hogy ez a szabályszerűség csak a nazálisok esetében érvényesül, vagy kiterjeszhető más mássalhangzókra is, illetve hangsúlyos szótagokban is megtalálható-e a jelenség. WHITE (2002) kísérleteiben hangsúlyos szótagokat vizsgált. Az eredmények azt mutatták, hogy bizonyos szavak (*main*, *mend* és *send*) esetében egyértelműen mérhető volt, hogy a megnyilatkozás kezdetén a szó első hangja rövidebb, mint megnyilatkozás közben. Ez az összefüggés hangsúlyos és hangsúlytalan helyzetben is mérhető volt. Más szavak (*port*, *compose*, *dispose*, *suppose* és *juice*) esetében viszont nem volt ilyen egyértelmű a helyzet. Ezen szavak kezdő hangjainak időtartamában ugyanis nem találtak mérhető változást hangsúlyos helyzetben, hangsúlytalan helyzetben viszont a közlés kezdetén az időtartamokban kicsi, de egyértelmű nyúlást figyeltek meg. A mássalhangzót követő magánhangzó időtartamára semmilyen esetben nem volt hatással az adott szó megnyilatkozásán belüli pozíciója. Az első szótag záró mássalhangzója egyes esetekben (*main*, *mend* és *send*) hosszabb volt, ha a szó a megnyilatkozás elején szerepelt, szemben a megnyilatkozás közepén lévő előfordulásokkal. Más esetekben (*port*, *compose*, *dispose*, *suppose* és *juice*) pedig ezen mássalhangzónál sem volt kimutatható interakció a megnyilatkozásbeli pozíció és a hangzó időtartama között. Nem következetes tehát, hogy a megnyilatkozás kezdetén a különböző beszédhangok időtartama hogyan alakul.

Az előző alfejezetben tárgyalt megnyilatkozás vagy intonációs frázis végi nyúláshoz hasonlóan nemcsak a megnyilatkozások határai rendelkeznek kitüntetett szereppel, hanem más

szintek kezdete is hatással van a beszédhangok időtartamára. Az intonációs frázis elején lévő mássalhangzók több nyelven hosszabbnak bizonyultak, mint az intonációs frázis belsejében, de szintén szó elején szereplő mássalhangzók (francia: FOUGERON 2001; koreai: CHO–KEATING 2001).

A magyarra végzett kutatások eredményei ellentmondásosak. MAGDICS (1966) szerint mind a magánhangzók, mind a mássalhangzók rövidebbek felolvasott mondatok kezdetén, mint hangsor belsejében. KASSAI (1979) szerint viszont a hangsor eleji hangzók hosszabbak a közlés belseji hangzóknál. GÓSY és KREPSZ (2016) tanulmányában szintén a hangsor eleji pozíció nyújtó hatását mutatta ki a hangsor belseji pozícióhoz képest magánhangzók esetében. Az eltérések fakadhatnak az eltérő anyagból és szemléletből. KASSAI (1979) a hangsor eleji pozíciót úgy határozta meg, hogy a felolvasások mondatkezdő vagy szünet utáni hangzója, míg más kutatásokban csak mondatkezdő pozíciókat vizsgáltak (MAGDICS 1966). GÓSY és KREPSZ (2016) viszont célszavak magánhangzóinak időtartamát vizsgálták oly módon, hogy a célszavak a hétszavas mondatoknak második, más esetben pedig negyedik szavaként valósultak meg. A hangsor eleji pozíció időtartamra gyakorolt hatása tehát nem teljesen egyértelmű a magyar nyelvben, de a hangsor végi lassítás meglétét a gépi felolvasórendszerek modelljei mellett a felolvasásokon végzett akusztikai mérések is megerősítik.

1.2.3. A hangsúly és az időtartam kapcsolata

A hangsúlyt egy szótagnak a környezetéhez képesti kitűnéseként szokás értelmezni (GÓSY 2004; FLETCHER 2010; MARKÓ 2012). A különböző elképzelések szerint a hangsúly a beszélőnek valamilyen nagyobb produkciós erőfeszítéséből származik, amely együtt járhat nagyobb intenzitással, az alaphangmagasság megemelkedésével, a hangsúlyos szótag előtt tartott szünettel, sőt az időtartamok megnövekedésével is (FÓNAGY 1958; FOX 2000; HONBOLYGÓ 2009). Ezen akusztikai jegyek azonban nem feltétlenül jelennek meg olyan helyeken, amelyeket a hallgatók kiugróbbnak, feltűnőbbnek tartanak a többinél (HONBOLYGÓ 2009). Máiig nincs teljes konszenzus abban, hogy a hangsúlynak mely nyelvekben milyen a fonetikai megvalósulása. Az sem teljesen tisztázott, hogy a hangsúly hogyan és milyen módszerekkel írható le a legmegfelelőbbben, mi a funkciója, milyen fokozatai vannak (FOX 2000). A következőkben azokat a fonetikai kutatásokat próbálom meg bemutatni, amelyek esetében sikerült igazolni, hogy a hangsúly – figyelembe véve a különböző elméleti kereteket, fokozatokat – hatással van a beszédhangok, szótagok, szavak időtartamára.

Az akusztikai vizsgálatokban a hangsúlynak általában két különböző szintjét szokták vizsgálni. Az egyik a szóhangsúly (lexical stress), a másik a hierarchia magasabb szintjén található frázishangsúly (phrasal stress) vagy másképpen mondathangsúly (sentence stress). A *Mary's cousin George* angol frázisban például a *George* a legkiemelkedőbb szó, amelyet frázishangsúlynak (phrasal stress, nuclear accent) vagy mondathangsúlynak (sentence stress) is szokás nevezni. A kifejezésben ugyanakkor több olyan szótag is található, amelyek

feltűnőbbek a többi szótagnál, de nem a frázis legfeltűnőbb elemei. Ezeket a részeket tekintik az ún. lexikai vagy szóhangsúlynak (lexical stress, word-stress). Az adott példában ilyen szótag a szókezdő *cou-* vagy *Ma(r)-* szótagok, amelyek kiemelkedőbbek ugyanezen szavak második szótagjához képest (HAYES 1983; TURK–SHATTUCK–HUFNAGEL 2014).

A lexikai hangsúly esetében több nyelvben (például angol: CRYSTAL–HOUSE 1988; holland: RIETVELD–KERKHOFF–GUSSENHOVEN 2004; svéd: HELDNER–STRANGERT 2001) kimutatták, hogy a hangsúlyos szótagok, illetve azok beszédhangjai (amelyek ugyanakkor nem frázishangsúlyok) hosszabbak, mint a hangsúlytalan szótagok. A frázishangsúlyos szótagok még hosszabbnak bizonyultak a lexikai hangsúlyos szótagoknál is (például angol: TURK–SAWUSCH 1997, TURK–WHITE 1999; holland: CAMBIER–LANGEVELD–TURK 1999; francia és német: BARRY–ANDREEVA 2001; finn: SUOMI 2007). A frázishangsúly esetében az időtartamok változása nemcsak egy szótagot érint, hanem kiterjedhet a következő szótagokra is (angol: TURK–WHITE 1999; holland: CAMBIER–LANGEVELD–TURK 1999; finn: SUOMI 2007). A szótagokon belül a magánhangzók és a mássalhangzók időtartama is megnövekedhet hangsúlyos helyzetben (CRYSTAL–HOUSE 1988; VAN SANTEN 1992; HELDNER–STRANGERT 2001; WHITE 2002). Igaz, a magánhangzók és a mássalhangzók időtartama nem feltétlen egyenlő mértékben nyúlik meg függően a nyelvtől, a szó felépítésétől, a beszédhangok minőségétől stb. (a részletes összefoglalót lásd FLETCHER 2010).

Habár meglehetősen sok nyelven sikerült kimutatni, hogy a hangsúlyos helyzetnek hatása van az egymást követő beszédhangok időtartamára, nyelvenként ez a mérték eltérő lehet, sőt, nem minden nyelvben jelenik meg a lexikai vagy éppen a frázishangsúly időtartamokra gyakorolt hatása (lásd FLETCHER 2010). A magyar nyelven végzett kutatások eredményei sem egyértelműen támasztják alá a hangsúlyos helyzet időtartamokra gyakorolt hatását. MAGDICS (1966) azt találta, hogy a magyar magánhangzók hangsúlyos helyzetben hosszabbak, mint hangsúlytalan helyzetben. Ezt az összefüggést tükrözi a gépi beszédelőállítás magyar nyelvre alkalmazott modelljének időtartam-szabálya is (OLASZY 2006). Más kutatási eredmények viszont nem támasztják alá ilyen egyértelműen a hangsúlyos helyzet időtartamra gyakorolt hatását. KASSAI (1982) megállapítja, hogy hangsúlytalan helyzetben a rövid és hosszú magánhangzók időtartamai jobban kiegyenlítődnek, mint hangsúlyos helyzetben. A kiegyenlítődés okaként pedig azt tünteti fel, hogy a rövid magánhangzók nyúlnak, a hosszúak pedig változatlanok maradnak hangsúlyos helyzetben. Arról viszont nincs pontos információnk, hogy milyen típusú hangsúlyos helyzetet vagy helyzeteket vizsgáltak ezekben a kutatásokban. KOVÁCS (2002, 2009) kutatásaiban három rövid magánhangzó [i, e, œ] szóhangsúlyos helyzetét vizsgálta nem mondathangsúlyos mondatokon (fókuszos mondatok kerülésével). A három különböző minőségű magánhangzón a hangsúlyos és hangsúlytalan helyzet nem mutatott szignifikáns eltérést. Vizsgálatai alapján arra a következtetésre jutott, hogy „a hangsúlyos/hangsúlytalan magánhangzókat nem állítja szembe fizikai időtartamuk”. (KOVÁCS 2009: 69). WHITE és MÁDY (2008) fókuszos mondatokban vizsgálta a frázishangsúly időtartamokra gyakorolt hatását, de az [o] és [o:] magánhangzókon végzett eredményeik azt mutatták, hogy sem a rövid, sem pedig a hosszú magánhangzók időtartama nem mutat szisztematikus eltérést a hangsúly hatására.

A magyar nyelven végzett percepciók kísérletek sem mutatnak egyértelmű összefüggést az észlelhető hangsúlyos helyzet és az időtartamok között (például MARKÓ 2012). Más percepciók vizsgálatok alapján az időtartam-eltérések kevésbé hatékonyan jelzik a hangsúlyt a hallgató számára, mint az alaphangmagasság- és az intenzitásváltozás (HONBOLYGÓ–KOLOZSVÁRI 2015). Nem meglepő tehát, ha a hangsúlyos helyzet nem feltétlenül jár együtt időtartambeli változással (FÓNAGY 1958).

1.2.4. A szóhosszúság hatása

A közlésbeli pozíciókon kívül más tényezők is kiválthatnak lassuló vagy gyorsuló szakaszokat a beszédben. A szóhosszúságnak a hangzoidőtartamokra gyakorolt hatását több nyelven (például angol: LEHISTE 1972, TURK–SHATTUCK–HUFNAGEL 2000; holland: NOOTEBOOM 1972; svéd: LINDBLOM 1968) is vizsgálták, és igazolták, hogy a szóhosszúság hatására csökken a szókezdő, hangsúlyos szótag időtartama. A jelenség azonban nem univerzális, a finnben például nem érvényesül ilyen tendencia (SUOMI 2007).

A szótagszám időtartamokra gyakorolt hatása elképzelhető, hogy nem pusztán a szavak szintjén érvényesül, hanem nagyobb prozódiai egységek (frázis, megnyilatkozás) szótagszámának növekedésével is megjelenne (LEHISTE 1975; WHITE 2002). RAKERD, SENNET és FOWLER (1987) azt találták, hogy ugyanazon szó időtartama csökken, ha a beszélőknek ugyanúgy kezdődő, de valamivel hosszabb mondatot kell felolvasniuk (például: *His first date aroused some anxiety. / His first date aroused some anxiety for obvious reasons.*). Ezen eredménynek többféle magyarázata is lehet, például elképzelhető, hogy a mondat frázisokra tagolódása másképpen alakul, másképp valósulnak meg a különböző mondatokban a hangsúlyviszonyok, az intonációs frázis hangsúlyviszonyai stb., de az sem kizárható, hogy a beszélők igyekeznek a különböző hosszúságú frázisokat vagy megnyilatkozásokat azonos idő alatt ejteni (WHITE 2002).

A magyar nyelvre vonatkozó vizsgálatok kimutatták, hogy a hangsúlyos és a hangsúlytalan magánhangzók egyaránt rövidülnek a hosszabb szavakban. MEYER és GOMBOCZ (1909) vizsgálataikban a *tát, tátog, tátogat, tátogatók, tátogatóknak*, valamint a *köt, kötet, kötetek* izolált szavak felolvastatásán vizsgálta a jelenséget. A magánhangzókra megfigyelhető volt, hogy a hosszabb szavak esetében időtartamuk tendenciaszerűen egyre rövidebb. TARNÓCZY (1974) megismételte a *tát* és *toldalékolt* alakjain végzett méréseket, és hasonló eredményeket kapott, bár az általa mért időtartamok jóval rövidebbek, mint a korábbi kutatásban. KASSAI (1979, 1982) mondatokba ágyazott szavakon vizsgálta a szótagszám hatását a hangidőtartamokra, és mind a magánhangzók, mind a mássalhangzók esetében érvényesnek találta a szótagszám időtartam-rövidítő hatását. KREPSZ (2015) spontán beszédben is igazolta, hogy szótagszám növekedésével csökken a magánhangzók időtartama.

A jelenség meglétét a magyar nyelvben azonban nem minden kutatás támasztotta alá. WHITE és MÁDY (2008) kutatásukban egy rövid és egy hosszú magánhangzót ([o] és [oː])

beszédhangokat) vizsgáltak egy, két, illetve három szótagú szavak első, frázishangsúlyos szótagjában. Eredményeik alapján nem találtak összefüggést a szótagszám és a beszédhangok időtartama között. Egy másik kutatásban ugyanígy a hátul képzett, középső nyelvvállású rövid és hosszú magánhangzók időtartamát vizsgálták, de a hangsúlytalan helyzetű második szótagban. Továbbá nemcsak három, hanem négy, öt, hat, hét és nyolc szótagos célszavakat is felolvastattak mondatokba ágyazva. Az eredmények azt mutatták, hogy a rövid magánhangzó időtartama nem rövidül a szóhosszúság növekedésével. A hosszú magánhangzó időtartama ugyanakkor csökkent, ha a szó hossza nemcsak egy-egy, hanem több beszédhanggal megnövekedett (KOHÁRI 2012). GÓSY és KREPSZ (2016) eredményei azt mutatták, hogy a rövid és a hosszú magánhangzók időtartama is rövidül a szóhosszúság növekedésével, de csak négy szótagig. A négy és öt szótagú szavak magánhangzóinak időtartamában nem találtak statisztikailag is igazolható különbséget.

Ellentmondásosak tehát az eredmények a tekintetben, hogy a magyar nyelvben hogyan, milyen hangokon, hány szótagig, milyen környezetben van hatása a szótagszámnak a beszédhangok időtartamára. Az eredmények egy részét, ahol egyértelműen igazolható volt a jelenség, izolált szavak felolvasásán mutatták ki (MEYER–GOMBOCZ 1909; TARNÓCZY 1974). Ezekben az esetekben nem szétválasztható, hogy ténylegesen a szavak eltérő szótagszámának hatása volt mérhető az időtartamokon, vagy egy nagyobb egység, a megnyilatkozás egészének növekedése indukálta az eltérő időtartam-mintázatot. Hasonlóan a nemzetközi szakirodalomban találtakhoz (RAKERD–SENNET–FOWLER 1987; WHITE 2002), ez a probléma felmerülhet a felolvasott mondatokon végzett vizsgálatok esetében is. FÓNAGY és MAGDICS (1960) végzett magyar nyelven méréseket azzal kapcsolatban, hogy a szónál magasabb szintű egység szótagszámának növekedése együtt jár-e az időtartamok csökkenésével. A szólamot választották alapegységnek, amelyet olyan fonetikai, ritmikai egységnek tekintettek, melyet „a nyomaték foglal egybe”, és amelyet nem választ szét szünet (FÓNAGY–MAGDICS 1960: 452). Eredményeik szerint a rövidebb szólamokat valóban lassan ejtik a beszélők, de ez az összefüggés csak 6-8 hangig (kb. 2-3 szótagig) volt kimutatható. Ezen kutatási eredmények viszont újabb kérdésekhez vezetnek, például vajon magasabb szintű egységek (frázisok, beszédszakaszok, megnyilatkozások) esetében is található-e ilyen összefüggés, és ez milyen viszonyban van a szószintű rövidülésekkel. Ezek megválaszolásához további célzott, szisztematikus kutatásokra volna szükség.

1.3. A beszédritmus időbeli dimenziója

Az előző fejezetben megismerkedhettünk az egymás melletti beszédhangok időtartamát befolyásoló ismertebb tényezőkkel. A kutatások elsősorban arra koncentráltak, hogy az egyes faktorok külön-külön hatással vannak-e az időtartamokra. A beszédritmus-kutatások egészen

más szemléletben közelítik meg az időzítés kérdéskörét. Ezek a különböző tényezők interakcióját, összejátszását egyidejűleg próbálják megragadni, a folyamatos beszédben megjelenő általános időzítési jellemzők, mintázatok feltárására törekszenek.

1.3.1. A beszédritmus fogalma

A beszédritmus nehezen megközelíthető jelenség, éppen ezért leírására sokféle elmélet és definíció is született. Hagyományosan a beszédritmus alatt a beszédben megjelenő különböző struktúrák vagy jellemzők rendszerű ismétlődését értik (lásd Fox 2000; GÖSY 2004; NOLAN–JEON 2014). A beszédben a beszédritmust létrehozó egyik lehetséges tényező az időbeli szerveződés. Több kutatásban szótagok, szavak és nagyobb egységek időtartamának, valamint artikulációs tempóváltozásának vizsgálatán keresztül próbálták megragadni a beszédritmus lényegét. A beszédritmus azonban nem ragadható meg csak egy paraméter alapján, feltehetően szorosan kötődik például a hangsúlyozáshoz is (GÖSY 2000a; ARVANITI 2009; OLASZY 2010b; CUMMING 2011). Vannak olyan elképzelések is a beszédritmusról, hogy kevésbé lényeges eleme lenne a periodicitás, inkább a nyelvi elemek egymáshoz képesti feltűnősége, maguk a hangsúlyviszonyok alkotnának egyfajta struktúrát, amely szintén érzékelhető a beszédhangok felszíni időszerkezetén és más akusztikai tényezőkön keresztül (TURK–SHATTUCK–HUFNAGEL 2013). A beszédritmus és az időzítés tehát két külön fogalom, de szorosan kapcsolódnak egymáshoz, sőt, egy újabb megközelítésben (lásd az 1.3.3. fejezetben) központi szerepet tulajdonítanak az időzítésnek a beszédritmus kialakításában (NOLAN–JEON 2014).

A beszédritmussal kapcsolatban további felmerülő kérdés, hogy milyen szinten, milyen nyelvi egységekben kereshetjük a ritmus nyomait. A 90-es évek előtt elsősorban a beszédhangoknál nagyobb egységeket, szótagokat, szavakat, beszédszakaszokat vizsgálták a beszédritmus jelei után kutatva, és feltételezték, hogy a beszédritmus a globális időzítésért felelős (NOLAN–JEON 2014). Később azonban a beszédhangok időtartamát is kapcsolatba hozták a beszédritmus fogalmával (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002).

A beszédritmus fogalma egyelőre nem tisztázott, de az elmúlt évtizedben a beszédritmus egyik klasszikus elméleti keretéhez sikeresen társítottak mérhető akusztikai paramétereket. Ezen elméletről, a beszédritmusosztályokról lesz szó a következő fejezetben.

1.3.2. Az időzítésen alapuló beszédritmusosztály-hipotézis

Az 1940-es években felállított modell szerint minden élő nyelvben a beszéd egyenlő időtartamú részekre tagolódik (izokronia). De a beszédritmus alapegysége eltérő a különböző nyelveknél. Először két nagy csoportot különítettek el egymástól: a hangsúly-időzítésű és a szótag-időzítésű nyelveket (PIKE 1945). A szótag-időzítésű nyelvek hangzását a puska ropogásához hasonlították, ezekben a nyelvekben ugyanis egyenletes ritmust véltek felfedezni. Ezzel szemben

a hangsúly-időzítésű nyelvek hangzásáról úgy tartották, hogy a Morse-kódra emlékeztet, mivel szabályos, de nem egyenletes lüktetést véltek felfedezni benne. Ebből a percepció (hallható, észlelhető) élményből kiindulva más beszédritmus-alapegységeket határoztak meg a különböző osztályokban. A szótag-időzítésű nyelvek esetében az észlelhető egyenletességet a szótagok időtartamához kapcsolták, és feltételezték, hogy a kiejtett szótagok időtartama közel azonos. Ezzel szemben a hangsúly-időzítésű nyelvekről úgy gondolták, hogy a szabályos, de nem egyenletes ritmusról más alapegység időtartambeli azonossága kelti. Feltételezték, hogy a hangsúlytól hangsúlyig terjedő szakaszok időtartama nagyjából megegyezik (PIKE 1945; ABERCROMBIE 1967). A hangsúly-időzítésű nyelvek közé sorolták például az angolt, a hollandot. Szótag-időzítésűnek gondolták a franciát, a spanyolt (lásd GRABE–LOW 2002). A későbbiekben e két kategória mellé egy harmadikat is felállítottak, az ún. moraidőzítésű nyelvekét (LADEFOGED 1975). A moraidőzítésű nyelvek esetében az egymást követő morák (azaz a szótag súlyát megadó fonológiai egységek) időtartamáról feltételezték, hogy közel azonos. Később több nyelvről megpróbálták megállapítani a percepció élmény, a szótagfelépítési szabályok és az adott nyelvben meglévő hangsúlyozási elvek alapján, hogy mely osztályba tartozhat. Az *I. táblázat*ban található egy összefoglalás a vizsgált nyelvek besorolásáról (GRABE–LOW 2002; MAIRANO–ROMANO 2007).

I. táblázat

*A nyelvek beszédritmus-kategóriák szerinti osztályozása**
(GRABE–LOW 2002 és MAIRANO–ROMANO 2007)

Beszédritmustípus	Nyelvek
hangsúly-időzítésű	arab, brit angol, cseh, holland, német, portugál, thai
szótag-időzítésű	francia, görög, olasz, spanyol, tamil
moraidőzítésű	japán
kevert	katalán, lengyel

*A legtöbb nyelv (észti, maláj, mandarin, román stb.) nem besorolt.

A beszédritmus osztályozásáról szóló feltételezéseket megpróbálták időtartammérésekkel is alátámasztani. Az akusztikai vizsgálatokkal azonban nem sikerült kimutatni, hogy a hangsúly-időzítésű nyelvekben az egymást követő hangsúlyos szótagok között valóban azonos időtartamnyi távolság lenne, vagy hogy a szótag-időzítésű nyelvekben a szótagok időtartama közel azonos lenne (ROACH 1982, 2002). A moraidőzítésű nyelvek esetében (japán) sem találtak az alapegységeket, ez esetben a morák időtartamát közel azonosnak (HOEQUIST 1983). Az elméletnek ellentmondó akusztikai eredmények a beszédritmus fogalmának, illetve a beszédritmus és a beszéd időzítésbeli kapcsolódásának újragondolását inspirálták.

DAUER (1983) más időtartamra ható tényezőkkel próbálta összefüggésbe hozni a három fő beszédritmustípus elkülönülését. Feltételezte, hogy az eltérő beszédritmusokat a különböző

nyelvekben a fonológiai, fonetikai, lexikai, szintaktikai tényezők eltérő kombinációja eredményezi. Az elképzelés szerint a szótagok felépítése az egyik olyan tényező, amely meghatározza a beszédritmust. A szótagtípusok nagyobb variabilitása a hangsúly-időzítésű nyelvekre jellemző, míg a kevesebb különböző szótagtípus a szótag-időzítésű nyelveknél figyelhető meg. A másik fontos tényező a magánhangzó-redukció. Feltételezte, hogy a magánhangzók időtartama hangsúlytalan helyzetben rövidebb, mint hangsúlyos helyzetben, és olykor formánszerkezetük svászerű magánhangzóra redukálódik. Ez a jelenség pedig a hangsúly-időzítésű nyelvekben gyakoribb, mint a szótag-időzítésű nyelvekben. DAUER (1987) szerint a különböző időtartamra ható tényezők hatásai összeadódnak, és így alakul ki egy nyelv beszédritmusa. Ebből adódik, hogy egy nyelv nem feltétlenül sorolható be a tisztán hangsúly-időzítésű vagy a szótag-időzítésű osztályba, hanem mindkettő jellegzetességeit mutathatja. Ezzel DAUER egy olyan elméletet alkotott meg, melyben a beszédritmus kontinuumként fogható föl, és a kontinuum két végpontja a tisztán szótag-időzítésű és hangsúly-időzítésű nyelvek osztálya (GRABE–LOW 2002). Több kutatásban átvették azt a gondolatot, mely szerint a beszédritmusosztályok között nincsenek éles határok, ezért azokban az esetekben, amikor a tanulmányok szemléletükben nem merev kategóriákként kezelik ezeket, akkor az eltérés jelzésére a későbbiekben az osztály helyett a *csoportosulás* terminust fogjuk használni.

Habár az akusztikai vizsgálatok nem támasztották alá az alapegységek időtartamáról szóló hipotézist, ez nem jelenti azt, hogy a felállított osztályok vagy csoportosulások nem is léteznek. A nyelvek beszédritmusuk szerinti csoportosulását támasztja alá többek között az a percepciókísérlet is, amelyben francia újszülötteket vizsgáltak. Úgy torzították a felvett hanganyagot, hogy csak az alacsony frekvenciákat tartották meg. Ez lényegében azt jelenti, hogy észlelés szempontjából előtérbe kerültek a beszéd szupraszegmentális eszközei (például beszédritmus, beszéddallam, tempó, hangerő stb.). A csecsemők reakcióit az alapján számszerűsítették, hogy milyen ütemben szívták a cumisüvegből az ételt. Az eredmények szerint az újszülöttek a felvételeket hallva különbséget tudtak tenni angol (hangsúly-időzítésű nyelv) és japán (moraidőzítésű nyelv) mondatok között, de a holland és az angol mondatok között, azaz két hangsúly-időzítésű nyelv között nem (lásd RAMUS–NESPOR–MEHLER 1999). A következő alfejezetben arról lesz szó, hogy az ismert beszédritmusosztályokhoz hogyan társítottak sikeresen mérhető akusztikai paramétereket.

1.3.3. A beszédritmus-mérőszámok meghatározása

A beszédritmus-mérőszámok alkalmazása nyelvek összehasonlítására

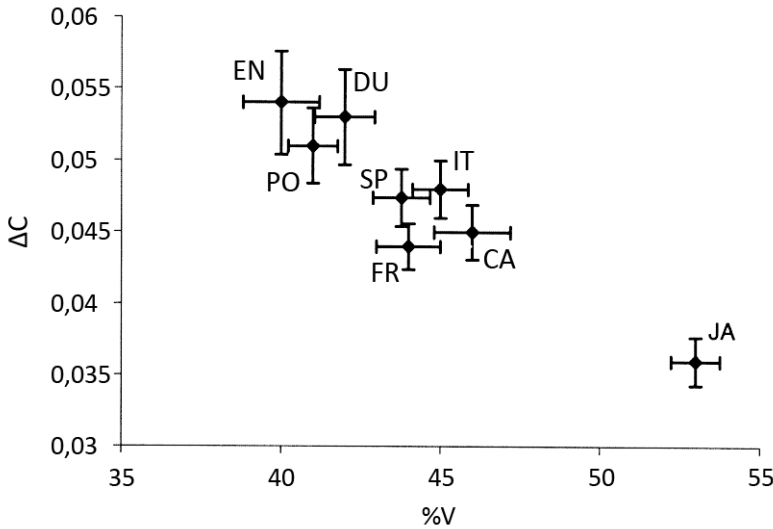
A beszédritmusosztályok (hangsúly-időzítésű, szótag-időzítésű, moraidőzítésű) létezéséről szóló elméletet nem sikerült akusztikai mérésekkel is igazolni, de a ritmikailag másképpen viselkedő nyelvek meglétét továbbra is feltételezték a percepciókísérletek eredményei miatt. A vizsgálatokban azonban a szótagok, hangsúlytól hangsúlyig terjedő szakaszok, illetve

a morák vizsgálata helyett egységes beszédritmus-alapegységet hoztak létre. Minden nyelven az ún. magánhangzós szakaszok és mássalhangzós szakaszok időtartamát vizsgálták. Mássalhangzós szakasz alatt az egymást közvetlenül követő mássalhangzók együttesét értjük, amelyeket nem választ szét magánhangzó. Így a hangsorok felépítésétől függően egy mássalhangzós szakasz állhat egy, kettő vagy akár több mássalhangzóból is. A mássalhangzós szakasz időtartamát pedig a benne lévő hangzók időtartamának összege adja. Az *ablaka* szóban például a [b] és a [l] együttesen alkot egy mássalhangzós szakaszt, mivel a két mássalhangzó közvetlenül egymás mellett van. E két hang időtartamának összege az adott mássalhangzós szakasz időtartama is egyben. A példában a [k] hangzó pedig önmagában alkot egy mássalhangzós szakaszt a definíció értelmében. A magánhangzós szakasz hasonló módon definiálható, mint a mássalhangzós szakasz. Egy magánhangzó vagy az egymást közvetlenül követő két vagy több magánhangzó együttese, amelyet nem választ szét mássalhangzó, alkot egy magánhangzós szakaszt. A magánhangzós és mássalhangzós szakaszok átívelnek a szavakon, sőt a szüneteken is oly módon, hogy a szünetek időtartamát nem veszik figyelembe. Tehát a beszéd teljes folyamatát úgy határozták meg, mint a magánhangzós és mássalhangzós szakaszok váltakozása, azaz a definíció alapján a magánhangzós szakaszt csak mássalhangzós szakasz követheti, a mássalhangzós szakaszt pedig csak magánhangzós szakasz, egészen addig, amíg véget nem ér a beszéd vagy a kiválasztott beszédrészlet (RAMUS–NESPOR–MEHLER 1999). Ez alapján ezt a példamondatot így tagolhatnánk: |A | n|a|gym|a|m|a| sp|e|c|ia||i|t|á|s|a| k|é|ts|é|g|e|t k|i|z|á|r|ó|a|n| a| b|a|r|a|ckb|e|f|ő|tt|. A magánhangzós szakaszok és mássalhangzós szakaszok időtartamának segítségével különböző mérőszámokat hoztak létre az egyes nyelvek sajátosságainak a mérésére (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002). A nyelveket – a klasszikus beszédritmusosztály szerinti besorolástól függetlenül – ugyanazzal a módszerrel tanulmányozták, és az így kapott eredményeket vetették össze.

RAMUS, NESPOR és MEHLER (1999) fejlesztették ki az egyik módszert. Megmérték egy adott nyelvre az összes mássalhangzós szakasz időtartamát, és meghatározták a szórását (a mérőszám jele: ΔC). Ez a mérőszám lényegében azt mutatja meg, hogy a mássalhangzós szakaszok időtartama mennyire változatos egy adott nyelven. Minél szélsőségesebb időtartamértékek jellemeznek egy nyelvet, ez a mérőszám annál nagyobb értéket mutat. Amennyiben az időtartamértékek nagyon hasonlók egymáshoz, ez a mérőszám alacsony értéket vesz fel. Az eredmények szerint a mássalhangzós szakaszok időtartamának szórása (ΔC) szoros összefüggésben van a korábbi elmélet szerint meghatározott beszédritmusosztályokkal. Magasabb értékeket mutat a klasszikusan hangsúly-időzítésűekhez sorolt nyelvekben (angol, holland), mint a szótag-időzítésűekhez sorolt nyelvekben (francia, olasz, spanyol).

A másik mérőszám, amely összefüggést mutatott a korábban meghatározott beszédritmusosztályokkal, a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának arányára épül. Összeadták az összes magánhangzós szakasz időtartamát, és kiszámolták, hogy a teljes időtartamnak hány százaléka a magánhangzós szakaszok időtartama (a mérőszám jele: %V). Ez a mérőszám azt mutatja meg, hogy mennyire kiegyensúlyozott a magánhangzós és a mássalhangzós szakaszok időtartamának aránya, tehát ha ez a mérőszám 50%-nál kisebb

értéket vesz fel, akkor a magánhangzók időtartamának összege kisebb a mássalhangzók időtartamának összegénél. Azt találták, hogy ez a mérőszám (%V) magasabb értékeket vesz fel a hagyományosan szótag-időzítésűekhez sorolt nyelvekben, mint a hangsúly-időzítésűekhez soroltakban. A két mérőszám ábrázolható egy koordináta-rendszerben (%V az x tengely, ΔC az y tengely az 1. ábrán), amelynek segítségével csoportosíthatók a nyelvek beszédritmusuk szerint. Jól látható, hogy a tipikusan hangsúly-időzítésűnek tartott nyelvek (például angol, holland) elkülönülnek a tipikusan szótag-időzítésű nyelvektől (például francia, olasz, spanyol), és a japán mint moraidőzítésű nyelv mindkét csoportosulástól távol esik (RAMUS–NESPOR–MEHLER 1999).



1. ábra

A nyelvek beszédritmus-mérőszámok alapján történő csoportosítása (RAMUS–NESPOR–MEHLER 1999)

(EN = angol, DU = holland, IT = olasz, FR = francia, SP = spanyol, CA = katalán, JA = japán, PO = lengyel)

RAMUS, NESPOR és MEHLER (1999) arra a következtetésre jutottak, hogy e két mérőszám összefüggésben lehet az adott nyelvben található szótagtípusokkal. Hiszen a különböző szótagtípusok nagyszámú előfordulása egy nyelvben vagy egy közlésben azt jelenti, hogy a mássalhangzók számának variabilitása is nagy. Tehát a mássalhangzós szakaszok állhatnak egy, kettő, három vagy akár több mássalhangzóból, ami a mássalhangzós szakaszok időtartamának szórását (ΔC) is növeli. Ugyanakkor ha több mássalhangzó-torlódás fordul elő egy közlésben, akkor a magánhangzós szakaszok időtartamának aránya a teljes hangzó beszédéhez képest (%V) kisebb lesz. A szótagtípusok számáról úgy tartják, hogy magas a hangsúly-időzítésű nyelvekben (például angol, holland) (DAUER 1983), ennek megfelelően a ΔC értéke ezen nyelveknél volt a legmagasabb a vizsgálatban. A %V mérőszám pedig az imént említett negatív

korreláció miatt ezekben az esetekben a legalacsonyabb. A ΔC és %V mérőszámok közt tehát van kapcsolat, de nem beszélhetünk lineáris összefüggésről. Bár a szerzők külön nem emelik ki, de a %V-t más tényezők is befolyásolják, például a magánhangzós szakaszokban szereplő magánhangzók mennyisége vagy közvetetten a minősége, hangkörnyezete is. A későbbiekben megvizsgálták, hogy a fonológiai felépítettségtől függetlenül is kimutatható-e eltérés a különböző nyelvek között ezen mérőszámok alapján. PRIETO és munkatársai (2010, 2012) három különböző anyagon tesztelték hipotézisüket. Egy részről olyan mondatokat olvastattak fel, amelyeknek felépítése – a lehetőségekhez mérten – csak CV szótagokból állt, illetve létrehoztak olyan mondatokat is, amelyek csak zárt szótagokból (CVC, CVCC) álltak. A harmadik csoportban csak olyan mondatok szerepeltek, amelyek az adott nyelvre jellemző fonológiai felépítést tükrözték, tehát zárt és nyílt szótagok keverve voltak. Ehhez RAMUS, NESPOR és MEHLER (1999) mondatait használták fel. Habár a fonológiai felépítettség egyértelműen hatással volt a %V értékeire a különböző nyelveken, az eredmények alapján különbség mutatkozott a katalán, az angol és a spanyol nyelvek között mindhárom mondat típus esetében. A mássalhangzós szakaszok időtartamának szórása (ΔC) csak a zárt és nyílt szótagokat is tartalmazó megnyilatkozások esetében mutatott jelentős eltéréseket a nyelvek között (PRIETO et al. 2010, 2012).

A harmadik meghatározó mérőszám a magánhangzós szakaszok időtartamának a szórása. Az eredmények azt mutatják, hogy a ΔV nagyobb a tipikusan hangsúly-időzítésű nyelvekben (például angol, holland), mint a jellemzően szótag-időzítésű nyelvek esetében (például francia, olasz, spanyol). A ΔV egy olyan mérőszám, amely érzékeny a magánhangzók időtartamára ható tényezőkre: a hangsúlyos helyzet nyújtó hatása, a magánhangzó-redukció, a környezet hatása, a közlésben elfoglalt hely, rövid-hosszú oppozíció fenntartásának igénye stb. Tehát a hangsúly-időzítésű nyelvek magas ΔV -értékei azt jelentik, hogy ezen tényezők (feltehetően elsősorban a hangsúlyviszonyok) szélsőséges hatással vannak a magánhangzós szakaszok időtartamára. Az imént felsorolt jelenségek azonban kevésbé fellelhetők vagy meghatározók a szótag-időzítésű nyelvekben (BARRY és ANDREEVA 2001). Egy több nyelvet összehasonlító beszédritmus-vizsgálatban például azt találták, hogy a lexikai hangsúly időtartamnyújtó hatása az angolban a legerőteljesebb, a németben valamivel kevésbé jelentős ez a hatás, a klasszikusan szótag-időzítésűnek tekintett olasz, spanyol és francia nyelvekben pedig még kevésbé volt jellemző (MAIRANO–ROMANO 2011). Egy másik kutatás szerint a különböző szintű hangsúlyok (lexikai, frázishangsúly), sőt a frázis végi nyújtás időtartamra gyakorolt hatása az angol nyelvben nagyobb volt, mint a katalán és spanyol nyelvekben (PRIETO et al. 2012).

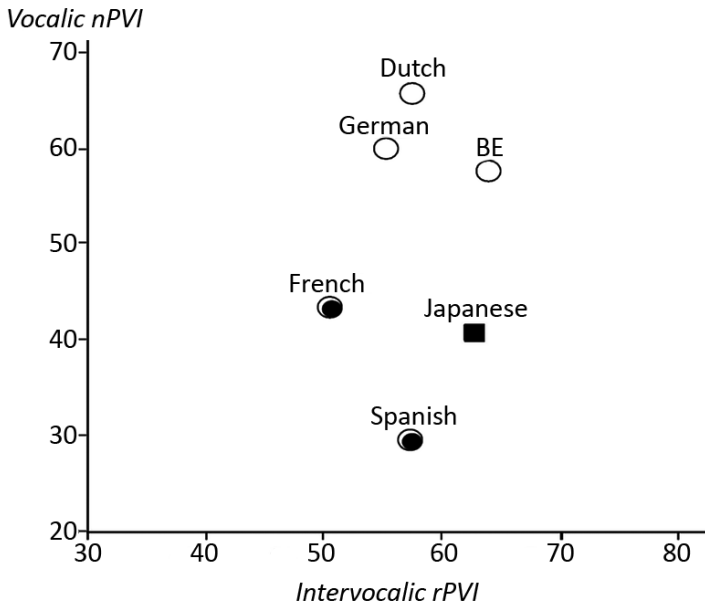
A magánhangzós és mássalhangzós szakaszokból számtalan más mérőszámot is létrehoztak, hogy minél jobban feltérképezhessék a beszédritmus és az időzítés kapcsolatát. A GRABE és LOW (2002) által felállított mérőszám típus (PVI) figyelembe veszi a magánhangzós és mássalhangzós szakaszok sorrendjét is. A PVI mérőszámok nem a különböző szakaszok tiszta időtartamából indulnak ki, hanem az egymást követő egységek időtartamainak különbségét veszik alapul (GRABE–LOW 2002), az időtartamokat tehát minden esetben a környező egységek

időtartamához képest határozzák meg. Külön mérőszámot alkottak a mássalhangzós és a magánhangzós szakaszokból: rPVI-C és nPVI-V. Az rPVI-C (raw Pairwise Variability Index) az egymást követő mássalhangzós szakaszok időtartamának variabilitását adja meg. Képlete a következő:

$$rPVI = \left(\sum_{k=1}^{m-1} |d_k - d_{k+1}| \right) / (m-1),$$

ahol m a szakaszok száma, a d_k pedig a k -edik szakasz időtartama. A másik mérőszám az nPVI (normalized Pairwise Variability Index), a PVI mérőszám beszédtempó szerint normalizált változata, azaz az rPVI osztva az összes szakasz időtartamának összegével. Ezt a mérőszámot a magánhangzós szakaszokra szokás alkalmazni (nPVI-V), ritkábban mássalhangzós szakaszok esetében is használják (nPVI-C).

$$nPVI = 100 \times \left(\sum_{k=1}^{m-1} |(d_k - d_{k+1}) / ((d_k + d_{k+1}) / 2)| \right) / (m-1)$$



2. ábra

A nyelvek PVI mérőszámok alapján történő csoportosítása

(○ a hangsúly-időzítésűekhez, ● a szótag-időzítésűekhez és ■ a moraidőzítésűekhez sorolt nyelvekre kapott PVI-eredmények (Grabe–Low 2002))

A két mérőszám alapján elkülönültek a tipikusan hangsúly-időzítésűnek tartott nyelvek (angol, holland) a klasszikusan szótag-időzítésűekhez sorolt francia és spanyol nyelvtől (2. ábra). Az nPVI-V mérőszám (hasonlóan a ΔV mérőszámhoz) magasabbnak mutatkozott a hangsúly-időzítésűekhez sorolt nyelvekben, mint a szótag-időzítésűekhez soroltakban. Az rPVI-C mérőszám értékei (hasonlóan a ΔC mérőszámhoz) alacsonyabbnak mutatkoztak a tipikusan szótag-időzítésű nyelvek esetében, mint a tipikusan hangsúly-időzítésű nyelveknél. Habár a PVI mérőszámok hasonló eredményekhez vezettek a nyelvek közti összehasonlításban, mégsem teljesen ugyanazon aspektusát tárják fel a nyelvek beszédritmusának, mint a korábban felállított mérőszámok. Míg a ΔV és ΔC mérőszámok egy adott szakaszban megjelenő szélsőséges időtartamértékekre érzékenyek, addig a PVI mérőszámok az egymást követő ritmusváltásoknak a nagyságát ragadják meg a beszédben. Minél egyenletesebb a tempó, azaz hasonlóbb az egymást követő egységek időtartama a beszédben, annál alacsonyabb ez a mérőszám. Minél nagyobb tempóbeli váltások vannak a beszédben, azaz hirtelen gyorsulások és hirtelen lassulások követik egymást, annál magasabb ez az érték. Többen úgy találták, hogy a PVI mérőszámok ezen tulajdonságuk miatt alkalmasak a prominenciajelölés (hangsúlyozás) időtartambeli megjelenésének mérésére (például FUCHS 2014). Ráadásul az artikulációs tempóra normalizált változata (nPVI-V) robusztusnak bizonyult a többi mérőszámhoz képest a fonológiai felépítettségével szemben is az angol, a spanyol és a katalán nyelvben (PRIETO et al. 2010, 2012). A különböző mérőszámok közti eltérés elsősorban olyan nyelvek esetében lehet érdekes, amelyekben az egymást közvetlenül követő szakaszokban nagy időtartambeli eltérések tapasztalhatók. Továbbá a nyelveken belüli vizsgálatokban is jelentősége lehet a különböző mérőszámok eltéréseinek. Ezekre fogunk példákat látni a következő két alfejezetben.

A beszédritmus-mérőszámok bírálati és a problémák kezelésének lehetséges módjai

A nyelvek beszédritmus-mérőszámok alapján történő csoportosításával kapcsolatban több probléma is felmerült. Az egyik kérdés, hogy a mérőszámok által felállított koordináta-rendszer mennyiben alkalmazható a nyelvek klasszikus három nagy beszédritmusosztályba sorolására. A mérőszámok alapján létrehozott koordináta-rendszerben a különböző típusú nyelvek nem különülnek el élesen egymástól. Találtak olyan nyelveket, amelyek semmilyen csoportosulásba nem sorolhatók. A görög, maláj, román, tamil, walesi nyelv, úgy tűnik, átmenetet képez a szótag-időzítésű és a hangsúly-időzítésű nyelvek között (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002).

További problémaként jelent meg a moraidőzítésű nyelvek csoportosulása. Ezek közül elsősorban a japán nyelvet vizsgálták, amely a kétféle különböző koordináta-rendszer (ΔC -%V, nPVI-V-rPVI-C) alapján eltérő sajátosságokat mutatott. Az egyik módszer szerint (RAMUS–NESPOR–MEHLER 1999) határozottan elkülönül a többi csoportosulástól, viszont a másik módszer alapján csak a klasszikusan hangsúly-időzítésű nyelvektől (GRABE–LOW 2002). Továbbá GRABE és LOW felhívták a figyelmet munkájukban arra is, hogy a thai és a tamil a RAMUS, NESPOR és MEHLER által felállított mérőszámok alapján szótag-időzítésű nyelveknek

tekinthetők. A PVI mérőszámok szerint viszont a hangsúly-időzítésűekhez közelítenek. Ugyanakkor a görög, a katalán és a walesi egyértelműen a hangsúly-időzítésű nyelvek közé sorolható a ΔC és a %V mérőszám alapján, de a PVI mérőszámok szerint átmenetet képeznek a hangsúly-időzítésű és a szótag-időzítésű csoportosulás között (GRABE–LOW 2002). A két mérés tehát nagyon hasonló egymáshoz, mégsem vezetnek ugyanarra az eredményre. Az eltérések egyik oka az lehet, hogy a ΔC és a %V mérőszámok a mondat vagy a közlés egészére vonatkoznak, míg a PVI mérőszámok figyelembe veszik a magánhangzós és a mássalhangzós szakaszok sorrendjét is, tehát a két mérőszám más-más aspektusát helyezi előtérbe a beszédritmusnak.

Egy másik problémát jelent, hogy a tipikusan szótag-időzítésű és tipikusan hangsúly-időzítésű nyelvek mérőszámértékei között nem minden esetben mutatható ki statisztikailag szignifikáns eltérés (vö. WHITE–MATTYS 2007; ARVANITI 2012). A probléma egyik lehetséges forrása, hogy a beszélők közti nagyfokú variabilitás elmossa a különbségeket az alapvetően eltérő beszédritmusú nyelvek között. Másrésztől a mérőszámok érzékenyek a vizsgált anyag felépítésére, beszédmódjára (szövegfelolvasás, mondatfelolvasás vs. spontán beszéd) is (ARVANITI 2012). Az eddigi kutatások eredményei szerint tehát nem teljesen egyértelmű, hogy a nyelvek közti beszédritmusbeli különbségek mennyire megbízhatóan határozhatók meg a mérőszámok alapján (ARVANITI 2012).

Ezek alapján felvetődik a kérdés, hogy léteznek-e egyáltalán valódi beszédritmusosztályok vagy -csoportosulások. A mérőszámok által kapott eredmények ugyanis nem különítenek el kategóriákat egymástól, sokkal inkább skálákat hoznak létre, amelyeken elhelyezhetők a különböző nyelvek egymás viszonylatában. Ebben a megközelítésben nem okoz problémát, hogy előfordulnak átmeneti nyelvek, vagy hogy a nyelvek között átfedés van, hiszen a mérőszámok által kijelölt beszédritmustérben folytonosan helyezkedhetnek a különböző nyelvek sajátosságaik alapján. A kérdés az, hogy vajon más mérőszámok képesek lennének-e ténylegesen kategorizálni a nyelveket, tehát csak az eddig fejlesztett mérőszámok nem alkalmasak erre, vagy ez a jelenség a nyelvek beszédritmus-sajátosságaiból fakad (vö. TURK 2013; NOLAN–JEON 2014).

Egy másik korlátja az imént tárgyalt beszédritmus-felfogásnak, hogy a beszédritmust a produkció felől közelíti meg. A beszédritmust azonban percepcióis élményként is értelmezhetjük (lásd ARVANITI 2009; BARRY–ANDREEVA–KOREMAN 2009; KOHLER 2009; CUMMING 2011). Ezen kiindulópontból könnyen belátható, hogy a beszéd többé-kevésbé rendszerszerű ismétlődésének érzetét nemcsak a szakaszok időtartambeli változatossága okozhatja a befogadóban, hanem számtalan más akusztikai tényező is indukálhatja. Az egyik lehetséges paraméter az intenzitás, melynek összefüggését a beszédritmussal spektrogram alapján, a szakaszokon alapuló mérőszámoktól függetlenül kísérelték meg mérni (TILSEN–JOHNSON 2008), de már született kísérlet az intenzitás PVI mérőszámokkal történő megragadására is (CICHOCKI–SELOUANI–PERREAULT 2014; FUCHS 2014). Az alaphangmagasság változása szintén kapcsolatba hozható a beszédritmussal, melynek akusztikai vonatkozását az intenzitáshoz hasonlóan a PVI mérőszámokkal vizsgálták (CUMMING 2011; FUCHS 2016).

A beszédritmus percepció felőli megközelítése további kérdéseket is felvet. A beszélő produkciójának vizsgálata során objektív adatokat kaphatunk a beszédritmusról. De ez nem mutatja meg, hogy ezek a paraméterek mennyire releváns információk a hallgató számára a beszédritmus szempontjából. A különböző anyanyelvű beszélők számára más-más akusztikai tényezőknek lehet jelentősége a beszédritmus észlelésében. CUMMING (2010) azt találta, hogy svájci német beszélők esetében az időtartamnak nagyobb szerepe volt a percepció ritmusítéletekben, mint az alaphangmagasságnak. A svájci francia és a franciaországi francia beszélők esetében viszont a két paraméternek hasonlóan fontos szerepe volt a percepció ítéletekben (CUMMING 2010). A beszédritmus-mérőszámok mért értékei ugyan kapcsolatba hozhatók az észlelt beszédritmussal (például WHITE et al. 2007), de jelenleg azt még nem tudjuk megbecsülni, mekkora különbség szükséges az objektív adatokban egy adott nyelven, hogy a hallgató számára is észlelhető legyen az eltérés.

Összefoglalva: a beszédritmus-mérőszámok a beszédritmusnak csak egyetlen aspektusával foglalkoznak, az időzítést mérik, és csupán ezt a dimenziót lehet velük vizsgálni. Ugyanakkor az időzítésen alapuló beszédritmus-mérőszámok nem alkalmazhatók beszédritmosztályok szigorú elkülönítésére, de egyfajta iránymutatók lehetnek a nyelvek közti beszédritmusbeli különbségek mérésében. A nyelvek közti különbségek eltéréseinek vizsgálata mellett a nyelven belüli időzítésbeli jellemzők feltárására is alkalmasak lehetnek. Felmerül a kérdés, hogy mennyiben megbízhatóak ezek a mérőszámok a nyelveken belül, illetve milyen tényezőkkel állhatnak kapcsolatban; ezekre a kérdésekre keressük a választ a következő alfejezetben.

1.3.4. A beszédritmus-mérőszámok és más tényezők kapcsolata egy-egy nyelven belül

Egy nyelven belül a beszédritmus időbeli dimenziója a mérőszámok alapján meglehetősen nagy változatosságot mutat (például ARVANITI 2012). A mérőszámok értékeit ugyanis számtalan tényező befolyásolhatja, mint például a beszélő dialektusa, neme, kora, a beszédmód (például felolvasás vs. spontán beszéd), az elhangzott beszéd szótagszintű felépítése, sőt, bizonyos mérőszámok összefüggést mutatnak az artikulációs tempóval is. A következőkben ezen tényezőknek és a beszédritmus-mérőszámoknak a kapcsolatát vesszük sorra.

Az artikulációs tempó és a beszédritmus-mérőszámok kapcsolata

A beszédritmus-mérőszámok egy részéről ismert, hogy összefügg az artikulációs tempóval. A ΔC , ΔV , rPVI-C egyértelműen negatívan korrelál az artikulációs tempóval, így ezen mérőszámok esetében nem szétválasztható, hogy vajon az artikulációs tempó vagy közvetlenül a beszédritmus (vagy legalábbis annak időzítésbeli dimenziója) mutat összefüggést a talált

jelenségekkel (DELLWO 2010). Mivel a kutatásokban más tényezők kapcsolatát is vizsgálni szeretnék volna, ezért az artikulációs tempótól függetlennek tekinthető mérőszámokat (VarcoV, VarcoC) hoztak létre. A VarcoC lényegében a ΔC , a VarcoV pedig a ΔV artikulációs tempóra normalizált változata.

$$\text{VarcoC} = \frac{\Delta C \cdot 100}{\text{mean}_c}$$

$$\text{VarcoV} = \frac{\Delta C \cdot 100}{\text{mean}_v}$$

A Varco-mérőszámok hasonló eredményeket hoztak a nyelvek közti különbségek esetében a korábbiakhoz képest (lásd DELLWO 2009), de használatukkal egyúttal könnyebbé vált a nyelveken belüli más tényezőkkel való összefüggések feltérképezése is, hiszen segítségükkel kizárható az artikulációs tempó hatása a vizsgált jelenségekre.

A PVI mérőszámok esetében az nPVI-V mérőszám definíciójából fakadóan szintén függetlennek tekinthető az artikulációs tempótól. A %V mérőszám artikulációs tempóval való összefüggése azonban nem teljesen egyértelmű. A mérőszám a definíciója szerint azt mutatja meg, hogy egy beszédrészlet szünetek nélküli időtartamának hány százaléka a magánhangzós szakaszok időtartama. Ez alapján tehát a %V – ahogy nevében is benne van – egy százalékos arány, azaz függetlenített a tényleges időtartamok nagyságától. Ebből az következne, hogy akár gyors, akár lassú valakinek az artikulációs tempója, ugyanúgy lehetne a magánhangzós és mássalhangzós szakaszok időtartamának aránya például 50-50%. A kutatási eredmények azonban ezt nem teljesen támasztják alá. Habár a %V mérőszámot meglehetősen robusztusnak találták a tempóval szemben más mérőszámokhoz (például ΔV , ΔC , rPVI-C) viszonyítva, mégis találtak valamiféle gyenge kapcsolatot közöttük néhány nyelvben. DELLWO (2010) a francia és cseh nyelvben nem talált összefüggést az artikulációs tempó (CV/s) és a %V között, angol és német nyelvben viszont igen. Ezekben a nyelvekben az artikulációs tempó növekedése tendenciaszerűen együtt járt a %V mérőszám növekedésével. A statisztikai tesztek viszont csak részben igazolták, hogy ez a tendencia jelentős lenne, a szerző éppen ezért elhanyagolhatónak tartja. RUSSO és BARRY (2008) 3 egyenlő részre (lassú, normál, gyors) osztották a spontán beszédben mért artikulációs tempóadatokat, és azt találták az olasz nyelvben, hogy a %V a tempó növekedésével (hang/s és szótag/s) csökkenő trendet mutat. Igaz, statisztikai próbával nem támasztották alá ezt az állítást. A német nyelvben ellentmondásosak az eredmények a tekintetben, hogy amikor a hangok számát osztották az időtartammal (hang/s), akkor a tempónövekedés együtt járt a %V kismértékű növekedésével, míg ha a szótagokat vették alapul, akkor csökkent a %V értéke a tempó növekedésével. Hangsúlyozzák, hogy a változás egyik esetben sem jelentős. A mérési eredmények tehát nem igazolják a %V és az artikulációs tempó közvetlen kapcsolatát.

A beszélők közti variabilitás

Egy nyelven belül a különböző beszélők nagy változatosságot mutathatnak a beszédritmus tekintetében. Egy beszélőnél a különböző beszédritmus-szabályszerűségek eltérő működését támasztja alá, hogy más-más beszédritmus-mérőszám vizsgálatok nem feltétlen ugyanannak a beszélőnek a beszédritmusa áll a legközelebb a nyelv átlagos beszédritmusához (ARVANITI 2012). Másrészről DELLWO, LEEMANN és KOLLY (2012) azt találták a svájci német nyelvet vizsgálva, hogy a %V- és az nPVI-V-értékek a különböző beszélők esetében nagyban eltérhetnek. Az rPVI-C azonban nem mutatott szignifikáns különbséget a beszélők között, tehát nem feltételezhető, hogy egy nyelven belül minden mérőszámra lényeges eltérések találhatóak a beszélők között. Felmerül a kérdés, hogy a beszélők közti különbségek milyen tényezőkkel mutatnak összefüggést.

Az egyik ilyen tényező a nemek közti eltérések. Az angol nyelven végzett vizsgálatokban a nőknél az egymást követő magánhangzós szakaszok időtartamának különbségei nagyobb szórást mutatnak, mint a férfiak esetében (mérőszáma: nPVI-V) (BENTON et al. 2007; TORGERSEN–SZAKAY 2012). A kínai nyelvre végzett vizsgálat során egy másik jellemzőt, a mássalhangzós szakaszok hosszúságának szórását (mérőszáma: ΔC) találták a nők esetében nagyobbak a férfi beszélőknél kapott értékekhez képest (BENTON et al. 2007). Német nyelvű felvételek elemzésekor lényeges eltéréseket találtak a különböző műfaji jellegű vizsgálati anyagok beszédritmikai mérőszámai között. Hírek felolvasásakor a férfiak beszédében az egymást követő szakaszoknak a szórása (nPVI-V és nPVI-C) mutatkozott nagyobbak a nőknél, míg a riportok vizsgálatában ugyanezen mérőszámok nők esetében kapott értéke bizonyult nagyobbak (GRAWUNDER et al. 2008). Az eredmények alapján tehát nem feltételezhető, hogy univerzális összefüggések állnának fenn a nemek és a beszédritmus-mérőszámok értékei között, de különbségek mutathatók ki a nők és férfiak beszédritmusa között.

A beszélők beszédritmusát az is befolyásolhatja, hogy milyen nyelvváltozatot használnak. A portugál nyelvet vizsgálva kimutatták, hogy a Brazíliában beszélt változat eltér a Portugáliában beszélt változattól (FROTA–VIGÁRIO 2001). Az eredmények alapján megállapították, hogy a brazíliai változat sokkal inkább a szótag-időzítésű nyelvekhez közelít, mint a portugál, amely inkább a hangsúly-időzítésű nyelvek felé tolódik el. A PVI mérőszámok egy változata alapján is találtak eltérést a két nyelvváltozat között, de nem különültek el élesen egymástól, átfedéseket mutattak (BARBOSA–VIANA–TRANCOSO 2009). Ugyanazon nyelvnek két változata tehát eltérhet egymástól, ez megerősíti azt a koncepciót, hogy a beszédritmus-mérőszámok által kimutatható különbségek nem pusztán egy nyelv fonológiai felépítettségének sajátosságaiból fakadnak. Az olasz nyelvben is mutattak ki eltéréseket a területi tagolódás szerint néhány mérőszám esetében (BARRY et al. 2003). Az angol nyelvben a beszélők lakóhelye szerint vizsgálódtak, és azt találták, hogy a belső londoni beszélők (Hackney-ben élők) által produkált beszédben az nPVI-V alacsonyabb volt, mint a külső londoni beszélők (Haveringben élők) esetében. Az eredmények feltehetően kapcsolatban vannak a multikulturális környezet, más nyelvek hatásával is. Ugyanebben a kutatásban az életkor is összefüggést

mutatott a beszédritmussal. A fiatalabb beszélők esetében magasabb nPVI-V-értékeket mértek, mint az idősebb beszélők esetében, de az eltérés nem volt jelentős (TORGERSEN–SZAKAY 2012). A beszélők neme, nyelvjárása, sőt kora is hatással lehet tehát arra, hogy a beszédritmus időbeli dimenziója hogyan realizálódik a beszédprodukciónban.

A vizsgált anyag kapcsolata a beszédritmus-mérőszámokkal

A beszédritmus-mérőszámokra hatással van, hogy a beszélők milyen szituációban, milyen beszédmódban nyilvánultak meg. A mért értékek igen eltérőek lehetnek egy nyelven belül attól függően, hogy felolvasásról vagy spontán beszédéről van szó. A mandarin nyelven végzett vizsgálat szerint a szövegfelolvasás során mért mérőszámok többsége (ΔC , nPVI-V, rPVI-C) tendenciaszerűen alacsonyabb volt a társalgás során mért mérőszámoknál. A %V volt az egyetlen mérőszám, amely esetében nem találtak hasonló összefüggést (LIN–WANG 2007). ARVANITI (2012) hat nyelven (angol, német, görög, olasz, koreai, spanyol) végzett kutatásában a spontán beszéd beszédritmusa szintén eltért a felolvasásától. Mind a mondatok, mind a szövegek felolvasása során mért mérőszám-értékek alacsonyabbak voltak a spontán beszédben mért értékeknél. Ezek a tendenciák azonban nem minden nyelvben, illetve mérőszám esetében érvényesültek. A ΔC mérőszám magasabb volt angol, német és görög nyelvben a spontán beszédben, mint felolvasásban, de ez a tendencia nem volt statisztikailag alátámasztható az olasz, spanyol és koreai nyelvben. Angol nyelvben különbség jelent meg a mondat- és szövegfelolvasás között is, az nPVI-V értékei szignifikánsan magasabbak voltak szövegfelolvasásban, mint mondatfelolvasásban. Olasz nyelvben ez a különbség a VarcoV mérőszámmal volt kimutatható. A VarcoC mérőszám esetében viszont nem a mondatokra kapott értékek voltak a legalacsonyabbak, hanem a szövegre számolt értékek, de olasz, koreai és spanyol nyelvekre ez a tendencia sem volt statisztikailag alátámasztható (ARVANITI 2012). Nem egyértelmű tehát, hogy a mondat- és szövegfelolvasások beszédritmusa között milyen különbségek vannak, sőt a különböző beszédmódok beszédritmusának eltérései meglehetősen nyelvfüggők az eredmények alapján.

Az eltéréseket befolyásolhatta az is, hogy a mondatok fonológiai felépítése szintén hatással van a beszédritmus-mérőszámokra, igaz, ez csak néhány mérőszámról mondható el (ΔC , rPVI-C, %V). PRIETO és munkatársai (2012) a beszédritmus-mérőszámokat három nagy csoportra osztják a nyelvek közötti összehasonlítások alapján. Az egyikben azok vannak, amelyek főként a szótag típusától függenek, ilyen mérőszám a ΔC és a VarcoC. A másik csoportban azok a mérőszámok vannak, amelyeket a frázisok prozódiai tényezői befolyásolnak, ilyenek a VarcoV és nPVI-V mérőszámok. A harmadik csoportba sorolják a %V, ΔV és az rPVI-C mérőszámokat, amelyek mind a fonológiai felépítettségtől, mind prozódiai tényezőktől függenek. Egy másik kutatás (ARVANITI 2012) szerint a különböző nyelvekben eltérő lehet, hogy mely mérőszámok érzékenyek a fonológiai felépítettségre. Egyrészt olyan mondatokat olvastattak fel, amelyekben a mondatok a szótagok szerkezetét nem kontrollálták. Másrészt olyan mondatokat választottak anyagnak, amelyek a hangsúly-időzítésű nyelvekben meglévő

szótagfelépítési sajátosságokkal rendelkeznek, tehát mássalhangzó-torlódásokat, geminátákat, diftongusokat stb. tartalmaznak, amennyire csak a nyelv ezt lehetővé teszi. A harmadik csoportba tartozó mondatokra pedig ezek teljes vagy majdnem teljes hiánya volt jellemző, ami a szótag-időzítésű nyelvekhez köthető tulajdonság. A %V mérőszám itt is minden esetben szoros összefüggést mutatott a szótagfelépítéssel a vizsgált nyelveken (angol, német, görög, olasz, koreai, spanyol) belül. A többi mérőszám azonban nem minden nyelvben mutatott szignifikáns különbséget attól függően, hogy kontrolláltan tipikusan „szótag-időzítésű nyelvre jellemző” vagy tipikusan „hangsúly-időzítésre jellemző” mondatokat olvastak fel. Az nPVI-V esetében nem volt jelentős különbség a két mondattípus között a koreai és a spanyol nyelvben, a VarcoV esetében a görög és a német nyelvben; az rPVI-C esetében a koreai nyelvben, a VarcoC esetében pedig a német és az olasz nyelvben. Fontos hozzátennünk, hogy a szélsőségesen kontrollált mondatokhoz képest a nem kontrollált mondatok nyelven belül egy köztes értéket vettek fel minden mérőszám esetében (ARVANITI 2012).

A fonológiai felépítettség befolyásának kiküszöbölésére további mérőszámot hoztak létre (CCI), amely az rPVI mérőszám módosított változata. Minden magánhangzós szakasz és minden mássalhangzós szakasz időtartamát elosztották a szakaszban lévő beszédhangok számával. A fonológiai megközelítés miatt a hosszú magánhangzókat és mássalhangzókat, illetve a diftongusokat két beszédhangnak tekintették. A szakaszokban lévő hangzók számába a töröltött beszédhangokat is beleszámították, mivel azokat kompenzációs jelenségként kezelték (BERTINETTO–BERTINI 2008, 2010; MAIRANO–ROMANO 2007).

A beszédritmus-mérőszámok módszerei, mint láthattuk, meglehetősen új keletűek, ezért a beszédritmus időbeli dimenziójáról, általános jellemzőiről és különböző tényezőkkel való összefüggéseiről még keveset tudunk. Néhány tényezőnek a beszédritmusra gyakorolt hatását (beszélők, beszédtempó, a közlés kommunikációs helyzete stb.) a beszédritmusra azonban már sikerült kimutatni. A módszer tökéletesítése folyamatosan zajlik, több problémát nem sikerült eddig megnyugtatóan megoldani. A megközelítés módszertana és alapegységei azonban hasznos eszközöknek tűnnek nyelveken belüli időbeli változások vizsgálatára.

1.3.5. A magyar nyelven végzett akusztikai beszédritmus-kutatások

Magyar nyelven végzett kutatásokban a beszédritmust komplex jelenségként értelmezve számtalan tényezővel próbálták összefüggésbe hozni. Az intenzitás, a beszédhang vagy az ezekkel összefüggésbe hozható hangsúly és a beszédritmus kapcsolatának vizsgálata a különböző verstanokban is nagy hagyományokkal rendelkezik (lásd GÁBOR 1925; KECSKÉS 1984). KECSKÉS (1966) a versritmussal foglalkozó tanulmányában objektív hangnyomás- és hangmagasság-méréseket is végzett az időtartammérések mellett. GÓSY (2000a) reppelt és normál ejtésű szövegekkel kapcsolatban is tárgyalja az intenzitás szerkezetet, a dallam és a beszédritmus viszonyát. A legtöbb kutatásban az időszerkezetet is vizsgálták, mint a beszédritmus egyik lehetséges megnyilvánulását. Habár nagy szakirodalma van a beszédritmus fonológiai

indíttatású és hangsúlyalapú vizsgálatának is a magyar nyelvben (VARGA 2002, 2005), ennek részletes ismertetésétől eltekintünk, és csak az időzítéssel kapcsolatos kutatásokra koncentrálnunk a következőkben.

A magyar nyelvet is megpróbálták besorolni a beszédritmusosztály-elmélet alapján. A magyar nyelvről feltételezték, hogy alapvetően a szótag-időzítésű nyelvek közé tartozik (SIPTÁR–TÖRKENCZY 2000; VARGA 1998, 2002, 2005). Ugyanakkor nem lehet tisztán ebbe a csoportba sorolni, mivel a szótagok hossza feltehetően változó a rövid és hosszú magánhangzók miatt (VARGA 2002). A magyarra végzett akusztikai mérések, hasonlóan a nemzetközi eredményekhez, nem támasztották alá a szótagok időtartamának egyenlőségét (KECSKÉS 1966; GÓSY 2000a). A versek beszédritmusának vizsgálata során KECSKÉS (1966) kimutatta, hogy a rövid és hosszú magánhangzók befolyásolják a szótagok időtartamát, valamint a mássalhangzó-torlódások megléte vagy hiánya is eltéréseket okoz a szótagok időtartamában. GÓSY (2000a) reppelt szöveg, időmértékes versek és normál ejtésű szövegek vizsgálatakor hasonló eredményre jutott. Kimutatta, hogy a szerkezetileg hosszú szótagok (zárt szótagok vagy hosszú magánhangzót tartalmazó szótagok) átlagosan hosszabbak a szerkezetileg rövid szótagoknál. Valamennyi szövegtípus elemzése azt mutatta, hogy az egymást követő szótagok időtartama váltakozik, mely egyfajta ritmikus lüktetésként fogható fel.

A magyar nyelv beszédritmusáról az elterjedt beszédritmus-mérőszámok alapján már érintőlegesen közöltek adatokat. WHITE, MATTYS és WIGET (2012) tanulmányában a magyar nyelvű felolvasott mondatokra kapott átlagértékek szerepelnek egy ábrán két mérőszám alapján (VarcoV és %V). A grafikonon szereplő adatok azt mutatják, hogy a magyar nyelv a tipikusan szótag-időzítésűekhez sorolt nyelvekhez hasonló tulajdonságokkal bír felolvasott mondatok alapján (WHITE–MATTYS–WIGET 2012). Láthattuk azonban, hogy a beszédritmus fonetikai megközelítései korántsem problémamentesek. Ugyanazon nyelvre más mérőszámok más eredményt hozhatnak, és nyelven belül igen nagy szórást eredményezhetnek különböző tényezők (beszédmód, beszélők közti variabilitás stb.). Éppen ezért több mérőszám bevonásával, a beszédritmus időbeli dimenzióját befolyásoló faktorok figyelembevételével további szisztematikus mérések elvégzésére van szükség a magyar nyelv beszédritmus-sajátságainak feltárásához.

1.4. Az artikulációs tempó és a beszédtempó

A beszéd kapcsán a *tempó* szó alatt sok mindent érthetünk. Az egyik legalapvetőbb megközelítés szerint elkülöníthetjük egymástól az objektív (mérhető) és a szubjektív (az egyéni érzékeléstől, észleléstől függő) tempót (GIBBON–KLESSA–BACHAN 2014). Az akusztikai méréseken belül figyelembe vehetjük a tempónál a szünetek időtartamát (beszédttempó), vagy figyelmen kívül is hagyhatjuk (artikulációs tempó). A vizsgált rész nagysága alapján pedig

elkülöníthetünk lokális vagy globális artikulációs tempót. A következőkben ezen fogalmak tisztázására kerül sor.

Az akusztikai méréseket sokféleképpen végezhetjük. Az egyik legfontosabb különbség, hogy figyelembe vesszük-e a szünetek időtartamát, vagy sem. Ez alapján megkülönböztetünk beszédtempót (*speaking rate*) és artikulációs tempót (*articulation rate*) (lásd TROUVAIN 2004). A nemzetközi szakirodalomban használják még a *gross* 'bruttó' vs. *net* 'nettó' kifejezéseket is a tempók megkülönböztetésére (TROUVAIN 2004; GIBBON–KLESSA–BACHAN 2014), de a továbbiakban a magyar nyelvben már használt terminológiához igazodunk. A beszédtempó egy időegység alatt elhangzó beszédjelek számát jelenti oly módon, hogy a beszéd idejébe a szünetek ideje is beleszámít. Az artikulációs tempó ezzel szemben azt adja meg, hogy a beszélők adott idő alatt hány hangot, szótagot vagy szót stb. ejtenek ki, ha a szünetek időtartamát nem vesszük figyelembe (GÓSY 2004). A tempót többféle mértékegységgel is szokták mérni: egyrészt az időegység mérése történhet ezredmásodpercben, másodpercben vagy percben, másrészt a beszédjelek számát meghatározhatjuk a hangok, a szótagok, a szavak vagy a mondatok számában is. Az eltérő mértékegység választása pedig különbségeket okozhat az eredményekben (vö. BÓNA 2009; GIBBON–KLESSA–BACHAN 2014).

1.4.1. Tempóingadozások a beszédben

A tempó egy másik megközelítése alapján megkülönböztethetünk globális és lokális tempót. Amíg a globális tempó a teljes szövegre vagy egy beszélő általános beszédmódjára vonatkozik, addig a lokális artikulációs tempó inkább a produkált szövegen belül megjelenő tempóvariációkat jelenti. Habár nem teljesen egyértelmű, hogy hol húzódik a határ a két kategória között (vö. FLETCHER 2010), megkülönböztetésük azért lehet iránymutató, mert egészen más eredményekhez vezethetnek. Hozzá kell tennünk, hogy mindkét típus vizsgálható akusztikai mérésekkel és percepciós tesztekkel egyaránt (PFITZINGER 1996, 1999; WAGNER–WINDMANN 2011; GIBBON–KLESSA–BACHAN 2014), de itt most csak az előbbire koncentrálnak. Kimutatták, hogy az időbeli mintázatok eltérőek lehetnek attól függően, hogy mekkora alapegységet vizsgálunk (lásd DE LOOZE 2010). Többek között DANKOVIČOVÁ (1997, 1999) mérései igazolták, hogy egy intonációs frázison (azaz hanglejtésük, hangsúlyozásuk szempontjából egy egységet alkotó beszédrésztelen) belül az artikulációs tempó nem egységes, hanem változik a szavak kiejtésének sebessége. A három-, négy- és ötszavas intonációs frázisok esetében a szavak artikulációs tempójában 6,7% és 56,6% között mozgott a tempóváltozás. A 10% alatti és 50% feletti változás igen ritkának volt mondható az anyagban. A szó helye a frázisban és a szavak hossza is befolyásolta a tempót, sőt, későbbi tanulmányában DANKOVIČOVÁ (1999) igazolta, hogy a funkciószavak gyorsabbak, mint a tartalmas szavak. Az adatok azt mutatták, hogy az intonációs frázison belül nem véletlenszerű a változás, hanem egyfajta mintázat, rendszerű szerveződés figyelhető meg. Az artikulációs tempóra kapott értékeket ordinális adatokká konvertálták, és csak azt vették figyelembe, hogy melyik szó valósult meg gyorsabban

a többihez képest, tehát sorrendet állítottak fel a szavak artikulációs tempója alapján. A cseh nyelvű adatok azt mutatták, hogy a kétszavas intonációs frázisokban lényegesen gyakoribb, hogy az első szó artikulációs tempója gyorsabb, mint a másodiké. Háromszavas intonációs frázisokban a szavak artikulációs tempója folyamatosan növekedett az esetek 49%-ában. Az esetek 39%-ában a második szó volt a leggyorsabb, az első valamivel lassabb, és a leglassabb tempójú az utolsó szó volt. Egyetlen esetben fordult elő, hogy az utolsó szó volt a leggyorsabb tempójú szó a frázison belül. A négyzavas intonációs frázisokon belül hasonló tendenciák jelentek meg. Az esetek 40%-ában az első vagy a második szó volt a leggyorsabb, 20%-ában pedig a harmadik, és nem fordult elő olyan, hogy az utolsó szó lett volna a leggyorsabb. Az esetek 68%-ában az utolsó szó volt a leglassabb. Öt- és hatszavas intonációs frázisok az alacsony darabszám miatt nehezen voltak vizsgálhatók. A lassuló mintázatokat részben a frázis végi nyúlás megjelenése magyarázhatja. Az intonációs frázisokkal ellentétben a beszédszakaszok nem mutattak szisztematikus mintázatot az artikulációs tempó tekintetében. Az intonációs frázison belüli artikulációstempó-ingadozást nem csak a cseh nyelvben igazolták. Angol felolvasásban és spontán szövegben hasonló lassuló tendenciákat találtak intonációs frázison belül (DANKOVIČOVÁ 1999).

DANKOVIČOVÁ (1997) módszereit magyar nyelvű szövegfelolvasásra is alkalmazták. A magyar nyelvben is kimutatható volt, hogy az artikulációs tempó nem tekinthető konstansnak. Ugyanis a kétszavas és a háromszavas intonációs frázisokon belül az első szóhoz képest a többi szó artikulációs tempója átlagosan lassabb, a háromszavas intonációs frázisokon belül pedig a második és a harmadik szó között is szignifikáns különbség mutatkozott. A négyzavas intonációs frázisokban az első és a második szó artikulációs tempójához képest a harmadik és a negyedik szó artikulációs tempója kimutathatóan lassabb volt. A mintázatok vizsgálatából kiderült, hogy a kétszavas intonációs frázisokban beszélőnként nagyjából fele-fele arányban fordulnak elő olyan esetek, amikor az első szó lassabb a másodiknál, illetve amikor a második szó lassabb az elsőnél (VÁRADI–BEKE 2013). Tehát a kétszavas intonációs frázisokban a magyar nyelvű adatok esetében nem volt egyértelmű lassuló tendencia. A háromszavas intonációs frázisok esetében az adatközlők háromnegyedénél az volt a leggyakoribb mintázat, hogy az első szó volt a leggyorsabb, a második szó a leglassabb, majd gyorsulás figyelhető meg. Két adatközlőnél az egész frázison végigvonuló gyorsulás volt a legjellemzőbb. Négy szóból álló intonációs frázisoknál 19 típusba sorolták az artikulációstempó-mintázatokat. A megvalósuló mintázatok háromnegyedében a leggyorsabb szó az intonációs frázis első vagy második szava volt. DANKOVIČOVÁ (1997) definíciója alapján az minősül egyértelműen lassuló tendenciájú intonációs frázisnak, amelyben az első vagy a második szó a leggyorsabb, és nincs egynél többször gyorsulás. Ez alapján kicsivel több, mint az esetek felénél sikerült ezt megállapítani. Az esetek 7%-ára volt jellemző a folyamatos gyorsulás, amelyekben az utolsó vagy az utolsó előtti szó volt a leggyorsabb, és nem fordult elő egynél több lassulás. Az esetek jelentős részét nem sikerült kategorizálni ezen csoportok alapján, ugyanis sem egyértelmű lassulást, sem egyértelmű gyorsulást nem tudtak megállapítani. Ezen úttörő kutatás megmutatta, hogy a magyar nyelvben az intonációs frázisokon belül egyfajta lassuló trend jellemző. A módszernek

viszont több korlátja is megmutatkozott. A háromnál több szavas intonációs frázisok nehezen ragadhatók meg ebben a keretben. A lassulás és gyorsulás, mint két külön kategória jelenik meg, miközben sok esetben nem teljesen egyértelmű az adott intonációs frázis tendenciája. Továbbá információt veszünk azzal, hogy nem vesszük figyelembe a tempóbeli változások mértékét, csak az egymáshoz képesti viszonyait, így nem vizsgálható, hogy mennyire jelentős a gyorsulások és lassulások mértéke. Mindamellett a módszerrel kapott eredmények egyértelműen mutatják, hogy a tempó nem kezelhető statikusan.

Nemcsak intonációs frázisokon belül mutatták ki az artikulációs tempó ingadozását, hanem nagyobb egységekben is megfigyelték a magyar nyelvben, hogy eltérő értéket vehet fel a szakasz nagyságtól függően. OLASZY (2006) beszédszakaszokat (szünettől szünetig terjedő beszédegységeket) vizsgált. (Bár munkájában erre a szerkezeti egységre a *prozódiai frázis* terminust alkalmazta, az átláthatóság érdekében továbbra is a *beszédszakasz* szakszót használom, mivel ebben az esetben megfeleltethetők egymásnak.) Azt találta, hogy az artikulációs tempó variabilitást mutatott a beszédszakaszokban is, igaz, ez kevésbé volt jelentős, mint az intonációs frázisokban. Hírolvasások vizsgálata azt is megmutatta, hogy a közlemény elején általában nagyobb az artikulációs sebesség, mint a közlemény végén. BÓNA (2007) azt vizsgálta, hogy vajon a tempóértékek közel hasonlónak tekinthetők-e attól függetlenül, hogy hány másodperces szakaszt veszünk figyelembe. Spontán narratívákon kimutatta, hogy eltérő eredményeket kapunk, ha egy hosszabb, 1,5 perces szakaszt vizsgálunk, vagy ha 10, illetve 1,5 másodperces szakaszokat vizsgálunk.

Az eredmények alapján úgy tűnik tehát, hogy megalapozott lenne szétválasztani a globális és lokális artikulációs tempót, még ha mérőmódszertanilag nem is teljesen egyértelmű, hogy hol húzható meg a határ a két típus között. Ha nem is szigorú kategóriákként különítjük el őket, mindenképpen érdemes figyelembe venni, hogy az adott artikulációs vagy beszédtempó mekkora szakaszokra jellemző. A kisebb szakaszokban lévő dinamikus temporális változások háttérben rengeteg jelenség állhat, például a frázis vagy megnyilatkozás végi nyúlás, a hangsúly időtartambeli hatása, a szóhosszúság hatása stb., ahogy azt korábban láthattuk (lásd részletesen a 1.2. fejezetben). Ugyanakkor hozzá kell tennünk, hogy a lassuló és a gyorsuló szakaszok éppen aktuális kiváltó okait nem feltétlen ismerjük. A következő alfejezetben viszont azokat a tényezőket vesszük sorra, amelyek a megnyilatkozás egésze alatt állandónak tekinthetők a beszédhelyzet miatt, és mégis mérhető hatással vannak a tempóra.

1.4.2. Az artikulációs és beszédtempóra ható általános tényezők

A különféle kommunikációs helyzetek eltérő tervezési nehézségek elé állítják a beszélőket, ami tempókülönbségekben is megnyilvánulhat. A különböző beszédmódokban eltérő artikulációs és beszédtempót mértek (lásd TROUVAIN 2004). Magyar nyelvben a kutatások többsége a felolvasást mérte a leggyorsabb tempójának a spontán beszédhez, a társalgáshoz vagy

a képleíráshoz képest (VÁRADI 2010; BÓNA 2013; MARKÓ 2014). A vizsgálatok azt is kimutatták, hogy a szöveg típusa is hatással van az artikulációs tempóra. A cselekményközpontú szövegek artikulációs tempóját például gyorsabbnak mérték az értékelésközpontúakénál (ANDÓ 2002). A beszédmód mellett a frázisok, mondatok vagy beszédszakaszok (szünettől szünetig terjedő részek) hosszától is függhet a tempó. A hosszabb frázisokat, illetve beszédszakaszokat tendenciaszerűen gyorsabban ejtjük, mint a rövidebb szakaszokat (QUENÉ 2007; JACEWICZ–FOX–WEI 2010).

Az artikulációs és beszédtempóban jelentős egyéni különbségek is előfordulnak (QUENÉ 2004, 2007, 2008; JACEWICZ et al. 2009; JACEWICZ–FOX–WEI 2010; GÓSY–BEKE–HORVÁTH 2011). A különböző életkorú beszélők tempója például jelentősen eltérhet egymástól (FÓNAGY–MAGDICS 1960; GOCSÁL 2000; MENYHÁRT 2000; VERHOEVEN–DE PAUW–KLOOTS 2004; BÓNA 2010). Az idős beszélők lassabban beszélnek, mint a fiatalabb felnőttek, ez a magyar nyelvben is kimutatható (GOCSÁL 2000; MENYHÁRT 2000; BÓNA 2010; BÓNA 2014, 2015). Ugyanakkor az egyik kutatás rámutat arra is, hogy az idősebb beszélők hajlamosabbak rövidebb frázisokat használni, és ez együtt járhat a tempó lassulásával, bár ezen tényező figyelembevételével is tapasztalható valamelyest az életkor hatása a tempóra (QUENÉ 2008). A nyelvelsajátítás során is megfigyelhetők különbségek a beszéd- és artikulációs tempóban. A gyermekek artikulációs és beszédtempója egyre gyorsabb az idő előrehaladtával (például HASELAGER–SLIS–RIETVELD 1991; WHITESIDE–HODGSON 2000; NEUBERGER 2013).

Az életkor mellett a nemek is okozhatják a beszélők közti temporális eltéréseket. Több kutatás azt igazolta, hogy a nők lassabban beszélnek a férfiaknál (BYRD 1994; VERHOEVEN–DE PAUW–KLOOTS 2004; JACEWICZ et al. 2009; JACEWICZ–FOX–WEI 2010). A magyar nyelvben talált eredmények viszont ellentmondásosak. VÁRADI (2009) kutatásában – hasonlóan a nemzetközi szakirodalomhoz – azt találta, hogy a férfiak gyorsabban beszélnek a nőknél. Nem minden kutatásban sikerült azonban alátámasztani ezt a különbséget, és többen arra a következtetésre jutottak, hogy magyar nyelvben a nők és a férfiak tempóértékeiben nincs, vagy szinte nincs eltérés egymástól (GOCSÁL 2001; GOCSÁL 2010; BÓNA 2015).

A beszélő érzelmi állapota is befolyásolhatja az artikulációs és beszédtempót (lásd SCHERER 2003; TROUVAIN 2004). SCHERER (1995) azt találta, hogy a beszédtempó gyorsabb félelem és felháborodás esetében, szomorúság és unalom esetében viszont lassabb. Magyar nyelvű beszédben is hasonló eredményeket találtak. Kimutatták, hogy a vidám érzelmi töl-tettel elmondott beszéd gyorsabb, mint a szomorú hangulatú (SZABÓ 2008). A percepciók kísérletek is azt támasztják alá, hogy a lelkiállapot és a beszéd időbeli szerveződése között kapcsolat áll fenn (MAGDICS 1964; GOCSÁL 2010). Továbbá nemcsak a beszélő hangulata befolyásolhatja az artikulációs és beszédtempót, hanem a személyiségjegyek is (például DEWAELE–FURNHAM 2000; GOCSÁL–HUSZÁR 2003; GOCSÁL 2010). Számos kutatás igazolta azt is, hogy a dialektus befolyással lehet az artikulációs és beszédtempóra (BYRD 1994; VERHOEVEN–DE PAUW–KLOOTS 2004; QUENÉ 2008; JACEWICZ et al. 2009; JACEWICZ–FOX–WEI 2010; MENYHÁRT 2010). Láthatjuk tehát, hogy a globális tempó mértékét rengeteg külső tényező együttes hatása hozza létre.

1.4.3. A szünetezés

Az artikulációs tempó és a beszédtempó között lévő fő különbség, hogy míg az előbbibe nem számítjuk bele a szünetek időtartamát, az utóbbi esetében figyelembe vesszük. A beszéd-szünet a következőképpen definiálható (GÓSY 2000b: 2): „a beszéd folyamatban jelentkező szünet olyan kismértékben akaratlagos beszédkimaradás, amely néma vagy jellel kitöltött, de független a beszédhang képzésétől. Funkcióját tekintve a beszédprodukciónban 1. biztosítja az artikulációt lehetővé tevő légáramot, 2. elősegíti a közlés értelmi tagolását, 3. a beszédtervezés során az ún. ellentmondások, téves utak stb. feloldására szolgál, 4. a mentális lexikonban történő keresési idő kitöltését biztosítja, illetőleg lehetőséget nyújt a nyelvi kódolás módosítására. Funkciói a beszédmegértésben: 1. az elhangzottak könnyebb feldolgozása, 2. az entrópia csökkentése és 3. a megértés és az értelmezés működési folyamatainak biztosítása”. Ezenfelül társalgásokban is megjelenhetnek jelkimaradások a beszédben két társalgási egység között, amely jelenségre MARKÓ (2005) a *hallgatás* terminusát vezeti be. A szünet ugyanakkor nem feltétlenül jelenti az akusztikai jel teljes hiányát (GÓSY 2000; TROUVAIN 2004; MARKÓ 2005; FLETCHER 2010). A jellel kitöltött szünetet *hezitálás*nak nevezzük, amelynek magyar beszédben megjelenő leggyakoribb formái az *öö* és az *mm*, illetve az *öm* stb. (HORVÁTH 2010; GÓSY et al. 2013). Gyakran a hezitálás előtt vagy után, vagy mindkét helyen megjelenhetnek néma szünetek is, akár folyamatosan egymást válthatják, ezeket *kombinált szünetek*nek nevezik (HORVÁTH 2010; BÓNA 2013a). A szünetektől el szokás különíteni a bizonytalanságból fakadó megakadásokat, mint például az ismétléseket és a téves kezdéseket (lásd GÓSY 2003, 2005).

A beszéd-szünetek gyakorisága és időtartama számos tényezőtől függ. Egyrésztől meghatározó lehet a mondat hossza és összetettsége, ugyanakkor hozzá kell tennünk, hogy a szünetek megjelenése nem feltétlenül esik egybe a szintaktikai, prozódiai határokkal, illetve ezeken a helyeken nem szükségszerűen jelenik meg szünet (VOLSKAYA 2003; TROUVAIN 2004; KRIVOKAPIC 2007). A szünetek megjelenése gyakran együtt jár azzal, hogy az előtte lévő beszédhangok időtartama megnövekedik, de a prozódiai határokkal való kapcsolat miatt nehéz lenne szétválasztani, hogy a szünettartás bekövetkezte vagy a szintaktikai/prozódiai határ indukálja ezt a jelenséget (vö. HARDCASTLE–LAVER–GIBBON 2010). A beszéd-szünetek megjelenésének gyakorisága és hossza továbbá meglehetősen egyénfüggő (MARKÓ 2005; BÓNA 2013a), hatással van rá a beszélő életkora (BÓNA 2013b), de fontos szerepet játszik a műfaj (DUEZ 1982; IMRE 2005; OLASZY 2006) és nem utolsósorban a beszédmód is (DUEZ 1982; MARKÓ 2005; VÁRADI 2010; BÓNA 2013a).

A különböző beszédmódokra más-más szünetezési mintázatok jellemzők. A szünetek előfordulásának aránya és hossza kimutathatóan nagyobb spontán beszédben, mint felolvasásban (VÁRADI 2010; BÓNA 2011; BÓNA 2013a). A kitöltött szünetek egyáltalán nem (VÁRADI 2010), vagy nagyon ritkán jelennek meg felolvasásban (BÓNA 2013a), míg spontán beszédben igen gyakran fordulnak elő. Ez kapcsolatba hozható azzal, hogy a felolvasásban nincs szükség nyelvi tervezésre, így az ebből fakadó időnyerési stratégiák következtében nem jelennek

meg szünetek (VÁRADI 2011). A szünetek előfordulásának aránya és hossza még a spontán beszédnél is nagyobb tartalom-visszamondáskor narratívákban, ami nem meglepő, hiszen a nyelvi tervezés ebben a beszédmódban tekinthető a legösszetettebb feladatnak a vizsgált beszédmódok közül (BÓNA 2013a). A nyelvi jelek folyamatos sorozata a beszédben tehát több okból is megszakadhat, éppen ezért nehéz megjósolni, hogy a beszélők hol és mennyi ideig tartanak szünetet.

1.4.4. Az artikulációs és beszédtempó szélsőséges értékeinek korlátai

Amint azt korábban kifejtettük (az *1.1. fejezetben*), a beszédhangokra jellemző egy sajátos időtartam, és szótagok, szavak szintjén megvalósulhatnak olyan időzítésbeli szabályszerűségek, mint a hangsúly nyújtó hatása vagy a közlés/megnyilatkozás végén mérhető nyúlás. A lassabb vagy gyorsabb tempó ezen szabályszerűségekkel összhangban kell hogy megvalósuljon (TURK–SHATTUCK–HUFNAGEL 2014). A beszédhangok esetében ismert, hogy nem egyformán tudnak nyúlni vagy rövidülni. A magánhangzók időtartamában például nagyobb változás mutatható ki általánosságban, mint a mássalhangzókéban (HOFHUIS–GUSSENHOVEN–RIETVELD 1995), bár a mássalhangzó képzésmódjától is jelentősen függ, hogy mennyiben nyúlhat meg: mint ahogy korábban tárgyaltuk, a spiránsok elaszticitása igen nagy a többi képzésmódú mássalhangzóhoz képest (például OLLER 1973).

A beszédhangok lehetséges nyúlása, illetve rövidülése ugyanakkor végesnek tekinthető. Ezt egyrésztől meghatározza, hogy legkevesebb mennyi idő alatt lehet kiejteni egy hangot, illetve mennyi ideig lehet fenntartani (lásd KLATT 1973; XU 2009; TURK–SHATTUCK–HUFNAGEL 2014). Mivel a hangok nem rövidülhetnek és nyúlhatnak a végtelenségig, így a hangok időtartamából számolt tempónak is korlátai vannak. Magyar nyelvben a mért adatok alapján 20–25 hang/ms is előfordulhat, és ez alapján OLASZY (2010a) a 30 hang/s-ot tekinti a felső határnak, így egy beszédhangra átlagosan 33 ms jutna. Bár hozzátézi, hogy ilyen sebességnél már jelentős torzulásokkal lehet számolni. Az artikulációs korlátok mellett viszont az is fontos szempont, hogy a beszéd a hallgató számára érthető maradjon. A kétféle szempontot ötvözi LINDBLOM (1990) H&H (hyper- and hypospeech) elmélete, amely szerint a beszélő a kommunikációs helyzetben kétféle célt követ. A beszélő egyrésztől törekszik a gazdaságos beszédprodukcióna, tehát minimalizálni próbálja az artikulációs befektetést, ezért az egyes hangokra jellemző artikulációs és akusztikai mintát (ún. célkonfigurációt) nem mindig éri el. A beszélő másik célja, hogy a hallgató vagy hallgatók számára érthető legyen, amit mond, ezért törekszik a beszédhangok, szavak megkülönböztetéséhez elegendő akusztikai jegyet produkálni. A két célt tekinthetjük egy skála két végpontjának. Amennyiben a gazdaságosság és a redundancia elkerülése kerül előtérbe, akkor alulartikulált beszédről van szó. Amennyiben az érthetőség kerül előtérbe, akkor túlartikulált beszéd jöhet létre. A hétköznapi életben a beszédprodukción nyilvánvalóan a skála két végpontja között helyezkedik el valahol.

1.5. A beszédtechnológia és az időszerkezet

Az elmúlt néhány évtizedben a beszédtechnológia különböző területein belül is megnőtt az érdeklődés a beszéd időbeli sajátosságai iránt. Mind a beszédpszintézisben, mind a beszédészlelésben központi kérdés a beszéd időbeli szerveződése és annak modellezése. A beszédtechnológia különböző területein a feladatok végrehajtásához akusztikai modelleket állítanak fel, amelyeknek értelemszerűen részét kell hogy képezze az időzítés is. Továbbá a modellek létrehozását gyakran akusztikai mérésekkel készítik elő, olykor a létrehozott modellek pontosságának ellenőrzését végzik időtartamméréseken keresztül, amely eredmények szintén szorosan kapcsolódnak a jelen kötet témájához. Több olyan időtartamra ható tényezőre, illetve különböző faktorok interakciójára irányították rá a figyelmet, amelyekkel a kontrollált akusztikai kísérletekben kevésbé foglalkoztak. Ezért a következőkben röviden áttekintjük ezen alkalmazások, illetve kutatások témánkhoz kapcsolódó részeit, különös tekintettel a magyar nyelvre vonatkozóakra.

A beszéd gépi előállítása során az a feladat, hogy leírt szövegből kiindulva meghatározzák a folyamatos hangzó szövegek beszédhangjainak időtartamát. A beszédpszintézisben az érthetőség mellett a természetesség is fontos szempont, amelyhez elengedhetetlen a szuprasegmentális eszközök (köztük a temporális sajátosságok) minél jobb ismerete és ezen tudás felhasználása az alkalmazásokban (HINTERLEITNER–NORRENBROCK–MÖLLER 2013; HENTER et al. 2016). A beszédpszintézis megvalósulásai között ugyan van olyan módszer (például elemkiválasztás-alapú technológiák), amelynek nincs kifejezetten időtartammodellje, de mi a következőkben csak az időzítést is figyelembe vevő technológiákra koncentrálnak.

Az akusztikai vizsgálatok alapján megállapítható ugyan egy-egy beszédhang időtartama a konkrét helyzetben, mivel azonban a beszédhangok időtartamára rengeteg tényező hat, nehezen jósolható meg előre akár egy szó vagy egy mondat, esetleg egy teljes szöveg időszerkezete. A beszéd gépi előállításához ezért olyan modelleket próbáltak meg létrehozni, amelyek meghatározzák a beszédhangok időtartamát bármely kontextusban. Az egyik első ilyen kísérlet KLATT (1976) szabályalapú modellje volt, amelyhez kiindulási alapot jelentettek korábbi fonetikai mérések eredményei. A beszédhangot tekintette alapegységnek, és minden beszédhangra megállapított egy átlagos időtartamot különböző kontextusban végzett mérések alapján. Meghatározta, hogy például a megnyilatkozás végi helyzet vagy a hangsúlyos pozíció arányaiban mennyivel nyújtja meg a hangzókat, és ezzel az aránnyal hosszabbította meg az eredetileg megállapított átlagos időtartamot. Olyan további szabályszerűségeket vett figyelembe, mint hogy a nem közlés végi hangzók rövidülnek az átlagos időtartamhoz képest, a hangsúlytalan magánhangzók nyúlnak zöngés mássalhangzók előtt stb. (Ezen szabályok mindegyike az *1.1.* és az *1.2. fejezet*ben már ismertetett akusztikai mérési eredményekkel összhangban vannak.) Egy beszédhangra értelemszerűen több szabály is vonatkozhat egyszerre ebben a modellben. Előfordulhat, hogy egy hangzónak a környezete és a pozíciója miatt több szabály alapján is arányosan rövidebbnek kell lennie a rá jellemző megállapított

időtartamnál, de a modell tartalmaz egy minimum korlátot. Egy bizonyos szint alá viszont nem csökkenhet a hangzók időtartama. Előfordulhat olyan helyzet is, hogy ugyanazt a hangot egy szabály nyújtaná, egy másik rövidítené, ilyenkor ezen szabályok kiolthatják egymás hatását. VAN SANTEN (1992) továbbfejlesztette a szabályalapú (főként a Klatt-féle) modellek alapvető elveit, és egy olyan megközelítést hozott létre, amelyben a beszédhangok – elsősorban a magánhangzók – időtartamára ható tényezők interakciójának hatását is jobban közelítette, továbbá kibővítette az időtartamra ható faktorok sorát. A rendelkezésére álló mérési adatok alapján azt találta az angol nyelvben, hogy a magánhangzók időtartamának variabilitását leginkább magyarázó tényezők közé tartozik a hangzó minősége, a szótag- és a frázishangsúly, a követő mássalhangzó zöngéssége, a szón és a megnyilatkozáson belüli pozíció. Számítalan hasonló szabályalapú modellt hoztak létre különböző nyelvekre (például német: KOHLER 1988, PORTELE–SENDLEMEIER–HESS 1990; francia: O’SHAUGHNESSY 1984, BARTKOVA–SORIN 1987), köztük magyarra is (OLASZY 2000, 2006).

Az 1990-es évek elején fejlesztett első magyar beszéd-szintézis-modell célkitűzése az volt, hogy egy meglévő szövegből szabályok révén meg tudja jósolni az egyes beszédhangok időtartamát a hangsorban. A modell kiindulópontja szintén az ún. specifikus időtartam, amelyről korábban már érintőlegesen szó esett. A specifikus időtartam definíciója szerint olyan alapidőtartamot jelent, amely a beszédhang artikulációs konfigurációjából és a közvetlen hangkörnyezetéből fakad (OLASZY 2010a). A specifikus időtartamot ebben a modellben percepciók ítéletek alapján határozták meg, majd az így kapott eredményeket szabályok révén módosították a végleges időtartam elérése céljából. A specifikus időtartamokat először a szó szintjén változtatták meg: a szavak első hat szótagjában rövidítették a beszédhangokat. A mondat szintjén vették figyelembe a hangsúlyozást: a hangsúlyos helyzetben a hangzókat nyújtották, hangsúlytalan szótagokban rövidítették, ami egyfajta gyorsítást-lassítást hozott létre az időszerkezetben (lásd OLASZY 2006). További szabály, hogy minél hosszabb egy mondat, annál jobban rövidítik a benne lévő hangok időtartamát. A modell a szó helyzetét is figyelembe vette a mondaton belül, a szünet előtti szót, különösen a szónak a végét lassították. A szünetet határoló elemnek tekintették az időszerkezet tekintetében, a szünet után ugyanis mindig új időszerkezeti frázis indul. Az ezen modell segítségével előállított beszéd időtartamadatai hasonló mintázatot mutattak a természetes felolvasásból származó időtartamadatokkal, ami a meghatározott szabályok, szabályszerűségek meglétére utal (OLASZY 2006, 2010b).

A gépi szövegfelolvasók egy másik nagy ága a statisztikai parametrikus beszéd-szintézis. A beszédből először paraméterek formájában kinyerik a beszéd különböző tulajdonságait (spektrális és prozódiai információkat), majd ezt az adathalmazt statisztikai, gépi tanuló eljárásokkal modellezzik. A gépi tanuló eljárás sokféle lehet. A két legismertebb ezek közül a rejtett Markov-modell (Hidden Markov-modell, HMM) és a neurális hálózat (lásd részletesen például TAYLOR 2009; RAO 2012; ZEN 2015). Ezeknek a modelleknek általában nincsenek olyan explicit időtartammodelljei, mint a szabályalapú beszéd-szintézisnél láthattuk, de a tanító-adatbázis címkézése és a tanulóalgoritmusok fejlesztése során hasonló nyelvi és

kontextusfüggő tényezőket használnak fel a gépi beszéd előállításához, mint a szabályalapú modelleknél. A HMM-alapú beszéd-szintézis során a tanító-adatbázist hang-, szó- és mondat-szinten is felcímkézik, a címkézés által pedig meghatározhatóvá válnak például egy hang és a hang környezetének (például megelőző és követő hang) sajátosságai is. Az összes mondat összes hangjához kiszámolják a környezetet leíró paramétereket, mivel azonban a lehetséges kombinációk száma túl nagy ahhoz, hogy megfelelően reprezentatív adatbázist tudjanak hozzá készíteni, ezért ún. döntési fákat alkalmaznak, lényegében osztályokba sorolják az adatokat. A döntési fák egyfajta hierarchiát állítanak fel a különböző tényezők között, azért, hogy kiderüljön, melyek jellemzik jobban a rendelkezésre álló adatokat, melyek kevésbé. Az egyik ilyen magyar nyelvű, rejtett Markov-moddellen alapuló gépi szövegfelolvasó modellje alapján a hangzók képzési konfigurációja kulcsszerepet játszik az időtartamok meghatározásában. A modell felépítéséhez további olyan tényezőket vettek figyelembe, mint például a szótag hangsúlyos vagy hangsúlytalan volta, a fonológiai felépítettség, az aktuális szótag pozíciója a szóban, a szó pozíciója a mondatban, a szavak száma a mondatban stb. (TÓTH–NÉMETH 2010; TÓTH 2013).

A gépi tanuló eljárások közül az utóbbi években a neurális hálózatokat is egyre többet használják. Ezen módszer előnye a döntési fákon alapuló, kontextusfüggő rejtett Markov-modellekkal szemben (másnéven: HMM), hogy hatékonyabban figyelembe tudják venni a különböző időtartamokra ható faktorok interakcióját, hiszen ezen módszerrel súlyozni lehet az egyes faktorok közti kapcsolatok erősségét. Ezáltal jobban tudják modellezni a beszéd időzítési mintázatait is (WATTS et al. 2016). NAGY és NÉMETH (2016) a beszédhangok időtartamának jóslásához mély neurális hálózatokat és környezetfüggő rejtett Markov-modellt is használtak. Azt tapasztalták, hogy az átlagos hosszúságú mondatok esetében mindkét módszer hasonlóan hatékonyan jósolta meg az időtartamokat. A rövid (egy, két és három szótagos) mondatok esetében viszont a neurális hálózattal készített modell alapján létrehozott időtartamok kevésbé tértek el a ténylegesen mért időtartamoktól, jobban korreláltak velük. BEKE és GÓSY (2014) a rövid és hosszú magánhangzók időtartamát modellezték, figyelembe véve a hangzó minőségén és hosszúságán kívül a szótag helyzetét is. Eredményeik szerint neurális hálózatokkal pontosabban lehet becsülni a magánhangzók időtartamát, mint rejtett Markov-modellekkel.

A beszéd gépi előállítása során a beszédhangok időtartamának meghatározásánál tehát felhasználják az akusztikai mérések eredményeit, hogy egyáltalán milyen tényezők hathatnak az időszerkezetre. Ismereteink azonban hiányosak a tekintetben, hogy a különböző tényezők hogyan hatnak együttesen, milyen általános szabályszerűségek fogalmazhatók meg folyamatos szövegekben. Habár a legújabb gépi tanuló eljárások (neurális hálózatok) egyre jobban jósolják a beszédhangok időtartamait megnyilatkozásokban, még ezen modellekben is jelentős mértékű időtartambeli eltéréseket tapasztalhatunk a jóslt és a valós adatok között (vö. NAGY–NÉMETH 2016). Bármilyen gépi előállítási módszerről is legyen szó, feltételezhető, hogy további időtartamra ható tényezők bevonása tovább csökkentheti az eltérések mértékét a jövőben (VAN SANTEN 1992; TÓTH 2013; NAGY–NÉMETH 2016).

A beszédtechnológia másik nagy területe – a beszéd-szintézis mellett – a gépi beszéd-észlelés, amelynek egyes megoldásaiban szintén felmerülő problémakör az időzítés kezelése és modellezése. A gépi beszéd-észlelés egyik fő feladata a beszéd szöveggé alakítása, azaz a beszéd gépi felismerése. A további lehetséges feladatok közé tartozik a beszélő személyének a felismerése, a beszéd detektálása, azaz annak megállapítása, hogy ki mikor beszél, a beszéd részekre (frázisokra) tagolása is (SZASZÁK–MIHAJLIK–FEGYÓ 2010).

A gépi beszéd-felismerők jelenleg a beszédet szöveggé alakítják anélkül, hogy a megnyilatkozások jelentését, tényleges üzenetét megértenék. Először a felismerni kívánt beszédből kinyerik a lényeges akusztikai tulajdonságokat, majd ezeket az információkat összehasonlítják referencia-adatbázisokkal, amelyeknek már ismert a lejegyzése. A lehető legjobban illeszkedő elemet próbálják megtalálni az adatbázisból, hogy megkapják a beszéd leírt változatát. A mintaillesztésre számos eljárás ismert, köztük statisztikai alapú mintaillesztési módszerek (rejtett Markov-modell, neurális hálózat) is. A felismerés során problémát jelent, hogy a beszédre nagy akusztikai változatosság jellemző, például ugyanazon szavakat nem ugyanolyan időtartamban ejtjük ki. Ez azt jelenti, hogy ugyanazon szó két különböző előfordulása során a beszédhangok nem ugyanolyan mértékben rövidebbek vagy hosszabbak az egyik vagy másik megjelenés esetében. Nem lehet tehát lineárisan megrövidíteni vagy meghosszabbítani az időtartamokat, hanem bonyolultabb eljárásra van szükség ezen probléma megoldásához. Az ún. dinamikus idővetemítés módszerét használják ilyenkor, amely során úgy transzformálják a két összehasonlítani kívánt szó jellemzőinek időtengelyét, hogy minél nagyobb egyezés jöjjön létre a tesztadat és a minta között. Úgy párosítják a szavak időbeli egymásutánban mért tulajdonságait, hogy az eltérések összege minimális legyen (SZASZÁK–MIHAJLIK–FEGYÓ 2010). Ez a módszer fontos a gépi beszéd-felismerés szempontjából, ám emiatt a rövid-hosszú fonológiai oppozíció nehezen ragadható meg a gépi felismerésben. Ennek ellenére a hangidőtartamokat nem feltétlenül szokták expliciten modellezni a beszéd-felismerőkben (MIHAJLIK 2010). Többen viszont (PYLKKÖNEN–KURIMO 2004; TÓTH–KOCOSOR 2005) hangsúlyozzák az időtartammodellek használatának jelentőségét, főként olyan nyelvek esetében, ahol az időtartamnak jelentésmegkülönböztető szerepe van.

A prozódia, azon belül pedig a temporális jellemzők (beszédrítmus, beszéd és artikulációs tempó) felhasználása a beszéd-felismerésben még kevésbé elterjedt. A beszéd karaktersorozattá alakításához nélkülözhetőnek tűnik, hiszen ezen információk nélkül is meglehetősen jó arányban ismerik fel az alkalmazások a szavakat (például GALES–YOUNG 2008; MIHAJLIK et al. 2010). Ugyanakkor prozódiai modell felhasználásával könnyebben meghatározható az írásjelek használata a beszéd-felismerésben, továbbá a beszédet könnyebben tagolhatjuk frázisokra (SZASZÁK–MIHAJLIK–FEGYÓ 2010). Magyar beszédben történtek már kísérletek arra, hogy a beszéd időszerkezete és más prozódiai tulajdonságok alapján megjósolják a hangsúlytól hangsúlyig terjedő szakaszok vagy a fonológiai frázisok kezdetét és végét. Az eddigi eredmények alapján viszont nem sikerült kimutatni, hogy az időtartamok hozzájárulnának a beszéd ezen egységeire való tagolódásában (vö. SZASZÁK–MIHAJLIK–FEGYÓ 2010; BEKE–GÖSY 2014).

Ebben a fejezetben láthattuk, hogy a folyamatos beszéd időszerkezetének modellezése meglehetősen nehéz feladat. A különböző beszédtechnológiai területeken alkalmazott megközelítések olyan megoldásokat tárnak elénk, amelyekben a beszédhangok, nagyobb egységek kontextusfüggő időviszonyait próbálják megragadni oly módon, hogy egyszerre veszik figyelembe az időtartamra ható összes tényezőt és azok interakcióját, valamint az időzítés véletlenszerű változatosságát is.

2. A kísérletsorozat kérdésfeltevései, célja, hipotézisei

Az itt bemutatott kutatás célja az volt, hogy akusztikai mérések segítségével első közelítést adjon arról, hogy a magyar beszédben különféle beszédmódokban milyen alapvető időzítésbeli mintázatok jellemzők.

I. A kötet egyik központi kérdésköre, hogy hogyan jellemezhető a magyar nyelv beszédritmusának időbeli dimenziója a nemzetközi szakirodalomból ismert beszédritmus-mérőszámokkal. Igazolható-e vajon akusztikai mérésekkel is, hogy a magyar nyelv beszédritmusa a szótag-időzítésű nyelvekéhez hasonló tulajdonságokkal rendelkezik beszédmódtól és beszélőtől függetlenül? A kutatás további célja, hogy összefüggéseket találjon a különböző beszélők, nemek szerinti csoportok, beszédmódok, a szünettartás és a hangzó beszéd időszerkezete között a beszédritmus-mérőszámok alapján.

Mivel a magyar nyelvet a szótag-időzítésű nyelvekhez szokták sorolni, így feltehetően a beszédritmus-mérőszám által mutatott akusztikai mérések eredményei azt igazolják majd, hogy a magyar nyelv a szótag-időzítésű nyelvekhez hasonló sajátosságokkal rendelkezik.

H1: Hipotézisem szerint a különböző beszédmódokban eltérő beszédritmus tapasztalható. A spontán beszédben feltehetően magasabb mérőszámértékek jellemzők a megjelenő szélsőségesebb időtartamértékek miatt, mint szöveg- vagy mondatfelolvasásban.

H2: Feltételezem, hogy a mondatfelolvasás és a szövegfelolvasás beszédritmusa eltér, mivel a folyamatos beszédben megjelenhetnek olyan átfogó időzítésbeli szabályszerűségek, amelyek egy-egy mondat felolvasásában nincsenek jelen.

H3: A beszédritmus feltehetően nemcsak a különböző beszédmódokban, hanem a nemek szerint is eltérést mutat.

H4: A különböző beszélők szünettartási szokásai tetten érhetők a beszédritmusban is. Hipotézisem szerint azon kísérleti személyek beszédében, akik gyakrabban tartanak szünetet, azaz jobban tagolják a szöveget, magasabb beszédritmusmérőszám-értékek mérhetők, mint azokéban, akik kevesebb szünetet tartanak.

II. A kötet másik központi kérdése, hogy hogyan jellemezhetőek a tempóingadozások a magyar beszédben. Meghatározható-e egy általános lassuló vagy gyorsuló tendencia, amely a beszédet jellemzi? A kutatás további célja, hogy feltárja a lassulások és a gyorsulások általános sajátosságait és kapcsolatát más tényezőkkel. Kérdés, hogy az általános trend vajon különböző nagyságú beszédrészekben azonos módon jelentkezik-e. Található-e eltérés a különböző beszédmódok általános trendjei, mintázatai között? Milyen eltérések vannak a férfiak és a nők ejtésében? Hogyan alakulnak a lassuló vagy gyorsuló beszédszakaszokban a beszédhangok időtartamai? Vajon hangról hangra egyre nő az időtartam lassulás esetén, illetve egyre csökken gyorsulás esetén, azaz eltörlődnek a specifikus időtartamok közti

különbségek, vagy az egyes beszédszakaszokban csak általánosságban érzékelhető a lassulás, illetve a gyorsulás? Ezen kérdések megválaszolásához mérés módszertani újítások alkalmazása szükséges.¹

A kutatás célja ezért továbbá olyan eljárás kidolgozása, amely egy adott beszédszakasról meghatározza valamilyen módon, hogy jellemzően lassuló vagy gyorsuló trendet mutat-e, továbbá a lassulásokhoz és gyorsulásokhoz valamiféle mértéket tud rendelni.

H5: Hipotézisem szerint a magyar beszédben általános lassuló trend jellemző, amely kimutatható tagmondatokban és tagmondatoknál nagyobb egységekben egyaránt. Mivel a beszéd nem tud folyamatosan lassulni, ezért a lassulás vélhetően kevésbé érzékelhető egy szöveg egészén, mint kisebb egységeken.

H6: Mivel felolvasás során a beszélőnek nem kell megterveznie a szöveget, koncentrálni a szöveg tagolására, ezért feltételezhető, hogy következetesebben jelenik meg a lassuló trend a szövegfelolvasásokban, mint a spontán beszédben.

H7: Hipotézisem szerint – hasonlóan a beszédritmus-mérőszámokhoz – a lassulás mértékében is tapasztalhatók lesznek eltérések a férfiak és a nők ejtésében, de várhatóan az általános trend mindkét csoportnál egyértelműen a lassulás lesz.

H8: Továbbá feltételezem, hogy található olyan szakaszok a beszédben, ahol a lassulás vagy a gyorsulás eltörli a beszédhangok specifikus időtartamai közti különbséget, és beszédhangról beszédhangra vagy szakaszból szakaszra fokozatosan nő vagy csökken az egységek időtartama.

III. Könyvem harmadik kérdésköre, hogy milyen összefüggésben van a megnyilatkozás végén a temporális szerkezet és az irreguláris, illetve a nem modális (irreguláris és/vagy leheletes) zöngemínőség megjelenése.

H9: Hipotézisem szerint a glottalizáció megjelenése összefügg a lassítások megjelenésével. Egyrészt elképzelhető, hogy a két jelenség jellemzően együtt fordul elő, közösen jelölve a megnyilatkozásvéget. Ugyanakkor ennek ellenkezője is feltételezhető, miszerint kiegészítő viszonyban vannak egymással, tehát a beszélők azokban az esetekben, amikor glottalizációval jelzik a határ közeledését, nem/kevésbé alkalmaznak lassítást a megnyilatkozás végéhez közeledve; és fordítva: tempólassítás alkalmazásakor a zöngéképzésük nem/kevésbé alakul irregulárisba.

H10: Feltételezem továbbá, hogy a különálló mondatok felolvasása esetén mindkét jelenség (a tempólassítás és a glottalizáció is) gyakoribb, illetve mértékük nagyobb a mondat végéhez közeledve, mint a szövegfelolvasás mondataiban. Ennek háttérében azt feltételezem, hogy a különálló mondatok önállóbb egységek, mint a szövegmondatok, így ezek a megnyilatkozásvégek nagyobb mértékben lehetnek jelöltek.

¹ A mérés módszertani újítások szükségességének háttéréről részletesebben lásd az *1.4.1. fejezetet*.

3. Általános módszertan

Ebben a fejezetben tárgyalt kutatások kísérleti személyeinek, felhasznált anyagának, valamint módszertanának általános bemutatására kerül sor. Részletesebb információk az egyes kutatásokkal kapcsolatos további fejezetekben találhatóak.

3.1. Kísérleti személyek és anyag

Az időzítés vizsgálatához 20 magyar anyanyelvű beszélő (10 nő és 10 férfi) spontán beszédét, szövegfelolvasását és mondatfelolvasását használtuk fel. Életkoruk 20 és 60 év között mozgott. Az összehasonlíthatóság érdekében a beszélők életkora nemek szerinti megoszlásban is hasonló volt: a férfiak átlagéletkora 36,9 év, a nőké 37,5 év (2. táblázat). Az adatközlők ép hallók és ép beszédűek voltak.

2. táblázat
A férfi és női beszélők életkora a használt anyagban

	Férfi beszélők	Életkor	Női beszélők	Életkor
	F1	20	N1	20
	F2	27	N2	28
	F3	29	N3	29
	F4	33	N4	31
	F5	35	N5	30
	F6	38	N6	34
	F7	40	N7	44
	F8	45	N8	45
	F9	45	N9	54
	F10	57	N10	60
Átlag		36,9		37,5

A felvételek a BEA adatbázisból származnak (Gósy 2008). A felvételek mindig azonos körülmények között készültek ugyanabban a csendesített helyiségben. A rögzítés digitális,

közvetlenül számítógépre történik a GoldWave hangeditáló szoftverrel, 44,1 kHz-es mintavételezéssel (tárolás: 16 bit, 86 kbyte/s, monó) (Gósy et al. 2012).

Az adatbázis tartalmaz spontán beszédet, mondat- és szövegfelolvasást, mondatok ismétlését, tartalomösszegzést is, melyekből csak az elsőként említett hármat tettem vizsgálataim tárgyává. Interjúhelyzetben rögzített spontán beszédből beszélőnként legalább 2 percnyi egybefüggő részletet választottam ki, amelyben az interjúztató nem szólalt meg. A beszélők saját munkájukról vagy éppen folyó tanulmányaikról monologikusan beszéltek a felvételeken. Az első lehetséges egybefüggő monologikus részt választottam ki, hogy az adatközlő lehetőleg minél inkább a rá jellemző tempóban, módon nyilvánuljon meg, kevésbé alkalmazkodjon az interjúztató beszédéhez.² A kutatásokhoz felhasználtam a beszélők szövegfelolvasását is. Ugyanazon szöveg felolvasása a különböző beszélők esetében eltérő időtartamban valósult meg, a legrövidebb szövegfelolvasás 107,7 másodperc, a leghosszabb 159,7 másodperc volt. A mondatfelolvasások esetében beszélőnként az adatbázisban szereplő mind a 25 mondatot felhasználtam. Vizsgálataimat ezen anyag alapján végeztem de egy-egy állítás igazolásához olykor csak egy részét használtam fel a rendelkezésemre álló adatoknak. A kutatási céloknak megfelelően választottam ki, hogy a meglévő adatokból mennyit elemzek.

A különböző beszédmódok összevetésének vizsgálatakor többféleképpen használtam fel a rendelkezésemre álló anyagot. Az egyik esetben a szövegek felolvasását vettem össze spontán monologikus szöveggel. (Mivel a beszélők meglehetősen tagoltan, nagy szünetekkel elválasztva valósították meg az egymás után következő mondatok felolvasását, ezért e vizsgálatból kihagytam ezen beszédmód vizsgálatát.) A spontán beszédrészletek kétpercesek voltak, ezért a szövegfelolvasásnak csak az első két percét elemeztem minden beszélő esetében. A 120 másodperc egy hozzávetőleges érték, az eltérés mértéke maximum +/- 11 mp. Ennek egyik oka, hogy öten néhány másodperccel rövidebb idő alatt olvasták fel a szöveget, illetve nem vágtam ketté a szavakat, tehát megkerestem a 120. másodpercnél lévő hangot, és az azt tartalmazó szó végéig vettem a szövegrészletet, mind a spontán beszédben, mind a szövegfelolvasásban.

Egy teljesen másik elemzési anyagot is használtam a beszédmódok összevetéséhez, a teljes egybefüggő szövegek vizsgálatán túl tagmondatnyi egységeken is megvizsgáltam a beszédmódok közti különbségeket. A beszédmódok összehasonlításába belevontam a mondatfelolvasásokat is. Csak azokat a mondatokat vettem figyelembe az összehasonlíthatóság érdekében, amelyek egy tagmondatból álltak. 20 darab ilyen mondat fordult elő az adatbázisban, amelyeknek szótagszáma 15 és 27 között volt, a mondatok átlagos szótagszáma 21,4.

A szöveg összesen 25 tagmondatra bontható, ám innen is csak 20 tagmondatot vizsgáltam. Mivel több beszélő nem olvasta fel a címet, így a szöveg esetében a tagmondatok számolását a szövegtest első mondatával kezdtem. A tagmondatok kiválasztásának alapját

² Ennek ellenére a beszédalkalmazkodás valamilyen mértékben befolyásolhatja az eredményeket, különösen, ha az interjúalany és az interjúztató ismerték egymást. A beszédalkalmazkodás hatása viszont bármilyen felvételhelyzetben nehezen szűrhető, ezért ettől a hatástól eltekinttem.

a szótagszám adta. A legkevesebb szótagszámú tagmondatok kizárásra kerültek, csak a 11 szótagnál hosszabbakat elemeztem. A spontán beszédből szintén 20 tagmondatot vizsgáltam, amelyek minimum 10 szótagúak voltak. A tagmondatok elkülönítése felolvasások esetében könnyebb, mint spontán beszédben. A tagmondathatárok meghatározásában az irányelv az volt, hogy az adott rész egy alany-állítmányi viszonyal rendelkező szövegegység legyen (például *Péter pizzát kért az étteremben.*). Két vagy több alany-állítmányi viszony pedig összetett mondatnak minősül (KESZLER 2000: 461), tehát több tagmondatra bontható. A halmozott állítmányi mondatrészeket egy tagmondathatárral tekintetem abban az esetben, ha az alanyon kívül nem volt más, nem közös bővítményük (hasonlóan KESZLER 2000). Ilyen halmozott állítmányi viszonyal rendelkező tagmondatra példa: *A vaníliás fagy finom és édes.; Evelin mosott, főzött és takarított.* A spontán beszédben gyakran fordultak elő néhány szótagos tagmondatok. Ezek szélsőségesen befolyásolhatták volna a beszédritmus-mérőszámok eredményeit, ezért kimaradtak az elemzésből. A minimum szótagszámot – közelítve a szövegfelolvasásban előforduló legalacsonyabb értékhez – 10 szótagban határoztam meg. A tagmondatokra tagolás egy másik nehézsége az volt, hogy tagmondathatáron előfordult olyan eset, amikor a beszélők különböző tagmondathoz tartozó hangzókat szünet nélkül összeolvadva ejtettek ki. Az egyik beszélő ejtésében fordult elő szövegfelolvasásnál, hogy például *A legtöbb háztartásban télen és tél végén kedvelt vitaminforrás a primőr fejes saláta, amely a szakhatóságok elmúlt éveiben végzett vizsgálatait alapján a legtöbb lemoshatatlan mérget tartalmazza.* mondatban a *saláta* és az *amely* szavak utolsó és első hangját szünet és glottalizáció nélkül valósította meg. Ebben és ehhez hasonló esetekben egyik tagmondathatárnál sem vettem figyelembe a hangokat és azok időtartamát.

Összegezve a fentebb leírtakat, a három beszédmódot (mondatfelolvasás, szövegfelolvasás, spontán beszéd) tagmondataikra bontva is megvizsgáltam, követve azt az irányelvet, hogy egy tagmondat egy alany-állítmányi viszonyal rendelkező szövegegység legyen. Ezzel szintaktikailag hasonló „méretű” egységeket tudtam vizsgálni, amelyeket szótagszám tekintetében is közelítettem egymáshoz egy alsó korlát meghatározásával.

3.2. Módszer, szegmentálási nehézségek

A hanganyagokat hangszinten annotáltam a magyar nyelvre is adaptált MAUS elnevezésű automatikusan szegmentáló szoftverrel (SCHIEL 1999). Jelöltem a beszédhangok minőségét, határaikat, valamint a szüneteket. Az így kapott hanghatárokat manuálisan ellenőriztem a Praat 5.1 szoftverben (BOERSMA–WEENINK 2009). A magánhangzók időtartamát a formánsszerkezet kezdetéhez és végéhez igazítottam az oszcillogram, a spektrogram és auditív információk segítségével, követve a nemzetközi szakirodalomban szokásos hangelhatárolási kritériumokat (PETERSON–LEHISTE 1960; GRABE–LOW 2002). A formánsstruktúra mellett figyelembe vettem

az oszcillogramon megjelenő amplitúdóváltozásokat és a spektrogramon látható intenzitás-eltéréseket is. A spiránsokat és a zöngétlen felpattanó zárhangokat követő magánhangzók elválasztását a formánsstruktúra – elsősorban a második formáns – megjelenéséhez igazítottam. Ehhez hasonlóan a magánhangzó és a zöngétlen spiráns vagy a zöngétlen felpattanó zárhang határát a formánsok (elsősorban a második formáns) végéhez igazítottam. A nazálisok és a magánhangzók határának eldöntésében elsődlegesen a formánsstruktúra változása játszott szerepet, de emellett az intenzitás változását is figyelembe vettem. A közlést kezdő és a szünet utáni zöngétlen felpattanó zárhangok zárszakaszának idejét úgy határoztam meg, hogy ugyanazon közlés következő, ugyanolyan képzésű hangzó zárszakaszának idejét hozzáadtam a zárhang zörejes részéhez. Mivel a frázis végi nyúlás nagymértékben befolyásolja a hangok időtartamát (lásd az *1.2.1. fejezetben*), ezért csak azokat az ugyanolyan képzésű, zöngétlen felpattanó zárhangokat vettem figyelembe, amelyek legalább három szótagonál nagyobb távolságra voltak a felolvasott mondatok végétől és a szünetektől. Az approximánsok és a magánhangzók határának meghatározásában az F1-ben és az F2-ben, valamint az F3-ban bekövetkező hirtelen formánsváltozások adták a kiindulópontot, továbbá az intenzitás hirtelen csökkenése is iránymutató volt a jelölésben az auditív információk mellett (hasonlóan: MACHAČ–SKARNITZL 2009). Azokban az esetekben, ahol a hanghatárok nem voltak egyértelműek az oszcillogram és a spektrogram alapján, tehát a fenti elhatárolási elveket egyáltalán nem lehetett érvényesíteni, a hanghatárt a formánsátmenet felénel helyeztem el. A magyar nyelvben igen gyakori, hogy a magánhangzók nem periodikus zöngével realizálódnak (MARKÓ 2014), ezeket a szakaszokat is a magánhangzók részeként jelöltem.

A hangszintű címkesor alapján kiszámítottam a magánhangzós és a mássalhangzós szakaszok időtartamát. Mássalhangzós szakasznak tekintetem a beszédnek azon részeit, amelyben a mássalhangzók egymást közvetlenül követik, és nem választja őket szét magánhangzó. A magánhangzós szakaszokat ehhez hasonlóan úgy határoztam meg, hogy az egymást közvetlenül követő magánhangzók alkottak egy egységet (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002 alapján). A hangsorok felépítésétől függően egy mássalhangzós szakasz állhat egy, kettő vagy akár több mássalhangzóból is, míg a magánhangzós szakaszok szintén tartalmazhatnak egy, kettő vagy több magánhangzót. A beszédben tehát mássalhangzós és magánhangzós szakaszok folyamatosan váltják egymást.

A szegmentumokat egy C++ nyelven írt program segítségével C és V szegmentumokra kódoltam. A közvetlenül egymás mellett lévő mássalhangzókat, amelyeket nem választott el magánhangzó egymástól, összevontam. A mássalhangzós szakaszok idejét pedig úgy számoltam ki, hogy a bennük lévő mássalhangzók időtartamát összeadtam. Ugyanezt a műveletet végeztem el a magánhangzókra is. Így kaptam meg külön a magánhangzós és a mássalhangzós szakaszok időtartamát. Közvetlenül egymás mellett állónak tekintetem a magánhangzókat, illetve a mássalhangzókat akkor is, ha szünet vagy szóhatár választotta el őket. A szünetek időtartamát nem számoltam bele a szakasz időtartamába (hasonlóan: GRABE–LOW 2002; WHITE–MATTYS 2007). A magánhangzós szakaszok és a mássalhangzós szakaszok egybeolvasztása a szüneteken keresztül azért indokolt, mert így hosszabb szövegek is vizsgálhatóvá

válnak (ez elsősorban az egymást követő szakaszok variabilitását vizsgáló mérőszámok szempontjából lényeges). A felolvasásban előforduló megakadásokat nem vettem figyelembe az elemzéskor. A hiátustöltő funkciójú /j/ realizációkat, mivel ezeket fonológiailag is mássalhangzóként kezelik (lásd É. KISS–SIPTÁR–KIEFER 2003), és definíciószerűen két magánhangzó között helyezkednek el, mássalhangzóként kezeltem, és a mássalhangzós szakasz részének tekintettem, hasonlóan az időzítéssel és beszédritmussal foglalkozó szakirodalmi forrásokhoz (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002).

A statisztikai elemzések egy részét az SPSS 20.0 szoftverrel végeztem. Mivel a statisztikai eljárások elvégzésének előfeltétele az adatok normál eloszlása, ezért a próbák elvégzése előtt normalitásvizsgálatot végeztem (Shapiro–Wilk-próba alapján). Az adatok nem normál eloszlása miatt nemparametrikus tesztek alkalmaztam (Mann–Whitney-próba, Wilcoxon-próba, Spearman-féle korrelációelemzés, χ^2 -próba), továbbá a nemparametrikus tesztek használatát indokolta a 4.2. és a 4.3. fejezetben az adatok nem arányskála jellege is. Az adatok felépítése és struktúrája miatt saját mérési módszer fejlesztésére is sor került (lásd a 4.2.1. és a 4.2.2. fejezetben).

4. Kísérletsorozatok az időzítési mintázatok vizsgálatára

4.1. Az időzítés mint a beszédritmus egyik dimenziója

A kutatás egyik központi témája, hogy hogyan jellemezhető a magyar nyelv beszédritmusának időbeli dimenziója a nemzetközi szakirodalomból ismert beszédritmus-mérőszámokkal. Az egyik alapvető kérdés: vajon igazolható-e akusztikai mérésekkel is, hogy a magyar nyelv beszédritmusa a szótag-időzítésű nyelvekéhez hasonló tulajdonságokkal rendelkezik beszédmódtól és beszélőtől függetlenül? Korábban a magyar nyelvről többen feltételezték, hogy alapvetően a szótag-időzítésű nyelvek közé tartozik (SIPTÁR–TÖRKENCZY 2000; VARGA 2002, 2005). Beszédritmus-mérőszámok tekintetében korábban WHITE, MATTYS és WIGET (2012) vizsgáltak magyar nyelvű felolvasott mondatokat. A szerzők két mérőszám (VarcoV és %V) értékét közölték, amelyek hasonlónak bizonyultak a hagyományosan szótag-időzítésűnek tekintett nyelvekre kapott adatokhoz. Megjegyzendő azonban, hogy az ilyen összevetéseket befolyásolhatja a vizsgált beszédritmus-mérőszámok megválasztása, valamint hogy egyéb tényezők (például beszédmód, beszélők közti variabilitás stb.) egy nyelven belül is jelentős eltéréseket eredményezhetnek (ARVANITI 2012). Ezért a magyar nyelv beszédritmus-sajátságainak pontosabb feltárásához több mérőszámot is tekintetbe vevő további mérésekre van szükség.

A beszédritmusra ható egyik fontos tényező a beszédmód. Több nyelven kimutatták, hogy a beszédritmus-mérőszámokra kapott értékek nagyban függenek attól, felolvasásról vagy spontán beszédre van szó (például LIN–WANG 2007; ARVANITI 2012). A mandarin nyelvre végzett vizsgálatok eredményei (LIN–WANG 2007) szerint például a beszédritmus-mérőszámok többsége (ΔC , nPVI-V, rPVI-C) jellemzően alacsonyabbnak bizonyult szövegfelolvasás esetében, mint társalgás során (a %V esetében azonban nem mutatkozott ilyen összefüggés). Hasonló következtetésekre jutott ARVANITI (2012) hat nyelvet (angol, német, görög, olasz, koreai, spanyol) elemezve: a mérőszám-értékek mind különálló mondatok, mind összefüggő szövegek felolvasásakor többnyire alacsonyabbnak bizonyultak, mint spontán beszéd során. Ezen átlagos tendenciák azonban nem jelentkeztek minden nyelv és minden mérőszám esetén. Az sem egyértelmű, hogy a mondat- és szövegfelolvasások beszédritmusa között milyen különbségek vannak. Ezért kutatásom további célja, hogy megvizsgáljam a beszédmód hatását a beszédritmus-mérőszám értékeire a magyar nyelvben; megvizsgáljam a spontán beszéd, a szöveg- és a mondatfelolvasás közti különbségeket.

Egy nyelven belül a különböző beszélők között is jelentős eltérések tapasztalhatók a beszédritmus szempontjából. Nem minden mérőszám esetében volt kimutatható jelentős

eltérés a beszélők között, továbbá nem mindig ugyanannak a beszélőnek a beszédritmusa állt legközelebb a nyelv átlagos beszédritmusához. Ezen eredmények alapján felmerül a kérdés, hogy a beszélők közti különbségek milyen tényezőkkel mutatnak összefüggést. A beszélőknél például meglehetősen változó, hogy milyen gyakran tartanak szünetet (GÓSY 2003). A szünetek megjelenése pedig gyakran együtt jár azzal, hogy az előtte lévő beszédhangok időtartama megnövekedik. A prozódiai határokkal való kapcsolat miatt azonban nehéz lenne szétválasztani, hogy a szünettartás bekövetkezte vagy a szintaktikai/prozódiai határ indukálja ezt a jelenséget (vö. FLETCHER 2010). Eltekintve az ok-okozati viszonytól, a beszédhangok időtartama kapcsolatban van a szünetezéssel, így feltehetően a beszédritmus-mérőszámok értékeivel is (közvetett vagy közvetlen) összefüggésben lehet. Ezért kutatásomban összefüggéseket kerestem a szünettartási szokások és a beszédritmus-mérőszám értékei között.

A beszélő neme szintén befolyásolhatja a beszédritmust. Angol, kínai és német nyelven végzett kísérletek során is kimutatták, hogy a beszédritmus több szempontból eltér a nők és a férfiak ejtésében (BENTON et al. 2007; GRAWUNDER et al. 2008; TORGERSEN–SZAKAY 2012). A beszédritmus nemek szerinti eltérései másként alakultak a különböző nyelvekben, ugyanis más-más beszédritmus-mérőszámok tértek el jelentősen a férfiak és a nők ejtésében nyelvtől függően. Az eredményeket a beszédmód is befolyásolta (GRAWUNDER et al. 2008). Feltehetően a magyar beszéd ritmusában is megnyilvánulnak a nemek között különbségek, ezért ezt a lehetséges összefüggést is megvizsgáljam a 4.1.2.3. fejezetben, figyelembe véve a különböző beszéd módokat.

4.1.1. Kísérleti személyek, anyag és módszer

Kísérleti személyek és anyag

A kutatás során azon kísérleti személyek anyagait elemeztem, amelyeket a 3.1. fejezetben már ismertettem. A főbb paraméterek a következők voltak: 10 férfi és 10 nő felolvasása került elemzésre a BEA adatbázisból (GÓSY 2008). A beszélők életkora 20 és 60 év között szóródott. A vizsgálatok elvégzéséhez spontán monologikus beszédrészeket, szövegfelolvasásokat és mondatfelolvasásokat is felhasználtam. A különböző beszéd módok összevetéséhez kétféleképpen vizsgáltam az anyagot: teljes egybefüggő szövegeken és tagmondatnyi egységeken (részletesen lásd a 3.1. fejezetben). A különböző beszéd módok felhasználását elsősorban az indokolta, hogy a beszédritmus szakirodalmában korábban más-más anyagon vizsgáldták. Míg RAMUS, NESPOR és MEHLER (1999) a beszédritmus-mérőszámokat külön-külön egy-egy mondat felolvasásán vizsgálták, addig GRABE és LOW (2002) egy teljes, egybefüggő szöveg felolvasására alkalmazták ugyanazokat a mérőszámokat. Mint korábban kifejtettem, eredményeik eltéréseket mutattak (lásd az 1.3.3. fejezetben).

A hanganyag annotálása és az akusztikai elemzésben használt módszerek

A hanganyag hangszintű annotálását, ahogy korábban részletesen kifejtettem a (3.2. fejezetben), a nemzetközi szakirodalomban található kritériumoknak megfelelően végeztem el, figyelembe véve a magyar beszédben lévő sajátos akusztikai realizációkat. A MAUS szoftverrel történő automatikus hangszintű annotálást (SCHIEL 1999; KISLER–SCHIEL–SLOETJES 2012) a Praat 5.1 szoftverben (BOERSMA–WEENINK 2009) ellenőriztem, majd C++ nyelven írt program segítségével meghatároztam a mássalhangzós és magánhangzós szakaszokat és időtartamukat. Az artikulációs tempót úgy számoltam ki, hogy a hangok számát osztottam az artikuláció tiszta idejére eső időtartammal (hang/s) a szakirodalomban lévő eljárásnak megfelelően (lásd GÓSY 2004). Az artikulációs tempó számításánál mindig akkora egységet vettem alapul, amekkora egység képezte a beszédritmus-mérőszám számításának alapját is. Tehát ha tagmondatokat vizsgáltam, akkor a tagmondatok artikulációs tempóját és beszédritmus-mérőszámait határoztam meg, amennyiben 2 perces szövegeket vettem figyelembe, akkor az abban lévő összes hangra számoltam ki mind az artikulációs tempót, mind a mérőszámokat.

Mivel a korábbi eredmények azt mutatták, hogy a különféle mérőszámok segítségével felállított csoportok kismértékben eltérhetnek, ezért többféle mérőszámmal is megvizsgáltam a beszédritmust. Minden egyes tagmondatra, szövegre kiszámoltam a ΔC , ΔV , %V, Varco-C, Varco-V, rPVI-C, nPVI-C és nPVI-V mérőszámokat.

A használt beszédritmus-mérőszámok definíciói és képletei összegezve a következők:

- ΔV : a magánhangzós szakaszok időtartamának szórása
- ΔC : a mássalhangzós szakaszok időtartamának szórása
- %V: a magánhangzós szakaszok idejének összege, osztva az összes magánhangzós és mássalhangzós szakaszok idejével és szorozva százzal
- VarcoC: a ΔC artikulációs tempóra normalizált változata
- VarcoV: a ΔV artikulációs tempóra normalizált változata (a számítási módja megegyezik)
- rPVI (raw Pairwise Variability Index): az egymást követő szakaszok időtartamkülönbségeinek átlaga

$$rPVI = \left(\sum_{k=1}^{m-1} |d_k - d_{k+1}| \right) / (m-1),$$

ahol m a szakaszok száma, a d_k pedig a k -edik szakasz időtartama

- nPVI (normalized Pairwise Variability Index): az rPVI osztva az összes szakasz időtartamának összegével (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002; WHITE–MATTYS 2007), azaz

$$nPVI = 100 \times \left(\sum_{k=1}^{m-1} |(d_k - d_{k+1}) / ((d_k + d_{k+1}) / 2)| \right) / (m-1).$$

Az nPVI mérőszámot elsősorban a magánhangzós szakaszokra szokás alkalmazni (nPVI-V), ritkábban mássalhangzós szakaszok esetében is használják (nPVI-C). Nyelven belüli vizsgálatok szempontjából hasznos lehet az artikulációs tempótól függetlenül is megvizsgálni az egymást követő mássalhangzós szakaszok időtartamának variabilitását, ezért az nPVI-C mérőszámot is meghatároztam minden esetben. A %V mérőszám lényegében a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának arányára reagál. Ha ugyanazt a mérőszámot kiszámolnánk a mássalhangzós szakaszokra, akkor az egyenlő lenne a 100%-ból kivont %V értékével. Így nincs értelme külön kiszámolni vagy feltüntetni a %C értékét.

A különböző mérőszámokat egymás függvényében is vizsgáltam, mivel a magánhangzós és a mássalhangzós szakaszokra épülő mérőszámok együttes vizsgálata több információval szolgál, mint ha önmagukban vizsgálnánk, hiszen jellemzően ezek a mérőszámok kiegészítik egymást. Kisebb módosításokkal követtem a megszokott ábrázolási módokat (vö. RAMUS–NESPOR–MEHLER 1999). Mivel nem hasonlítottam össze különböző nyelveket, hanem elsősorban a nyelven belüli tényezők vizsgálata volt a cél, ezért a mássalhangzós és magánhangzós szakaszok mérőszámait párba állítottam. Az egymásnak megfeleltethető mérőszámokat ábrázoltam egymás függvényében. A ΔC mérőszámot a ΔV függvényében ábrázoltam, a VarcoC mérőszámot pedig a VarcoV függvényében, ahogy a szakirodalomban is található. Az nPVI mérőszámoknál szokás a mássalhangzós szakaszok esetében a nem normált alakot használni (rPVI-C), de ezen mérőszám nem független az artikulációs tempótól, amely befolyásolhatja az eredményeket. Ezért az nPVI-C-t ábrázoltam az nPVI-V függvényében. Mivel a %V mérőszám jellemzően másképpen viselkedett, mint a többi mérőszám, ezért ezt a mérőszámot teljesen külön is megvizsgáltam. A magyar nyelvre kapott eredményeinket összevettem más nyelvekre mért szakirodalmi beszédritmus-adatokkal is (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002), ahol a megszokott koordináta-rendszereket alkalmaztam (ΔC –%V és rPVI-C–nPVI-V). A statisztikai elemzések elvégzése előtt normalitásvizsgálatot végeztem. Mivel a beszédritmus-mérőszámok által kapott adatok nem tekinthetők normál eloszlásúnak (Shapiro–Wilk-próba alapján), ezért nemparaméteres próbákat alkalmaztam (Spearman-féle korrelációanalízis, Mann–Whitney-próba, Wilcoxon-próba). Az összefüggő mintákat párosított próbával vizsgáltam, ahol lehetőség nyílt véletlenszerű párosításra, ott is ezen próbákat alkalmaztam, független minták esetében pedig az ennek megfelelő Mann–Whitney-próbát. A statisztikai elemzéseket az SPSS 20.0 program segítségével végeztem. Több ábra elkészítéséhez a Gnuplot 4.6-os verzióját használtam.

4.1.2. Eredmények³

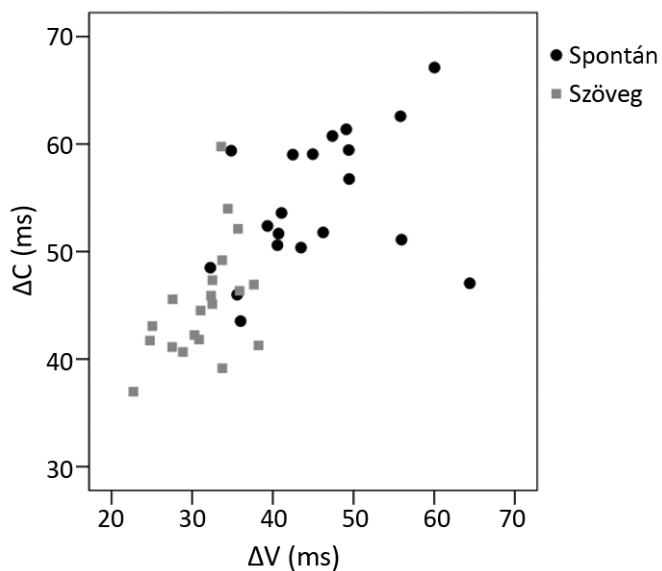
4.1.2.1. Különböző beszédmodok általános időzítési jellemzőinek összehasonlítása a beszédritmus-mérőszámok alapján

A különböző beszédmodok időbeli szerkezete egyértelmű eltérést mutatott a beszédritmus-mérőszámok alapján. Az az általános trend volt megfigyelhető, hogy a spontán beszédben nagyobb variabilitás jellemzi a magánhangzós és a mássalhangzós szakaszok időtartamát. A következőkben azt vesszük végig, hogy a különböző beszédmodok mely jellemzők, konkrét mérőszámok esetében különültek el egyértelműen egymástól időzítési mintázatuk alapján.

Először 2 percnyi összefüggő spontán monológ és 2 percnyi szövegfelolvasás időbeli struktúráját vizsgáltam meg, hogy általános képet kapjunk a beszédmodok közti eltérésekről. Egy beszélő teljes monológjára, illetve szövegfelolvasására kiszámoltam a különböző beszédritmus-mérőszámokat, és összehasonlítottam a különböző beszédmodokra kapott eredményeket. A magánhangzós szakaszok szórása (ΔV) az összes beszélő átlagát tekintve nagyobb volt spontán beszédben (átlag: 45,5 ms, szórás: 8,7 ms), mint felolvasásban (átlag: 31,5 ms, szórás: 4,3 ms). Az eltérés jelentősnek mondható a Wilcoxon-próba alapján ($Z = -3,920, p \leq 0,001$). A magánhangzós szakaszok szórásához (ΔV) hasonlóan a mássalhangzós szakaszok szórásának értéke (ΔC) majdnem minden beszélőnél (3 adatközlő kivételével) nagyobb volt spontán beszédben, mint szövegfelolvasásban. Ezzel összhangban az összes beszélő átlagát tekintve is nagyobb volt a ΔC értéke spontán beszédben (átlag: 54,6 ms, szórás: 6,3 ms), mint felolvasásban (átlag: 45,2 ms, szórás: 5,4 ms). A statisztikai próba itt is alátámasztotta, hogy a két beszédmodra kapott értékek szignifikánsan eltérnek egymástól ($Z = -3,920, p \leq 0,001$). A két beszédmod egyértelmű elkülönülését e két mérőszám alapján a 3. ábra szemlélteti.

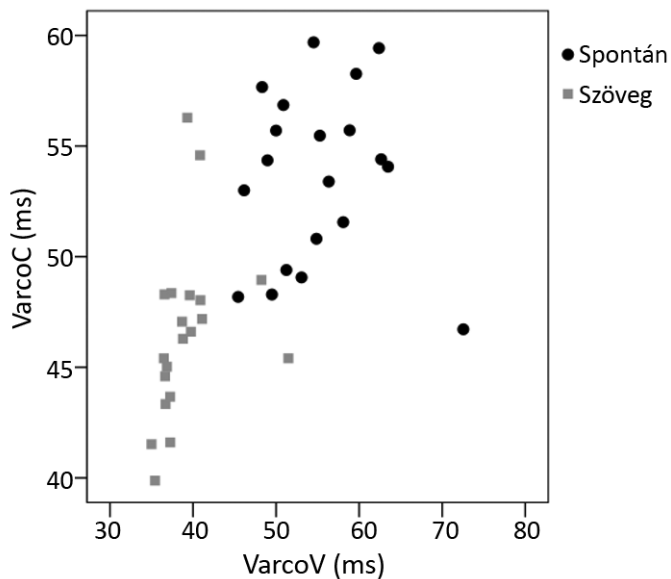
A ΔC és ΔV mérőszámokról azonban ismert, hogy az artikulációs tempó hatással van értékükre, ami a vizsgált anyagban is kimutatható volt a Spearman-féle korrelációs teszttel ($\rho \geq -0,562, p \leq 0,001$). A szövegfelolvasás és a spontán beszéd artikulációs tempója viszont nem tért el szisztematikusan egymástól (szövegfelolvasásban: átlag: 13,5 hang/s, szórás: 1,1 hang/s; spontán beszédben átlag: 13,1 hang/s, szórás: 1,1 hang/s). A statisztikai próba sem jelzett szignifikáns eltérést a különböző beszédmodokban ($Z = 1,680, p \geq 0,093$). A korrelációs összefüggés miatt azonban megvizsgáltam az artikulációs tempótól függetlennek tekinthető Varco mérőszámokkal is, hogy befolyásolják-e az eredményeket az artikulációs tempó ingadozásai.

³ A fejezet alapjául több korábbi tanulmányom (vö. KOHÁRI 2013a, 2013b, 2014a) szolgál.



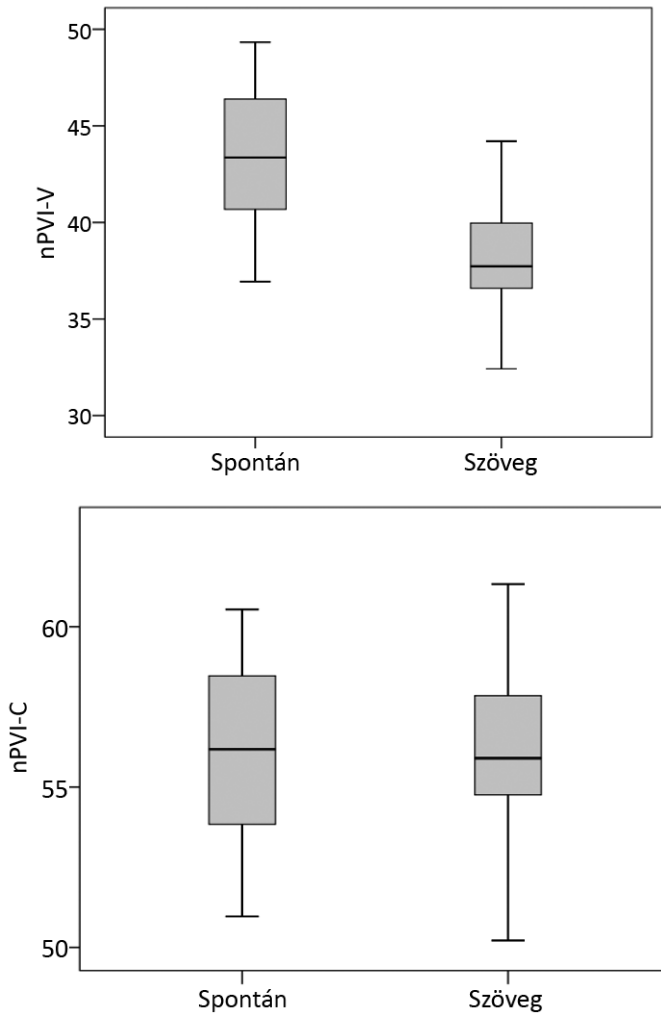
3. ábra

A 20 beszélő 2 perces spontán monológjának (fekete pont) és 2 percnyi szövegfelolvasásának (szürke négyzet) ΔV és ΔC mérőszámokra kapott értékei



4. ábra

A 20 beszélő 2 perces spontán monológjának (fekete pont) és 2 percnyi szövegfelolvasásának (szürke négyzet) VarcoV és VarcoC mérőszámokra kapott értékei



5. ábra

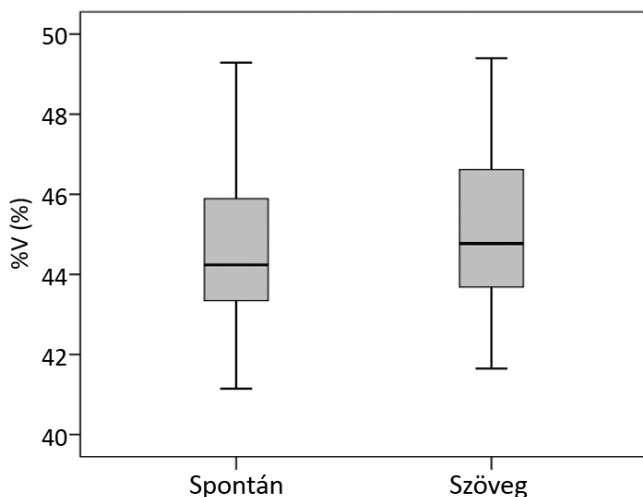
A 2 percnyi szövegfelolvasásra és spontán monológpra számolt nPVI-V (fent) és nPVI-C (lent) értékek

A Varco mérőszámok eredményei is azt mutatták, hogy a két beszédmód között egyértelmű eltérés van. Az egyes beszélők teljes monológjára, illetve szövegfelolvasására kiszámolt VarcoV-értékek minden adatközlő esetében nagyobbak voltak spontán beszédben, mint szövegfelolvasásban. A VarcoC esetében is csak egyetlen beszélőnél mértem magasabb értéket szövegfelolvasásban, mint spontán beszédben, de az értékek közel estek egymáshoz (szövegfelolvasás: 56,3 ms, spontán beszédben: 56,8 ms). A 4. ábra mutatja, hogyan különülnek el egymástól a két beszédmódra kapott értékek. A VarcoV átlaga 55,1 ms (szórás: 6,8 ms) volt spontán beszédben, és csak 39,2 ms (szórás: 4,1 ms) szövegfelolvasásban. A különbség statisztikailag is igazolódott

($Z = -3,920, p \leq 0,001$). A VarcoC átlaga spontán beszéd esetében 53,6 ms (szórás: 3,9 ms), szövegfelolvasás esetében pedig valamivel kisebb, 46,5 ms (szórás: 4,0 ms) volt. A statisztikai próba itt is megmutatta, hogy szignifikáns az eltérés a két beszédmód között ($Z = -3,845, p \leq 0,001$).

Az eddigi mérőszámok azt mutatták meg, hogy a teljes monológon belül, illetve szövegfelolvasásban mekkora kitérések lehetnek az időtartamokban. Az egymás mellett lévő hangok, így az egymást követő mássalhangzós és magánhangzós szakaszok időtartamai kapcsolatban vannak, valamilyen irányelv szerint rendeződnek, mintázatot alkotnak (például a megnyilatkozás végi lassulás nemcsak egy, hanem több hangzót is érint egyszerre). Ezért megvizsgáltam a különböző beszédmódokat az nPVI mérőszámokkal is (5. ábra), amelyek figyelembe veszik a magánhangzós, illetve a mássalhangzós szakaszok sorrendjét. Az nPVI-V mérőszám a többihez hasonlóan szisztematikusan minden beszélő esetében nagyobb értéket vett fel spontán beszédben, mint szövegfelolvasásban. Az nPVI-V átlaga spontán beszédben 43,6 (szórás: 3,6), szövegfelolvasásban pedig 38,1 (szórás: 2,9) volt. Az eltérés itt is jelentősnek mondható ($Z = -3,920, p \leq 0,001$).

Az nPVI-C esetében nem található ilyen összefüggés: a beszélőnként számolt értékek hol spontán beszédben, hol a szövegfelolvasásnál voltak nagyobbak. A különböző beszédmódokra kapott átlagok is közel vannak egymáshoz (spontán beszédben: átlag: 56,3, szórás: 3,5; szövegfelolvasásban: átlag: 56,0, szórás: 3,2). A statisztikai próba pedig nem mutatott szignifikáns eltérést ($Z = -0,411, p = 0,681$).



6. ábra

A %V értékei 2 pernyi spontán monológ és szövegfelolvasás esetében

A két beszédmód eltérést mutatott mind a magánhangzók, mind a mássalhangzók tekintetében. Felmerül a kérdés, hogy vajon a magánhangzós szakaszok időtartamának és a mássalhangzós szakaszok időtartamának aránya egymáshoz képest eltolódik-e, vagy valamilyen kiegyenlítő-dés jellemzi. Ez az összefüggés megragadható a %V mérőszámmal, amely a magánhangzós

szakaszok arányát adja meg a teljes beszédidőhöz képest (6. ábra). A %V átlagosan 44,6% volt (szórás: 2,1%) spontán beszédben, ez az arány valamivel magasabb volt a szövegfelolvasásokban (átlag: 45,1%, szórás: 2,3%). A mérőszám eltérései viszont a két beszédmódban nem voltak jelentősek a statisztikai próba alapján ($Z = 1,419, p = 0,156$).

4.1.2.2. A szünetek és megakadások kiküszöbölésével kapott értékek

A spontán beszéd és a szövegfelolvasás között eltérés mutatkozott a szünettartásban és a hezitálások, kitöltött szünetek megjelenésében. A szünetek és a hezitálások előtti pozícióban a szótagok, a hangok időtartama megnyúlhat, ezért felmerül a kérdés, hogy vajon a gyakoribb szünettartás okozza-e a két szövegtípus beszédritmus-mérőszámokkal mérhető időzítésbeli eltéréseit.

3. táblázat

*Az összes szünet és az összes megakadás száma a két vizsgált beszédmódban
(A félkövér betűtípus a szövegfelolvasásban előforduló nagyobb gyakoriságot jelöli.)*

A beszélők	A szünetek száma spontán beszédben	A szünetek száma szövegfelolvasásban	A megakadások száma spontán beszédben	A megakadások száma szövegfelolvasásban
F1	43	32	8	9
F2	64	35	30	3
F3	51	34	22	3
F4	53	38	16	1
F5	48	22	15	2
F6	32	36	28	4
F7	53	34	9	2
F8	72	67	26	3
F9	34	34	43	0
F10	59	42	12	3
N1	52	33	17	6
N2	52	40	15	0
N3	69	52	19	6
N4	71	38	25	7
N5	52	49	15	1
N6	44	38	31	8
N7	59	32	20	8
N8	57	54	13	1
N9	43	61	9	3
N10	53	56	14	2

A spontán beszédben a beszélők gyakrabban tartottak szünetet, gyakrabban hezitáltak vagy akadtak el, mint a szövegfelolvasásban. Beszélőkre lebontva a különbség a szünettartásban nem következetes a két beszédmód között (lásd 3. táblázat), de jellemzően a spontán beszédben gyakoribb. A hezitálások és a jelölt megakadások száma egyértelműen több a spontán beszédben, mint a szövegfelolvasásban. Ezért megvizsgáltam, hogy a szünetek és a megakadások környékének kihagyása esetén mutatkozik-e eltérés a két beszédmód között a beszédritmus-mérőszámok alapján.

Elsőként azt vizsgáltam meg, hogy a szüneteket és a megakadásokat megelőző egy, két vagy három magánhangzós/mássalhangzós szakasz kihagyása esetén is érvényesül-e a két beszédmód közti különbség. Az éppen vizsgált mérőszám alapegységének megfelelően figyelmen kívül hagytam a szünet vagy megakadás előtti utolsó egy, kettő vagy három szakaszt, tehát a ΔV és a VarcoV esetében az utolsó egy, kettő vagy három V-s szakaszt, a ΔC és a VarcoC esetében az utolsó egy, kettő vagy három C-s szakaszt. Mivel a %V a mássalhangzók és a magánhangzók időtartamának arányát méri, ezért ebben az esetben az utolsó egy, kettő vagy három szakaszt hagytam figyelmen kívül, függetlenül a szakasz típusától. Eltekintettem a szünetek hosszának, funkciójának vizsgálatától, mivel ha az adott szünet nyújtó hatása a megelőző beszédhangokon, szakaszokon elhanyagolható, akkor kihagyása nem változtat érdemben az időtartamok variabilitását mérő beszédritmusértékeken. Fordított esetben viszont a nagyobb időtartamértékek nagymértékben befolyásolhatnák az eredményeket.

Az eredmények azt mutatták, hogy a szünet és a megakadások előtti utolsó három szakasz kihagyása esetén is mérhető a különbség a két beszédmód között. A ΔV esetében akár az utolsó egy vagy kettő vagy három egységet hagyjuk figyelmen kívül a szünetek és a megakadások előtt, az eltérés jelentősnek mondható a Wilcoxon-próba alapján ($Z \leq -3,584$, $p < 0,001$). A mássalhangzós szakaszok szórása (ΔC) – a magánhangzós szakaszok szórásához (ΔV) hasonlóan – statisztikailag alátámaszthatóan nagyobb volt spontán beszédben, mint felolvasásban, függetlenül a figyelmen kívül hagyott szakaszok számától ($Z \leq -2,651$, $p < 0,008$). A VarcoV és VarcoC mérőszámok esetében a statisztikai próba szintén alátámasztotta, hogy a két beszédmódra kapott értékek szignifikánsan eltérnek egymástól, akár mennyi egységet hagyunk figyelmen kívül a szünetek, illetve a megakadások előtt ($Z \leq -2,875$, $p \leq 0,04$). Ugyanakkor megfigyelhető az a tendencia, főként a Varco mérőszám értékeinél, hogy a két beszédmódra kapott értékek nemcsak közelednek egymáshoz, hanem keverednek is. Ha a beszélőkre külön-külön lebontva nézzük, akkor is többnyire érvényesül az a tendencia, hogy spontán beszéd esetében nagyobbak a beszédritmusértékek, mint szövegfelolvasásnál. A VarcoV esetében akár az utolsó egy, kettő vagy három szakaszt hagyjuk el, minden beszélőnél következetesen látható ez az összefüggés. A többi mérőszám (ΔV , ΔC , VarcoC) esetében a figyelmen kívül hagyott egységek számától függően kettő vagy három, de maximum négy beszélőnél fordult elő, hogy a szövegfelolvasás értékei magasabbnak bizonyultak a spontán beszédre számolt értékeknél.

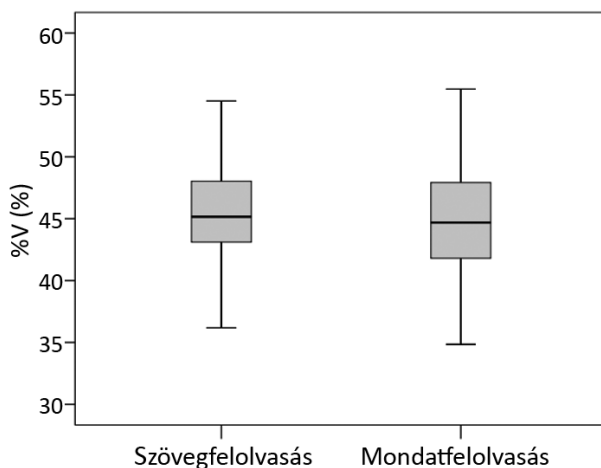
A szünetek és a megakadások hatással lehetnek a követő beszédhangok időtartamára is. Ezért megvizsgáltam, hogy vajon a szünetet vagy megakadást megelőző egy, kettő vagy három szakasz és az azt követő egy, kettő vagy három szakasz figyelmen kívül hagyásával együttesen van-e különbség a spontán beszéd és a felolvasás között. A statisztikai próbák alátámasztották, hogy így is szignifikáns az eltérés a spontán beszéd és a szövegfelolvasás között, akárhány egységet hagyunk figyelmen kívül (ΔV : $Z \leq -3,323$, $p \leq 0,001$; ΔC : $Z \leq -2,352$, $p \leq 0,019$; VarcoV: $Z \leq -3,285$, $p \leq 0,001$; VarcoC $Z \leq -2,875$, $p \leq 0,004$). Ha beszélőkre külön-külön lebontva nézzük, akkor is többnyire érvényesül az a tendencia, hogy spontán beszéd esetében nagyobbak a beszédritmusértékek, mint szövegfelolvasásnál, de itt sem következetesen mindenkire érvényesül. A figyelmen kívül hagyott egységek számától függően egy, kettő vagy három, de maximum négy beszélőnél fordult elő, hogy a szövegfelolvasás értékei magasabbnak bizonyultak a spontán beszédre számolt értékektől.

Mivel a PVI mérőszámok a szakaszok sorrendjét is figyelembe veszik, az egymást követő szakaszok különbségén alapulnak, így e mérőszámok esetében nem volt értelme megvizsgálni a szünetek és a megakadások befolyását ezzel a módszerrel. Szintén eltekintettem a %V mérőszám vizsgálatától, mivel ezen mérőszám egyszerre vizsgálja a magánhangzós szakaszokat és a mássalhangzós szakaszokat, és az ismertetett eljárás alkalmazása problematikus lenne a fonológiai felépítettség miatt.

A mondatfelolvasás és más beszédmodok

Az eddigiekben hosszabb, egybefüggő szövegeket, monológokat vizsgáltunk. Felmerül a kérdés, hogy vajon van-e különbség az egybefüggő szövegek és a különálló mondatok időzítési tulajdonságai között. Az esetleges eltérések a két beszédmód között nagyobb kiterjedésű, átfogó időbeli struktúrára utalhatnak a szövegekben, ezért összevettem a két beszédmód beszédritmusát. A vizsgálatot kiterjesztettem a spontán beszédre is, hogy több oldalról megvizsgálhassuk a spontán monológok és a felolvasások közti különbséget. Az összehasonlításhoz a legalább 10 szótagú tagmondatokat választottam alapegységnek, amely minden beszédmódban viszonylag könnyen detektálható (bővebben lásd a 3.2-es módszertani fejezetben). Ezen tagmondatnyi egységekre számoltam ki a beszédritmus-mérőszámokat. A felolvasott mondatok közül csak azok kerültek bele a vizsgálatba, amelyek önmagukban egy tagmondatnyi hosszúak, tehát nem tartalmaztak alá- vagy mellérendelést (bővebben lásd a 4.1.1-es módszertani fejezetben).

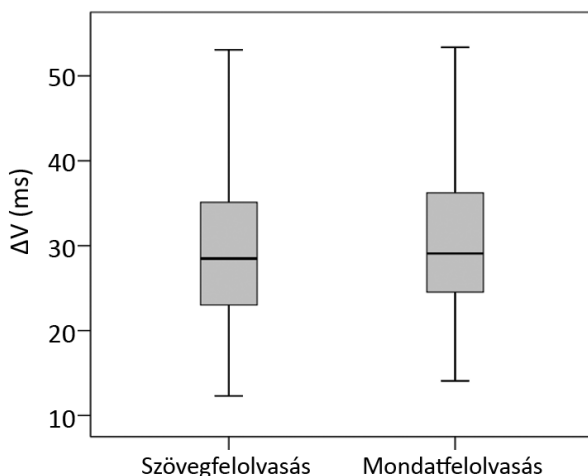
A szöveg tagmondatai és a különálló mondatok tagmondatai több mérőszám esetében mutattak szignifikáns eltérést. Az egyik ilyen mérőszám a %V volt (7. ábra), amely esetében a különbség statisztikailag alátámaszthatóan is megjelent (Wilcoxon-próba: $Z = -4,726$, $p < 0,001$). A szövegfelolvasás esetében a tagmondatok %V-értéke átlagosan 45,6% volt (szórás: 3,6%), ugyanakkor a mondatok felolvasásánál ez az arány kisebbnek bizonyult (átlag: 44,8%, szórás: 4,1%). A magánhangzós szakaszok aránya a teljes beszédidőhöz képest tehát a szövegfelolvasásban nagyobb a mondatfelolvasáshoz képest.



7. ábra

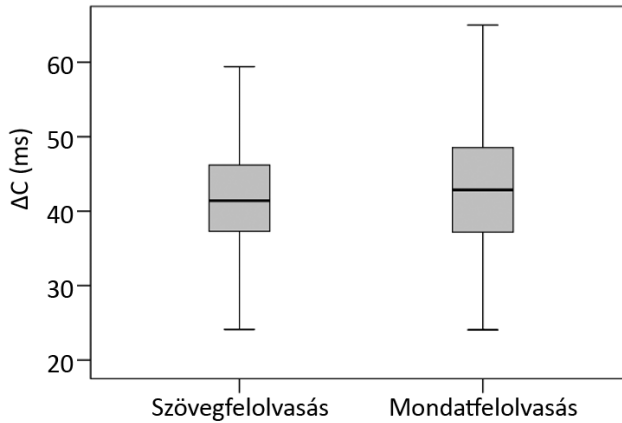
A %V értékei a mondatfelolvasás és a szövegfelolvasás tagmondatai alapján

Egy másik mérőszám, amely esetében a statisztikai próba szignifikáns különbséget mutatott ($Z = -2,250, p = 0,024$), a ΔV mérőszám volt (8. ábra). A ΔC mérőszám (9. ábra) esetében szintén statisztikailag is igazolódott a különbség a két beszédmód között ($Z = -2,652, p = 0,008$). Mindkét mérőszámnak az értéke átlagosan magasabb volt mondatfelolvasásban, mint szövegfelolvasásban (lásd 4. táblázat). Mind a magánhangzós szakaszok, mind a mássalhangzós szakaszok esetében tehát jellemzően szélsőségesebb időtartamértékek jelentek meg mondatfelolvasásban, mint szövegfelolvasásban.



8. ábra

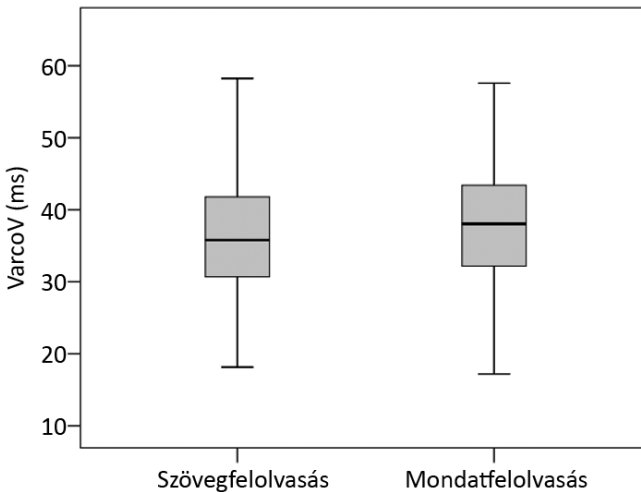
A ΔV értékei a mondatfelolvasás és szövegfelolvasás tagmondatai alapján



9. ábra

A ΔC értékei a mondatfelolvasás és szövegfelolvasás tagmondatai alapján

A VarcoV mérőszám értéke átlagosan szintén magasabbnak mutatkozott mondatfelolvasásban, mint szövegfelolvasásban (10. ábra). A statisztikai próba alátámasztotta a két beszédmód közti eltérést ($Z = -3,297, p = 0,001$).



10. ábra

A VarcoV értékei a mondatfelolvasás és szövegfelolvasás tagmondatai alapján

A többi mérőszám esetében viszont a statisztikai próbák nem mutattak valódi eltérést a két beszédmód között ($Z \geq -1,399, p \geq 0,162$). Az átlagok nem függték össze rendszerszerűen a két beszédmóddal, a VarcoC és nPVI-V mérőszámok átlagos értéke szövegfelolvasásban, a nPVI-C átlagos értéke pedig a különálló mondatok esetében volt magasabb (lásd 4. táblázat).

Az artikulációs tempó hasonló volt a szövegfelolvasásban (átlag: 13,5 hang/s, szórás: 1,3 hang/s) és a mondatfelolvasásban (átlag: 13,4 hang/s, szórás: 1,3 hang/s). A köztük lévő különbség statisztikailag nem támasztható alá ($Z = -1,193, p = 0,233$). Az artikulációs tempó tehát közvetlenül nem befolyásolhatta a kapott eredményeket.

4. táblázat

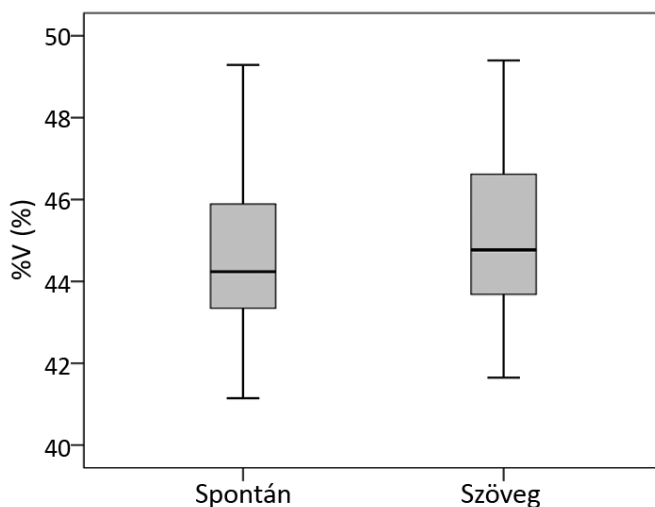
A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek (A félkövér a legmagasabb értéket jelöli.)

	Spontán monológ		Szövegfelolvasás		Különálló mondatok	
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
%V	44,9	4,7	45,6	3,6	44,8	4,1
ΔV	41,2	23,5	30,3	10,8	30,8	9,0
ΔC	49,6	15,5	42,7	8,3	43,6	8,8
VarcoV	47,7	17,7	37,0	9,5	38,3	8,4
VarcoC	49,9	11,4	44,8	7,0	44,5	6,3
nPVI-V	42,7	11,0	38,3	8,0	37,2	7,7
nPVI-C	56,2	12,3	54,7	8,2	55,4	9,6

A spontán beszéd és a különálló mondatok tagmondatait megvizsgálva minden mérőszám átlagosan magasabb értéket vett fel, mint különálló mondatok esetében (4. táblázat). A statisztikai tesztek majdnem minden mérőszámra szignifikáns eltérést mutattak ($Z \leq -5,928, p \leq 0,001$). Ez alól két kivétel található. Az egyik az nPVI-C mérőszám, amely ugyan a többi mérőszámhoz hasonlóan magasabbnak bizonyult átlagosan spontán beszédben, mint mondatfelolvasásban, de az eltérés nem volt jelentős ($Z = -0,528, p = 0,598$). Hasonló a helyzet a %V mérőszám esetében, ahol a statisztikai próba nem igazolta, hogy a két beszédmód lényegesen eltérne ezen időzítéssel tulajdonság szempontjából ($Z = -0,358, p = 0,720$). Az artikulációs tempó különbségei feltehetően nem befolyásolhatták az eredményeket. Az artikulációs tempó valamivel kisebb volt ugyan a spontán beszédben (átlag: 13,3 hang/s, szórás: 1,9 hang/s), mint a mondatfelolvasások esetében (átlag: 13,4 hang/s, szórás: 1,3 hang/s), de statisztikailag ez a különbség nem bizonyult jelentősnek ($Z = -1,456, p = 0,145$).

Legvégül a spontán monológban vizsgált tagmondatok és a szövegfelolvasásban lévő tagmondatok mérőszámértékeit hasonlítottam össze. Az eredmények hasonlóak ahhoz, amit a 2 percnyi spontán monológ és a szövegfelolvasás összehasonlításakor kaptam. A spontán monológot – a %V kivételével – minden mérőszám esetében átlagosan magasabb értékek jellemezték, mint a szövegfelolvasást (4. táblázat). Az eredmények szignifikánsnak mondhatók ($Z \leq -6,410, p \leq 0,001$), csak az nPVI-C mérőszám képez kivételt ez alól ($Z = -1,898, p = 0,058$).

A %V mérőszám – a 2 percnyi felolvasások és spontán monológok vizsgálatához hasonlóan – a szövegfelolvasásban átlagosan valamivel magasabbnak bizonyult, mint a spontán monológban (11. ábra). A két beszédmód szignifikánsan különbözött egymástól ($Z = -3,631$, $p < 0,001$), tehát ebben eltér a tagmondatok vizsgálata a korábbiaktól. Az artikulációs tempó átlagosan 13,3 hang/s (szórás: 1,9 hang/s) volt spontán beszédben, és átlagosan valamivel magasabb artikulációs tempót, 13,5 hang/s-ot (szórás: 1,3 hang/s) mértem szövegfelolvasásokban. A %V mérőszám ugyanakkor negatívan és gyengén korrelált az artikulációs tempóval ($\rho = -0,382$, $p = 0,001$). A két beszédmód tagmondatainak artikulációs tempójában nem volt lényeges eltérés a statisztikai próba alapján ($Z = -1,893$, $p = 0,058$), így ez nem befolyásolhatta az adatokat.



11. ábra

A %V értékei a spontán monológ és a szövegfelolvasás tagmondatai alapján

A statisztikai összehasonlítást az eddigiekben úgy végeztem el, hogy több beszélő több tagmondatát vizsgáltam meg egyszerre. Felmerülhet, hogy az egyes beszélőkön belüli változottság torzíthatja az eredményeket, ezért beszélőnként vettem a mérőszámértékek mediánját (figyelembe véve a nem normál eloszlást), és ezen adatokat hasonlítottam össze párosított próbával. Ezen eredmények egyetlen kivétellel alátámasztották, hogy ahol az eddigiekben szignifikáns eltérést találtam, ott ezen mérések is szignifikáns eltérést mutattak ($Z \leq -2,016$, $p \leq 0,044$). Ahol nem találtam eltérést a beszédmódok között az összes tagmondat alapján, ott a mediánok alapján sem mutatkozott eltérés ($Z \geq -1,419$, $p \geq 0,156$). A kivétel a %V mérőszám volt a mondatfelolvasás és a szövegfelolvasás összehasonlításában, ebben az esetben ugyanis nem találtam szignifikáns eltérést a mediánok alapján ($Z = -1,531$, $p = 0,126$). Ez esetben tehát nem tekinthető lényegesnek a különbség a mondat- és a szövegfelolvasás között.

A különböző beszédmódok időzítésének beszélőnkénti eltérései

A spontán beszéd és a felolvasások összehasonlításában eltérést jelent az anyagban, hogy míg a mondat- és szövegfelolvasásokban a beszélők ugyanazt olvasták fel, addig a spontán beszédben változatos fonológiai felépítettségű szövegek hangoztak el. Felmerül a kérdés, hogy vajon ez mennyiben befolyásolhatta a spontán beszéd vs. mondatfelolvasás és a spontán beszéd vs. szövegfelolvasás tagmondatainak összehasonlítása alapján kapott eredményeket. Ennek megvizsgálásához az egyes beszélők mintáin belül vettem össze a különböző beszédmódokat, hogy a spontán beszédben a fonológiai felépítettség miatt megjelenő esetleges nagyobb szórás hatása kizárható legyen. A beszélőnkénti elemzés lehetővé teszi, hogy beszédmódonként 20 teljesen különböző fonológiai felépítettségű tagmondat⁴ kerüljön összehasonlításra.

Elsősorban a spontán beszéd és a mondatfelolvasás közti különbséget vizsgáltam meg közelebbről, mivel ezen két beszédmód összehasonlítására csak a tagmondatok alapján volt lehetőség. Továbbá összehasonlítottam a spontán beszéd és a szövegfelolvasás beszédritmusát is, hogy a korábban talált eredmények mennyire tekinthetők következetesnek egy beszélőn belül, és mennyiben jelentős az eltérés a két beszédmód között.

Először a spontán beszédet és a mondatfelolvasást vettem össze. A magánhangzós szakaszok időtartamára épülő mérőszámok (ΔV , VarcoV, nPVI-V) átlagukat tekintve szisztematikusan magasabbak voltak spontán beszédben, mint mondatfelolvasásban, majdnem minden beszélő esetében (lásd 5., 7. és 9. táblázat), eltekintve egyetlen esettől (N5-ös beszélő nPVI-V-átlaga). Ez összhangban van a beszélőket együttesen vizsgáló korábbi eredményekkel. A statisztikai próbák azonban nem igazolták, hogy ezek az eltérések a legtöbb beszélő esetében jelentősek lennének. A ΔV mérőszám esetében 8 beszélőnél (F3, F4, F5, F7, N4, N7, N8 és N10) volt igazolható, hogy a spontán beszédre és a mondatfelolvasásra kapott értékek lényegesen eltérnek egymástól (Wilcoxon-próba: $Z \leq -1,972$, $p \leq 0,049$). A VarcoV mérőszám esetében szintén a beszélők majdnem felénél (F3, F4, F5, F7, N3, N6, N7 és N10) volt igazolható, hogy a spontán beszéd és a mondatfelolvasás beszédritmusa különbözik egymástól ($Z \leq -2,091$, $p \leq 0,037$). Az nPVI-V esetében csak 5 beszélőnél (F4, F5, F6, F7 és N8) mondható jelentősnek a különbség a spontán beszéd és a mondatfelolvasások között ($Z \leq -1,979$, $p \leq 0,048$).

A mássalhangzós szakaszok időtartamára épülő mérőszámok átlagukat tekintve kevésbé mutatnak olyan rendszerszerű képet, mint a magánhangzós szakaszok időtartamára épülő mérőszámok. A ΔC mérőszám esetében négy beszélőnél (F6, N2, N6, N9) is a mondatfelolvasásnál mérhető magasabb érték a spontán beszéddel szemben. Igaz, ezen négy beszélő egyikénél sem szignifikáns az eltérés a két beszédmód között ($Z \geq -1,493$, $p \geq 0,135$). A beszélők kevesebb mint felénél (F3, F8, F10, N1, N4, N7, N10) viszont statisztikailag alátámaszthatóan spontán beszédben figyelhetők meg magasabb értékek ($Z \leq -2,165$, $p \leq 0,030$). A VarcoC

⁴ A tagmondatok legalább 10 szótagból álltak (részletesen lásd 3.1. fejezet).

mérőszám esetében csak két beszélőnél (F6, N3) fordult elő, hogy a mondatfelolvasásnál magasabb átlagos érték volt, mint spontán beszédben. Ezek egyike sem volt szignifikáns eltérés (F6: $Z = -0,971$, $p = 0,332$; N3: $Z = -0,336$, $p = 0,737$). A többi beszélőnél átlagosan a spontán beszédben mért értékek voltak magasabbak, de csak egy részüknél (F2, F3, F8, F10, N1, N2, N4, N5, N7) volt szignifikáns az eltérés a két beszédmód között ($Z \leq -2,091$, $p \leq 0,037$). Az nPVI-C mérőszám esetében hol spontán beszédben, hol mondatfelolvasásban fordultak elő átlagosan magasabb értékek (lásd 10. táblázat). Mindössze három beszélő (F3, N4, N6) esetében volt szignifikáns az eltérés a két beszédmód között (N4: $Z \leq -1,979$, $p \leq 0,048$). A %V mérőszám esetében szintén változó volt, hogy a spontán beszédben mért értékek vagy a mondatfelolvasásban mért értékek voltak magasabbak (lásd 11. táblázat). A különbségeket azonban nem tekinthetjük jelentősnek a statisztikai próba alapján, kivéve egy beszélőt, akinél a mondatfelolvasásban mértem magasabb értékeket (N10: $Z = 103,0$, $p = 0,009$), és egy beszélőt, akinél a mondatfelolvasásban tapasztaltam magasabb átlagot (N10: $Z = -2,315$, $p = 0,021$).

A tagmondatok alapján a spontán beszédet a szövegfelolvasással is összevettem. Habár egy korábbi fejezetben már sikerült kimutatni, hogy a beszélők teljes szövegfelolvasás esetében magasabb mérőszámértékeket produkálnak spontán beszédben, mint szövegfelolvasásban, az eltérések jelentősége beszélőn belül még nem került tárgyalásra. Ha a mérőszámok átlagait tekintjük, összecsengenek a korábban talált eredményekkel, azaz a spontán beszédben nagyobb időtartambeli szórások, variabilitások jellemzők a beszélőkön belül, mint szövegfelolvasásban. A statisztikai próbák azonban itt sem mutattak mindenhol szignifikáns eltérést beszélőn belül, hasonlóan a mondatfelolvasás és a spontán beszéd összevetésekor találtakhoz.

A magánhangzós szakaszok időtartamára épülő mérőszámok (ΔV , VarcoV, nPVI-V) átlagukat tekintve szisztematikusan magasabbak voltak spontán beszédben, mint szövegfelolvasásban majdnem minden beszélő esetében (lásd 5., 7. és 9. táblázat) eltekintve két esettől (F6-os beszélő ΔV -átlaga és N5-ös beszélő nPVI-V-átlaga). Habár az átlagokban való eltérés következetesnek tűnik a beszélőknél, a beszédmódok közti különbség nem minden beszélőnél jelentős. A ΔV mérőszám esetében mindössze 9 beszélőnél (F3, F4, F5, F8, N2, N3, N4, N7 és N8) volt igazolható, hogy a spontán beszédre és szövegfelolvasásra kapott értékek lényegesen eltérnek egymástól ($Z \leq -2,053$, $p \leq 0,040$). A VarcoV mérőszám esetében tizenhárom beszélőnél (F1, F3, F4, F5, F8, N1, N2, N4, N5, N6, N7, N8 és N10) volt igazolható, hogy a spontán beszéd és a szövegfelolvasás beszédritmusa különbözik egymástól ($Z \leq -1,979$, $p \leq 0,048$). Az nPVI-V esetében mindössze csak négy beszélőnél (F7, F9, N1 és N4) mondható jelentősnek a különbség a spontán beszéd és a szövegfelolvasások között ($Z \leq -2,016$, $p \leq 0,044$).

A mássalhangzós szakaszokra épülő mérőszámok átlagukat tekintve nem mutattak olyan rendszerszerű különbséget a spontán beszéd és a szövegfelolvasás között, mint a magánhangzós szakaszokra épülő mérőszámok. A ΔC mérőszám esetében négy beszélőnél (F6, N2, N6, N9) is a szövegfelolvasásnál mérhető magasabb érték a spontán beszéddel szemben, hasonlóan

a mondatfelolvasásban megmutatkozott eredményekhez. A VarcoC mérőszám esetében csak két beszélőnél (F6, N9) fordult elő, hogy magasabb átlagos érték volt a mondatfelolvasásban, mint a spontán beszédben. Egyik esetben sem szignifikáns az eltérés a két beszédmód között ($Z \geq -1,120, p \geq 0,263$). A többi beszélőnél átlagosan a spontán beszédben mért értékek voltak magasabbak, de csak egy részükénél volt szignifikáns az eltérés a két beszédmód között (ΔC esetében: F2, F3, F7, F8, N1, N4, N5, N7, N10 $Z \leq -1,979, p \leq 0,048$; VarcoC esetében: F2, F3, F8, F10, N1, N4, N5, N7 $Z \leq -2,651, p \leq 0,008$). Az nPVI-C mérőszám esetében hol spontán beszédben, hol szövegfelolvasásban fordultak elő átlagosan magasabb értékek beszélőtől függően (lásd 10. táblázat). Öt beszélőnél (F1, F3, N1, N4, N7) volt igazolható statisztikailag is, hogy magasabb mérőszámértékek jellemzők a spontán beszédben, mint szövegfelolvasásban ($Z \leq -1,979, p \leq 0,048$). Egy beszélőnél (N6) pedig ennek az ellenkezőjét támasztotta alá a statisztikai próba ($Z = -2,091, p = 0,037$). A %V mérőszám esetében szintén változó volt, hogy a spontán beszédben mért értékek vagy a szövegfelolvasásban mért értékek voltak magasabbak (lásd 11. táblázat). Az F1, F4 és N5 beszélőknél jelentős volt az eltérés, és az F4-es beszélő kivételével a szövegfelolvasásban voltak magasabb értékek, mint spontán beszédben ($Z \leq -2,091, p \leq 0,037$).

A tagmondatok alapján a szövegfelolvasást összevettem a mondatfelolvasással is. Beszélőnként változó, hogy mondat- vagy szövegfelolvasásban magasabbak az értékek. Egyik mérőszám esetében sem találtam szisztematikus eltérést a két beszédmód között (5–11. táblázat). A legtöbb beszélő (F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, N1, N2, N3, N6, N7) esetében egyik mérőszám sem tért el szignifikánsan a két beszédmódban. A ΔV mérőszám esetében mindössze egy beszélőnél (N4) találtam jelentős eltérést ($Z = -2,651, p = 0,008$). Az N4-es beszélő mondatfelolvasásban átlagos nagyobb értékeket produkált, mint szövegfelolvasásban. Szintén a mondatfelolvasás átlagos értéke volt nagyobb a szövegfelolvasásnál a VarcoV mérőszámánál az N4-es és N5-ös beszélő esetében, amely különbségeket a statisztikai próbák is alátámasztották (N4: $Z = -2,912, p = 0,004$; N5: $Z = -2,352, p = 0,019$). A %V mérőszám esetében öt beszélőnél (N5, N6, N8, N9, N10) is jelentős eltéréseket találtam a mondatfelolvasás és a szövegfelolvasás között ($Z \leq -2,277, p \leq 0,023$). Mindegyik adatközlőnél a szövegfelolvasásban tapasztaltam átlagosan magasabb értékeket.

5. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként a ΔV esetében (A félkövér a legmagasabb értéket jelöli.)

		ΔV (ms)					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	29,8	10,0	27,8	8,6	25,8	4,4
	F2	32,5	16,2	28,7	11,5	27,6	7,9
	F3	41,8	20,0	24,8	5,8	27,9	6,3
	F4	55,1	32,6	32,3	22,3	28,4	7,9
	F5	38,5	18,5	26,4	5,6	29,0	5,6
	F6	29,8	12,9	33,4	13,4	29,7	7,6
	F7	36,1	27,0	24,2	6,1	24,6	5,8
	F8	38,9	16,2	29,5	8,1	32,7	8,6
	F9	38,6	21,2	31,7	9,7	35,0	8,2
	F10	40,4	22,3	33,3	12,0	30,9	7,3
	N1	26,7	6,8	22,9	6,3	25,9	5,3
	N2	48,3	19,9	38,0	11,4	39,7	8,8
	N3	44,2	30,2	26,4	10,1	28,7	8,2
	N4	47,1	16,7	31,6	6,4	36,9	8,7
	N5	36,9	17,0	30,7	7,9	33,4	9,7
	N6	40,2	14,2	35,5	11,7	35,0	8,5
	N7	41,3	18,3	32,2	10,3	30,2	8,7
	N8	54,2	27,1	32,6	9,0	33,8	8,5
	N9	33,6	15,9	32,1	8,9	29,6	7,3
	N10	35,9	13,3	31,8	10,3	27,7	7,1

6. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként a ΔC esetében (A félkövér a legmagasabb értéket jelöli.)

		ΔC (ms)					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	43,4	10,6	39,1	6,1	39,9	7,5
	F2	46,9	9,4	40,7	7,5	42,6	6,8
	F3	58,0	17,9	40,3	5,2	42,9	8,6
	F4	43,5	9,5	40,4	6,2	41,3	10,1
	F5	45,2	15,0	39,6	6,0	39,0	8,4
	F6	37,7	10,8	42,8	10,2	42,4	9,0
	F7	49,5	9,2	41,6	5,8	45,1	7,9
	F8	57,4	15,5	40,9	6,4	43,0	6,3
	F9	52,9	19,5	44,7	5,5	46,5	7,8
	F10	57,2	18,6	43,7	7,2	44,0	7,1
	N1	43,6	6,8	35,5	5,1	36,9	7,2
	N2	44,4	11,1	45,1	6,2	48,1	9,6
	N3	52,9	20,3	42,8	7,8	50,3	10,1
	N4	56,0	15,1	44,3	6,3	42,0	7,4
	N5	49,9	17,1	40,6	5,2	43,4	10,4
	N6	47,0	9,1	48,8	8,1	49,9	7,7
	N7	59,3	16,9	45,1	14,0	44,2	9,5
	N8	50,6	20,0	44,1	10,2	46,2	7,7
	N9	43,8	12,8	45,6	5,7	44,6	9,4
	N10	48,5	11,9	38,0	5,4	39,2	5,7

7. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként a VarcoV esetében (A félkövér a legmagasabb értéket jelöli.)

		VarcoV					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	41,8	10,5	36,0	7,8	35,9	5,5
	F2	40,4	13,9	33,6	12,1	32,7	7,6
	F3	50,4	18,4	36,5	6,3	39,1	7,6
	F4	58,9	25,2	41,8	20,5	39,4	8,8
	F5	51,3	18,2	39,3	7,9	41,4	6,6
	F6	43,5	14,0	42,2	11,7	40,0	7,8
	F7	47,4	26,9	35,1	7,7	34,3	6,5
	F8	45,2	14,4	37,2	7,4	39,1	7,8
	F9	46,9	14,2	39,4	9,1	43,2	8,7
	F10	43,4	18,8	37,7	10,5	35,2	6,4
	N1	42,0	7,9	34,3	7,1	39,1	7,4
	N2	50,2	14,9	38,2	8,0	40,3	7,6
	N3	44,6	18,8	33,8	9,6	33,6	8,1
	N4	51,3	15,6	36,0	6,2	42,7	8,0
	N5	49,2	14,4	36,8	6,1	43,2	8,3
	N6	44,7	11,4	35,4	8,5	36,0	7,4
	N7	48,0	15,1	36,9	8,8	35,9	7,9
	N8	53,2	18,0	38,4	6,4	41,6	7,2
	N9	41,9	15,7	36,0	6,8	38,3	7,3
	N10	38,9	11,3	32,9	8,4	31,7	7,0

8. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként a VarcoC esetében (A félkövér a legmagasabb értéket jelöli.)

		VarcoC					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	48,7	7,8	45,5	6,4	46,0	5,2
	F2	51,5	7,7	43,9	6,3	44,6	5,9
	F3	55,9	14,5	44,7	5,7	46,8	7,5
	F4	45,3	7,8	45,1	5,7	43,9	8,6
	F5	51,0	14,5	46,6	5,3	44,9	5,3
	F6	44,0	8,6	45,9	8,3	45,8	6,8
	F7	49,6	7,3	44,6	6,2	46,9	6,0
	F8	54,9	12,4	45,6	6,3	44,0	5,0
	F9	52,3	14,7	46,3	5,4	46,7	4,9
	F10	53,9	11,4	43,0	6,5	42,9	5,5
	N1	50,6	8,1	40,4	5,7	41,6	6,2
	N2	47,3	10,2	47,0	5,9	47,3	6,2
	N3	46,2	12,7	41,2	5,9	46,3	8,7
	N4	57,8	10,8	47,9	5,1	45,3	4,9
	N5	49,7	11,0	42,7	5,3	43,8	6,8
	N6	47,8	7,8	46,7	5,7	45,6	4,2
	N7	55,4	11,8	42,8	8,8	41,3	5,8
	N8	46,9	9,7	45,8	9,5	45,4	6,0
	N9	41,8	9,4	42,2	4,7	40,9	5,6
	N10	44,3	10,0	39,6	5,0	40,3	5,8

9. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként az nPVI-V esetében (A félkövér a legmagasabb értéket jelöli.)

		nPVI-V					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	40,6	7,8	37,4	8,5	35,5	6,6
	F2	39,2	9,7	33,7	8,6	33,1	7,3
	F3	42,7	9,3	38,6	6,7	39,4	6,7
	F4	47,1	14,0	40,0	7,1	35,8	4,6
	F5	48,6	9,2	42,8	6,5	39,2	6,5
	F6	42,5	11,5	40,0	9,4	35,9	6,6
	F7	43,5	8,6	36,4	5,9	33,3	7,0
	F8	41,0	11,3	40,8	6,6	40,7	7,9
	F9	45,3	10,6	39,6	7,4	39,0	7,5
	F10	39,1	11,0	36,4	11,4	34,8	6,5
	N1	44,7	7,8	38,4	7,5	40,9	8,8
	N2	45,9	14,4	40,4	7,3	38,5	9,4
	N3	42,8	13,4	35,6	10,0	35,8	8,5
	N4	44,9	11,0	38,8	6,7	40,4	8,4
	N5	37,5	9,4	38,8	7,1	39,6	9,0
	N6	43,0	11,0	36,8	6,5	35,8	8,0
	N7	38,9	8,8	38,0	8,7	35,6	7,0
	N8	47,3	11,7	42,5	7,6	40,4	6,5
	N9	37,2	8,4	36,9	5,5	36,8	7,5
	N10	35,7	6,4	32,1	6,7	32,5	5,4

10. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként az nPVI-C esetében (A félkövér a legmagasabb értéket jelöli.)

		nPVI-C					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	59,8	11,7	54,7	7,4	56,2	8,4
	F2	56,1	10,5	54,0	8,2	57,0	9,7
	F3	63,6	11,6	55,2	7,0	56,3	9,9
	F4	51,8	11,0	55,7	8,8	52,6	10,2
	F5	55,7	9,9	58,2	7,4	55,1	7,2
	F6	54,1	10,4	53,3	8,2	55,9	9,0
	F7	60,7	11,9	55,8	6,6	56,9	10,8
	F8	57,1	11,9	56,4	7,3	57,7	11,1
	F9	60,4	10,1	58,3	9,7	59,0	10,1
	F10	54,4	8,8	53,8	8,5	53,1	9,2
	N1	61,3	12,2	48,9	9,2	54,3	10,7
	N2	55,1	10,2	58,5	7,7	57,3	10,0
	N3	50,7	9,4	51,0	7,4	54,5	8,1
	N4	66,0	13,8	58,3	6,4	56,6	10,7
	N5	53,4	12,2	53,8	7,6	56,8	11,4
	N6	49,2	10,1	56,4	7,7	56,9	8,4
	N7	59,7	15,1	49,1	6,4	53,8	10,3
	N8	55,4	14,2	52,5	8,5	53,6	8,4
	N9	51,0	15,2	56,0	8,8	53,5	10,1
	N10	48,7	7,8	50,4	7,5	51,4	8,6

4. Kísérletsorozatok az időzítési mintázatok vizsgálatára

11. táblázat

A spontán monológ, szövegfelolvasás és különálló mondatok tagmondataira kapott mérőszámértékek beszélőnként a %V esetében (A félkövér a legmagasabb értéket jelöli.)

		%V					
		Spontán		Szövegfelolvasás		Mondatfelolvasás	
		Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
Beszélő	F1	44,3	4,3	46,5	2,8	45,3	3,7
	F2	45,5	3,6	47,5	2,5	46,5	3,8
	F3	43,2	4,5	42,5	2,4	43,5	3,7
	F4	47,8	4,8	44,8	3,2	43,2	3,9
	F5	45,2	4,6	43,9	2,5	44,7	4,5
	F6	43,7	4,1	45,1	3,7	44,4	3,8
	F7	41,9	4,0	42,1	2,1	42,4	3,8
	F8	44,2	3,8	46,3	2,9	45,6	4,3
	F9	42,8	4,8	44,8	3,3	44,7	3,9
	F10	45,9	3,9	45,7	2,9	45,6	3,7
	N1	41,6	3,3	42,6	2,2	42,6	3,2
	N2	49,4	3,7	50,1	2,6	48,9	3,5
	N3	43,5	4,9	42,0	2,7	43,4	3,7
	N4	48,2	2,7	48,5	2,1	47,9	3,4
	N5	41,4	4,1	46,0	2,6	43,4	4,2
	N6	46,5	3,3	48,4	3,3	46,8	3,3
	N7	43,7	3,4	44,8	2,4	43,7	3,4
	N8	47,9	4,9	46,2	2,9	44,0	3,6
	N9	43,3	4,2	44,5	2,3	41,2	2,8
	N10	45,0	3,4	49,6	2,7	46,9	2,9

A különböző beszédmódokra kapott eredmények összegzése és a levonható következtetések

Ebben a fejezetben három beszédmód időzítésen alapuló beszédrítmusának összevetését végeztem el. Összefoglalva a különböző mérőszámok és mérési módszerek eredményeit azt mondhatjuk, hogy a spontán beszédben mind a mássalhangzós szakaszok, mind a magánhangzós szakaszok időtartamában általánosságban nagyobb variabilitás található, mint a szöveg- vagy mondatfelolvasásban – hasonlóan más nyelvekben mért eredményekhez (LIN–WANG 2007, ARVANITI 2012). Ezt támasztják alá a kétperces szöveg-, illetve monológ-részleteken alapuló mérések és részben a tagmondatok beszédrítmusának vizsgálata is. Ha azonban beszélőkre lebontva vizsgáljuk meg a különbségeket a két beszédmód között, árnyaltabb képet kapunk. Nem minden beszélő esetében mondható el, hogy a felolvasások beszédrítmusa jelentősen eltérne a spontán beszédétől. Továbbá eltérés mutatkozik a magánhangzós és a mássalhangzós szakaszokra épülő mérőszámok között. A kétféle mérőszám-csoportban ugyanis eltér, hogy a beszélőknél mennyire következetesen valósul meg a felolvasás és a spontán beszéd közti különbség. A magánhangzós szakaszokra épülő mérőszámok (ΔV , VarcoV, nPVI-V) esetében a teljes szövegre kapott beszédrítmus-mérőszámok értékei minden beszélőnél következetesen nagyobbak voltak spontán beszédben, mint felolvasásban. A legalább 10 szótagú tagmondatok vizsgálatakor is majdnem minden beszélő esetében elmondható, hogy a magánhangzós szakaszok időtartamára épülő mérőszámok átlagosan magasabbak voltak spontán beszédben, mint mondat- vagy szövegfelolvasásban, de a statisztikai próbák sokszor nem mutatták jelentősnek a különbségeket. A mássalhangzós szakaszokra épülő mérőszámok (ΔC , VarcoC, nPVI-C) esetében azonban még az átlagok tekintetében sem realizálódott következetesen az eltérés a felolvasás és a spontán beszéd között. Akadtak olyan beszélők, akiknél a mérőszámértékek a spontán beszédben voltak magasabbak, mint a szöveg- vagy a mondatfelolvasásban, másoknál azonban a felolvasásokban mérhető értékek bizonyultak nagyobbak. Igaz, ez utóbbi esetben az eltérések nem voltak jelentősek. Több esetben viszont statisztikailag is igazolható volt, hogy a spontán beszédben mért értékek magasabbak a felolvasásban mértéknél. Általánosságban tehát megfogalmazható egy olyan tendencia, hogy a spontán beszédben magasabb értékek jellemzők, mint felolvasásban, de beszélőnként változó, hogy ez az eltérés valóban realizálódik-e a két beszédmód között.

Felmerül a kérdés, hogy mi okozhatja a felolvasás és a spontán beszéd közti különbséget. A talált eltérések nem hozhatók összefüggésbe a globális artikulációs tempó különbségeivel, függetlennek tekinthetők attól. Ezt támasztja alá, hogy több, az artikulációs tempóval nem korreláló mérőszám értékeiben is megfigyelhető a beszédmódok közti eltérés. Habár a szünetek és a megakadások gyakorisága eltérő a két beszédmódban – a spontán beszédben jóval gyakoribbak, mint szövegfelolvasásban –, a jelenség nem magyarázható önmagában ezen eltérésekkel az eredmények alapján. A szünetek és megakadások közvetlen környékén lévő magánhangzós és mássalhangzós szakaszok időtartamainak kihagyásával éppúgy található

eltérés a két beszédmód között. Ha beszélőkre külön-külön lebontva nézzük, akkor a figyelmen kívül hagyott egységek számától függően egy, kettő vagy három, de maximum négy beszélőnél fordult elő, hogy a szövegfelolvasás értékei magasabbnak bizonyultak a spontán beszédre számolt értékektől. A különböző beszélők esetében tehát eltérő, hogy a szünetek és a megakadások számának mekkora jelentősége van a felolvasás és a spontán beszéd közti eltérésekben. A két beszédmód beszédritmusbeli eltéréseinek feltehető további oka, hogy a spontán beszédben a beszéd tervezése és az ily módon megfogalmazott szöveg meghangosítása egyidejűleg zajlik, míg felolvasáskor a beszélőnek csak a korábban megtervezett szöveg megszólaltatására kell koncentrálnia (GÖSY 2005). A spontán beszédben lévő többletfeladat, illetve annak nehézségei pedig az időviszonyok szélsőséges ingadozásához vezethetnek (vö. FLETCHER 2010).

A spontán beszéd és a felolvasás beszédritmusa nem minden tekintetben tért el egymástól. Két mérőszám is eltérő mintázatokat mutatott a többi mérőszámtól. Az egyik ilyen mérőszám az nPVI-C, amely nem mutatott különbséget a két beszédmódban. Ez a mérőszám az egymást követő mássalhangzós szakaszok időtartamának variabilitását adja meg oly módon, hogy a mássalhangzós szakaszok sorrendjét is figyelembe veszi. A sorrendet nélkülöző mérőszámok esetében, amelyek a mássalhangzós szakaszok időtartamának szórását határozzák meg (ΔC , VarcoC), viszont igazolható volt az eltérés. Ez azt jelenti, hogy a mássalhangzós szakaszoknak általánosságban szélsőségesebb időtartamértékei jelennek meg a spontán beszédben, mint felolvasásban, de az egymást követő mássalhangzós szakaszok időtartamának variabilitása mindkét beszédmódban hasonló. Ez arra utal, hogy spontán beszédben a nagyon hosszú vagy nagyon rövid időtartamú mássalhangzós szakaszok nem közvetlenül egymást követik, hanem valamilyen módon struktúrákba rendeződnek. Ha a szélsőséges időtartamok egymást követnék a beszédben, akkor egyértelműen mérhető lenne a különbség a spontán beszéd és a szövegfelolvasás beszédritmusa között.

A mérőszámok közül a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamarányát mutató mérőszám (%V) szintén eltérő mintázatokat mutatott a többi mérőszámtól. Míg a spontán beszédben a többi mérőszámérték magasabb volt a felolvasáshoz képest, addig ezen mérőszám értékei a szövegfelolvasásban voltak magasabbak a hosszú szövegek alapján. Habár a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamértékei egyaránt nagyobb variabilitást mutattak spontán beszédben, mint szövegfelolvasásban, a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamértékei mégsem ugyanolyan vagy hasonló mértékben tértek el a különböző beszédmódokban. Erre utal a %V mérőszám eltérése, amely azt mutatta, hogy a szövegfelolvasásban nagyobb a magánhangzós szakaszok (leegyszerűsítve a magánhangzók) időtartamának aránya a mássalhangzós szakaszokhoz (illetve a mássalhangzókhoz) képest, mint spontán beszédben. Beszélőkre lebontva viszont változó volt, hogy a spontán beszédben mért értékek vagy a szövegfelolvasásban mért értékek voltak magasabbak. A különbségek pedig a legtöbb esetben nem is bizonyultak jelentősnek a statisztikai próba alapján. A %V mérőszám ugyanakkor korrelációt mutatott az artikulációs tempóval, így feltehetően ez az összefüggés befolyásolta a beszédmódok közti

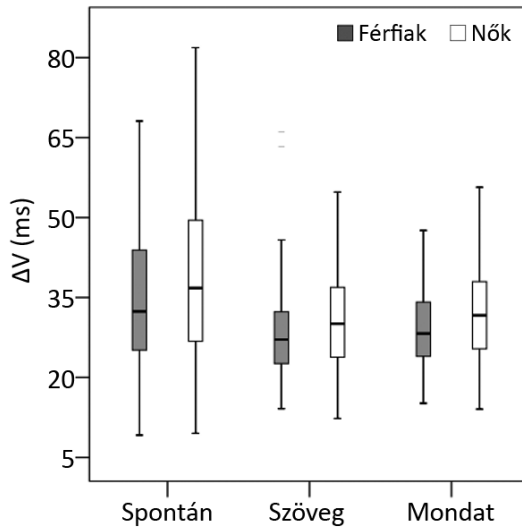
különbségekkel kapcsolatos eredményeket. Mivel a szakirodalomban található mérési eredmények szerint nem igazolható, hogy a %V és az artikulációs tempó közvetlen kapcsolatban lenne, illetve összefüggésüket elhanyagolhatónak találták (RUSSO–BARRY 2008; DELLWO 2010), ezért a magánhangzók és a mássalhangzók időtartamarányának kapcsolatát az artikulációs tempóval a 4.1.2.5. fejezetben részletesebben is megvizsgáljuk.

Láthattuk, hogy a spontán beszéd és a felolvasás beszédritmusának időbeli dimenziója között megjelennek különbségek. A mondatfelolvasás és a szövegfelolvasás beszédritmusában nem található szisztematikus eltérés. A magánhangzós és a mássalhangzós szakaszokra épülő mérőszámok átlagosan hol az egyik, hol a másik beszédmódban mutatkoztak magasabbnak. A különbségek mértéke többnyire nem volt jelentős (kivétel: VarcoV, ΔV és ΔC), főként, ha beszélőnként vizsgáltuk az eltéréseket. A mérőszámok alapján tehát nem található jelentős eltérés a szöveg részét képező tagmondatok és a teljesen különálló mondatot alkotó tagmondatok beszédritmusa között.

4.1.2.3. A nemek hatása a beszédritmus-mérőszámok értékeire

A különböző nemű beszélők beszédritmusa több tekintetben is eltért egymástól. A vizsgálat-hoz összevettem a férfiak és a nők tagmondatokra kapott beszédritmusmérőszám-értékeit külön-külön az eltérő beszédmódokban. (A teljes szövegfelolvasásra és a spontán beszédre kapott mérőszámértékek vizsgálatától eltekintettem az al csoportokban lévő alacsony elem-szám miatt.)

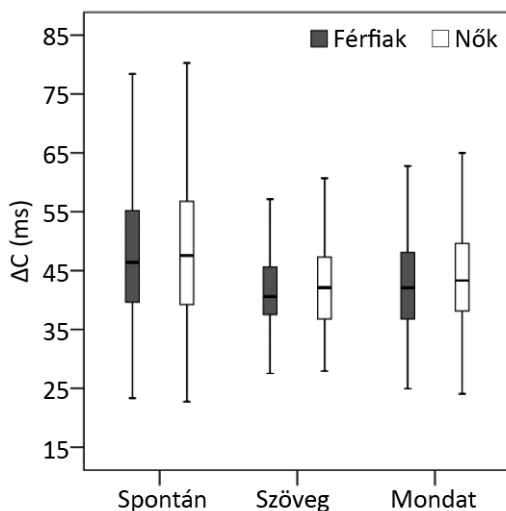
A magánhangzós szakaszok időtartamának szórása (ΔV) mindhárom beszédmódban különbséget mutatott a nők és férfiak ejtésében (12. ábra). A nők tagmondatainak átlagos ΔV -értéke 31,4 ms (szórás: 11,6 ms) volt a szövegfelolvasás tagmondatait alapul véve, amely magasabb a férfiak átlagos ΔV -értékénél (átlag: 29,2 ms, szórás: 10,0 ms). A Mann–Whitney-próba szignifikáns eltérést mutatott a két csoport értékei között ($Z = -2,887$, $p = 0,003$). A mondatfelolvasások vizsgálata hasonló eredményeket hozott. A ΔV értéke átlagosan nagyobb volt a nők esetében (átlag: 32,1 ms, szórás: 9,0 ms), mint a férfiak esetében (átlag: 29,2 ms, szórás: 7,5 ms), amely eltérés szintén jelentősnek tekinthető a statisztikai próba alapján ($Z = -3,312$, $p = 0,001$). A spontán beszéd tagmondatnyi egységeit is megvizsgáltam nemek szerint, és a felolvasásokban talált eredményekhez hasonló tendenciák mutatkoztak. A férfiak átlagos ΔV -értéke alacsonyabb volt (átlag: 38,2 ms, szórás: 21,3 ms), mint a nőké (átlag: 40,8 ms, szórás: 20,1 ms). A statisztikai próba szignifikáns eltérést jelzett a nagy szórás ellenére ($Z = -2,019$, $p = 0,043$).



12. ábra

A ΔV értékeinek nemek közti eltérései a spontán monológ, a szövegfelolvasás és a mondatfelolvasás tagmondatai alapján

A magánhangzós szakaszok időtartamának szórásához hasonlóan a nők beszéde átlagosan szintén nagyobb variabilitást mutatott a mássalhangzós szakaszok időtartamának tekintetében is a férfiakéhoz képest (13. ábra). Szövegfelolvasásban a mássalhangzós szakaszok szórása (ΔC) a nőknél átlagosan 43,0 ms volt (szórás: 8,6 ms), és csak 41,4 ms (szórás: 6,8 ms) a férfiaknál. A statisztikai próba alapján viszont ez az eltérés nem mondható jelentősnek ($Z = -1,651, p = 0,099$). A mondatfelolvasásban is megfigyelhető az az eltérés, hogy a nőknél átlagosan magasabb értékek jellemzők (átlag: 44,5 ms, szórás: 9,4 ms), mint a férfiaknál (átlag: 42,7 ms, szórás: 8,1 ms). Spontán beszédben is látszik a nemek közti eltérés. A férfiak esetében a ΔC értéke átlagosan 49,2 ms volt (szórás: 15,4 ms), nők esetében valamivel magasabbnak mutatkozott ez az érték (átlag: 49,6 ms, szórás: 15,3 ms). A szövegfelolvasáshoz hasonlóan azonban sem a mondatfelolvasásban ($Z = -1,627, p = 0,104$), sem a spontán beszédben nem volt jelentős az eltérés ($Z = -0,279, p = 0,780$). Tehát csak a magánhangzós szakaszok időtartamának szórása (ΔV) esetében található számottevő eltérés a nemek között, a mássalhangzós szakaszok időtartamának szórása (ΔC) esetében nem mutatható ki ilyen kapcsolat.

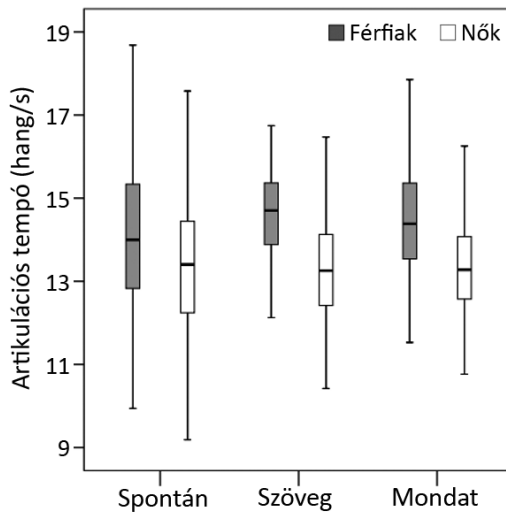


13. ábra

A ΔC értékeinek nemek közti eltérései a spontán monológ, a szövegfelolvasás és a mondatfelolvasás tagmondatai alapján

Az imént vizsgált két mérőszámban (ΔV és ΔC) közös, hogy összefüggést mutatnak az artikulációs tempóval (DELLWO 2009). Annak érdekében, hogy megtudjuk, a két csoport eltérései mennyiben csak az artikulációs tempó szisztematikus eltéréséből fakadnak, megvizsgáltam a nemeknek az artikulációs tempóra gyakorolt hatását, valamint a mérőszámok és az artikulációs tempó korrelációját.

Az artikulációs tempó eltérően alakult a férfiaknál és a nőknél. Mindhárom vizsgált beszédmódban tendenciózusan gyorsabban beszéltek a férfiak, mint a nők (14. ábra). A spontán beszédben a nők átlagos artikulációs tempója 13,0 hang/s (szórás: 1,9 hang/s), a férfiaké pedig 13,7 hang/s (szórás: 1,9 hang/s) volt. Ugyanez a tendencia látható a felolvasásokban is. A szövegfelolvasásban a nők átlagos artikulációs tempója 12,9 hang/s (szórás: 1,3 hang/s), míg a férfiaké 14,1 hang/s (szórás: 1,0 hang/s) volt. A mondatfelolvasásban hasonlóan alakultak az artikulációs tempó átlagai: a nőknél 12,9 hang/s (szórás: 1,2 hang/s), a férfiaknál pedig 14,0 hang/s (szórás: 1,2 hang/s) volt a tempó. A két csoport közti eltérés szignifikánsnak bizonyult a Mann–Withney-próba alapján mind a szövegfelolvasásban ($Z = -9,509, p = 0,001$), mind a mondatfelolvasásban ($Z = -8,159,0, p = 0,001$), és a spontán beszédben is ($Z = -3,810, p = 0,021$).



14. ábra

Az artikulációs tempó nemek közti eltérései a spontán monológ, a szövegfelolvasás és a mondatfelolvasás tagmondatai alapján

A Spearman-féle korrelációelemzés megmutatta, hogy az artikulációs tempó és az eddig vizsgált mérőszámok szoros összefüggésben vannak mindegyik beszédmódban. Az artikulációs tempó közepesen negatívan korrelált mind a ΔV ($\rho = -0,591, p < 0,001$), mind a ΔC mérőszámmal ($\rho = -0,501, p < 0,001$) spontán beszédben. Ez azt jelenti, hogy minél kisebb volt az artikulációs tempó, azaz lassabban valósult meg egy tagmondat, annál nagyobb volt a magánhangzós és a mássalhangzós szakaszok időtartamának szórása. A szövegfelolvasásban szintén szignifikáns összefüggés volt az artikulációs tempó és a ΔV ($\rho = -0,553, p < 0,001$), illetve a ΔC ($\rho = -0,322, p < 0,001$) között. Mondatfelolvasásban hasonló a helyzet a többi beszédmóddhoz. A statisztikai próba itt is negatív korrelációt mutatott az artikulációs tempó és a két beszédritmus-mérőszám között (a ΔV esetében: $\rho = -0,447, p < 0,001$; a ΔC esetében: $\rho = -0,345, p < 0,001$).

Mivel az eddig vizsgált beszédritmus-mérőszámok nemenkénti különbségei feltehetően az artikulációs tempó szisztematikus eltéréseiből fakadnak, ezért megvizsgáltam azokat a beszédritmus-mérőszámokat (VarcoV, VarcoC, nPVI-V, nPVI-C) is, amelyek definíciójukból következően nem függenek az artikulációs tempótól. A Spearman-féle korrelációelemzés megmutatta, hogy az artikulációs tempó és ezen mérőszámok egyike sem korrelált mondatfelolvasásban ($p \geq 0,118$). Szövegfelolvasásban sem volt igazolható jelentős összefüggés az artikulációs tempó és egyik mérőszám között sem. A statisztikai próba ugyan a VarcoV mérőszám esetében szignifikáns korrelációt jelzett ($p = 0,001$), de az összefüggés mértéke igen gyengének mondható ($\rho = -0,163$). A többi mérőszám esetében nem volt semmilyen összefüggés alátámasztható ($p \leq 0,001$). Spontán beszédben a VarcoV, a VarcoC és az nPVI-V mérőszám is

gyenge negatív korrelációt mutatott az artikulációs tempóval ($-0,311 \leq \rho \leq -0,132, p \leq 0,008$), az nPVI-C mérőszám viszont semmilyen mértékben nem függött össze az artikulációs tempóval ($p = 0,467$).

A nemek közti különbségek nem mindenhol jelentek meg az artikulációs tempótól független vagy függetlennek tekinthető mérőszámok esetében. Eltérő mintázat volt tapasztalható a mérőszámértékekben attól függően, hogy a mássalhangzós szakaszokra vagy a magánhangzós szakaszokra épülnek.

A magánhangzós szakaszokra épülő mérőszámok esetében semmilyen jelentős különbség nem volt megfigyelhető a nemek ejtésében. A magánhangzós szakaszokra épülő mérőszámok (VarcoV és nPVI-V) átlagai ugyan a férfiaknál voltak magasabbak spontán beszédben és szövegfelolvasásban, míg mondatfelolvasásban a nők ejtésében realizálódtak nagyobb értékek (lásd 12. táblázat), a statisztikai próbák nem igazolták a két csoport beszédritmusának eltérését. A VarcoV mérőszám sem a szövegfelolvasásban ($Z = -1,717, p = 0,086$), sem a mondatfelolvasásban ($Z = -0,395, p = 0,693$), sem a spontán beszédben ($Z = -6,402, p = 0,521$) nem mutatott szignifikáns különbséget nemek szerint. A VarcoV mérőszámhoz hasonlóan az nPVI-V mérőszám értékei sem tértek el a férfiak és a nők ejtésében, bármelyik beszédmódot is vettem figyelembe (szövegfelolvasás: $Z = -1,200, p = 0,230$; mondatfelolvasás: $Z = -1,130, p = 0,258$; spontán beszéd: $Z = -0,965, p = 0,334$).

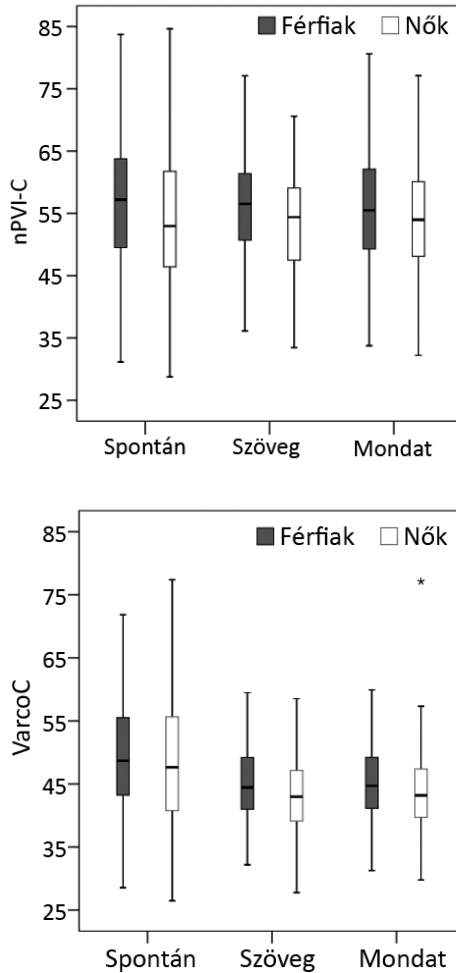
12. táblázat

A beszédritmusmérőszám-értékek nemek szerinti átlaga és szórása a spontán monológ, a szövegfelolvasás és a mondatfelolvasás tagmondatai alapján (A félkövér betűtípus azt jelöli beszédritmusként, hogy a mérőszámok értéke átlagosan a férfiaknál vagy a nőknél nagyobb.)

	Spontán beszéd				Szövegfelolvasás				Mondatfelolvasás			
	Férfi		Nő		Férfi		Nő		Férfi		Nő	
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás
VarcoV	46,9	18,4	46,4	15,0	37,9	10,9	35,9	7,7	38,0	7,9	38,2	8,3
VarcoC	50,7	11,5	48,8	11,0	45,1	6,2	43,6	6,8	45,2	6,2	43,8	6,5
nPVI-V	43,0	10,7	41,8	11,0	38,6	8,2	37,8	7,7	36,7	7,1	37,6	8,2
nPVI-C	57,3	11,1	55,0	13,2	55,5	7,9	53,5	8,3	56,0	9,6	54,9	9,7
%V	44,5	4,5	45,0	4,6	44,9	3,2	46,3	3,7	44,6	4,0	44,9	4,1

A mássalhangzós szakaszokra épülő mérőszámok esetében viszont jelentős eltérés mutatkozott a nemek közt. A VarcoC mérőszám és az nPVI-C mérőszám értékei a férfiak esetében voltak magasabbak mindhárom beszédmódban (lásd 12. táblázat). A nemek közti eltérés a VarcoC mérőszám értékeiben szignifikánsnak bizonyult szövegfelolvasásban ($Z = -2,670$,

$p = 0,008$) és mondatfelolvasásban ($Z = -2,403, p = 0,016$) is. Spontán beszédben viszont nem volt kimutatható különbség a két csoport között ($Z = -1,621, p = 0,105$). Az nPVI-C értékeiben megjelenő különbség nemek szerint szövegfelolvasásban ($Z = -2,474, p = 0,013$) és spontán beszédben ($Z = -2,482, p = 0,013$) is szignifikáns a statisztikai próbák alapján, viszont az eltérés mondatfelolvasásban nem tekinthető jelentősnek ($Z = -1,428, p = 0,153$). A férfiak tehát átlagosan magasabb VarcoC- és nPVI-C-értékeket produkáltak a nőknél, azaz a mássalhangzós szakaszok időtartama szélsőségesebben realizálódott a férfiak ejtésében, mint a nőkben (15. ábra), és ezen eltérések igazolhatóan szignifikánsak voltak a legtöbb esetben.



15. ábra

A VarcoC és az nPVI-C mérőszám nemek közti eltérései a spontán monológ, a szövegfelolvasás és a mondatfelolvasás tagmondatai alapján

A magánhangzós szakaszok és a mássalhangzós szakaszok időtartamának arányára reagáló %V mérőszámot is megvizsgáltam, hogy található-e eltérés a nemek közt. A nők a szövegfelolvasásban és a mondatfelolvasásban is átlagosan magasabb értékeket produkáltak, mint a férfiak, spontán beszédben viszont a férfiak ejtésében realizálódtak magasabb értékek (lásd *12. táblázat*). A nemek közti különbség csak szövegfelolvasásban tekinthető jelentősnek ($Z = -3,789$, $p = 0,001$). A statisztikai próba ugyanis nem támasztotta alá, hogy a férfiak és a nők %V értékei szignifikánsan eltérnének mondatfelolvasásban ($Z = -0,763$, $p = 0,445$) vagy spontán beszédben ($Z = -1,176$, $p = 0,239$). A %V mérőszám ugyanakkor közepes negatív korrelációt mutatott az artikulációs tempóval. Ez az összefüggés mindhárom beszédmódban megjelent (spontán beszédben: $\rho = -0,316$, $p < 0,001$; szövegfelolvasásban: $\rho = -0,485$, $p < 0,001$; mondatfelolvasásban: $\rho = -0,313$, $p < 0,001$). Az eredményeket tehát ezen mérőszám esetében befolyásolhatta, hogy a férfiak és a nők artikulációs tempója között szisztematikus eltérések találhatók.

Következtetések

A különböző nemű beszélők időzítésbeli sajátosságai több tekintetben is eltértek egymástól. A magánhangzós és mássalhangzós szakaszok időtartamának szórása (ΔV és ΔC) esetében számottevő eltérés található a nemek között. A magánhangzós és a mássalhangzós szakaszok időtartama a nők ejtésében átlagosan nagyobb variabilitást mutatott a férfiakéhoz képest mindhárom vizsgált beszédmódban (mondatfelolvasás, szövegfelolvasás, spontán beszéd). Az eredményeket azonban befolyásolhatta az artikulációs tempó. A magyar nyelvben talált korábbi eredmények ellentmondásosak a tekintetben, hogy az artikulációs tempó eltér-e jelentősen nemek szerint. Eredményeim –VÁRADI (2009) mondatfelolvasáson, valamint mondatvisszamondáson végzett kísérleteihez hasonlóan – azt támasztják alá, hogy a férfiaknál magasabb artikulációstempó-értékek jellemzők, azaz gyorsabban beszélnek, mint a nők. Az artikulációs tempónak ez a nemek szerinti eltérése mindhárom beszédmódban megjelent. A nőknél megjelenő lassabb artikulációs tempó egyben azt is jelenti, hogy a magánhangzós és mássalhangzós szakaszok abszolút időtartama szélesebb skálán változhat, nagyobb szórást tapasztalhatunk, mint a gyorsabban beszélő férfiaknál.

Nem pusztán az artikulációs tempóban van különbség a nemek között, hanem más időzítésbeli tényezők is eltérést mutattak. Az artikulációs tempóval nem korreláló, mássalhangzós szakaszokra épülő beszédritmus-mérőszámok (VarcoC, nPVI-C) értékei a férfiak esetében voltak átlagosan magasabbak mindhárom beszédmódban. A férfiak esetében tehát a mássalhangzós szakaszok időtartamértékei nagyobb variabilitást mutattak a nőkéénél, ha függetlenítettem az adatokat az artikulációs tempótól. Továbbá ha a mássalhangzós szakaszok sorrendjét is figyelembe vettem, akkor szintén a férfiaknál tapasztaltam szélsőségesebb megvalósulásokat az egymást követő szakaszok időtartamában. Hozzá kell tennünk, hogy nem minden beszédmódban volt szignifikáns az eltérés a nemek között. A VarcoC mérőszám értékei szövegfelolvasásban és mondatfelolvasásban, az nPVI-C mérőszám értékei pedig szövegfelolvasásban és spontán beszédben tértek el jelentősen a férfiak és a nők ejtésében.

A különbségek tehát felolvasásban is realizálódtak, amikor ugyanazon felépítésű mássalhangzós szakaszok kerültek kimondásra. Egyrészt tehát előfordulhat, hogy az eltérések oka a mássalhangzó-torlódásokban fellépő hangzótörlésekben keresendő, másrésztől maguk a mássalhangzók időtartamai is szélsőségesebben valósulhattak meg a férfiak ejtésében. A szélsőségesebb időtartamértékeket sok minden implikálhatja, mint például erőteljesebb lassítás megnyilatkozás vagy frázis végén, gyakoribb szünettartás tagmondaton belül, esetleg a hangsúlyos vagy fókuszpozíció gyakoribb vagy jelentősebb időtartamra gyakorolt hatása stb. A PVI mérőszámban talált eltérés arra utal, hogy a férfiak dinamikusabb időtartambeli váltásokat hajtanak végre; kérdés, hogy vajon ez a nagyobb egységek tempóváltásaiban is tetten érhető-e, vagy csak az egymást követő mássalhangzós szakaszok időtartamaiban realizálódnak. (A tempóváltásokban megjelenő esetleges nemek közti különbségekre a 4.2. fejezetben még visszatérünk.)

A magánhangzós szakaszokra épülő és az artikulációs tempótól függetlennek tekinthető mérőszámok (VarcoV és nPVI-V) esetében nem sikerült nemek közti különbséget statisztikailag is igazolni. A magánhangzós szakaszokra épülő mérőszámok átlagai a férfiaknál voltak magasabbak spontán beszédben és szövegfelolvasásban, míg mondatfelolvasásban a nők ejtésében realizálódtak nagyobb értékek, de ezek a különbségek nem tekinthetők jelentősnek. A %V mérőszám esetében a nők szövegfelolvasásban átlagosan magasabb értékeket produkáltak, mint a férfiak, a többi beszédmód azonban nem mutatott szignifikáns eltérést nemek szerint. A %V mérőszám ugyanakkor közepesen és negatívan korrelált az artikulációs tempóval minden beszédmódban, feltehetően ez befolyásolta az eredményeket.

Összességében elmondható, hogy a férfiak és a nők beszédritmusában talált különbségek a mássalhangzós szakaszok időtartamának variabilitásában realizálódnak, de ezen eltérések sem következetesen valósulnak meg minden beszédmódban. A nemzetközi kutatások alapján nem feltételezhető, hogy univerzális összefüggések állnának fenn a nemek és a beszédritmus-mérőszámok értékei között: hol a nőknél, hol a férfiaknál mértek magasabb értékeket nyelvtől, mérőszámtól és beszédmódtól függően (BENTON et al. 2007; GRAWUNDER et al. 2008; TORGERSEN–SZAKAY 2012). Nyitott marad a kérdés, hogy vajon az általam vizsgált anyagban, ahol 10 nőt és 10 férfit hasonlítottam össze, az eredmények mennyiben tulajdoníthatók valóban a nemek közti eltérésnek, és mennyiben csak a különböző beszélők sajátosságai mutatkoznak meg véletlenszerűen.

4.1.2.4. A különböző beszélők szünettartási szokásainak hatása a beszédritmusra

A különböző beszélők nagy változatosságot mutathatnak a beszédritmus tekintetében (ARVANITI 2012; DELLWO 2012). A beszélők közti beszédritmusbeli eltérések egyik lehetséges oka, hogy mások a szünettartási szokásaik, a beszélők gyakrabban vagy ritkábban tagolhatják a szöveget szünetekkel. Annak megvizsgálásához, hogy ez kapcsolatban állhat-e a beszédritmus-mérőszámokkal, elemeztem minden beszélő teljes szövegfelolvasását, hogy a különböző beszélők hány szünetet tartanak ugyanazon szöveg felolvasásakor.

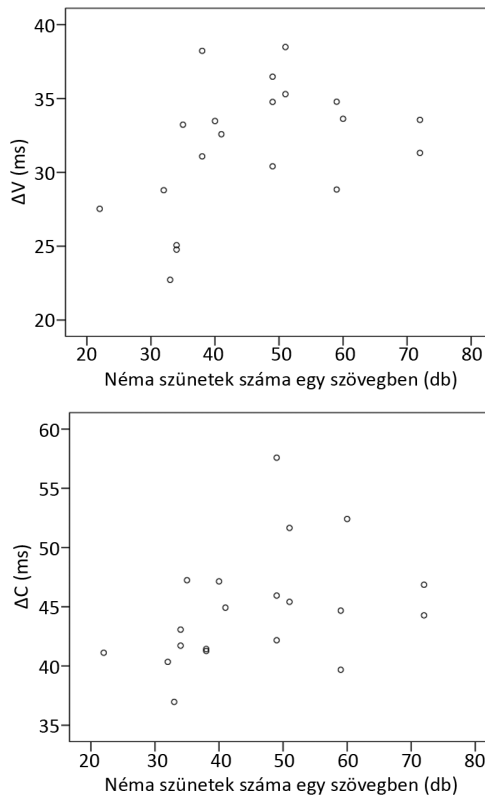
A beszélők a 12 összetett mondatból álló szöveget átlagosan 46 olyan szünettel tagolták, amely előtt vagy után nem jelent meg megakadás, tehát egyedülálló néma szünetek voltak. Hozzá kell tennünk, hogy a beszélők átlagosan 4 alkalommal olvastak félre vagy hezitáltak, ami általában néma szünettel párosult. Ezeket a megakadásokat és a néma szünettel megvalósuló kombinációjukat egy kategóriaként kezeltem. (Ebben a kategóriában a megakadás bármilyen formában szakította meg a beszédet, mindig egynek tekintettem, akár néma szünet nélkül, akár néma szünettel valósult meg.) Összességében a beszélők a szöveget átlagosan 50 alkalommal szakították meg valamilyen módon. A 13. táblázatban látható, hogy a különböző beszélők esetében hogyan alakultak a szünettartási szokások. (Fontos megjegyezni, hogy ez a táblázat eltér a korábban közölt 3. táblázat adataitól, itt ugyanis a teljes szövegfelolvasás néma szüneteit és megakadásait vizsgáltam módszertani megfontolásokból.)

13. táblázat

A különböző beszélők szünettartása szövegfelolvasásban

Beszélő	Néma szünet (db)	Megakadások (néma szünettel) (db)	Összesen (db)
F1	32	9	41
F2	38	3	41
F3	34	3	37
F4	38	1	39
F5	22	2	24
F6	51	8	59
F7	34	2	36
F8	72	3	75
F9	35	0	35
F10	49	3	52
N1	33	6	39
N2	49	1	50
N3	59	9	68
N4	41	8	49
N5	49	1	50
N6	51	8	59
N7	40	9	49
N8	60	1	61
N9	72	5	77
N10	59	4	63

A Spearman-féle korrelációelemzések azt mutatták, hogy a beszélők szünettartásainak, illetve tagolásainak száma és a beszédritmus-mérőszámok értéke között a legtöbb esetben nincs összefüggés. A teszt nem támasztotta alá, hogy a Varco mérőszámok (VarcoV, VarcoC) vagy az nPVI mérőszámok (nPVI-V, nPVI-C) értékei kapcsolatban lennének a szünettartás gyakoriságával. Ez az összefüggés nem volt alátámasztható a néma szünetek esetében ($-0,215 \leq \rho \leq 0,121$, $p \geq 0,363$). Ha az összes beszédmegszakítást vettem figyelembe, a beszédritmus-mérőszámok akkor sem mutattak korrelációt a szünetek mennyiségével ($-0,351 \leq \rho \leq 0,341$, $p \geq 0,129$). A %V mérőszám sem a néma szünetek gyakoriságával ($\rho = 0,372$, $p = 0,106$), sem az összes megszakítás darabszámával ($\rho = 0,362$, $p = 0,116$) nem korrelált. A magánhangzós szakaszok időtartamának szórása (ΔV) viszont pozitívan korrelált a néma szünetek darabszámával ($\rho = 0,561$, $p = 0,010$). A mássalhangzós szakaszok időtartamának szórása (ΔC) és a néma szünetek között szintén kapcsolatot jelzett a statisztikai próba ($\rho = 0,492$, $p = 0,028$). A pozitív korreláció a 16. ábrán tekinthető meg. Minél több szünet volt egy szövegfelolvasásban, annál nagyobb volt a magánhangzós és a mássalhangzós szakaszok időtartamának szórása.



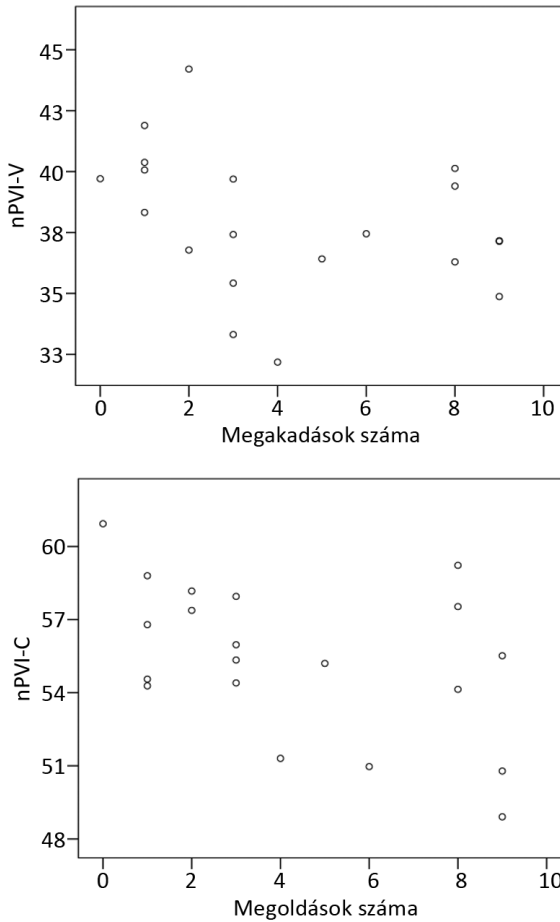
16. ábra

A magánhangzós (fent) és a mássalhangzós szakaszok (lent) időtartamának szórása a néma szünetek előfordulási gyakoriságának függvényében

A magánhangzós és mássalhangzós szakaszok időtartamának szórása (ΔV és ΔC) – mint korábban szóba került – negatívan korrelál az artikulációs tempóval. A beszélőnként vett teljes szövegek artikulációs tempója az ugyanilyen módon mért ΔV mérőszámmal közepes, negatív korrelációt mutatott ($\rho = -0,707, p < 0,001$). A ΔC mérőszám esetében hasonló összefüggés tapasztalható, a statisztikai próba itt is megmutatta az artikulációs tempóval való szoros kapcsolatot ($\rho = -0,674, p = 0,001$). A néma szünetek gyakorisága is korrelált az artikulációs tempóval ($\rho = -0,710, p = 0,001$). Minél több szünet fordult elő egy szövegben, annál kisebb volt az artikulációs tempó (a ténylegesen kiejtett hangok sebessége, tehát nem a szünetekkel együtt mért tempó), azaz a beszélők annál lassabban artikuláltak. Mivel az artikulációs tempó mind a néma szünetek gyakoriságával, mind a ΔV és ΔC mérőszámmal összefüggött, ezért nem zárható ki, hogy a statisztikai próba valójában csak az artikulációs tempóval való korreláció miatt mutatott szignifikáns összefüggést ezen két mérőszám és a szünetezés között.

A szöveg egészében megjelenő összes beszédmegszakítás darabszáma szintén összefüggött a magánhangzós szakaszok időtartamának szórásával (ΔV). A statisztikai próba szignifikáns, pozitív korrelációt mutatott ki közöttük ($\rho = 0,465, p = 0,039$). A szövegben megjelenő összes szünet darabszáma szintén korrelált az artikulációs tempóval ($\rho = -0,661, p = 0,001$), tehát itt sem zárható ki, hogy csak az artikulációs tempóval való szoros összefüggés okozta a jelenséget. A mássalhangzós szakaszok időtartamának szórása (ΔC) nem mutatott összefüggést a megakadások és a néma szünetek összesített számával ($\rho = 0,366, p = 0,112$).

A megakadások és az olyan megakadások, amelyeket néma szünet előz meg, vagy néma szünet követ, beszélőnként igen változó számban fordultak elő (lásd *13. táblázat*). A megakadások és néma szünettel megvalósuló kombinációjuk száma korrelációt mutatott a PVI mérőszámokkal. A magánhangzós szakaszok időtartamaira épülő nPVI-V közepesen, negatívan korrelált a megakadások szövegenként megjelenő darabszámával ($\rho = -0,483, p = 0,031$). Minél több megakadás, hezitálás jelent meg tehát egy szövegben, annál alacsonyabb mérőszámérték volt jellemző az adott beszélő szövegfelolvasása esetében. Hasonló összefüggés mutatható ki az nPVI-C ($\rho = -0,453, p = 0,042$) mérőszám esetében is (*17. ábra*). Egyik PVI mérőszám sem korrelált az artikulációs tempóval (nPVI-V: $\rho = 0,325, p = 0,162$; nPVI-C: $\rho = 0,045, p = 0,850$), tehát ez nem befolyásolhatta az eredményeket. A megakadások száma sem mutatott korrelációt az artikulációs tempóval ($\rho = -0,089, p = 0,709$). A statisztikai próba nem mutatott korrelációt a többi beszédritmus-mérőszám ($\Delta V, \Delta C, \text{VarcoV}, \text{VarcoC}, \%V$) és a megakadások száma között ($-0,170 \leq \rho \leq -0,065, p \geq 0,473$).



17. ábra

A magánhangzós (fent) és a mássalhangzós (lent) szakaszok időtartamára épülő PVI mérőszámok időtartamának szórása a megakadások számának függvényében

Következtetések

A vizsgálatban arra a kérdésre kerestem a választ, hogy a beszélőkre jellemző szünettartási szokások milyen összefüggésben vannak a beszédritmus időbeli dimenziójával szövegfelolvasásban. A központi kérdés az volt, hogy vajon egy teljes szöveg szakaszokra tagolása milyen hatással van a beszédritmus-mérőszámok értékeire. A szünetezésnek csak egyetlen aspektusával, a szünetek mennyiségével foglalkoztunk. Az eredmények azt mutatták, hogy a magánhangzós szakaszok időtartamának szórása (ΔV) és a mássalhangzós szakaszok időtartamának szórása (ΔC) is pozitívan korrelált a néma szünetek darabszámával. Azaz minél

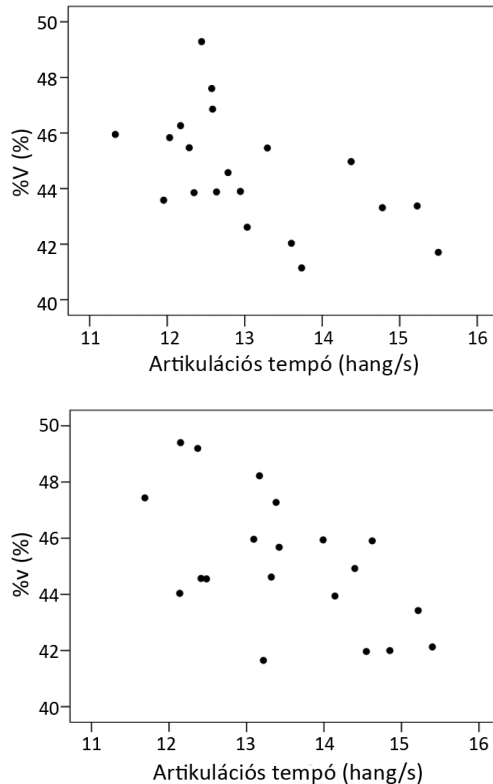
több szünetet tartott egy beszélő a szövegben, annál több szélsőséges időtartamértéket produkált. Ugyanakkor minél több szünet fordult elő egy szövegben, annál kisebb volt az artikulációs tempó, azaz a beszélők annál lassabban beszéltek. Mivel az artikulációs tempó mind a néma szünetek gyakoriságával, mind a ΔV és ΔC mérőszámmal összefüggött, ezért nem zárható ki, hogy a statisztikai próba valójában csak az artikulációs tempóval való korreláció miatt mutatott szignifikáns összefüggést ezen két mérőszám és a szünetezés között. A legtöbb mérőszám esetében nem találtam összefüggést a tagolás és a beszédritmus-mérőszám értékei között. A megakadások megjelenésének gyakorisága azonban összefüggést mutatott a PVI mérőszámokkal. Minél több megakadás, hezitálás jelent meg egy szövegben, annál inkább alacsonyabb mérőszámérték (nPVI-V, nPVI-C) volt jellemző az adott beszélő szövegfelolvasása esetében. Ez azt jelenti, hogy a kevesebbet megakadó vagy hezitáló beszélők esetében az egymást követő mássalhangzós, illetve magánhangzós szakaszok időtartama tendenciaszerűen nagyobb variabilitást mutattak, mint a többet megakadó beszélőknél. Ennek egyik lehetséges magyarázata, hogy a szövegfelolvasásban kevesebbet megakadó beszélőknek kevésbé okozott problémát a hangos olvasás, éppen ezért dinamikusabban tudták felolvasni az adott szöveget a feladatot nehezebben teljesítő beszélőknél. Habár a szünetek és a beszédritmus-mérőszámok között nem sikerült további összefüggéseket kimutatni a rendelkezésemre álló anyag felépítése miatt, ettől függetlenül jóval komplexebb kapcsolatot tételezhetünk fel a szünetezés és a beszédritmus között, ami további kutatások tárgyát képezheti.

4.1.2.5. A magánhangzók és a mássalhangzók időtartamának az aránya és a globális, valamint a lokális artikulációs tempó összefüggései

A korábbi fejezetek vizsgálatai során érintőlegesen megmutatkozott, hogy a magánhangzós szakaszok időtartamának százalékos aránya a teljes időtartamhoz képest (%V) összefügghet az artikulációs tempóval. Eközben a szakirodalmi adatok szerint a %V mérőszám robusztus az artikulációs tempóval szemben (DELLWO 2010). Ezért ebben a fejezetben külön részletesebben is kifejtem azokat az eredményeket, amelyek korábban már más vizsgálatokban módszertani szempontok miatt érintőlegesen említésre kerültek, és alaposabban is megvizsgálom, hogy van-e közvetlen kapcsolat az artikulációs tempó és a %V mérőszám között. Mivel a %V mérőszám a magánhangzós szakaszok időtartamának az arányát adja meg a teljes időtartamhoz képest, lényegében a magánhangzós és a mássalhangzós szakaszok időtartamának arányára reagál egy adott szövegben vagy szövegrészletben. Ezen mérőszám esetében tehát nincs jelentősége, hogy a magánhangzók vagy mássalhangzók szakaszokat alkotnak-e, így a következőkben egyszerűsítve a magánhangzók és a mássalhangzók időtartamának arányaként fogok rá hivatkozni.

Az artikulációs tempó többféleképpen is összefügghet a %V mérőszámmal. Egyrésztől a különböző beszélők globális artikulációs tempója, azaz a rájuk jellemző tempó összefüggést mutathat azzal, hogy milyen arányban valósítják meg a magánhangzós és a mássalhangzós szakaszok (leegyszerűsítve a magánhangzók és a mássalhangzók) időtartamát. Másrésztől egy beszélőn belül is összefügghet a lassabb vagy gyorsabb tempó a %V mérőszámmal.

A különböző beszélők eltérő globális artikulációs tempója korrelációt mutatott a %V mérőszámmal. A vizsgálathoz 20 beszélő 2 perces monológját és 2 perces szövegfelolvasását használtam fel. A beszélők teljes szövegekre és spontán monológokra mért artikulációs tempója beszédmódtól függetlenül közepesen negatívan korrelált az ugyanilyen módon számolt %V mérőszámmal ($\rho = -0,509$, $p = 0,001$). Minél gyorsabban beszélt egy beszélő, annál inkább csökkent a magánhangzós szakaszok (tehát a magánhangzók) időtartamának aránya a teljes beszédidőtartamhoz képest. Megvizsgáltam külön-külön beszédmódonként, hogy a 20 beszélő globális artikulációs tempója valóban összefügg-e a %V mérőszámmal (18. ábra). A spontán beszédben statisztikailag is alátámasztható volt a korreláció (Spearman-féle korrelációelemzés: $\rho = -0,623$, $p = 0,003$). A szövegfelolvasásban szintén kimutatható volt ez az összefüggés ($\rho = -0,532$, $p = 0,016$). A magánhangzók időtartamának százalékos aránya a teljes időtartamhoz képest tehát kisebb a gyorsabban beszélők esetében, és nagyobb a jellemzően lassabb artikulációs tempójú kísérleti személyek esetében.



18. ábra

A %V mérőszám összefüggése az artikulációs tempóval spontán beszédben (fent) és szövegfelolvasásban (lent) (Az ábrán látható pontok egy-egy beszélő 2 perces spontán monológja, illetve szövegfelolvasásra számított értékeit mutatják.)

Felmerül a kérdés, hogy vajon csak különböző beszélők globális artikulációs tempója és a %V között van összefüggés, vagy a beszélőkön belüli tempóingadozás is kapcsolatban áll a magánhangzók és mássalhangzók időtartamarányára reagáló mérőszámmal. Ennek megvizsgálásához a beszélők 20-20 tagmondatát elemeztem külön-külön három beszédmódban (szövegfelolvasás, mondatfelolvasás és spontán beszéd). Egy-egy beszélőn belül beszédmódonként összevettem a 20 tagmondatra kapott %V értékeket az ugyanazon tagmondatokra számolt artikulációs tempó értékeivel.

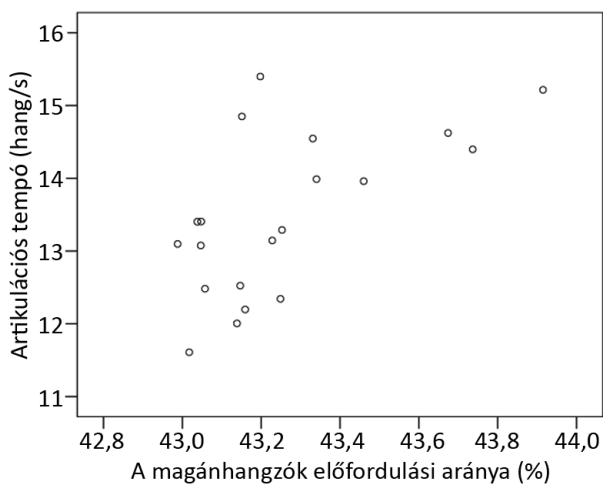
Az eredmények azt mutatták, hogy előfordulnak olyan esetek, amikor egy beszélő lokális artikulációs tempói korrelálnak a %V mérőszámmal, de ez az összefüggés elszórtan, kevésbé szisztematikusan jelenik meg, és az esetek többségében nem mutatott összefüggést a statisztikai próba. Nem volt olyan beszélő, akinél mindhárom beszédmódban korrelált volna az artikulációs tempó és a %V mérőszám. Spontán beszédben mindössze két beszélőnél (F4 és F9) volt szignifikáns az összefüggés (F4 esetében: $\rho = -0,554$, $p = 0,010$; F9 esetében: $\rho = -0,511$, $p = 0,021$). Ugyanezen két beszélőnél szövegfelolvasásban is kimutatható volt a korreláció (F4 esetében: $\rho = -0,695$, $p = 0,001$; F9 esetében: $\rho = -0,444$, $p = 0,050$). Mondatfelolvasásban viszont egyik beszélő esetében sem találtam összefüggést ($p \geq 0,230$). Egyetlen olyan beszélő (N10) volt, akinél az artikulációs tempó és a %V mérőszám mind szövegfelolvasásban ($\rho = -0,629$, $p = 0,003$), mind mondatfelolvasásban ($\rho = -0,468$, $p = 0,038$) összefüggést mutatott. A többi esetben vagy csak a szövegfelolvasásban, vagy csak a mondatfelolvasásban volt korreláció az artikulációs tempó és a %V között a három beszédmód közül. Mondatfelolvasásban összesen csak 4 beszélőnél (F8, N5, N8, N10) volt megtalálható a jelenség ($-0,457 \leq \rho \leq -0,705$, $p \leq 0,043$). Szövegfelolvasásban viszont 8 beszélőnél (F1, F10, N1, N3, N6, N10) korrelált a %V az artikulációs tempóval ($-0,444 \leq \rho \leq -0,695$, $p \leq 0,050$). Az esetek többségében azonban nem sikerült korrelációt kimutatni ($p \geq 0,056$). A %V mérőszám tehát a beszélőkön belül kevésbé függ össze az artikulációs tempóval, mint a beszélők között.

A magánhangzók és a mássalhangzók gyakoriságának és átlagos időtartamának összefüggése a globális artikulációs tempóval

A %V mérőszám – mint láhattuk – korrelációt mutatott a globális artikulációs tempóval. Minél gyorsabban beszélt egy beszélő, annál kisebb volt a magánhangzós szakaszok időtartama a mássalhangzós szakaszok időtartamához képest. Ez az összefüggés többféleképpen is megvalósulhat. Egyrészt előfordulhat, hogy a magánhangzók és a mássalhangzók előfordulásának gyakoriságában térnek el a beszélők. Ha a gyorsabban beszélők több magánhangzót törölnek, szemben a lassabb beszélőkkel, akik minden magánhangzót kiejtenek, akkor ennek következtében a magánhangzók időtartamának aránya a teljes beszéidőhöz képest kisebb lenne a gyorsabban beszélőknél. Másrészt az is elképzelhető, hogy a beszélők a mássalhangzók és a magánhangzók időtartamát tendenciózusan másképp valósíthatják meg. A gyorsabban beszélők a magánhangzók időtartamát vagy jobban rövidítik, mint a mássalhangzókét, vagy

csak a magánhangzókét rövidítik, a mássalhangzókét nem. Ezen kérdéskör alaposabb tanulmányozásához a teljes szövegfelolvasásokat és a spontán monológokat vizsgáltam meg.

Spontán beszédben a beszédmódból fakadóan a beszélők különböző számú szótagot ejtettek ki, ezért az előfordulási gyakoriságokat úgy vizsgáltam meg a %V mérőszám mintájára, hogy kiszámoltam, az összes beszédhang hány százalékában fordult elő magánhangzó. (A magánhangzók darabszámát elosztottam a magánhangzók és a mássalhangzók számának összegével, és megszoroztam százzal.) A magánhangzók darabszámának az összes beszédhanghoz viszonyított aránya nem mutatott korrelációt a %V mérőszámmal spontán beszédben ($\rho = 0,180$, $p = 0,446$). A magasabb, illetve alacsonyabb %V mérőszám tehát nem függött össze azzal, hogy a magánhangzók és a mássalhangzók milyen gyakran fordulnak elő egymáshoz képest egy monológ során. Ez az arány az artikulációs tempóval sem korrelált spontán beszédben ($\rho = 0,128$, $p = 0,591$). Szövegfelolvasásban a magánhangzók darabszámának az összes beszédhanghoz viszonyított aránya szintén nem mutatott összefüggést a %V mérőszámmal ($\rho = -0,089$, $p = 0,710$), az artikulációs tempóval viszont pozitívan korrelált ($\rho = 0,580$, $p = 0,007$). Ez azt jelenti, hogy minél több magánhangzó fordul elő a szövegfelolvasásban a mássalhangzókhoz képest, annál nagyobb volt az artikulációs tempó, tehát gyorsabban beszélt a beszélő. Ez az eltérés azonban nem érintette a %V mérőszám értékeit.



19. ábra

A magánhangzók előfordulási arányának összefüggése az artikulációs tempóval szövegfelolvasásban

A másik tényező, amely a %V és az artikulációs tempó összefüggéseit segíthet jobban megérteni, a beszédhangok időtartama. A magánhangzók és a mássalhangzók időtartama is tendenciózusan eltért a lassabb és a gyorsabb beszélőknél. Beszélőnként külön a kétperces

spontán monológ egészére és külön a teljes szövegfelolvasásra megvizsgáltam a magánhangzók és a mássalhangzók időtartamát. Ezen időtartamok átlagát vetettük össze az adott beszélő globális artikulációs tempójával. A magánhangzók átlagos időtartama erősen és negatívan korrelált az artikulációs tempóval szövegfelolvasásban ($\rho = -0,941, p < 0,001$) és spontán beszédben ($\rho = -0,932, p < 0,001$) is. A mássalhangzók átlagos időtartama szintén erős, negatív korrelációt mutatott az artikulációs tempóval mind spontán beszédben ($\rho = -0,812, p < 0,001$), mind szövegfelolvasásban ($\rho = -0,874, p < 0,001$). Minél gyorsabban beszélt valaki, jellemzően annál inkább kisebb volt a magánhangzók és a mássalhangzók időtartama is a beszédben. A %V mérőszámról viszont tudjuk a korábbi eredmények alapján, hogy minél gyorsabban beszélt egy beszélő, annál kisebb értékeket mutatott, azaz a magánhangzók időtartamaránya a teljes beszédhez képest csökkent. Ez úgy lehetséges, hogy a magánhangzók időtartama nagyobb mértékben csökken, mint a mássalhangzók időtartama a gyorsabb beszélőknél, összevetve a lassabb beszélőkkel. A magánhangzók átlagos időtartama erős, pozitív korrelációt mutatott spontán beszédben ($\rho = 0,808, p < 0,001$) és szövegfelolvasásban is ($\rho = 0,714, p < 0,001$) a %V mérőszámmal. Nem meglepő módon, minél nagyobb volt átlagosan a magánhangzók időtartama, annál nagyobb volt a beszédben a magánhangzók időtartamaránya is. A mássalhangzók esetében azonban nem igazolódtott ennek ellenkezője. A mássalhangzók átlagos időtartama nem függött össze a %V mérőszám értékeivel sem spontán beszédben, ($\rho = 0,177, p = 0,457$) sem szövegfelolvasásban ($\rho = 0,074, p = 0,758$).

Következtetések

A magánhangzós szakaszok időtartamának százalékos aránya a teljes időtartamhoz képest (%V) eredményeim szerint összefüggést mutatott az artikulációs tempóval. Korábban viszont nem sikerült kimutatni jelentős összefüggést a %V mérőszám és az artikulációs tempó között (DELLWO 2010; RUSSO–BARRY 2008). Méréseik alapján a %V mérőszámot robusztusnak tekintették az artikulációs tempóval szemben, mivel a kimutatott összefüggéseket nem sikerült statisztikailag igazolni. A magyar beszédben az általam talált eredmények a klasszikusan szótag-időzítésűnek tartott olaszéhoz hasonlóan (vö. RUSSO–BARRY 2008) azt a trendet mutatták, hogy az artikulációs tempó növekedése jellemzően együtt járt a %V mérőszám csökkenésével. Minél gyorsabban beszélt egy beszélő, annál inkább csökkent a magánhangzós szakaszok (tehát a magánhangzók) időtartamának aránya a teljes beszédidőtartamhoz képest. Hozzá kell tennünk, hogy az eredmények alapján a %V és az artikulációs tempó közti összefüggés nem általános érvényű. A különböző beszélők globális artikulációs tempója összefüggést mutatott a %V mérőszám értékeivel, tehát az, hogy ki milyen lassan vagy gyorsan beszél, kapcsolatban van azzal, hogy mi a magánhangzók és a mássalhangzók időtartamának az aránya, de a beszélőkön belüli tempóingadozás nem mutatott ilyen egyértelmű összefüggést ezzel az aránnyal. Az eredmények azt mutatták, hogy előfordulnak ugyan olyan esetek, amikor egy beszélő lokális tempóértékei korrelálnak a %V mérőszámmal, de ez az

összefüggés elszórtan, kevésbé szisztematikusan jelenik meg, és az esetek többségében nem kimutatható. Nem volt olyan beszélő, akinél mindhárom beszédmódban korrelált volna az artikulációs tempó és a %V mérőszám. A %V mérőszám tehát a beszélőkön belül kevéssé függ össze az artikulációs tempóval. Hangsúlyoznunk kell, hogy beszélőkön belül a %V mérőszám és az artikulációs tempó kapcsolatát számtalan tényező elfedheti, ezért elképzelhető, hogy az egyénenkénti nagyobb tagmondatszámmal vagy célzott anyaggal sikerülne kimutatni szisztematikus összefüggéseket. Jelen anyag alapján azonban csak annyit állíthatunk, hogy a különböző beszélők globális artikulációs tempója van összefüggésben a magánhangzók és a mássalhangzók időtartamának arányával.

Felmerül a kérdés, hogy a mássalhangzók és a magánhangzók időtartamának aránya hogyan tolódhat el a gyorsabb vagy a lassabb artikuláció hatására, mi okozza az arány eltéréseit a különböző tempójú beszélőknél. A magánhangzó-időtartamok ilyen arányú csökkenésének kézenfekvő a magyarázata lehetne, hogy a gyorsabb artikulációs tempó esetében a beszélők hajlamosabbak törölni a magánhangzókat, és ez mutatkozna meg az időtartamarányokban is. Szövegfelolvasásban minél nagyobb volt az artikulációs tempó, tehát gyorsabban beszélt a beszélő, annál több magánhangzó fordult elő a mássalhangzókhoz képest. Spontán beszédben viszont ez a tendencia nem érvényesült. Továbbá a %V mérőszám nem függött össze azzal, hogy a magánhangzók és a mássalhangzók milyen gyakran fordulnak elő egymáshoz képest. Így arra következtethetünk, hogy a törlés, illetve a magánhangzók előfordulásának az aránya a mássalhangzókhoz képest kevéssé játszik szerepet a %V mérőszám és az artikulációs tempó közt feltárt összefüggésben.

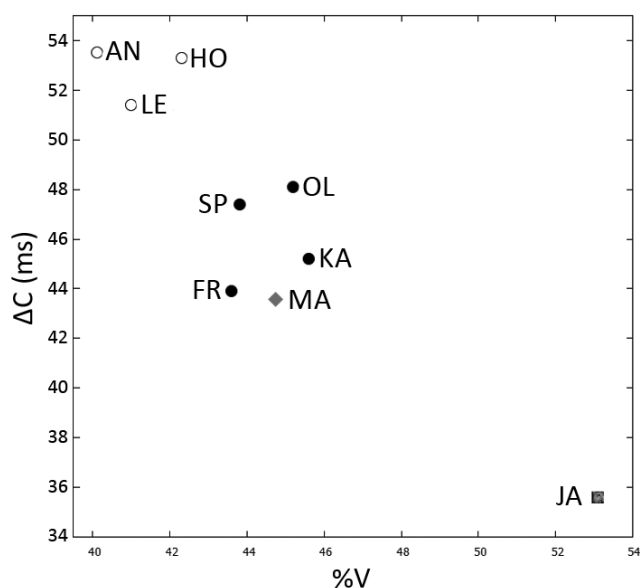
A másik lehetséges kézenfekvő magyarázat, hogy a gyorsabb, illetve lassabb beszélők a mássalhangzók és a magánhangzók időtartamát tendenciózusan másképp valósíthatják meg. Az eredmények nem meglepő módon azt mutatták, hogy minél gyorsabban beszélt valaki, jellemzően annál inkább kisebb volt a magánhangzók és a mássalhangzók időtartama is a beszédben. A gyorsabb beszélők tehát rövidebben ejtik ugyan a magánhangzókat és a mássalhangzókat is, de tendenciózusan jobban rövidítik a magánhangzókat a mássalhangzóknál. Így jöhet létre a magánhangzók és mássalhangzók időtartamarányának eltolódása gyorsabb vagy lassabb beszédben. Az eredmények azt is feltárták, hogy minél nagyobb volt átlagosan a magánhangzók időtartama, annál inkább nagyobb volt a beszédben a magánhangzók időtartamaránya is. A mássalhangzók átlagos időtartama viszont nem függött össze a %V mérőszám értékeivel sem spontán beszédben, sem szövegfelolvasásban. A magánhangzók időtartama flexibilisebben változtatható, mint a legtöbb képzésmódú mássalhangzóé (vö. HOFHUIS–GUSSENHOVEN–RIETVELD 1995), így nem meglepő, hogy a különböző beszélők következetesebben tudják variálni a magánhangzók időtartamát, mint a mássalhangzókéét. Elsősorban tehát a magánhangzók rövidebb vagy hosszabb ejtése áll annak hátterében, hogy a gyorsabb beszélőknél a magánhangzó időtartamaránya a mássalhangzókéhoz képest kisebb, mint a lassabb beszélőknél.

4.1.3. A beszédritmus-mérőszámok eredményei alapján a magyar beszédre levonható általános következtetések

A korábbi alfejezetekben láthattuk, hogy a magyar beszédben a beszédritmus időzítésbeli dimenzióját a különböző tényezők hogyan befolyásolják. Összefüggéseket találtunk a különböző beszédmodok (felolvasás és spontán beszéd), a nemek, a különböző beszélők szünettartási szokásai és a beszédritmus között 20 beszélő beszéde alapján. A következőkben áttekintjük a magyar beszédre kapott beszédritmusmérőszám-értékeket, és összehasonlítjuk a szakirodalomban szereplő, más nyelvekre kapott eredményekkel.

Az eredmények alapján elmondhatjuk, hogy a magyar beszéd beszédritmusa a szótag-időzítésű nyelvekéhez hasonló tulajdonságokat mutat mindkét ismert koordináta-rendszer (%V függvényében a ΔC és a PVI mérőszámok) alapján. A kétféle módszer anyaga és módszertana valamelyest eltér egymástól. RAMUS, NESPOR és MEHLER (1999) különböző mondatok felolvasására számolta ki a beszédritmus-mérőszámokat, és azok nyelvenként vett átlagát közölte. GRABE és LOW (2002) viszont több mondatból álló, felolvasott szövegekre számolta ki a PVI mérőszámokat. Az előbbi eljárással kapott adatokat a magyar nyelvben a felolvasott mondatokra kapott értékek átlagával, az utóbbit pedig a teljes szövegfelolvasásra kapott eredményekkel vetjük össze.

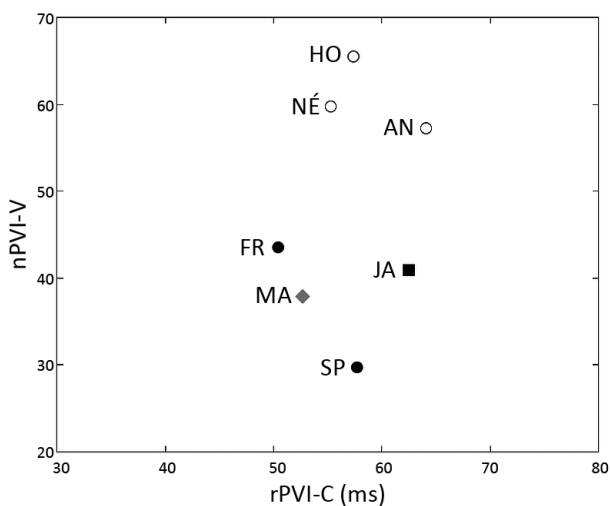
Elsőként a RAMUS, NESPOR és MEHLER által felállított koordináta-rendszerben helyeztem el a magyar nyelvre kapott értékeket (20. ábra). A ΔC -t a %V függvényében ábrázolva egyértelműen látszik, hogy a magyar nyelvben a mondatfelolvasásokra kapott értékek a klasszikusan szótag-időzítésű nyelvekéhez hasonló értékeket mutatnak. A magánhangzó-intervallumokon alapuló %V érték (44,7) magasabb, mint a tipikusan hangsúly-időzítésű nyelvek esetében (40,1–42,3), és alacsonyabb a tipikusan moraidőzítésű japán nyelvnél (53,1). A magyar nyelvre mért ΔC átlagos értéke (43,6 ms) szintén a hagyományos szótag-időzítésűnek tekintett nyelvekéhez közelít (43,9–47,5 ms). A klasszikusan hangsúly-időzítésűekhez sorolt nyelveknél (51,4–53,5 ms) viszont alacsonyabb, a japán beszédritmusához (35,6) képest pedig magasabb értéket vesz föl. Szintén érdekes lehet a magánhangzós szakaszok időtartamának a szórása a nyelvek közti összehasonlításban. A magyarra mért átlagos ΔV értéke (30,6 ms) szintén a szótag-időzítésű nyelvek beszédritmusához hasonló (33,2–40,1 ms), ugyanakkor alacsonyabb a hangsúly-időzítésű nyelvek beszédritmusánál (42,3–46,4 ms).



20. ábra

Klasszikusan ○ a hangsúly-időzítésűekhez, ● a szótag-időzítésűekhez és ■ a moraidőzítésűekhez sorolt nyelvekre kapott eredmények (RAMUS–NESPOR–MEHLER 1999), kiegészítve a ◆ magyar nyelvre kapott eredménnyel
(AN = angol, HO = holland, OL = olasz, FR = francia, SP = spanyol, KA = katalán, JA = japán, LE = lengyel, MA = magyar.)

A GRABE és LOW (2002) által felállított koordináta-rendszerben is elhelyeztem a magyar beszédre kapott értékeket (21. ábra). Az rPVI-C-t az nPVI-V függvényében ábrázolva szintén egyértelműen látszik, hogy a magyar nyelvben a teljes szövegfelolvasásokra kapott értékek átlaga a szótag-időzítésű nyelvekéhez hasonló értékeket mutat. [A magyar nyelvre kapott eredményt csak a klasszikusan szótag-időzítésűekhez és a hangsúly-időzítésűekhez sorolt nyelvekkel vettem össze, mivel a moraidőzítésű japán nyelvre kapott eredmények ebben a megközelítésben nem különülnek el a klasszikusan szótag-időzítésűnek tekintett nyelvekétől (részletesen lásd GRABE–LOW 2002).] A magyarra mért átlagos nPVI-V értéke (37,9) a klasszikusan szótag-időzítésűekhez sorolt nyelvek beszédritmusához hasonló (29,7–43,5), ugyanakkor alacsonyabb a hangsúly-időzítésű nyelvekre mért átlagos adatoknál (57,2–65,8). A magyarra mért rPVI-C átlagos értéke 52,6 ms, ami alacsonyabb a tipikusan hangsúly-időzítésű nyelvek értékeinél (55,3–64,1 ms); a klasszikusan szótag-időzítésűekhez sorolt francia (50,4 ms) és spanyol (57,7 ms) nyelv között helyezkedik el. Hozzá kell tennünk, hogy az rPVI-C tekintetében a klasszikusan szótag-időzítésűekhez sorolt nyelvek és a hangsúly-időzítésűekhez sorolt nyelvek átfedést mutatnak, de a két PVI mérőszám együttesen nyújt iránymutatást a különböző nyelvek beszédritmusbeli sajátosságairól.



21. ábra

Klasszikusan ○ a hangsúly-időzítésűekhez, ● a szótag-időzítésűekhez és ■ a moraidőzítésűekhez sorolt nyelvekre kapott eredmények (RAMUS–NESPOR–MEHLER 1999), kiegészítve a ◆ magyar nyelvre kapott eredménnyel (AN = angol, HO = holland, NÉ = német, FR = francia, SP = spanyol, JA = japán, MA = magyar.)

A klasszikus beszédritmusosztály-elmélet alapján a magyar nyelvről többen feltételezték, hogy alapvetően a szótag-időzítésű nyelvek közé tartozik (GÓSY 2000; SIPTÁR–TÖRKENCZY 2000; VARGA 2002, 2005). Akusztikai méréseink bizonyos értelemben megerősítik ezt a hipotézist. A magyar beszédben található időzítésbeli sajátosságok egy része a klasszikusan szótag-időzítésű (spanyol, francia) nyelvek sajátosságaihoz hasonló. Ez azonban nem azt jelenti, hogy a szótagok időtartama közel azonos lenne a beszédben, és ez határozná meg az időzítést. A korábbi, magyar nyelvre végzett akusztikai mérések nem is támasztották alá a szótagok időtartamának egyenlőségét, ehelyett nagy változatosságukat sikerült kimutatni (KECSKÉS 1966; GÓSY 2000). Egészen másfajta időzítésbeli sajátosságok alapján mutatható ki a magyar nyelv szótag-időzítésű nyelvekhez való hasonlósága. Egyrésztől meghatározó a magánhangzók és a mássalhangzók időtartamának aránya a beszédben, másrésztől, hogy mennyire változatos időtartamban fordulnak elő a mássalhangzók és a mássalhangzó-torlódások. Továbbá fontos tulajdonság, hogy a folyamatos beszédben az egymást követő szakaszok (elsősorban a mássalhangzós és magánhangzós szakaszok) időtartama mekkora variabilitást mutat (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002). A magyar nyelv ezen időzítésbeli tulajdonságok alapján hasonlít a klasszikusan szótag-időzítésűnek tekintett nyelvekhez. Nyilvánvalóan ezeket a tulajdonságokat meghatározzák olyan tendenciák, hogy a magyar beszédben a beszédhangok időtartamát befolyásoló különféle tényezők egy része nem vagy kevésbé fejti ki hatását.

Kimutatták például, hogy a hangsúly időtartamra gyakorolt hatása, illetve a magánhangzó-redukció jelensége nem vagy kevésbé erőteljesen jelenik meg a klasszikusan szótag-időzítésűekhez sorolt nyelvekben, mint a hangsúly-időzítésűekhez soroltakban (BARRY–ANDREEVA 2001). Ilyen tényező lehet például a frázis végi nyúlás időtartamra gyakorolt hatása is, ami az angol nyelvben nagyobb volt, mint a katalán és a spanyol nyelvekben (PRIETO et al. 2012). De befolyással lehet a szótagtípusok előfordulási aránya is (DAUER 1983; RAMUS–NESPOR–MEHLER 1999; PRIETO et al. 2012). Igen összetett tehát, hogy mennyi minden határozhatja meg a nyelvek beszédritmusának és időzítésének különbségét. Az időtartamokban is mérhető különbségek egyrésztől nem feltétlen jelentősek, másrésztől nem biztos, hogy minden beszélőnél, mindenféle beszédmódban realizálódnak (vö. ARVANITI 2012). A korábbi fejezetek alapján a magyar beszédben is tapasztalható volt a beszélők közti nagy szórás és eltérések a beszédmódok (első-sorban a felolvasás és a spontán beszéd) között. A beszédritmus-mérőszámok alapján a különböző nyelvek átfedést mutatnak egymással, a különböző típusú nyelvek nem különülnek el élesen egymástól (RAMUS–NESPOR–MEHLER 1999; GRABE–LOW 2002; WHITE–MATTYS 2007; ARVANITI 2012). Sokkal inkább skálákat hoznak létre, amelyeken elhelyezhetők a különböző nyelvek egymás viszonylatában. A mérőszámok nem alkalmazhatók beszédritmusosztályok vagy -csoportosulások szigorú elkülönítésére, csak egyfajta iránymutatók lehetnek a nyelvek közti beszédritmusbeli vagy időzítésbeli különbségek mérésében. Ennek értelmében az általunk vizsgált két koordináta-rendszer 20 köznyelvi beszélőtől származó adatai azt mutatják, hogy a magyar nyelv a szótag-időzítésű nyelvekéhez hasonló tulajdonságokkal rendelkezik, a felállítható beszédritmusskálán ezen nyelvekhez áll a legközelebb.

4.2. A lassulások és a gyorsulások a magyar beszédben

Az előző fejezetekben a beszédritmus-mérőszámokkal vizsgáltuk meg a magyar beszéd időzítésbeli sajátosságait. Ezek a mérőszámok (ΔV , ΔC , VarcoV, VarcoC, nPVI-V és nPVI-C) számszerű értéket rendelnek a folyamatos beszédben megjelenő tempóingadozásokhoz, ritmusváltásokhoz. A mérőszámok ugyanakkor nem tesznek különbséget a lassulások és gyorsulások között, hanem csak az időtartamokban bekövetkező változások, változatosság mértékét mérik. Ebben a fejezetben kísérletet teszek a beszédben megjelenő tempóingadozások jellemzésére oly módon, hogy külön kezeltem a lassuló, illetve gyorsuló szakaszokat. A vizsgálatok során megpróbálok meghatározni, hogy a beszédet általános lassuló vagy gyorsuló tendencia jellemzi. A kutatás további célja, hogy feltárja a lassulások és a gyorsulások általános sajátosságait és kapcsolatát más tényezőkkel. Ezen kérdések megválaszolásához egy olyan eljárást dolgoztam ki, amely egy adott beszédszakasról meghatározza valamilyen módon, hogy jellemzően lassuló vagy gyorsuló trendet mutat-e, továbbá a lassulásokhoz és gyorsulásokhoz valamilyen mértékét tud rendelni.

4.2.1. Kísérleti személyek, anyag és módszer

Kísérleti személyek és anyag

A kutatáshoz azon kísérleti személyek anyagait elemeztem, amelyeket a 3.1. fejezetben már ismertettem. Röviden a fontosabb adatok a következők voltak: 20 magyar anyanyelvű beszélő (10 nő és 10 férfi) spontán beszédét és szövegfelolvasását elemeztem a BEA adatbázisból (Gósy 2008). Életkoruk 20 és 60 év között mozgott (a férfiak átlagéletkora: 41,6 év, a nőké: 38,0 év). Az adatközlők ép hallók és ép beszédűek voltak. Interjúhelyzetben rögzített spontán beszédükből beszélőnként legalább 2 percnyi részletet választottam ki, amelyben az interjúztató nem szólalt meg. A beszélők saját munkájukról vagy éppen folyó tanulmányaikról monologikusan beszéltek a felvételeken. Továbbá megvizsgáltam mindegyik beszélő szövegfelolvasását is. Ugyanazon szöveg felolvasása a különböző beszélők esetében eltérő időtartamban valósult meg, a legrövidebb szövegfelolvasás 107,7 másodperc, a leghosszabb 159,7 másodperc volt.

A hanganyagok annotálása, az akusztikai elemzésben használt módszerek

A hanganyagok hangszintű annotálását, ahogy korábban részletesen kifejtettem (a 3.2. fejezetben), a nemzetközi szakirodalomban található kritériumoknak megfelelően végeztem el, figyelembe véve a magyar beszédben lévő sajátos akusztikai realizációkat. Először a MAUS szoftverrel automatikusan, hangszinten annotáltam az anyagot (SCHIEL 1999, KISLER–SCHIEL–SLOETJES 2012), majd a Praat 5.1 szoftverben (BOERSMA–WEENINK 2009) történo kézi ellenőrzés után egy C++ nyelven írt program segítségével meghatároztam a mássalhangzós és a magánhangzós szakaszokat, illetve ezek időtartamát.

Statisztikai elemzések

A temporális változások vizsgálatához az ún. lépéssziszatika módszert alkalmaztuk (GYÜRE–BARTOS–JÁNOSI 2007). A korábban más tudományágban használt módszer célja, hogy hosszú idősorok sziszimetriáját vagy asziszimetriáját felfedje. Az eredetileg hőmérsékletingadozás vizsgálatára használt eljárást alkalmaztuk a beszéd folyamatban lévő lassulások és gyorsulások arányának feltárására.

A módszer a következőképpen használható temporális változások kimutatására a beszédben: az egymást követő egységek (beszédhangok, magánhangzós szakaszok és mássalhangzós szakaszok) időtartamának különbségét rendre kiszámoljuk. Minden egyes különbség egy-egy ún. lépést jelent az egymás mellett lévő egységek között. Amennyiben a követő egység időtartama hosszabb, mint a megelőző egységé, akkor a különbséget növekvő lépésnek tekintjük, mivel ez egyfajta lassulást jelez. Amennyiben a követő egység időtartama rövidebb, mint a megelőző egység időtartama, akkor ez felfogható gyorsulásnak, ezért csökkenő

lépésként kezeljük a továbbiakban. A 14. táblázat azt szemlélteti, hogy az egymást követő egységek időtartamának különbségeit hogyan minősítjük növekvő vagy csökkenő lépéseknek. A példában az alapegységek az egymást szükségszerűen felváltva követő mássalhangzós szakaszok és magánhangzós szakaszok (CV-szakaszok). Az első egység egy magánhangzós szakasz, amelynek időtartama (50 ms) hosszabb a követő mássalhangzós szakasz időtartamánál (40 ms), tehát az első egység időtartamához képest a következő egység rövidebb, így az első lépés típusa csökkenő lesz. A második egység (mássalhangzós szakasz) időtartamához képest a harmadik egység (magánhangzós szakasz) időtartama hosszabb, így az egy növekvő lépés lesz. A kategorizálás után megszámloljuk külön a növekvő (N_I) lépések és külön a csökkenő lépések (N_D) darabszámát. A növekvő lépések darabszámát elosztjuk a csökkenő lépések darabszámával, tehát a kettő aránya (N_I/N_D) adja az ún. lépésszámarányt. A ms-ban mért időtartamokat kilenc tizedesjegyre vettem figyelembe, és a teljesen megegyező időtartamokat (3 darab) a növekvő időtartamokhoz számítottam. Mivel a különbségek nagyságukat tekintve rendkívül kicsik is lehetnek, ezért önmagában a lépésszámarányok nem szolgálnának információval a trendek relevanciájáról. Ezért nemcsak a lépések számát, hanem a nagyságát (az időtartam-különbségek abszolút értékét) is figyelembe vesszük. Ehhez kiszámoljuk külön a növekvő lépések átlagos nagyságát ($\langle \Delta t_I \rangle$), valamint a csökkenő lépések átlagos nagyságát ($\langle \Delta t_D \rangle$). A növekvő lépések átlagos nagyságát elosztjuk a csökkenő lépések átlagos nagyságával, tehát a kettő aránya határozza meg az átlagos lépésnagyságarányt ($\langle \Delta t_I \rangle / \langle \Delta t_D \rangle$). Az átlagolt szakasz hosszának a teljes szöveget tekintettem, illetve összevetésképpen megvizsgáltam külön a tagmondatokra kapott értékeket is.

14. táblázat

*Példa a csökkenő és növekvő lépések meghatározására
(I_v a magánhangzós szakaszt, I_c a mássalhangzós szakaszt jelöli.)*

Egymást követő alapegységek	I_v	I_c	I_v	I_c
Mért időtartam (ms)	50	40	60	45
A különbségek	-10	+20	-15	
Lépés típusa	N_D	N_I	N_D	

Az átlagos lépésnagyságarány és a lépésszámarány ábrázolható egy koordináta-rendszerben (22. ábra). Ez a két arány olyan paramétersíkot jelöl ki, amelyben jellemezhetővé válik a gyorsulások és lassulások számának és mértékének egymáshoz viszonyított aránya. Ha mindkét arányszám pontosan 1, akkor az idősor tökéletesen szimmetrikus. Ez azt jelenti, hogy ugyanannyi növekvő és csökkenő lépés található a vizsgált szakaszon, tehát számuk megegyezik, és a növekvő és csökkenő lépések átlagos nagysága is tökéletesen megegyezik. Sem lassulásról, sem gyorsulásról nem beszélhetünk, akár a lépések számát, akár a lépések nagyságát vesszük figyelembe. A lassulás és a gyorsulás másképp is lehet kiegyenlített. Elképzelhető olyan eset,

amikor a lépések száma kompenzálja a lépések mértékét, és fordítva. Például ha egy adott szakaszon kétszer annyi növekvő lépés fordul elő, mint csökkenő lépés, de a növekvő lépésekből csak feleannyi van, mint csökkenő lépésekből, akkor a növekvő lépések és a csökkenő lépések egyfajta kiegyenlítetttségéről beszélhetünk. Hiszen a növekvő lépések nagy mennyiségét ellensúlyozni tudja a csökkenő lépések nagysága, tehát sem lassulásról, sem gyorsulásról nem beszélhetünk, egyik oldalra sem billen a mérleg nyelve. Értelmezve ezt a jelenséget a koordináta-rendszeren meghatározhatunk egy olyan görbét, amelynek pontjai esetében igaz az, hogy a gyorsulások és a lassulások kiegyenlítik egymást, mértékük és gyakoriságuk fordítottan arányos. Másképpen megfogalmazva: azon pontok halmaza a paramétersíkon, amelyek a lépésszám- és lépésméretarányok olyan viszonyával jellemezhetők, hogy csökkenő vagy növekvő trend nélküli idősort írnak le, azok a $\langle \Delta t_1 \rangle / \langle \Delta t_D \rangle = N_D / N_1$ görbén (azaz egy $y = 1/x$ jellegű hiperbolán) helyezkednek el. A paramétertér ezen tartományát stacionárius görbének nevezzük. A 22. ábrán a stacionárius görbét vastag fekete vonallal jelöltem. A görbe fölött elhelyezkedő pontok lassulást mutatnak, míg a görbe alatt elhelyezkedő pontokra a gyorsulás jellemző. Az alábbiakban végigvesszük a különböző lehetséges eseteket.

Az *A* esetben a több növekvő lépés kompenzálja a kevesebbszer előforduló, ám átlagosan nagyobb csökkenő lépéseket, tehát a lassulás és a gyorsulás kiegyenlítettnek tekinthető. A *B* esetben a sokkal gyakoribb növekvő lépéseket a csökkenő lépések nagysága már nem tudja ellensúlyozni, így az ezen a területen szóródó adatok lassuló trendet mutatnak. Amennyiben mindkét mérőszám nagyobb egynél (*C* eset), akkor a növekvő lépések gyakorisága és nagysága is nagyobb a csökkenő lépéseknél, tehát ez esetben egyértelmű lassuló trendről beszélhetünk. Előfordulhat az is, hogy a lépésszámarány kisebb 1-nél, de a lépésméretarány nagyobb, mint 1 (*D*, *E*, *F* eset). *D* esetben az átlagosan nagyobb növekvő lépéseket a gyakoribb csökkenő lépések nem tudják kompenzálni, így itt lassulás következik be. Fordított a helyzet az *F* esetben, ahol a gyorsulás a jellemző, ugyanis a gyakoribb csökkenő lépéseket nem tudja ellensúlyozni az átlagosan nagyobb növekvő lépések. *E* esetben a gyakoribb, ámde átlagosan kisebb csökkenő lépések kompenzálják a ritkább, de átlagosan nagyobb növekvő lépéseket, tehát hasonlóan az *A* esethez, itt sincs sem lassuló, sem gyorsuló trend. A *G* esetben egyértelmű gyorsulás figyelhető meg, mert a csökkenő lépések átlagos mértéke és gyakorisága is nagyobb a növekvő lépéseknél. *H* esetben az átlagosan nagyobb csökkenő lépéseket a gyakoribb növekvő lépések nem tudják kompenzálni, ezért itt is gyorsulásról beszélhetünk. A leírt módszer tehát alapvetően három nagy csoportot hoz létre: a lassuló trendhez tartozók, a gyorsuló trendhez sorolhatók és a trend nélküliek csoportját.

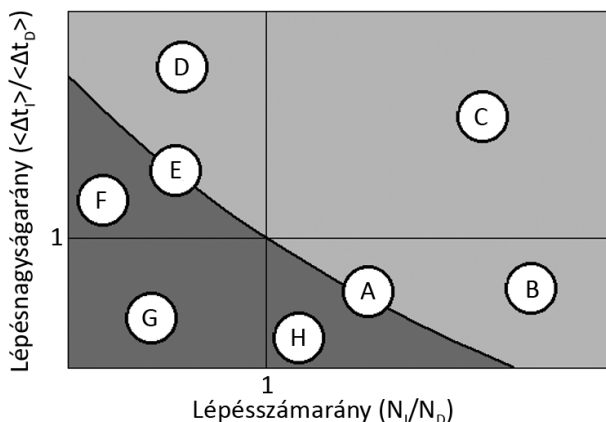
Az elsődleges kérdés az volt, hogy a beszédben mérhető adatok vajon mutatnak-e valamilyen trendet, tehát eltérnek-e jelentősen a lassulás vagy gyorsulás irányába. A kiszámolt lépésszámstatisztikai módszer csak azt mutatja meg, hogy milyen trend jellemzi a vizsgált beszéd-részletet, de az eltérés jelentőségét nem adja meg. Ennek meghatározásához a következő eljárást választottam: mindenekelőtt nullhipotézisként feltételeztem, hogy az egymást követő időtartamadatok (beszédhangok vagy mássalhangzós, magánhangzós szakaszok) sorrendjének jelentősége van, struktúrárt alkotnak. Ha megváltoztatjuk az időtartamadatok sorrendjét,

akkor vélhetően megbontjuk a felépítés logikáját is. Ezért először megváltoztattam a mért időtartamadatok sorrendjét, véletlenszerűen összekevertem egy nyílt forráskódú C++ programmal (VINCZE 2014). A program LCG (Linear Congruential Generator) algoritlussal (PRESS et al. 2002) egyenletes eloszlású, kvázi véletlen számokat generál, amelyeket hozzárendel a meglévő idősorokhoz. Így az idősorok permutációit állítja elő, azaz ugyanazokat az értékeket más sorrendbe rendezi. Ezzel a módszerrel hoztam létre kevert idősorokat, amelyekre szintén kiszámoltam a lépésstatisztikát. Mivel a kevert adatsorok az időtartamok sorrendjének a lehetséges kombinációit mintázták, ezért többször is elvégeztem a keverést. Tagmondatonként tíz kevert adatsort hoztam létre, hogy meggyőződjek arról, nem csak véletlen egy-egy kevert adatsorra kapott lépésstatisztikai eredmény. A kevert adatsorokra számolt lépésstatisztikát összevettem a ténylegesen mért adatokra számolt lépésstatisztikával, hogy mennyiben és milyen sajátosságokban térnek el egymástól.

Elsőként meghatároztam a kevert adatsorokra kapott lépésstatisztikai eredmények stacionárius görbéhez viszonyított eltérését. Megbecsültem az euklideszi távolságot minden kevert adatsorra számolt lépésstatisztika esetében, és kiszámoltam a szórását, azaz az átlagos négyzetes eltérést (σ), majd a stacionárius görbét mindkét irányban eltoltam ezzel az értékkel. A szórás kétszeres, illetve háromszoros értékét is kiszámoltam, és ezen értékekkel is eltoltam a stacionárius görbét mindkét irányba. Így olyan három sáv jött létre, amellyel meghatározható, hogy a mért adatpontok lassulásának vagy gyorsulásának mértéke mennyire jelentős. Amennyiben valamely adatpont koordinátái a görbét leíró $\langle \Delta t_i \rangle / \langle \Delta t_d \rangle$ értékek $\pm \sigma$ szélességű környezetébe estek, akkor a stacionárius görbétől való eltérést nem tekintettem jelentősnek. Önmagában egy-egy adatpont elhelyezkedése ezen a sávon belül tekinthető véletlennek, ezért sem jelentős lassulást, sem jelentős gyorsulást nem feltételeztem ezekben az esetekben. Az e sávon kívül eső adatpontokat tekintettem úgy, hogy önmagukban biztosan valamilyen trenddel (gyorsulás vagy lassulás) rendelkeznek. A két, illetve három szórás által meghatározott sávon kívül eső adatpontok esetében pedig egyértelműen jelentős lassulásról vagy gyorsulásról beszélhetünk.

A kevert adatsorok lépésstatisztikai eredményeit összehasonlítottam a ténylegesen mért adatokra kapott eredményekkel is. Összevettem a mért és a kevert adatsorok lépésszám-arányát és a lépésméretarányát is. Továbbá kiszámoltam az adatpontok euklideszi távolságát a stacionárius görbétől, hogy mennyire távolodnak el a mért és a kevert adatsorokra kapott adatpontok, hogy vajon ugyanolyan mértékben gyorsulnak vagy lassulnak a tagmondatnyi egységek. Az euklideszi távolságokat előjellel is elláttam. Ha a stacionárius görbe fölött helyezkedett az adatpont, akkor pozitívnak tekintettem, ha alatta, akkor negatívnak. Pozitív előjelet tehát a lassuló tagmondatnyi egységek euklideszi távolsága kapott, negatívát pedig a gyorsulóké. Ily módon nemcsak a távolságok mértékét tudtam összehasonlítani, hanem azt is, hogy az adatpontok szimmetrikusan rendeződnek-e el a stacionárius görbe mentén, vagy található-e eltolódás a lassulás vagy a gyorsulás felé. A kevert adatsorok lépésszám-arányát, lépésméretarányát, euklideszi távolságát, illetve előjeles euklideszi távolságát a stacionárius görbétől összehasonlítottam a ténylegesen mért adatokra kapott eredményekkel a Wilcoxon- és

a Mann–Whitney-próba segítségével. A nemparaméteres próbák használatát az adatok nem normális eloszlása indokolta (22. ábra), és több esetben az adatok nem arányskála jellege is indokoltta tette. A statisztikai elemzéseket az SPSS 20.0 program segítségével végeztem, az ábrák elkészítéséhez pedig a Gnuplot 4.6-os verzióját használtam.



22. ábra

Szemléltető ábra a lépésstatistikai módszer lényegéről; lassuló trendet mutató adatok tartománya (világosszürke), gyorsuló trendet mutató adatok tartománya (sötétszürke), stacionárius görbe (fekete vonal)

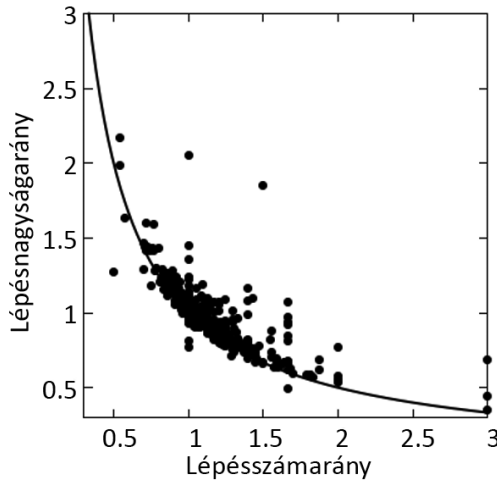
A lépésstatistikát többféle alapegységre alkalmazhatjuk. Az egységek, amelyeknek időtartam-különbségeit kiszámoljuk, lehetnek az egymást követő magánhangzós és mássalhangzós szakaszok, ahogyan a példában (14. táblázat) is szerepelt. Mint korábban láttuk, a magánhangzós szakaszokra, valamint a mássalhangzós szakaszokra épülő mérőszámok olykor eltérő eredményekre vezettek, a két szakasztípust tehát érdemes külön-külön is vizsgálni. A lépésszámokat és lépésméreteket tehát kiszámoltam csak az egymást követő magánhangzós szakaszokra is, és külön ugyanezt megtettem az egymást követő mássalhangzós szakaszok esetében is. A beszédritmus időbeli dimenziójának vizsgálatakor a magánhangzós és a mássalhangzós szakaszok ugyan hasznos alapegységnek bizonyultak korábban, de elképzelhető, hogy az időzítés jobban megragadható a beszédhangok időtartamai alapján. Ezért a beszédhangok időtartamaira is alkalmaztam az eljárást.

A lépésstatistika különböző nagyságú nyelvi egységekre is kiszámolható. Tetszőleges hosszúságú beszédrészletek választhatók, mint például tagmondat vagy teljes szövegfelolvasás, esetleg teljes monológ. Mindegyik típusra kiszámoltam a lépésstatistikát, és összevetettem a kapott eredményeket.

4.2.2. Eredmények⁵

4.2.2.1. Tempóbeli változások általános trendje a szövegfelolvasások tagmondataiban

A szövegfelolvasások tagmondatainak tempóingadozását vizsgáltam meg először a lépésstatisztikai módszerrel. A tagmondatokra jellemző trend meghatározásához a lépésnagyságarányt és a lépésszámárányt egymás függvényében ábrázoltam, ami a 23. ábrán látható. A lépésstatisztika alapegységeként először a magánhangzós és a mássalhangzós szakaszokat (CV-szakaszok) együttesen vizsgáltam.



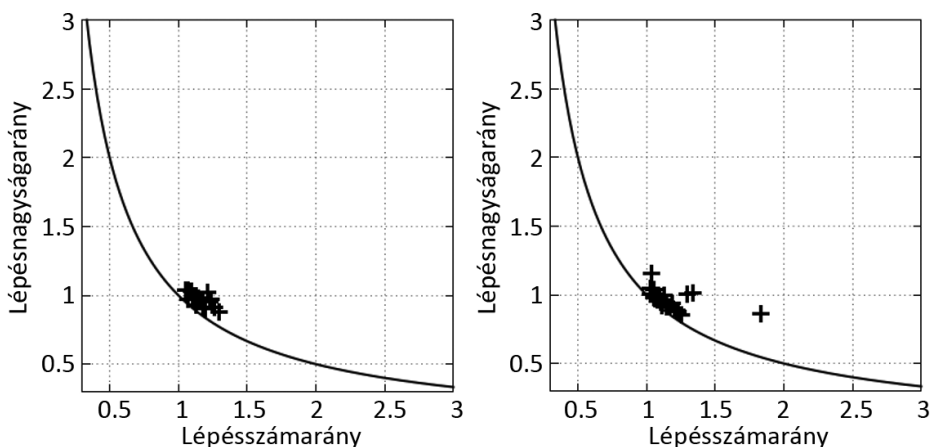
23. ábra

A CV-szakaszok tagmondataira kapott lépésstatisztikai eredmények

A tagmondatok igen nagy szórást mutatnak a lépésstatisztika alapján. Vannak olyan tagmondatok, amelyeket inkább a lassulás, másokat inkább a gyorsulás jellemez (23. ábra). Az összes tagmondatot figyelembe véve azonban kirajzolódik egy általános lassuló trend. Az egyes beszélők tagmondatainak átlagai is ezt támasztják alá, az átlagokat jelölő adatpontok mindegyike ugyanis az ún. stacionárius görbe fölött helyezkedik el (24. ábra). A növekvő lépések száma több, mint a csökkenő lépések száma, tehát a lépésszámárány nagyobb 1-nél. A beszélőnkénti átlagértékek a lépésszámárány esetében 1,057 és 1,300 között mozognak. A beszélőnkénti átlagértékek lépésnagyságaránya 1 körül szóródik (0,876–1,041). A csökkenő lépések átlagos nagysága tehát hol nagyobb, mint a növekvő lépéseké, hol kisebb. Kétféle esetről van tehát szó. Az egyik esetben mind a lépésszámárány, mind a lépésnagyságarány nagyobb 1-nél, tehát egyértelműen lassuló trendről van szó. A másik esetben a növekvő lépések

⁵ Ezen alfejezet alapjául több korábbi tanulmányom (KOHÁRI 2013c, 2014b) szolgált.

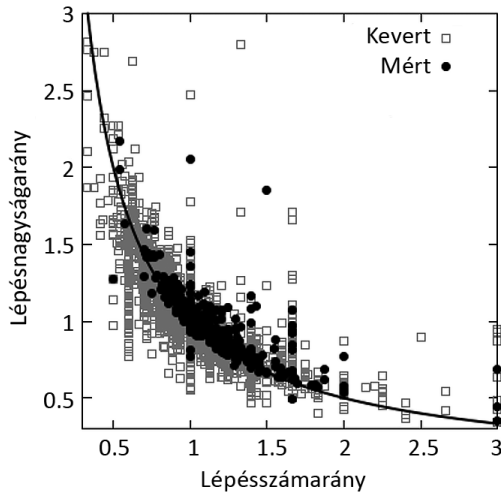
száma nagyobb, mint a csökkenőké, de a csökkenő lépések nagysága meghaladja a növekvő-két. A növekvő lépések nagysága ezekben az esetekben hiába kisebb átlagosan a csökkenő lépéseknél, gyakoribb előfordulásuk elbillenti a mérleget a lassulás felé. A szövegfelolvasás tagmondatainak többsége tehát a beszélőnkénti átlagok alapján inkább lassuló tendenciát mutat. Ugyanezt a mintázatot láthatjuk, ha a tagmondatonként kapott átlagértékeket vizsgáljuk meg. Ezen értékek mindegyike a stacionárius görbe fölött helyezkedik el, azaz lassuló trendet mutatnak. A különböző tagmondatoknál a lépésszamarány nagyobb 1-nél (1,026–1,833). A növekvő lépések száma tehát általánosságban több, mint a csökkenő lépések száma. A tagmondatonkénti átlagértékek lépésmagyságaránya 1 körül szóródik (0,855–1,155). A csökkenő lépések átlagos nagysága hol nagyobb, hol kisebb, mint a növekvő lépéseké. Hasonló a mintázat, mint a beszélőkre kapott átlagok esetében. Ahol a növekvő lépések nagysága nagyobb, és száma is nagyobb, ott mindkét paraméter szerinti lassulásról beszélhetünk. A növekvő lépések nagysága olykor hiába kisebb átlagosan a csökkenő lépéseknél, gyakoribb előfordulásuk ezekben az esetekben is elbillenti a mérleget a lassulás felé, mivel minden tagmondat átlaga a stacionárius görbe fölött helyezkedik el.



24. ábra

A CV-szakaszok tagmondataira kapott lépésstatistikai eredmények beszélőnként (baloldalt) és tagmondatonként (jobbaldalt) kapott átlagai

Felmerül a kérdés, hogy az eddigi lépésstatistikai eredmények mennyiben tekinthetők valóban jelentős eltérésnek a stacionárius görbétől. Ehhez a mért időtartamokat sorrendjét véletlenszerűen összekevertem, majd kiszámoltam az így kapott idősorokra a lépésstatistikát, és összevettem a mért adatokra számolt lépésstatistikai eredményekkel. A 25. ábrán látható, hogy a mért adatsorokra kapott eredmények eltérnek a kevert adatsorokra kapott eredményektől; nincsenek fedésben egymással.

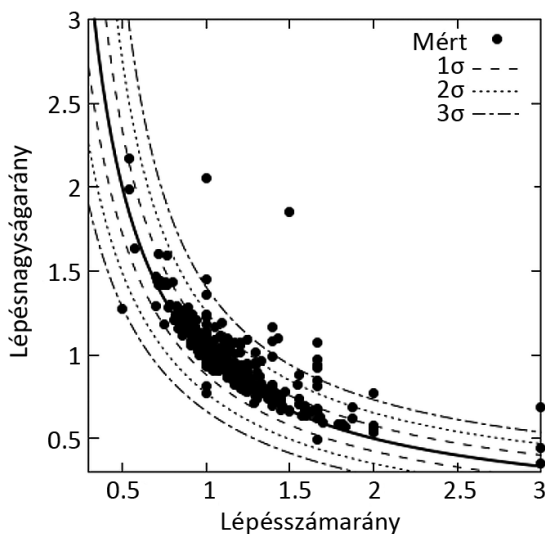


25. ábra

A szöveg tagmondataira számolt lépésstatistikai eredmények a CV-szakaszok alapján összehasonlítva kevert megfelelőikkel

A kevert adatsorok eredményei alapján meghatározható az egyes pontok távolsága a stacionárius görbétől. Az összes kevert tagmondatra kapott adatpontok távolsága alapján meghúzhatók különböző sávok, amelyek a görbétől való eltérések jelentőségét mutatják (lásd részletesen a 4.2.1. fejezetben). A sávokat úgy határoztam meg, hogy először kiszámoltam az adatpontok stacionárius görbétől lévő távolságának egyszeres, kétszeres, illetve háromszoros szórását (σ , 2σ , 3σ), majd a görbét mindkét irányba eltoltam ezekkel az értékekkel. Az így kapott görbék a 26. ábrán láthatók. A mért adatpontra számolt lépésstatistikai eredmények egy jelentős része a kevert adatsorra számolt eredmények alapján meghúzott sávokon belül helyezkedik el. Ha a kevert adatsorok görbétől való egységnyi szórását ($\sigma = 0,061$) vizsgáljuk, akkor a felolvasott 498 tagmondattól összesen 76 darab tagmondat, azaz az adatoknak mindössze 15,3%-a mutat önmagában bármilyen trendet. Az egy egységnyi szórással eltoltsáv felett 72 tagmondat helyezkedik el, azaz ezen tagmondatok általános lassuló trend jellemzi, míg a sáv alatt 4 tagmondat helyezkedik el, ezeket gyorsuló tagmondatoknak tekinthetjük. Ha a kevert adatsorokra számolt lépésstatistikai eredmények stacionárius görbétől való távolságának kétszeres szórását ($\sigma = 0,121$) vesszük, jóval kevesebb tagmondat, mindössze 5,0%, amely a sávon kívül helyezkedik el. Ebből 23 lassuló és 2 gyorsuló tagmondat. Háromszoros szórás ($\sigma = 0,182$) esetén olyan sáv húzható meg, amelyben a mért adatok 2,6%-a nem található meg. Ezek közül 12 lassuló és 1 gyorsuló tagmondat. A kevert adatsorok lépésstatistikájának segítségével meghúzható sávok alapján tehát az egyes tagmondatok általános lassuló trendje nem jelentős. Meglehetősen kis arányban fordulnak elő olyan tagmondatok, amelyek nagymértékben eltávolodnának a stacionárius görbétől. Ezen

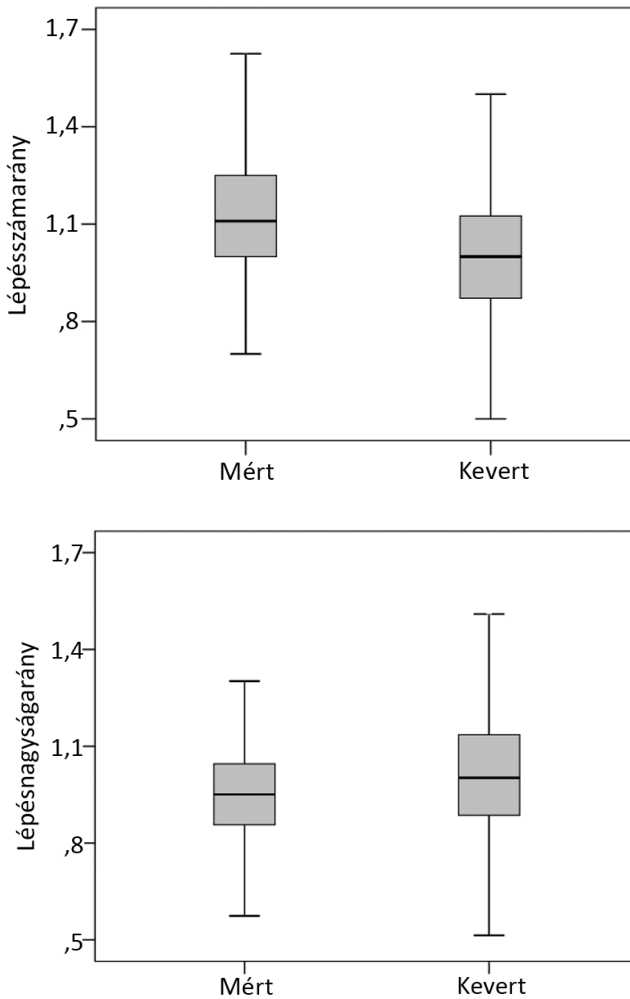
esetek közül is igen ritka, hogy egy tagmondat jelentős gyorsuló trendet mutatna. Ha ezzel a módszerrel valamilyen egyértelmű, jelentős mértékű trend kimutatható tagmondatokra, akkor az többnyire a lassuló trend. Külön-külön az egyes tagmondatok esetében ugyan többnyire nem mutatható ki jelentős mértékű lassulás vagy gyorsulás, de az összes tagmondatot egyszerre figyelembe véve látható az eltolódás a lassulás felé. A kevert adatsorokra számolt lépésstatisztikai eredmények vizsgálata viszont nem mutatott ilyen eltérést a lassuló és gyorsuló tagmondatok között. Az adatpontok a stacionárius görbétől láthatóan szimmetrikusan szóródtak (25. ábra). Ezért a következőkben a kevert adatsorok és a mért adatsorok eltérő tulajdonságait vizsgáljuk meg alaposabban.



26. ábra

A szöveg tagmondataira számolt lépésstatisztikai eredmények a CV-szakaszok alapján és a kevert adatok stacionárius görbétől való távolságának szórása

A lépésszámarány és a lépésnagyságarány is egyértelműen eltért a kevert és a mért adatsorok esetében (27. ábra). Az átlagos lépésszámarány az összes kevert adatsort figyelembe véve 1,021 volt (szórás: 0,253), míg a mért adatsoroknál ez az arány átlagosan magasabbnak bizonyult (átlag: 1,160, szórás: 0,319). A növekvő lépések száma általában tehát gyakoribb volt a mért és a kevert adatoknál is, mint a csökkenő lépések száma, de a mért adatok esetében a növekvő lépések száma még gyakoribbnak bizonyult. Az egymást követő magánhangzós és mássalhangzós szakaszok időtartama arányaiban gyakrabban nőtt a mért adatok esetében, mint a kevert adatok esetében.



27. ábra

A szöveg tagmondataira számolt lépésszám arány (fent) és lépésmagyságarány (lent) a CV-szakaszok alapján összehasonlítva kevert megfelelőikkel

A lépésmagyságarány szintén eltérést mutatott a mért és a kevert adatsorok esetében. A mért adatoknál ez az arány nem érte el átlagosan az 1-et (átlag: 0,963, szórás: 0,198), azaz a növekvő lépések átlagos nagysága kisebb volt, mint a csökkenő lépések átlagos nagysága tagmondatonként. Az egymást követő mássalhangzós és magánhangzós szakaszok közti időtartamkülönbségek tehát nagyobbak voltak abban az esetben, ha egy rövidebb szakasz következett, mint amikor egy hosszabb. A kevert adatsorok esetében a lépésmagyságarány viszont átlagosan magasabb volt a mért adatoknál (átlag: 1,029, szórás: 0,242). A növekvő lépések átlagos

nagysága nagyobb volt, mint a csökkenő lépések átlagos nagysága. A kevert adatsoroknál tehát az egymást követő mássalhangzós és magánhangzós szakaszok közti időtartam-különbségek rendszerint kisebbek voltak, ha egy rövidebb szakasz következett, mint akkor, amikor egy hosszabb.

A kevert és a mért adatsorok elemszáma nagymértékben eltért, ezért keverésenként külön-külön is összehasonlítottam az adatokat. Az egyes keverések átlagos adatai ugyanazokat a trendeket mutatták, mint amikor az összes keverést hasonlítottam össze a mért adatokkal (lásd 15. táblázat). A lépésszámarány a mért adatok esetében volt nagyobb, míg a lépésnagyságarány a kevert adatoknál vett fel magasabb értékeket. A lépésszám és a lépésnagyságarány esetében is előfordult olyan adatsor, amely átlagosan kisebb volt egynél, de ez az előbb említett tendenciát nem befolyásolta. A statisztikai próbákat úgy végeztem el, hogy az összes tagmondatra kapott értéket összehasonlítottam a kevert változatokra kapott értékekkel, és ezt megismételtem annyiszor, ahány keverés volt. A statisztikai próbák minden esetben szignifikáns eltérést mutattak mind a lépésnagyság-, mind a lépésszámarány tekintetében (a Wilcoxon teszt alapján: $Z \leq -2,924$, $p \leq 0,003$). A szöveg tagmondaitban az egymást követő magánhangzós és mássalhangzós szakaszok időtartam-különbségei tehát egyértelműen eltérnek a kevert adatsorok időtartam-különbségeitől több tekintetben is. A lépésszámarány statisztikailag igazolhatóan magasabb értékeket vesz fel a mért adatsorokban a kevertékhez képest, a lépésnagyságarány viszont a kevert adatsorok esetében magasabb.

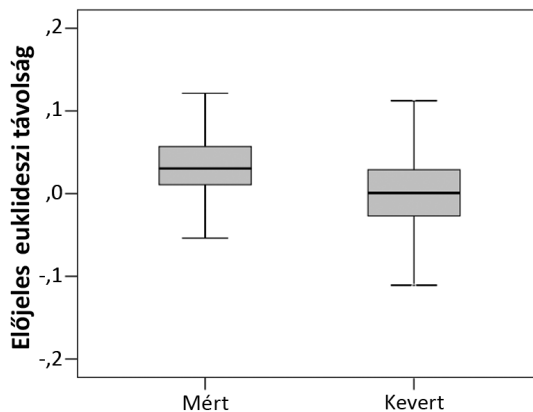
15. táblázat

A kevert és mért adatsorokra kapott lépésstatisztikai eredmények, különválasztva a 10, különböző keverésekre kapott adatokat

	Lépésszám- arány		Lépésnagyság- arány		Euklideszi távolság		Előjeles euklideszi távolság		
	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás	Átlag	Szórás	
Mért	1,160	0,319	0,963	0,198	0,054	0,077	0,046	0,082	
Kevert	1	1,026	0,244	1,019	0,213	0,044	0,049	-0,001	0,066
	2	1,028	0,282	1,031	0,242	0,046	0,072	0,002	0,085
	3	1,011	0,290	1,060	0,306	0,057	0,111	0,012	0,124
	4	1,029	0,238	1,011	0,203	0,055	0,065	0,000	0,085
	5	1,030	0,256	1,009	0,224	0,045	0,047	-0,007	0,065
	6	1,043	0,239	1,009	0,256	0,061	0,074	0,001	0,096
	7	0,983	0,277	1,080	0,260	0,047	0,063	0,002	0,079
	8	1,045	0,245	0,998	0,204	0,049	0,068	-0,001	0,084
	9	0,983	0,206	1,058	0,248	0,043	0,046	-0,003	0,063
	10	1,036	0,236	1,014	0,239	0,034	0,043	0,000	0,055

A kevert és a mért adatsorok eltérése a lépésszámárányban és a lépésmagyságarányban is megmutatkozott – felmerül a kérdés, hogy a stacionárius görbétől való távolságban is eltérnek-e a kevert és a mért adatsorok, azaz mutatnak-e eltérést abban, hogy milyen trend jellemzi inkább őket. Az egyik kérdés, hogy milyen messze szóródnak az adatpontok a stacionárius görbétől, azaz milyen mértékben térnek el a tagmondatnyi egységek a trend nélküli állapottól. A másik kérdés, hogy mennyire szimmetrikusan rendeződnek el a görbe körül, azaz a gyorsulás vagy a lassulás a jellemzőbb trend, vagy teljesen kiegyenlítően jelenik meg mindkettő. Előbbit az adatpontok stacionárius görbétől számolt euklideszi távolsága adja meg, utóbbit pedig az euklideszi távolságok előjeles változata. (Ha a stacionárius görbe fölött helyezkedett el az adatpont, akkor pozitív előjelet kapott a távolság értéke, ha alatta, akkor negatívát.) A korábbi adatok (25. ábra) alapján látható volt valamiféle eltérés a kevert és a mért adatok távolságában a stacionárius görbétől, a következőkben ezen különbségeket próbálom meg számszerűsíteni.

A mért adatsorok átlagos távolsága (átlag: 0,054, szórás: 0,077) csak valamivel volt nagyobb, mint az összes kevert adatsorra számolt átlagos távolság (átlag: 0,048, szórás: 0,067). A különböző kevert adatsorok átlagosan hol nagyobbak, hol kisebbek voltak a mért adatsorokhoz képest. A statisztikai próba több esetben (3-as, 4-es, 5-ös és 6-os számú kevert adatsor) nem támasztotta alá, hogy eltérés lenne az euklideszi távolságokban ($Z \geq -1,654$, $p \geq 0,098$). A többi esetben viszont jelentős volt az eltérés ($Z \leq -2,259$, $p \leq 0,024$). A mért és kevert adatsorok közti különbségek tehát véletlenszerűnek tekinthetők az euklideszi távolságok alapján. Ha viszont a stacionárius görbétől való elhelyezkedés irányát is figyelembe vesszük, egyértelmű eltéréseket tapasztalhatunk (28. ábra).



28. ábra

Az összes kevert adatsorra és a mért adatsorokra kapott előjeles euklideszi távolság a stacionárius görbétől a tagmondatokban

A kevert adatsorokra számolt lépésmagyságarány és lépésszámárány kiegyenlített, az adatpontok ugyanis a stacionárius görbéhez képest szimmetrikusan helyezkednek el. Az összes kevert adatsorra számolt előjeles euklideszi távolság átlaga 0,001 (szórás: 0,001), azaz szinte

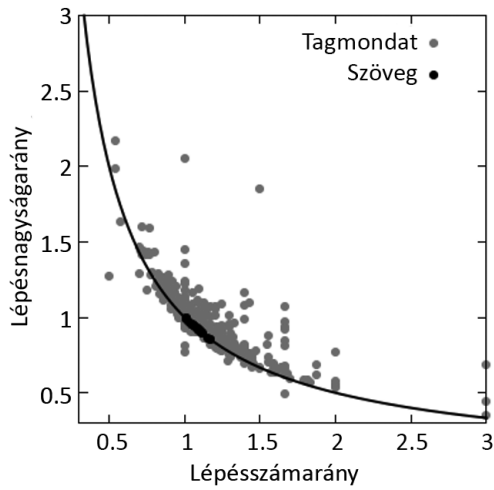
nincs is eltérés semmilyen irányban a stacionárius görbétől. A kevert adatsorok tehát sem a lassulás, sem a gyorsulás irányába nem tolódnak el. Ha egyenként vizsgáljuk a kevert adatsorokat, hol negatív, hol pozitív irányba tapasztalunk elmozdulást, de ezek az eltérések nem jelentősek (lásd 15. táblázat). A szövegfelolvasás tagmondataira mért adatok egészen mást mutatnak. Az előjeles euklideszi távolság a stacionárius görbétől átlagosan 0,046 (szórás: 0,004), tehát az adatpontok általában a stacionárius görbe fölött szóródnak, azaz nem szimmetrikusan helyezkednek el a görbéhez képest. Többségében tehát lassulás jellemzi a tagmondatokat, és kevésbé a gyorsulás, mint ahogyan azt korábban is kimutattam. A statisztikai próba minden esetben megerősítette, hogy a kevert adatsorok eltérnek a mért adatsortól ($Z \leq -8,636$, $p \leq 0,001$). Míg előbbi esetben igazolható, hogy a stacionárius görbéhez képest szimmetrikusan szóródnak az adatok, utóbbinál a lassulás mutatkozik meg általános trendként.

A mért és a kevert adatsorok összehasonlítása alapján tehát elmondható, hogy a mért adatsorok általánosságban megjelenő lassuló trendje nem véletlenszerű. Az egyes tagmondatokra külön-külön csak kevés esetben sikerült kimérni, hogy jelentős mértékű lassuló trendet mutatnának, de az összes tagmondatot együttesen vizsgálva statisztikailag igazolható volt az általánosan megjelenő lassuló trend.

4.2.2.2. Lassulás és gyorsulás a szakasz nagyságtól függően

A tagmondatokról tehát sikerült megállapítani, hogy egyfajta általános lassuló trendet mutatnak. Az artikulációs tempóról azonban ismert, hogy jellemzői változnak a vizsgált szakasz nagyságtól függően. Ezért felmerül a kérdés, hogy a tagmondatnyi egységekre jellemző kismértékű lassuló trend hosszabb szakaszok esetében hogyan realizálódik. Amennyiben egy sok tagmondatból álló szöveget próbálnánk vizsgálni, akkor a benne lévő tagmondatok általános lassuló trendjéből nem következne, hogy a teljes szöveg is lassuló trendet mutat, így megvizsgáltam a teljes szöveg lépésstatisztikáját is, és az eredményeket összehasonlítottam a tagmondatokéval.

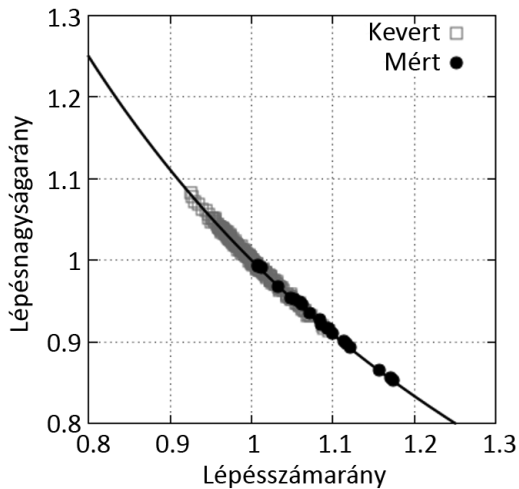
A szövegek nem mutattak olyan nagy szórást a lépésstatisztika alapján, mint a tagmondatok (29. ábra). A szöveg egészére számolt lépésstatisztikai eredmények ugyanakkor hasonlóak voltak a tagmondatokra kapott eredményekkel. A különböző beszélők teljes szövegfelolvasására kapott lépésszámarány minden esetben nagyobb volt 1-nél (1,008–1,174). A növekvő lépések száma tehát minden szövegben többnek adódott, mint a csökkenő lépések száma. A lépésnagyságarány minden esetben 1 alatt volt (0,853–0,994). A növekvő lépések tehát átlagosan kisebbek voltak, mint a csökkenő lépések. Az adatpontok mindegyike a stacionárius görbéhez egészen közel szóródott, 1%-os hibahatáron belül a görbén helyezkedtek el. (A lépésszámarány reciproka két tizedesjegyre kerekítve megegyezett a lépésnagyságaránnyal. Mivel ez az eltérés elhanyagolhatónak tűnik, a későbbiekben nem foglalkozunk a stacionárius görbétől való euklideszi távolságokkal.) A növekvő lépések nagysága tehát átlagosan kisebb volt a csökkenő lépéseknél, de gyakoribb előfordulásukkal kompenzálták az eltolódást. A szövegfelolvasások egészére kapott eredmények azt mutatták, hogy sem a lassulás, sem a gyorsulás nem jellemző rájuk. Sokkal inkább egyfajta kiegyensúlyozott állapot mutatkozott meg bennük.



29. ábra

A szöveg egészére és a tagmondatokra kapott lépésstatistikai eredmények

A szövegfelolvasásokra kapott eredmények tehát egyfajta mintázatot mutatnak. Felmerül a kérdés, hogy vajon a stacionárius görbén lévő eltolódás mennyiben tekinthető jelentősnek. Ezen kérdés megválaszolásához összehasonlítottam a mért adatokat a kevert megfelelőikkel. A lépésszámarány és a lépésnagyságarány is egyértelműen eltért a kevert és a mért adatsorok esetében (30. ábra).



30. ábra

A teljes szövegre mért és a kevert megfelelőkre kapott lépésstatistikai eredmények a CV-szakaszok alapján

Az átlagos lépésszámarány az összes kevert adatsort figyelembe véve alig tért el az 1-től (átlag: 1,001, szórás: 0,034), míg a mért adatsoroknál ez az arány átlagosan magasabbnak bizonyult (átlag: 1,083, szórás: 0,050). A növekvő lépések száma általában tehát gyakoribbnak bizonyult a mért adatoknál, mint a csökkenő lépések száma, de a kevert adatoknál a növekvő és csökkenő lépések száma kiegyensúlyozott volt. A mért adatok esetében tehát az egymást követő magánhangzós és mássalhangzós szakaszok időtartama gyakrabban nőtt, mint csökkent, míg a kevert adatoknál ugyanannyi esetben nőtt és csökkent az egymást követő szakaszok időtartama. A keverésenként lefuttatott statisztikai próbák alátámasztották, hogy a kevert és a mért adatsorok jelentősen eltérnek a lépésszámarány tekintetében ($Z \leq -3,435$, $p \leq 0,001$). Az egyes keverések átlagos adatai (lásd 16. táblázat) ugyanúgy 1-hez közeli lépésszámarányt mutattak, mint ahogyan azt az összes keverésre számolt adatok esetében láthattuk.

16. táblázat

A teljes szövegre kapott lépésstatisztikai eredmények és a 10 különböző keverésre kapott átlagos értékek

		Lépésszámarány		Lépésnagyságarány	
		Átlag	Szórás	Átlag	Szórás
Mért		1,083	0,050	0,927	0,043
Kevert	1	1,001	0,032	1,000	0,032
	2	0,993	0,033	1,009	0,033
	3	0,995	0,038	1,006	0,039
	4	1,001	0,041	1,000	0,041
	5	0,999	0,029	1,002	0,028
	6	1,007	0,036	0,994	0,035
	7	1,003	0,031	0,998	0,030
	8	0,997	0,025	1,003	0,025
	9	1,006	0,037	0,995	0,035
	10	1,004	0,040	0,997	0,040

A lépésnagyságarány szintén eltérést mutatott a mért és a kevert adatsorok esetében. Mint ahogyan az a 16. táblázatban megfigyelhető, a mért adatoknál ez az arány nem érte el átlagosan az 1-et, azaz a növekvő lépések átlagos nagysága kisebb volt, mint a csökkenő lépések átlagos nagysága a szöveg egészében. Az egymást követő mássalhangzós és magánhangzós szakaszok közti időtartam-különbségek tehát nagyobbak voltak abban az esetben, ha egy rövidebb szakasz következett, mint amikor egy hosszabb. A kevert adatsorok esetében a lépésnagyságarány viszont mindig magasabb volt a mért adatoknál. Az összes kevert adatsorra kapott lépésnagyság átlaga pontosan 1,000 volt, szórása pedig 0,034. A kevert adatsoroknál

tehát az egymást követő mássalhangzós és magánhangzós szakaszok közti időtartam-különbségek rendszerint ugyanakkorák voltak, akkor is, ha egy rövidebb szakasz vagy egy hosszabb szakasz következett. A keverésenként lefuttatott statisztikai próbák a lépésnagyságarány tekintetében is alátámasztották, hogy a kevert és a mért adatsorok jelentősen eltérnek egymástól ($Z \leq -3,360$, $p \leq 0,001$).

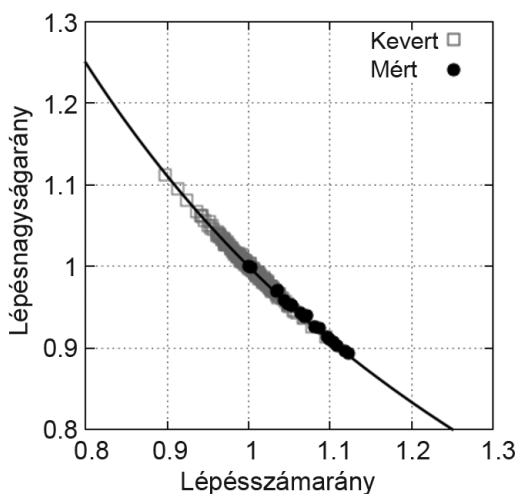
A teljes szövegekben tehát a kevert és mért adatsorok eltérése a lépésszámárányban és a lépésnagyságarányban is megmutatkozott hasonlóan ahhoz, ahogy a tagmondatok vizsgálatakor láthattuk. A mért adatsorok lépésszámárányai statisztikailag igazolhatóan magasabb értékeket vettek fel a kevert adatsorokban kapottakhoz képest, a lépésnagyságarány viszont a kevert adatsorok esetében volt magasabb. Ezek a különbségek a tagmondatok esetében és a teljes szövegfelolvasás esetében is megállapíthatók. A mért adatsorok – akár tagmondatnyi egységekről, akár teljes szövegegységekről van szó – egy jellegzetes mintázatot mutattak. A növekvő lépések száma valamelyest nagyobb volt a csökkenő lépéseknél, de a csökkenő lépések átlagos nagysága nagyobb volt a növekvő lépéseknél. Ez azt jelenti, hogy az egymást követő magánhangzós és mássalhangzós szakaszok időtartam-különbségei nagyobbak, ha egy rövidebb szakasz következik, mint amikor egy hosszabb, továbbá az időtartamok gyakrabban növekednek a beszéd folyamán, mint csökkennek. A tagmondatok és a teljes szövegfelolvasások közötti eltérés abban is megmutatkozott, hogy milyen általános trend jellemzi őket. Míg a tagmondatnyi egységek esetében nem nagymértékű, de következetesen megjelenő általános lassuló trend volt megfigyelhető, addig a teljes szöveg felolvasása esetében a növekvő lépések gyakoribb előfordulását a csökkenő lépések nagyobb mértéke teljes mértékben kompenzálta, ily módon a gyorsulások és lassulások egyensúlya jelent meg.

4.2.2.3. A lépésstatisztikai eredmények különböző alapegységekre számolva

Az előző alfejezetben láthattuk, hogy felváltva a magánhangzós szakaszok és a mássalhangzós szakaszok időtartamára számolt lépésstatisztika alapján az időzítésben egy jellegzetes mintázat található. A magánhangzók és a mássalhangzók időtartama viszont nem törvényszerűen ugyanúgy módosul különböző hatásokra: általában nagyobb változatosság mutatható ki a magánhangzók időtartamában, mint a mássalhangzókéban (vö. HOFHUIS–GUSSENHOVEN–RIETVELD 1995). A magánhangzók időtartama általánosságban könnyebben rövidíthető vagy nyújtható, mint ahogy arra a beszédrítmus-mérőszámokkal (%V) kapcsolatos eredmények is rámutattak a 4.1.2.5 fejezetben. Felmerül tehát a kérdés, hogy a mássalhangzós és a magánhangzós szakaszok hasonló eredményekre vezetnek-e a korábban ismertetett lépésstatisztikai módszerrel. További kérdés, hogy a CV-szakaszok vizsgálata helyett a magánhangzók és a mássalhangzók váltakozásának elemzése más eredményekhez vezet-e. Ezért megváltoztattam az eddig használt módszer alapegységét. Megvizsgáltam külön az egymást követő magánhangzós szakaszok időtartam-különbségeit, külön a mássalhangzós szakaszokét és a beszédhangokét is.

Mindegyik alapegység esetén hasonló eredményeket mutat a lépésstatisztika a teljes szövegfelolvasásokban abban a tekintetben, hogy sem gyorsulás, sem lassulás nem jellemzi őket.

Alapegységtől függően azonban nem teljesen azonos mintázatot találtam. A beszédhangok időtartamából számolt lépésstatisztika hasonló eredményekhez vezetett, mint amilyeneket a CV-szakaszok vizsgálatakor láthattunk. A beszédhangokra kapott lépésszámarány a teljes szövegfelolvasásra 1,000 és 1,123 között mozgott. Majdnem minden beszélő esetében elmondható tehát, hogy a növekvő lépések száma nagyobb, mint a csökkenő lépéseké. Ez azt jelenti, hogy az egymást követő beszédhangok esetében is érvényesül az a tendencia, miszerint gyakrabban nő egy-egy beszédhang időtartama, mint csökken. A lépésmagyságarány egy kivétellel minden esetben 1 alatt volt (0,894–0,999). A növekvő lépések tehát átlagosan kisebbek voltak, vagy ugyanakkorák, mint a csökkenő lépések. Vagyis, amikor a következő beszédhang hosszabb volt, akkor az időtartam-különbség mértéke tendenciaszerűen kisebbnek bizonyult, mint amikor a követő beszédhang rövidebben valósult meg. Az N6-os beszélő esetében a lépésmagyságarány csak minimálisan tért el az 1-től: 1,001-et mértem. Ugyanezen beszélő esetében a lépésszámarány 1,000 volt. Lényegében az egymást követő beszédhangok időtartamában semmilyen trend nem volt kimutatható ennél a beszélőnél. A beszédhangok alapján számolt lépésstatisztika adatpontjainak mindegyike a stacionárius görbéhez egészen közel szóródott, 2%-os hibahatáron belül a görbén helyezkedett el. (A lépésszámarány reciproka két tizedesjegyre kerekítve megegyezett a lépésmagyságaránnyal, vagy legfeljebb egy tizeddel eltért attól.) Hasonlóan a CV-szakaszokhoz, a beszédhangokat alapul véve szintén elmondható, hogy sem a lassulás, sem a gyorsulás nem jellemző a szövegfelolvasások egészére. Ugyanazt a mintázatot találtuk a beszédhangokra, mint a CV-szakaszokra. A növekvő lépések nagysága átlagosan kisebb volt a csökkenő lépéseknél, de gyakoribb előfordulásukkal kompenzálták az eltolódást. Egy beszélő esetében azonban a növekvő lépések és a csökkenő lépések száma, valamint nagysága is megegyezett.



31. ábra

A teljes szövegre mért és a kevert megfelelőikre kapott lépésstatisztikai eredmények a beszédhangok alapján

A beszédhangokra kapott eredmények szignifikánsan eltértek a kevert megfelelőikre kapott eredményektől (31. ábra). Az átlagos lépésszámarány az összes kevert adatsort figyelembe véve alig tért el az 1-től (átlag: 1,002, szórás: 0,030), míg a mért adatsoroknál ez az arány átlagosan magasabbnak bizonyult (átlag: 1,067, szórás: 0,035). A lépésnagyságarány szintén eltérést mutatott a mért és a kevert adatsorok esetében. A mért adatoknál ez az arány nem érte el átlagosan az 1-et (átlag: 0,940, szórás: 0,031). A kevert adatoknál viszont átlagosan az 1-hez közelített (átlag: 0,999, szórás: 0,030). A keverésenként lefuttatott statisztikai próbák alátámasztották, hogy a kevert és a mért adatsorok jelentősen eltérnek mind a lépésszámarány, mind a lépésnagyságarány tekintetében ($Z \leq -3,211$, $p \leq 0,001$). Az egyes keverések átlagos adatait a 17. táblázatban közlöm.

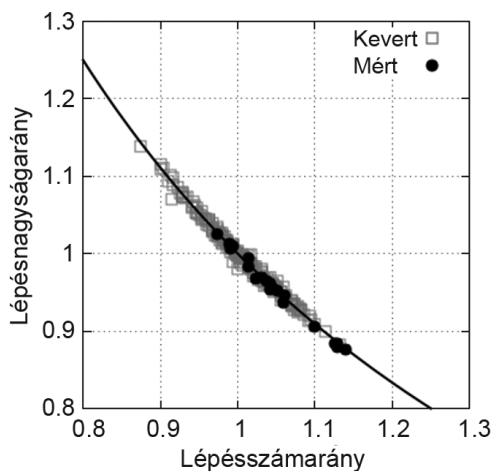
17. táblázat

A teljes szövegre kapott lépésstatisztikai eredmények és a 10 különböző keverésre kapott átlagos értékek, a beszédhangokat véve alapul

		Lépésszámarány		Lépésnagyságarány	
		Átlag	Szórás	Átlag	Szórás
Mért		1,067	0,035	0,940	0,031
Kevert	1	1,005	0,031	0,997	0,032
	2	1,007	0,027	0,995	0,027
	3	0,997	0,029	1,005	0,030
	4	0,997	0,028	1,004	0,028
	5	0,992	0,032	1,009	0,033
	6	1,014	0,018	0,987	0,018
	7	0,999	0,037	1,003	0,038
	8	0,991	0,029	1,010	0,030
	9	1,010	0,029	0,991	0,028
	10	1,007	0,029	0,993	0,028

Mind a CV-szakaszokra, mind a beszédhangokra kapott lépésstatisztikai eredmények eltértek tehát kevert megfelelőiktől. Láttuk, hogy a beszédhangok esetében és a CV-szakaszok esetében hasonló mintázatot mutattak az eredmények: a lépésszámarány tendenciózusan nagyobb volt 1-nél, míg a lépésnagyságarány kisebb. Felmerül a kérdés, hogy az alapegység megválasztása vajon hatással van-e az eltérések mértékére. Ezért összehasonlítottam a CV-szakaszokra és a beszédhangokra kapott lépésstatisztikai eredményeket a Wilcoxon-próbával. Sem a lépésszámarány ($Z = -1,459$, $p = 0,144$), sem a lépésnagyságarány ($Z = -1,409$, $p = 0,159$) esetében nem volt jelentős eltérés a próbák alapján.

Az egymást követő magánhangzós szakaszok időtartamából számolt lépésstatistika több szempontból hasonló eredményekhez vezetett, mint amit a beszédhangok vagy a CV-szakaszok és a beszédhangok esetében láthattunk. Néhány szempontból azonban eltérések is megmutakoztak. A lépésszamarány a teljes szövegfelolvasásra 0,974 és 1,140 között mozgott. A legtöbb beszélő esetében elmondható, hogy a növekvő lépések száma nagyobb volt, mint a csökkenő lépéseké. Azaz az egymást követő magánhangzós szakaszok időtartama valamivel gyakrabban nőtt a szövegben, mint csökkent. Négy beszélő (F1, F8, N1, N5) esetében viszont a lépésszamarány kisebb volt 1-nél, tehát az egymást követő magánhangzós szakaszok időtartama gyakrabban csökkent, mint nőtt. A lépésnagyság aránya szintén 1 körül szóródott (0,876–1,026). A beszélők többségénél a növekvő lépések átlagosan kisebbek voltak, mint a csökkenő lépések. Ugyanazon négy beszélőnél (F1, F8, N1, N5), akiknél a lépésszamarány eltérő tendenciát mutatott a többi beszélőtől, a lépésnagyságarány nagyobb volt 1-nél. Azaz az egymást követő magánhangzós szakaszok időtartam-különbsége nagyobbak bizonyult, ha nőtt a szakasz időtartama, mint amikor csökkent. Az adatpontok mindegyike – beszélőtől függetlenül – a stacionárius görbéhez közel szóródott, 2%-os hibahatáron belül a görbén helyezkedtek el. A többi alapegységhez hasonlóan sem a lassulás, sem a gyorsulás nem volt jellemző a szövegfelolvasások egészére, ha a magánhangzós szakaszok időtartamát vettem alapul. A lassulások és gyorsulások egyfajta kiegyensúlyozott állapota mutatkozott meg tehát a szövegekben a magánhangzós szakaszok alapján is. Mivel nem minden beszélő esetében volt megfigyelhető az a jellegzetes mintázat, hogy a lépésszamarány nagyobb 1-nél, a lépésnagyságarány pedig kisebb, felmerül a kérdés, hogy tekinthető-e jelentősnek a stacionárius görbén történő eltolódás a magánhangzós szakaszok esetében. Ezért összevettem a mért adatsorokra kapott eredményeket a kevert adatsorokra kapott eredményekkel (32. ábra).



32. ábra

A teljes szövegre mért és a kevert megfelelőikre kapott lépésstatistikai eredmények a magánhangzós szakaszok alapján

A magánhangzós szakaszokra kapott eredmények is eltértek a kevert megfelelőikre kapott eredményektől. Az átlagos lépésszámarány az összes kevert adatsort figyelembe véve alig tért el az 1-től (átlag: 1,004, szórás: 0,048), míg a mért adatsoroknál ez az arány átlagosan magasabbnak bizonyult (átlag: 1,049, szórás: 0,051). A lépésnagyságarány szintén eltérést mutatott a mért és a kevert adatsorok esetében. A mért adatoknál ez az arány nem érte el átlagosan az 1-et (átlag: 0,953, szórás: 0,031). A kevert adatoknál viszont általában az 1-hez közeli (18. táblázat). Átlagosan 0,998 volt a kevert adatokra kapott lépésnagyságarány (szórás: 0,048). A mért adatsorokra kapott eredmények – egy kivételével – jelentősen eltértek a kevert adatsoroktól a lépésnagyságarány tekintetében ($Z \leq -2,016$, $p \leq 0,044$). A 8. keverés esetében azonban nem mutatott szignifikáns eltérést a statisztikai próba ($Z = -1,867$, $p = 0,062$). A lépésszámarány esetében is többnyire kimutatható volt az eltérés a kevert és a mért adatsorok között ($Z \leq -2,296$, $p \leq 0,022$), az 1. és a 8. keverés esetében azonban itt sem volt szignifikáns az eltérés ($Z \geq -1,942$, $p \geq 0,052$). A magánhangzós szakaszok esetében tehát nem következetes az eltérés kevert megfelelőiktől.

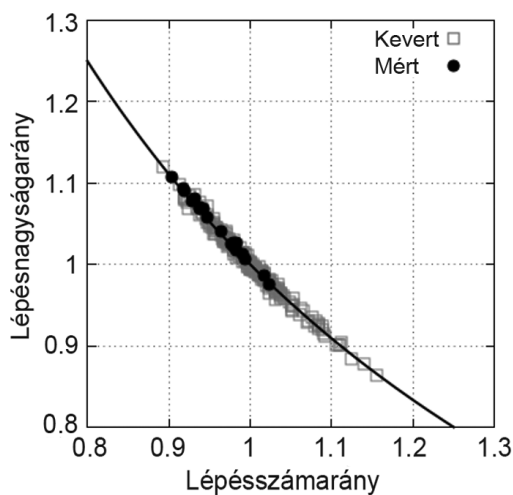
18. táblázat

A teljes szövegre kapott lépésstatisztikai eredmények és a 10 különböző keverésre kapott átlagos értékek, a magánhangzós szakaszokat véve alapul

		Lépésszámarány		Lépésnagyságarány	
		Átlag	Szórás	Átlag	Szórás
Mért		1,049	0,051	0,953	0,047
Kevert	1	1,012	0,054	0,991	0,052
	2	0,999	0,045	1,002	0,045
	3	1,004	0,046	0,998	0,044
	4	1,002	0,051	1,001	0,052
	5	1,000	0,051	1,002	0,049
	6	1,005	0,037	0,995	0,038
	7	1,013	0,045	0,989	0,046
	8	1,012	0,063	0,990	0,061
	9	1,010	0,048	0,993	0,048
	10	0,984	0,038	1,018	0,040

A beszédhangok és a magánhangzós szakaszok után a mássalhangzós szakaszokra is kiszámoltam a lépésstatisztikát. Az eredmények egészen eltérő mintázatot mutattak a többi alapegységhez képest. Abban hasonlóak voltak, hogy a lépésszámarány és a lépésnagyságarány

eltérései kompenzálták egymást minden beszélő esetében (33. ábra), és az adatpontok mindegyike – beszélőtől függetlenül – a stacionárius görbén helyezkedett el 2%-os hibahatáron belül. A lépésszámarány és a lépésmagyságarány viszont másfajta mintázatot mutatott. A lépésszámarány a teljes szövegfelolvasásra 0,903 és 1,024 között mozgott. A legtöbb beszélő esetében elmondható, hogy a növekvő lépések száma kisebb volt, mint a csökkenő lépéseké. Azaz az egymást követő mássalhangzós szakaszok időtartama valamivel gyakrabban csökkent a szövegben, mint nőtt. Két beszélő (F6, N3) esetében viszont a lépésszámarány nagyobb volt 1-nél, tehát az egymást követő mássalhangzós szakaszok időtartama gyakrabban nőtt, mint csökkent. A lépésmagyság aránya szintén 1 körül szóródott (0,975–1,108). A beszélők többségénél a növekvő lépések átlagosan nagyobbak voltak, mint a csökkenő lépések. Ugyanazon két beszélőnél (F6, N3), akiknél a lépésszámarány eltérő tendenciát mutatott a többi beszélőtől, a lépésszámarány esetében a lépésmagyságarány kisebb volt 1-nél. A mássalhangzós szakaszok esetében is összevettem a mért adatsorok eredményeit a kevert adatsorokéval (33. ábra).



33. ábra

A teljes szövegre mért és a kevert megfelelőikre kapott lépésstatistikai eredmények a mássalhangzós szakaszok alapján

A mássalhangzós szakaszokra kapott eredmények is eltértek a kevert megfelelőikre kapott eredményektől. Az átlagos lépésszámarány az összes kevert adatsort figyelembe véve majdnem pontosan 1 (átlag: 1,001, szórás: 0,047), míg a mért adatsoroknál ez az arány átlagosan kisebbnek bizonyult (átlag: 0,960, szórás: 0,034). A lépésmagyságarány szintén eltérést mutatott a mért és a kevert adatsorok esetében. A mért adatoknál ez az arány átlagosan nagyobb volt 1-nél (átlag: 1,046, szórás: 0,037). A kevert adatoknál 1 körül szóródott (19. táblázat).

Átlagosan 1,001 volt a kevert adatokra kapott lépésmagyságarány (szórás: 0,046). A mért adatsorokra kapott eredmények három kivétellel jelentősen eltértek a kevert adatsoroktól lépésszámarány tekintetében ($Z \leq -2,614, p \leq 0,009$). Az 1., a 3. és az 5. keverés esetében azonban nem mutatott szignifikáns eltérést a statisztikai próba ($Z \geq -1,831, p \geq 0,067$). A lépésmagyságarány esetében is többnyire kimutatható volt az eltérés a kevert és a mért adatsorok között ($Z \leq -1,998, p \leq 0,046$), szintén a 3. és az 5. keverés esetében nem volt szignifikáns az eltérés ($Z \geq -1,774, p \geq 0,076$). A mássalhangzós szakaszok esetében sem következetes tehát az eltérés a kevert adatsoroktól.

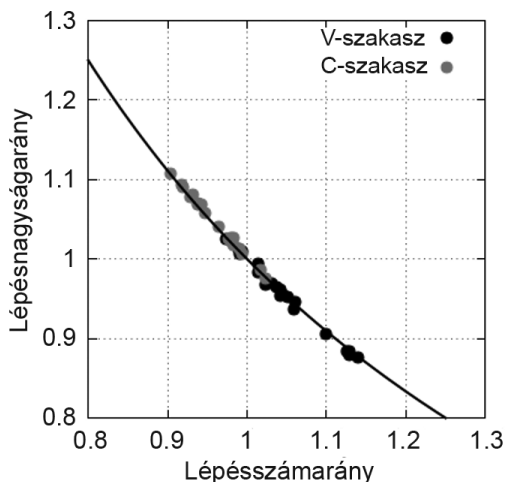
19. táblázat

A teljes szövegre kapott lépésstatisztikai eredmények és a 10 különböző keverésre kapott átlagos értékek, a mássalhangzós szakaszokat véve alapul

		Lépésszámarány		Lépésmagyságarány	
		Átlag	Szórás	Átlag	Szórás
Mért		1,049	0,051	0,953	0,047
Kevert	1	1,012	0,054	0,991	0,052
	2	0,999	0,045	1,002	0,045
	3	1,004	0,046	0,998	0,044
	4	1,002	0,051	1,001	0,052
	5	1,000	0,051	1,002	0,049
	6	1,005	0,037	0,995	0,038
	7	1,013	0,045	0,989	0,046
	8	1,012	0,063	0,990	0,061
	9	1,010	0,048	0,993	0,048
	10	0,984	0,038	1,018	0,040

Láthattuk, hogy a mássalhangzós szakaszok és a magánhangzós szakaszok eltérő mintázatot mutattak (34. ábra). Míg a magánhangzós szakaszok esetében a lépésszámarány jellemzően nagyobb volt 1-nél, addig a mássalhangzós szakaszok esetében kisebb. Míg az egymást követő mássalhangzós szakaszok időtartama gyakrabban csökkent, addig a magánhangzós szakaszoké gyakrabban nőtt. Ha összevetjük a mért adatsorokra kapott eredményeket a különböző alapegységeknél, akkor a különbség szignifikánsnak mondható a statisztikai próba alapján ($Z = -3,883, p < 0,001$). A lépésmagyságarány jellemzően nagyobb volt 1-nél a mássalhangzós szakaszok esetében, tehát az egymást követő mássalhangzós szakaszok időtartamában nagyobb volt a különbség, ha nőtt a szakasz hossza az előzőhöz képest, mint amikor csökkent.

A magánhangzós szakaszok esetében viszont akkor volt nagyobb a különbség, ha csökkent. A magánhangzós szakaszokra és a mássalhangzós szakaszokra kapott eredmények a lépés-nagyságarány szerint is jelentősen eltérnek ($Z = -3,883, p < 0,001$).



34. ábra

A teljes szövegre mért és a kevert megfelelőkre kapott lépésstatistikai eredmények, külön a mássalhangzós szakaszok és a magánhangzós szakaszok alapján

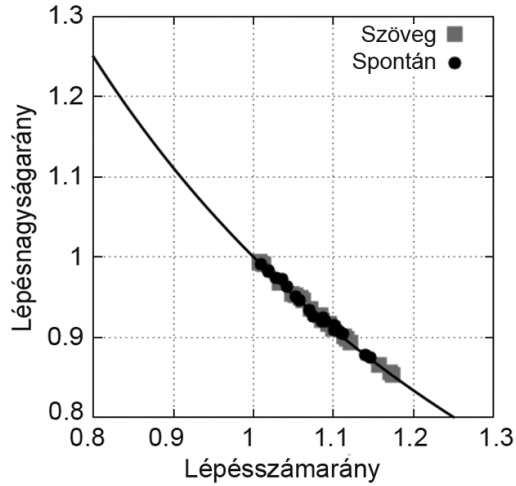
A szövegfelolvasások egészére kapott eredmények tehát azt mutatták, hogy sem a lassulás, sem a gyorsulás nem jellemző rájuk. Sokkal inkább egyfajta kiegyensúlyozott állapot mutatkozott meg bennük. A lassulások és a gyorsulások mértéke, illetve gyakorisága kiegyenlített. A szövegfelolvasások azonban nem teljesen trend nélküliek. Attól függően, hogy milyen alapegységet választottunk, eltérő mintázatokat figyeltünk meg szövegfelolvasásban.

4.2.2.4. A lépésstatistikai eltérések beszédmód szerint

Mint korábban láthattuk, a beszéd időzítését nagymértékben meghatározta a beszédmód (lásd 4.1.2.1. fejezet), ezért a lépésstatistikai módszerrel is összevettem a kétperces spontán monológokat és a teljes szövegfelolvasásokat. Az eredmények szerint több alapegység esetében is kimutatható eltérés a két beszédmód között.

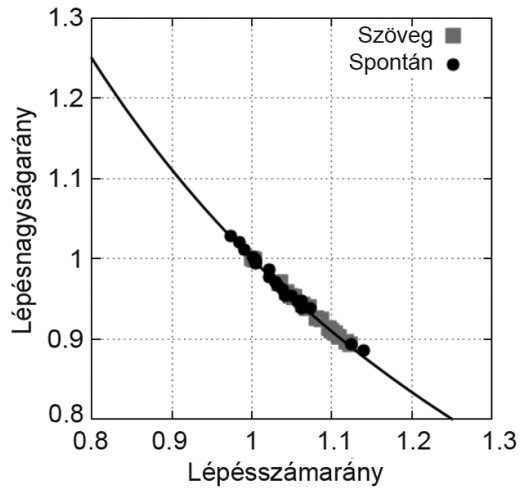
A spontán beszédben – a szövegfelolvasáshoz hasonlóan – a lépésszámárány és a lépés-nagyságarány eltérései kompenzálták egymást minden beszélő esetében (35., 36., 37. és 38. ábra). Az adatpontok mindegyike a stacionárius görbéhez közel található. A beszélőnként kapott eredmények a stacionárius görbén helyezkednek el 2%-os hibahatáron belül, a CV-szakaszokat, a mássalhangzós szakaszokat és a beszédhangokat alapul véve. A magánhangzós szakaszok esetében az eredmények 3%-os hibahatáron belül esnek a görbére.

Lényegében tehát a lassulások és a gyorsulások kiegyenlítettek, bármelyik alapegységet is vizsgáljuk.



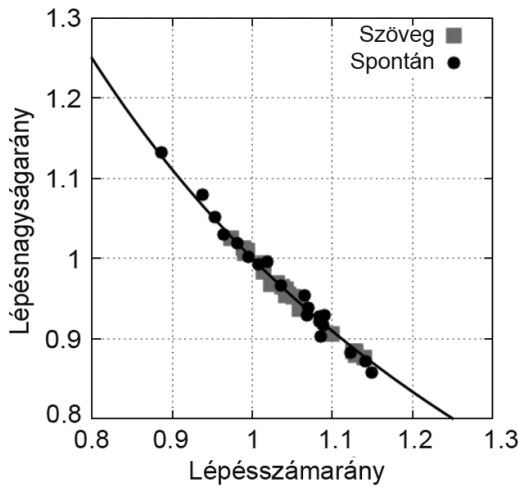
35. ábra

A spontán beszédre és szövegfelolvasásra kapott lépésstatistikai eredmények CV-szakaszok alapján



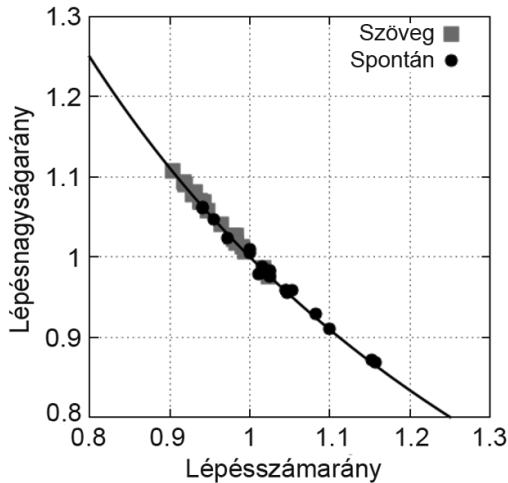
36. ábra

A spontán beszédre és szövegfelolvasásra kapott lépésstatistikai eredmények beszédhangok alapján



37. ábra

A spontán beszédre és szövegfelolvasásra kapott lépésstatistikai eredmények magánhangzós szakaszok alapján



38. ábra

A spontán beszédre és szövegfelolvasásra kapott lépésstatistikai eredmények mássalhangzós szakaszok alapján

A spontán beszédben bármelyik alapegységet is vettem figyelembe, a lépésszámarány átlagosan meghaladta az 1-et (20. táblázat), azaz gyakrabban nőtt az egymást követő egységek időtartama, mint csökkent. A lépésszámarány viszont kisebb volt 1-nél minden alapegység esetében. Az egymást követő egységek időtartam-különbsége tehát kisebb

volt, ha nőtt az egységek időtartama, mint amikor csökkent. Ez a tendencia spontán beszédben minden alapegységnél érvényesült. A mássalhangzós szakaszok esetében viszont pont ellenkező összefüggéseket találtam, mint amelyeket korábban megfigyelhettünk szövegfelolvasásban.

20. táblázat

A szövegfelolvasásokra és a kétperces spontán monológokra kapott lépéssziszatizikai eredmények átlagértékei és szórásai

		Lépésszámárány		Lépésmagyságarány	
		Átlag	Szórás	Átlag	Szórás
CV-szakasz	Spontán	1,074	0,040	0,934	0,035
	Szöveg	1,083	0,050	0,927	0,043
Beszédhang	Spontán	1,036	0,043	0,968	0,038
	Szöveg	1,067	0,035	0,940	0,031
V-szakasz	Spontán	1,042	0,071	0,965	0,072
	Szöveg	1,049	0,051	0,953	0,047
C-szakasz	Spontán	1,030	0,060	0,976	0,054
	Szöveg	0,960	0,034	1,046	0,037

A lépésszámárány átlagosan kisebb volt spontán beszédben, mint szövegfelolvasásban a CV-szakaszok, a beszédhangok és a magánhangzós szakaszok esetében is (lásd 20. táblázat). Ugyanezen egységeket alapul véve a lépésmagyságarány átlagosan spontán beszédben volt magasabb, mint szövegfelolvasásban. A teljes trend nélküli állapotól való eltolódás tehát jobban jellemző volt a szövegfelolvasásokra, mint a spontán beszédre. A statisztikai próbák a CV-szakaszok és a magánhangzós szakaszok esetében nem támasztották alá, hogy jelentős lenne az eltérés a két beszédmód között. Nem volt jelentős az eltérés sem a lépésszámáránynál (CV-szakasz: $Z = -0,859$, $p = 0,391$; magánhangzós szakasz: $Z = -0,299$, $p = 0,765$), sem a lépésmagyságaránynál (CV-szakasz: $Z = -0,672$, $p = 0,502$; magánhangzós szakasz: $Z = -0,336$, $p = 0,737$). A beszédhangok esetében viszont a lépésszámárány ($Z = -2,128$, $p = 0,033$) és a lépésmagyságarány ($Z = -2,203$, $p = 0,028$) is szignifikánsan eltért a két beszédmódban. A mássalhangzós szakaszokat alapul véve különböző mintázatot láthattunk a két beszédmódban. Így nem meglepő, hogy a statisztikai próbák mind a lépésszámárány ($Z = -3,472$, $p = 0,001$), mind a lépésmagyságarány ($Z = -3,547$, $p < 0,001$) esetében igazolták, hogy a spontán beszéd és a szövegfelolvasás jelentősen eltér egymástól. A szövegfelolvasás és a spontán beszéd között tehát találtam eltéréseket attól függően, hogy milyen alapegységet választottam.

4.2.2.5. A nemek közti különbségek a lépésstatistikában

A férfiak és a nők ejtésében a beszédritmus eltérő értékeket mutatott a mérőszámok alapján (lásd 4.1.2.3. fejezet), ezért megvizsgáltam a nemek közti eltéréseket a lépésstatistikai módszerrel kapott adatokon is. Összehasonlítottam külön a teljes szövegfelolvasásban, valamint külön a kétperces spontán monológokban kapott lépésszámarányt és lépésnagyságarányt a férfiak és nők ejtésében. Mind a négy eddig vizsgált alapegység (CV-szakaszok, beszédhangok, magánhangzós szakaszok, mássalhangzós szakaszok) esetében kiszámoltam az arányszámokat, és elvégeztem az összevetést.

Az eredmények nem mutattak jelentős különbségeket a férfiak és a nők ejtése között. A férfiak szövegfelolvasásában átlagosan magasabb lépésszámarányt mértem, mint a nőkében minden alapegység esetében (21. táblázat). A statisztikai próbák nem mutattak szignifikáns eltérést a nemek között egyik alapegységnél sem (Mann–Whitney-próba alapján, CV-szakasz esetében: $Z = -0,113$, $p = 0,912$; beszédhang: $Z = -1,399$, $p = 0,165$; magánhangzós szakasz: $Z = -1,512$, $p = 0,143$; mássalhangzós szakasz: $Z = -0,076$, $p = 0,971$). A lépésnagyságarány átlagosan kisebb volt a férfiaknál, mint a nőknél, de a statisztikai próbák ennél az aránynál sem igazolták, hogy a két csoport közti különbség szignifikáns lenne (CV-szakasz: $Z = -0,151$, $p = 0,912$; beszédhang: $Z = -1,324$, $p = 0,190$; magánhangzós szakasz: $Z = -1,587$, $p = 0,123$; mássalhangzós szakasz: $Z = -0,378$, $p = 0,739$).

21. táblázat

A szövegfelolvasásokra kapott lépésstatistikai eredmények átlagértékei és szórásai nemek szerint

		Lépésszámarány		Lépésnagyságarány	
		Átlag	Szórás	Átlag	Szórás
CV-szakasz	Férfi	1,087	0,053	0,924	0,045
	Nő	1,080	0,050	0,930	0,043
Beszédhang	Férfi	1,079	0,029	0,929	0,025
	Nő	1,055	0,038	0,952	0,034
V-szakasz	Férfi	1,066	0,058	0,938	0,053
	Nő	1,033	0,039	0,969	0,036
C-szakasz	Férfi	0,962	0,039	1,044	0,043
	Nő	0,958	0,030	1,049	0,032

A spontán beszédben a szövegfelolvasáshoz hasonlóan szintén nem találtam jelentős eltéréseket a férfiak és a nők ejtésében. Spontán beszédben a lépésszámarányt hol a férfiaknál, hol a nőknél volt átlagosan magasabb (22. táblázat). A statisztikai próbák itt sem mutattak

szignifikáns eltérést a nemek között, bármelyik alapegységet vizsgáltam (CV-szakasz esetében: $Z = -0,227$, $p = 0,853$; beszédhang: $Z = -0,983$, $p = 0,353$; magánhangzós szakasz: $Z = -1,058$, $p = 0,315$; mássalhangzós szakasz: $Z = -1,664$, $p = 0,105$). A lépésmélységarány esetében is az alapegységtől függött, hogy a férfiak vagy a nők ejtésében volt átlagosan nagyobb. A statisztikai próbák ennél az aránynál sem igazolták, hogy a két csoport közti különbség szignifikáns lenne, bármelyik alapegységet is vizsgáltam (CV-szakasz: $Z = -0,227$, $p = 0,853$; beszédhang: $Z = -1,058$, $p = 0,313$; magánhangzós szakasz: $Z = -1,285$, $p = 0,218$; mássalhangzós szakasz: $Z = -1,512$, $p = 0,143$). A vizsgált adatok alapján tehát a nők és a férfiak ejtésében nem találtam különbségeket sem szövegfelolvasásban, sem spontán beszédben.

22. táblázat

A kétperces spontán monológokra kapott lépésszisztematikai eredmények átlagértékei és szórásai nemek szerint

		Lépésszámarány		Lépésmélységarány	
		Átlag	Szórás	Átlag	Szórás
CV-szakasz	Férfi	1,069	0,043	0,939	0,036
	Nő	1,079	0,040	0,929	0,034
Beszédhang	Férfi	1,044	0,042	0,961	0,036
	Nő	1,028	0,044	0,976	0,041
V-szakasz	Férfi	1,029	0,050	0,978	0,050
	Nő	1,054	0,089	0,951	0,090
C-szakasz	Férfi	1,009	0,064	0,994	0,058
	Nő	1,051	0,048	0,957	0,043

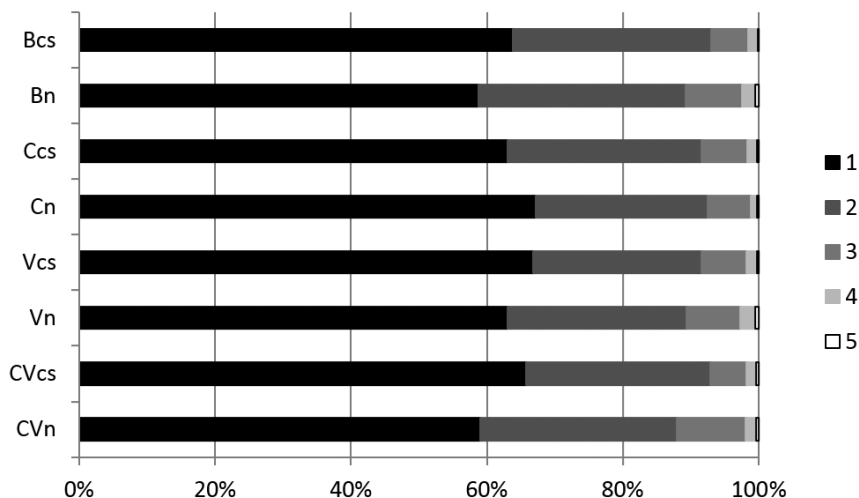
4.2.2.6. Tömbösödés a növekvő és csökkenő lépések alapján

Az előző alfejezetekben láthattuk, hogy a szövegfelolvasásokban és a spontán beszédben megjelenik egy jellegzetes mintázat. A lépésszámarány jellemzően meghaladta az 1-et, azaz a növekvő lépések száma gyakoribb volt, mint a csökkenő lépések száma. Ebben a fejezetben azt vizsgáljuk meg közelebbről, hogy a növekvő vagy a csökkenő lépések hogyan helyezkednek el egymáshoz képest a folyamatos beszédben. A csökkenő és a növekvő lépések egymást többé-kevésbé váltva találhatók, vagy tömböket alkotnak, tehát egymás után sokszor fordul elő, hogy az egymást követő szakaszok vagy a beszédhangok időtartama folyamatosan nagyobb vagy éppen folyamatosan kisebb, mint a megelőző egységé.

Ennek megvizsgáláshoz egybeolvasztottam azokat az eseteket, amelyeknél az egymást követő szakaszok rendre rövidebbek az utána követő szakasznál, valamint egybeolvasztottam

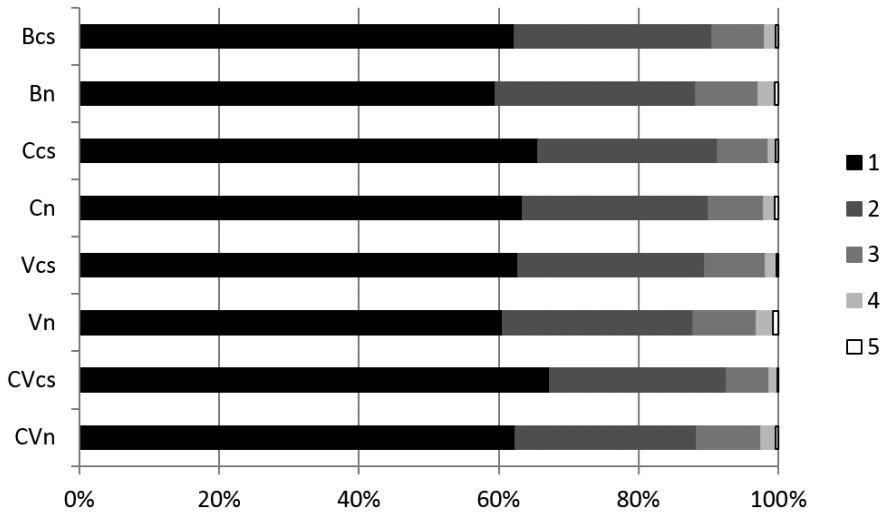
azokat az eseteket is, amikor az egymást követő szakaszok megszakítás nélkül sorra hosszabbak az utána követő szakasznál. Ezek után megszámláltam, hogy hány olyan eset van, amely csak egyetlen magánhangzós, mássalhangzós szakaszból vagy beszédhangból áll, továbbá, hány olyan eset van, amely kettő, három, négy, illetve több szakaszból vagy beszédhangból áll, különválasztva a növekvő lépésekből és a csökkenő lépésekből származó tömböket. (A beszélőkre kapott adatokat együttesen közlöm, külön számítva a szövegfelolvasásokat és a beszédmódokat.)

A szövegfelolvasásban arányaiban minden alapegységre hasonló eredményeket kaptam (39. ábra), akár a növekvő lépésekből álló tömböket vettem figyelembe, akár a csökkenő lépéseket. A tömbök többsége (61,7±2,9%) egyetlen egységből állt, azaz az esetek több mint felében, amikor a beszédhangok vagy a magánhangzós, illetve a mássalhangzós szakaszok rövidebbek voltak az azt megelőző egységtől, akkor a követő egység viszont hosszabb volt, vagy fordítva. A tömbök közel negyede (27,5±1,8%) mindössze két tagból állt. Jellemzően az esetek közel egy tizedében (8,5±1,1%) álltak három egységből a tömbök. Ritkán (1,9±0,5%) fordult elő, hogy a tömbök négy egységet tartalmaztak. Elhanyagolhatóan (0,5±0,2%) tekinthető azon tömbök száma, amelyek öt vagy több egységből álltak. Spontán beszédben hasonló eloszlásokat tapasztaltam minden alapegység esetében, amelyek a 40. ábrán figyelhetők meg. Nem jellemző sem a szövegfelolvasásban, sem a spontán beszédben, hogy az egymást követő egységek időtartama folyamatosan nőne vagy csökkenne hosszabb beszédrészeteken keresztül.



39. ábra

A szakaszok tömbösödése szövegfelolvasásban aszerint, hogy az egymást követő szakaszok hány egymást követő esetben rövidebbek vagy hosszabbak a követő szakasznál a különböző alapegységek vizsgálatakor (A használt rövidítések: B – beszédhang, CV – magánhangzós és mássalhangzós szakaszok együttesen, C – mássalhangzós szakaszok, V – magánhangzós szakaszok, n – növekvő lépések, cs – csökkenő lépések.)



40. ábra

A szakaszok tömbösödése aszerint, hogy az egymást követő szakaszok hány egymást követő esetben rövidebbek vagy hosszabbak a követő szakasznál spontán beszédben a különböző alapegységek vizsgálatakor (A használt rövidítések: B – beszédhang, CV – magánhangzós és mássalhangzós szakaszok együttesen, C – mássalhangzós szakaszok, V – magánhangzós szakaszok, n – növekvő lépések, cs – csökkenő lépések.)

Az eloszlások alapvetően hasonló arányokat mutatnak olyan szempontból, hogy a nagy elemszámú tömbök nem jellemzők, ám eltérések is megmutatkoztak aszerint, hogy növekvő vagy csökkenő lépéseket vizsgáltunk. Mind szövegfelolvasásban, mind spontán beszédben érvényesült az a tendencia, hogy több olyan egy egységből álló tömb található bennük, amely rövidebb az előtte lévő egységnél, mint olyan, amelyik hosszabb. Nagyobb arányban fordultak elő az olyan, legalább két egységet magukba foglaló tömbök, amelyek növekvő lépéseket tartalmaztak, mint azok, amelyek csökkenőket (lásd 23. táblázat). Ez az eltérés statisztikailag is alátámasztható volt mindkét beszédmódban majdnem minden alapegységben (χ^2 -próba alapján $\chi^2(4) \geq 17,478$, $p \leq 0,004$). Kivételt képeztek ez alól spontán beszédben a mássalhangzós szakaszok tömbösödései, amelyeknél nem igazolta a statisztikai próba a két eloszlás eltérést ($\chi^2(4) = 6,357$, $p = 0,174$). Az egyetlen eset, amely más eloszlást is mutatott a többi alapegységtől, az a szövegfelolvasásokban lévő mássalhangzós szakaszok tömbösödése. Gyakrabban voltak egyedülálló növekvő lépések, mint csökkenő lépések, azaz gyakrabban fordult elő olyan mássalhangzós szakasz, amely hosszabb volt a megelőző és a következő mássalhangzós szakasznál is, mint amelyek rövidebbek voltak mindkettőnél. A két vagy több mássalhangzós szakaszt tartalmazó tömbökben az egységek időtartama gyakrabban csökkent folyamatosan, mint nőtt. A növekvő és a csökkenő lépéseken alapuló tömbök gyakorisági eloszlásai közti különbséget a statisztikai próba is alátámasztotta

($\chi^2(4) = 17,708, p = 0,001$). A mássalhangzós szakaszok esetében tehát valamelyest más mintázat figyelhető meg, mint a többi alapegység esetében, hasonlóan a lépésszamarány és lépésmagyság vizsgálatához.

23. táblázat

Az egy tömbben lévő egységek számának gyakorisági előfordulásai különböző alapegységek esetében szövegfelolvasásban (fent) és spontán beszédben (lent)

(A használt rövidítések: B – beszédhang, CV – magánhangzós és mássalhangzós szakaszok együttesen, C – mássalhangzós szakaszok, V – magánhangzós szakaszok, n – növekvő lépések, cs – csökkenő lépések.)

		CVn	CVcs	Vn	Vcs	Cn	Ccs	Bn	Bcs
Az egy tömbben lévő egységek száma	1	59,0%	65,8%	63,0%	66,8%	67,2%	63,0%	58,7%	63,9%
	2	29,0%	27,1%	26,4%	24,8%	25,2%	28,5%	30,6%	29,1%
	3	10,0%	5,3%	7,9%	6,6%	6,4%	6,8%	8,2%	5,5%
	4	1,6%	1,5%	2,2%	1,6%	1,0%	1,4%	2,0%	1,3%
	5	0,4%	0,4%	0,6%	0,2%	0,2%	0,2%	0,5%	0,2%

		CVn	CVcs	Vn	Vcs	Cn	Ccs	Bn	Bcs
Az egy tömbben lévő egységek száma	1	62,4%	67,3%	60,6%	62,8%	63,4%	65,6%	59,5%	62,3%
	2	25,9%	25,2%	27,2%	26,8%	26,6%	25,7%	28,7%	28,2%
	3	9,2%	6,2%	9,1%	8,7%	7,9%	7,3%	8,9%	7,6%
	4	2,1%	1,1%	2,3%	1,6%	1,5%	1,1%	2,3%	1,5%
	5	0,4%	0,2%	0,8%	0,2%	0,5%	0,4%	0,5%	0,4%

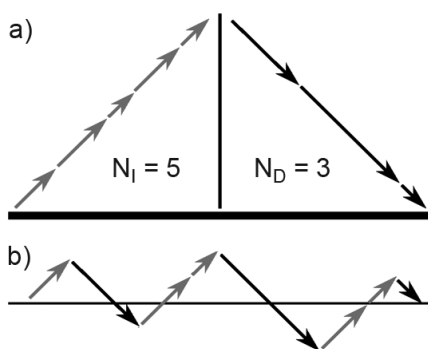
4.2.3. A lépésstatistikai eredmények értelmezése és a levonható általános következtetések

Ebben a fejezetben kísérletet tettem a beszédben megjelenő tempóingadozások jellemzésére egy átfogó, 20 beszélő beszédét elemző vizsgálat alapján. Egy olyan eljárást dolgoztam ki, amellyel mérhetővé vált a különböző nagyságú beszédrészletekben a lassulás és a gyorsulás mértéke. Kísérletet tettem egy olyan módszer létrehozására, amely tetszőleges nagyságú beszédrészletről meghatározza, hogy jellemzően lassuló vagy gyorsuló trendet mutat-e. Az eredmények azt mutatták, hogy a vizsgált szakasz nagyságától eltérő trend jellemzi a beszédet.

Elsőként a tagmondatok tempóingadozását vizsgáltam meg. A tagmondatoknál előfordulnak olyan esetek, amelyeket inkább a lassulás, másokat inkább a gyorsulás jellemez.

Az összes tagmondatot figyelembe véve kirajzolódik egy általános lassuló trend. A szöveg-felolvasás tagmondatainak többsége a beszélőnkénti átlagok, valamint a tagmondatonkénti átlagok alapján is inkább lassuló tendenciát mutatott. Statisztikailag is igazolható volt, hogy többször fordult elő lassuló, mint gyorsuló tagmondat a vizsgált szövegben. A lassulások és gyorsulások mértéke azonban önmagában nem volt jelentős egy-egy tagmondatban. Eredményeim összhangban vannak korábbi vizsgálatok mérési adataival (DANKOVIČOVÁ 1999; VÁRADI–BEKE 2013). Ezekben a kutatásokban a szavak mutattak általában lassuló trendet az intonációs frázisokon belül.

A teljes szövegfelolvasásokban talált eredmények eltértek a tagmondatokban találtakétól. A szövegfelolvasásokban sem a lassulás, sem a gyorsulás nem volt jellemző. Egyfajta egyensúly jelent meg ezekben a hosszabb beszédrészekben. Ez feltehetően arra vezethető vissza, hogy az artikulációs tempónak is vannak korlátai. A beszélő célja, hogy megértésért, ezért a beszédhangoknak érthetőnek kell maradniuk (LINDBLÖM 1990). Nem gyorsulhat vagy lassulhat folyamatosan a beszéd, mert szélsőséges tempó esetében befogadhatatlanná válna. Továbbá a beszédhangok rövidíthetősége és nyújthatósága is véges. A beszédhangok időtartamát korlátok közé szorítja, hogy legkevesebb mennyi idő alatt lehet kiejteni, illetve mennyi ideig lehet maximum fenntartani őket (lásd KLATT 1973; XU 2009; TURK–SHATTUCK–HUFNAGEL 2014). Mivel a beszédhangok nem rövidülhetnek és nyúlhatnak a végtelenségig, így a hangok időtartamából számolt tempó mértéke is behatárolt. Míg a tagmondatnyi nagyságú beszédrészekben mérhetők lassulások és gyorsulások, addig a 25 mondatos szövegfelolvasásban és a 2 perces spontán monológban már nem állapítható meg általános trend. Ilyen hosszú idő alatt feltehetően megtörténik az esetleges tempóingadozások visszarendeződése egy normál állapotba, az esetleges kilengések kiegyenlítődnek. A lassulások és a gyorsulások egyensúlya azonban nem jelentette a teljes trendnélküliséget. Egyfajta sajátos mintázatot találtunk a tempóingadozásban. A növekvő lépések száma valamelyest nagyobbak bizonyult a csökkenő lépéseknél, de a csökkenő lépések átlagos nagysága nagyobb volt a növekvő lépéseknél. Ez azt jelentette, hogy a beszédben az egymást követő magánhangzós és mássalhangzós szakaszok időtartam-különbségei általában nagyobbak voltak, ha egy rövidebb szakasz következett, mint amikor egy hosszabb. Továbbá az időtartamok gyakrabban növekedtek a beszéd folyamán, mint csökkentek. A növekvő lépések gyakoribb előfordulását a csökkenő lépések nagyobb mértéke ellensúlyozta, ahogyan azt a 41. a) szemléltető ábra mutatja. A beszéd folyamatában az egymást követő magánhangzós és mássalhangzós szakaszok azonban nem alkotnak folyamatosan növekvő vagy folyamatosan csökkenő tömböket. Adataink szerint a rövidebb szakaszokat jellemzően hosszabb követi, a hosszabbakat pedig rövidebb. Csak ritkán fordulnak elő folyamatosan növekvő vagy csökkenő szakaszok egymás után. A növekvő és csökkenő lépések tehát jellemzően váltogatják egymást (41. b) ábra). A beszéd folyamatában tehát úgy írhatnánk le az időtartamok változását, hogy gyakran a szakaszok időtartama többször egymás után kismértékben növekszik, majd egyszer csak ugrásszerűen nagyobb mértékben csökken, azután ismét nő, és ismét csökken.



41. ábra

Az a) a növekvő lépésekre (szürke, felfelé mutató nyíl) és a csökkenő lépésekre (fekete, lefelé mutató nyíl) talált mintázat egy lehetséges példája, a b) ugyanannak a mintázatnak időbeli sorrendben megvalósuló példája

Ez a mintázat beszédmódtól és alapegységtől függetlenül (beszédhang, magánhangzós szakasz, mássalhangzós szakasz és CV-szakasz) megjelent. Egyetlen kivétel volt ez alól: szövegfelolvasásban a mássalhangzós szakaszok vizsgálata más mintázathoz vezetett. Az egymást követő mássalhangzós szakaszok időtartam-különbségei átlagosan nagyobbak voltak szövegfelolvasásban, ha egy hosszabb szakasz következett, mint amikor egy rövidebb, továbbá az időtartamok gyakrabban csökkentek a beszéd folyamán, mint nőttek. Ugyanakkor spontán beszédben a mássalhangzós szakaszok esetében is hasonló mintázatot találtunk, mint ami a többi alapegységénél megfigyelhető volt. Hozzá kell tennünk, hogy a mássalhangzós szakasz és a magánhangzós szakasz vizsgálatoknál talált mintázat nem olyan következetesen valósult meg, mint azt a beszédhangoknál vagy a CV-szakaszoknál láthattuk. Feltehetően a mintázatot okozó tényezők hatásköre többnyire csak néhány hangra, egy-két szótagra koncentrálódik. Amennyiben csak a magánhangzós vagy mássalhangzós szakaszokat vesszük figyelembe, mindig átugrunk egy soron következő szakaszt, így valószínűleg ezen alapegységek vizsgálatoknál kevésbé mérhető hatásosan a talált mintázat.

Két tényező hatását is megvizsgáltuk a lépésstatistikai módszerrel kapott eredményekre. A nemek ejtésében nem találtunk különbséget, beszédmód szerint azonban néhány eltérést figyelhettünk meg. A szövegfelolvasásokban a beszédhangokat vizsgálva jobban jellemző volt a teljes trend nélküli állapottól való eltolódás, mint a spontán beszédben. Ez azt jelentette, hogy szélsőséesebb érték jelent meg a lépésszám- és lépésmagyságarányokban is. Mind a szövegfelolvasásban, mind spontán beszédben gyakrabban nőtt az egymást követő beszédhangok időtartama, mint csökkent, de szövegfelolvasásnál ez még inkább jellemző volt. A csökkenő lépések nagysága nagyobbak bizonyult, mint a növekvő lépéseké mindkét beszédmódban, de szövegfelolvasásban a növekvő és csökkenő lépések átlagos nagysága még nagyobb különbséget mutatott, mint spontán beszédben. A mássalhangzós szakaszok esetében ellenkező mintázatot találtunk a két beszédmódban. Spontán beszédben a többi alapegységhez

hasonlóan viselkedtek, szövegfelolvasásban viszont az egymást követő mássalhangzós szakaszok gyakrabban rövidültek, mint nyúltak. Szövegfelolvasásban arányaiban kevesebbszer fordult elő, hogy az egymást követő mássalhangzós szakaszok folyamatosan nőttek volna, mint az, hogy folyamatosan csökkentek. A magánhangzók és a mássalhangzók időtartamának eltérő viselkedését többen kimutatták. A megnyilatkozás végi helyzetben például eltérő mértékben nyúlhatnak (vö. HOCKEY–FAGYAL 1999; KOVÁCS 2009), de hogy pontosan mi okozhatja az eltérést a mássalhangzós szakaszok esetében szövegfelolvasásoknál, továbbra is nyitott kérdés. Összességében a szövegfelolvasás és a spontán beszéd között ugyan kimutatható volt valamiféle különbség, de a talált jellemző mintázat a legtöbb esetben azonosnak mutatkozott.

A lépésstatistikai módszer alapján tehát azt az általánosnak tekinthető mintázatot találtuk 20 köznyelvi beszélő beszéde alapján, hogy a beszédben az egymást követő egységek (beszédhangok vagy magánhangzós és mássalhangzós szakaszok) gyakrabban növekednek egymás után, mint csökkennek. Felmerül a kérdés, hogy mennyiben általános vagy csak lokálisan egy-két egységre ható tényezők mozdítják el ezt az arányt a lassulás felé. Amennyiben csak a magánhangzós szakaszokat vagy csak a mássalhangzós szakaszokat vizsgáltuk, láthatuk, hogy az eltolódás kevésbé következetesen és jelentősen valósul meg beszélőnként. Ez arra utalhat, hogy csupán néhány egymás mellett lévő egységre ható tényező állhat a háttérben, mint például megnyilatkozás végi nyúlás vagy esetleg a hangsúly időtartamokra gyakorolt hatása (habár utóbbi nem feltétlen érvényesül a magyar beszédben, vö. például KOVÁCS 2009; WHITE–MÁDY 2008). Kérdés, hogy ezen ismert és kevésbé ismert tényezők hatásai elegendők-e az általunk talált mintázat megformálásához. Az is elképzelhető, hogy a mintázat háttérben nem a beszédhangok nyújtásával történő jelölés áll. Ismert egy olyan percepció illúzió, amely szerint a befogadók rövidebbnek érzélik azt a beszédrészletet, amely egy időtartamában hasonló nagyságú, de valamivel rövidebb szakasz után következik. Az egymást követő egyre növekvő szótagokat a hallgatók hajlamosak ugyanakkorának hallani akkor is, ha ezek a szótagok nem megnyilatkozás vagy frázis végén helyezkednek el (WAGNER–WINDMANN 2011). Ezt a jelenséget kapcsolatba hozták az ismert beszédritmusosztályok percepció alapján történő elkülönítésével. A klasszikusan szótag-időzítésűnek tekintett francia nyelvben azt találták négy szótagos egységeket (hangsúlytól hangsúlyig terjedő szakaszokban) vizsgálva, hogy az első szótag rövid volt a többihez képest, a követő szótagok pedig egyre hosszabbak. A klasszikusan hangsúly-időzítésűekhez sorolt angol nyelvben viszont az első szótag volt hosszabb, míg a többi jellemzően inkább rövidült. A francia beszédben talált mintázat esetében az imént ismertetett percepció illúzió miatt a szótagokat a befogadó közel azonos időtartamban észlelheti, ami részben vagy egészben magyarázhatná a szótag-időzítésűnek tekintett nyelvekkel kapcsolatban kialakult percepció élményeket (WAGNER–WINDMANN 2009). A francia nyelvben talált eredmények párhuzamba állíthatók az általam magyar nyelvben mért adatokkal, amelyekben az egymást követő növekvő időtartamú beszédhangok, illetve magánhangzós és mássalhangzós szakaszok egyfajta felülreprezentáltsága jelent meg. Ezek alapján pedig elképzelhető, hogy az általam kapott mintázat egy, a magyar nyelvre is jellemző általános időzítési szabályszerűség eredményeképpen jön létre.

4.3. A temporális szerkezet és a zöngeminőség kapcsolata a megnyilatkozás végén

A nagyobb prozódiai egységek határai kitüntetett szerephez jutnak a beszédprodukción és a beszédfeldolgozásban is, ezért a beszélők igyekeznek (nem feltétlenül tudatosan) valamilyen módon jelölni ezeket. A hangzó megnyilatkozások szerkezetének egyik elsődleges velejárója az egységek határán megjelenő temporális mintázat, amelynek a legjellemzőbb megnyilvánulása a határon megjelenő nyúlás (szupraszegmentális nézőpontból tempólassulás) (lásd *1.2.1. fejezet*). Az irreguláris zöngeminőség (glottalizáció) ugyancsak gyakran tölt be határjelző funkciót, mind a frázis elején, mind a végén (lásd MARKÓ 2013 áttekintő összefoglalását). Jelen fejezetben ezt a két jelenséget vizsgáljuk magyar nyelvű felolvasásokban, a megnyilatkozások végén. Feltételezzük, hogy kapcsolat van a két jelenség között, ezért elemezzük, hogy milyen arányban jelentkezik irreguláris, illetve nem modális zöngeminőség (a lehetséges zöngét is beleértve), valamint mennyire jellemző és milyen mértékű a lassulás magyar felolvasott mondatokban. Azt is megvizsgáljuk, hogy a két jelenség előfordulása között kimutatható-e valamilyen összefüggés.

Az irreguláris zöngeminőség (glottalizáció) frázis/menyilatkozás végi határjelző szerepét több nemzetközi kutatás igazolta. A szakirodalom (elsősorban SLIFKA 2007) alapján feltételezhető, hogy a glottalizáció határjelző funkciójának fiziológiai háttere is van: a zöng felépülése és lecsengése nagyobb mértékben idézi elő a zöng irreregularitását a közlésegyeségek kezdetén és végén. HENTON és BLADON (1988) megnyilatkozás végi helyzetben mutatta ki a brit angol presztízsváltozatában, illetve az amerikai angolban az irreguláris hangszalagrezgés szintén a megnyilatkozás végét jelző akusztikai kulcsok egyike (SLIFKA 2006). Svéd olvasott beszéd frázishatárain is megjelenik az irreguláris zöng (FANT–KRUCKENBERG 1989), akárcsak a finnben, a csehben és a „szerbhorvát”-ban (LEHISTE 1965, idézi GORDON–LADEFOGED 2001). Prozódiai egységek kezdetén és végén a szlovénban is gyakori (JURGEČ é. n.).

A glottalizáció határjelző funkcióját magyar felolvasott és spontán szövegekben több különböző méretű beszédegység szintjén vizsgálták: beszédszakaszokban, intonációs frázisokban, megnyilatkozásokban és társalgási egységekben (MARKÓ 2013, 2014). A beszédszakaszok (szünetekkel körülhatárolt szövegegységek) és az intonációs frázisok határán nem volt kimagasló arányú az irreguláris zöng jelenléte: sem a vizsgált szövegegységek elején, sem a végén nem volt gyakoribb a glottalizáció, mint az egységek belsejében. (Ez arra enged következtetni, hogy a fent említett fiziológiai motiváció hatása itt nem jelentős.) Ezzel szemben felolvasott szöveg mondatvégein (korábbi kutatások – BÖHM–UJVÁRY 2008; MARKÓ 2010 – eredményeivel is egybehangzóan) gyakran jelentkezik glottalizáció, akárcsak a spontán beszéd megnyilatkozásainak végén. A glottalizációs határjelzés kifejezetten a szerkezet végén (ahhoz közeledve, az utolsó néhány szótagon) jelent meg, és olykor áthúzódott a következő egység elejére. A frázis eleji glottalizáció ugyanakkor a magyarban nem annyira jellemző, mint más nyelvekben (lásd fent), kivéve a V#V határokat (vö. MARKÓ 2013). A BEA

szövegfelolvasásaiban a glottalizáció szempontjából azt figyeltük meg, hogy a beszélők jellemzően a megnyilatkozások utolsó öt szótagján glottalizálnak (vö. MARKÓ 2013). A szakirodalom is megerősíti, hogy a megnyilatkozás végi glottalizáció sokszor nemcsak a legutolsó szótagon jelentkezik. Sőt, azoknak az eseteknek a száma sem elhanyagolható, amikor a glottalizáció nem a megnyilatkozás utolsó szótagján, hanem az azt megelőző két-három szótagon jelentkezik, és az utolsó szótagra regulárisra változik a zöngé minősége. Ezeket az eseteket is megnyilatkozás végi glottalizációnak tekintik a szerzők (például HENTON–BLADON 1988; SLIFKA 2006; BÖHM–UJVÁRY 2008).

Az irreguláris zöngé megjelenési gyakorisága erősen beszélőfüggő (lásd például HENTON–BLADON 1988; BÖHM–UJVÁRY 2008; MARKÓ 2013). Az eredmények szerint ugyanakkor a megnyilatkozathatár glottalizációval való megjelölése még azoknál a beszélőknél is relatíve gyakori, akiknek a zöngéképzése egyébként ritkán vált át irregulárisba. Sőt, kimondható, hogy minél kevesebbet glottalizál egy beszélő, annál valószínűbb, hogy ezt megnyilatkozást záró frázishatáron teszi. (Kivételt képeznek ez alól a rádióbemondók, akik kifejezetten kerülnek ezt a megoldást, feltehetően beszédtanári instrukció alapján, vö. MARKÓ 2013.)

A nem modális zöngeminőség másik viszonylag gyakori formája egyes nyelvekben a leheletes zöngé (vö. például KOHLER 2000; ISHI–ISHIGURO–HAGITA 2010), amelynek sajátossága, hogy zöngéképzés közben a hangszalagok között kiszökik a levegő, ezáltal turbulens zaj keletkezik, amelynek a lenyomata „zörejjel kevert zöngéként” látszik a regisztrátumon (GORDON–LADEFOGED 2001). Egy korábbi kutatás eredménye alapján az is felmerült, hogy a glottalizáció és a leheletes zöngé funkciókörében lehet átfedés (lásd MARKÓ 2013), illetőleg nemzetközi kutatások alapján egyes esetekben kizáró disztribúcióban állnak egymással (KOHLER 2000). Ezért indokoltnak tekinthető a két jelenség együttes elemzése.

4.3.1. Kísérleti személyek, anyag és módszer⁶

Kísérleti személyek és anyag

A kutatáshoz a 3.1. fejezetben ismertetett anyagnak egy része került felhasználásra, amelynek főbb tulajdonságai a következők voltak: a BEA adatbázis (GÖSY et al. 2012) tíz beszélőjének a mondat- és szövegolvasása képezte a vizsgálat anyagát.⁷ Az adatközlők fele nő, fele férfi, nem dohányoznak, nincs zöngéképzési vagy más beszédzavaruk, és nincs ismert hallásproblémájuk. Életkoruk 20 és 60 év között szóródik, az átlagéletkor 41,3 év.

⁶ Ezen alfejezet egy korábbi kutatásunk (KOHÁRI–MARKÓ 2015) alapján készült.

⁷ Ez a vizsgálat a kutatás egy korábbi fázisában készült, amikor még csak 10 adatközlőn készültek a vizsgálatok. Mivel ez a mérés volt a legösszetettebb, mind az annotálást, mind számításokat tekintve, az adatközlők számának a bővítésére itt már nem volt lehetőség. Másrészt mivel a glottalizációs gyakoriság erősen egyéni függő jelenség, feltehetően az adatközlők számának megkésztetése sem vezetett volna erősebb összefüggésekhez.

A korábbi kutatás (vö. MARKÓ 2013) alapján a megnyilatkozás vége nagyjából 12 beszédhangban határozható meg, ezért ez a vizsgálat minden felolvasott mondat esetében az utolsó 12 elhangzott beszédhangot elemezte. A tíz beszélőnek 374 olyan megnyilatkozása (246 önálló mondat és 128 szövegmondat) volt, amelynek utolsó 12 hangjában nem fordult elő nyelvbottlás, félreolvasás. Így összesen 4488 beszédhang zöngeminőségének és időtartamának elemzésére került sor. A lassuló vagy gyorsuló trend megállapításához nemcsak a mondatok utolsó 12 hangja képezte az elemzések tárgyát, hanem összevetésként az azokat megelőző 12 hang is, így további 4488 hang időtartama került be a vizsgálatba.

A hanganyagok annotálása és a statisztikai módszerek

A hanganyag hangszintű annotálását, ahogy korábban részletesen kifejtettem (a 3.2. fejezetben), a nemzetközi szakirodalomban található kritériumoknak megfelelően végeztem el, figyelembe véve a magyar beszédben lévő sajátos akusztikai realizációkat. A MAUS szoftverrel végzett automatikus hangszintű annotálás (SCHIEL 1999; KISLER–SCHIEL–SLOETJES 2012) után a kézi ellenőrzés a Praat 5.1 szoftverben (BOERSMA–WEENINK 2009) történt.

A hangszinten annotált hanganyagon minden beszédhang kapott egy zöngeminőséget jelző címkét a következőképpen: külön jelölés szerepelt, ha a hangzó zöngeminősége nem volt vizsgálható (zöngétlen obstruens), illetve ha nem lehetett megítélni – ide tartozott az esetleges zaj (a felvételek végéhez közeledve papírzörgés), valamint egységesen ide kerültek a zöngés obstruenseket és a pergőhang realizációi, mivel az akusztikus kép alapján ezeknél nem mindig dönthető el a zöngeminőség. A glottalizációra lehetőséget adó szegmentumok (magánhangzók és szonor mássalhangzók) esetében jelölésre került, ha a zöngé modális, ha irreguláris, illetve ha leheletes volt. Ha az adott beszédhang bármely kis részletében irreguláris periódusok jelentek meg, a hang „glottalizált” címkét kapott. A glottalizált realizációk annotálása a korábbi kutatások módszertanához igazodva (például DILLEY–SHATTUCK–HUFNAGEL–OSTENDORF 1996; BÓHM–UJVÁRY 2008) a vizuális és auditív információk együttes figyelembevételével történt. Akusztikai szempontból glottalizálnak minősült az adott beszédhangrészlet, ha az alapperiódusok időtartama vagy amplitúdója hirtelen jelentősen megváltozott, vagy ha hirtelen a beszélő normál/szokásos hangterjedelme alá csökkent az alapfrekvencia. Mindemellett percepciósszempontként szerepelt, hogy a szegmentum hangszínezete jól hallhatóan érdes, rekedtes. Akkor kapott „glottalizált” címkét egy beszédhangrészlet, ha az akusztikai lenyomatban látható és ezzel egyidejűleg auditív úton is észlelhető volt az irregularitás. A leheletes zöngé jelölési kritériuma szintén az volt, hogy a kiszökő levegő zöreje a zöngelenyomatban látható, valamint hallható legyen.

Fontos megjegyezni, hogy a glottalizáció többfunkciós jelenség. Az a tény tehát, hogy egy adott beszédhang irreguláris zöngével realizálódott, nem feltétlenül jelenti azt, hogy az adott szegmentumon a glottalizációnak az a szerepe, hogy a megnyilatkozás végének közeledtét jelezze. Ennek megfelelően csak akkor kapott „glottalizált” címkét egy szegmentum, ha az biztosan nem két szomszédos magánhangzó határát és/vagy fráziskezdetet jelölt (vö. MARKÓ 2013).

Mindezek alapján a megnyilatkozásvégek beszédhangjai körében a zöngeminőséget elemezve két körben zajlott a vizsgálat. Az egyik esetben a 12 vizsgált beszédhang közül meghatározásra került azoknak a beszédhangoknak a száma, amelyeknek a zöngeminősége vizsgálható (azaz a magánhangzók és a szonor mássalhangzók tartoztak ide). Ezeknek a beszédhangoknak a számát 100%-nak tekintve megmutatható, hogy ezek milyen arányban realizálódtak ténylegesen glottalizáltan. A másik esetben az utolsó 12 hang közül csak a magánhangzókat figyelembe véve az elemzés arra fókuszált, hogy ezek közül mennyi volt glottalizált. Az elemzések nemcsak az utolsó 12 hang glottalizáltságának arányára terjedtek ki, hanem külön az utolsó 6 és külön az utolsó 3 hangéra is.⁸ Különböző egységek vizsgálatára azért volt szükség, mert nem teljesen egyértelmű, hol jelenhet meg a glottalizáció a megnyilatkozás végének jelöléseként. Előfordul ugyanis olyan eset, amikor a frázis vége előtt az utolsó előtti szótagon megjelenik a glottalizáció, míg az utolsó szótag visszamodálisodik (HENTON–BLADON 1988). Mindkét mérés módszertani megközelítésben szerepel olyan összesítés is, ahol a leheletes és az irreguláris zöngeminőség egy kategóriát alkot, vagyis mindkettő a nem modális zöngéhez tartozik.

A temporális szerkezet vizsgálatának alapját a beszédritmus-mérőszámokat alkalmazó szakirodalomhoz hasonlóan a hangzók időtartama képezte (GRABE–LOW 2002). A lassulás és a gyorsulás mérésének az alapját az egyik beszédritmus-mérőszám (PVI) és az ún. lépésstatistikai módszer képezte. Ennek lényege, hogy a vizsgált egység időtartamát összehasonlítjuk a követő hangzó időtartamával, ily módon lehetővé válik a folyamatos beszédben lévő hangsor temporális viszonyainak feltárása. Az adott hang időtartamából kivonjuk a követő hang időtartamát, a különbség előjele megadja, hogy a két egység közötti viszony lassuló vagy gyorsuló. Ha a követő hang időtartama nagyobb, lassulásról beszélünk, ha rövidebb, akkor gyorsulásról. A különbségek mértéke pedig megadja a lassulás vagy gyorsulás mértékét. A módszer kiterjeszhető nagyobb egységek vizsgálatára és összehasonlítására is, így a több hangon átívelő jelenségek mérése ugyancsak lehetővé válik. A különböző egységek összehasonlításához azonban nem a beszédhangok időtartama, hanem az adott egységben lévő hangok darabszáma lett elosztva a beszédhangok teljes időtartamával, azaz a kiértékelés alapját az artikulációs tempó képezte. Az utolsó 12 hangra kapott artikulációs tempó az azt megelőző 12 hangra kapott artikulációs tempó értékéből lett kivonva. A különbség előjele alapján eldönthető, hogy a megnyilatkozás vége lassult vagy gyorsult. A lassulások és gyorsulások elemzéséhez azonban nem állapítható meg optimális ablakméret (lásd DE LOOZE 2010), ezért az elemzés különböző nagyságú egységek megvizsgálására is kiterjedt a glottalizációhoz hasonlóan.⁹ Az elemzés magában foglalta az utolsó 3 és 6 beszédhangra számolt

⁸ A glottalizáció mérésében igen gyakran szótagszinten címkéznek, ám a temporális összehasonlítás miatt jelen kutatásban a beszédhangok lettek az alapegységek. Mivel egy szótagot átlagosan három beszédhang alkot, ezért a nagyságrendileg egy, két és a megnyilatkozás végétől távolabb eső négy szótagnak tekinthető egységek képezték a vizsgálat tárgyát.

⁹ Az utolsó és az utolsó előtti szótagok hangzóinak megnyúlását többen igazolták megnyilatkozás végi helyzetben (lásd *1.2.1. fejezetben*), de előfordulhat, hogy még nagyobb szakasz érintett a lassulásban.

artikulációs tempó összevetését is az azt megelőző, ugyanakkora egységek artikulációs tempójával, amelyek szintén lassuló vagy gyorsuló besorolást kaptak a különbség alapján. Az összetett időzítési mintázatok ezen módszer vizsgálatából származó eredményei nem kezelhetők szigorú kategóriákként, hiszen például az egyes beszédhangok ún. intrinzikus időtartama eltérő (vö. például GÓSY 2004; OLASZY 2006), de az időtartamok viszonyainak ily módon történő leegyszerűsítése nagyobb trendek, összefüggések kimutatására alkalmas lehet. A különféle egységekre kapott esetleges azonos trendmegjelölések pedig mutathatják az eredmények megbízhatóságát is.

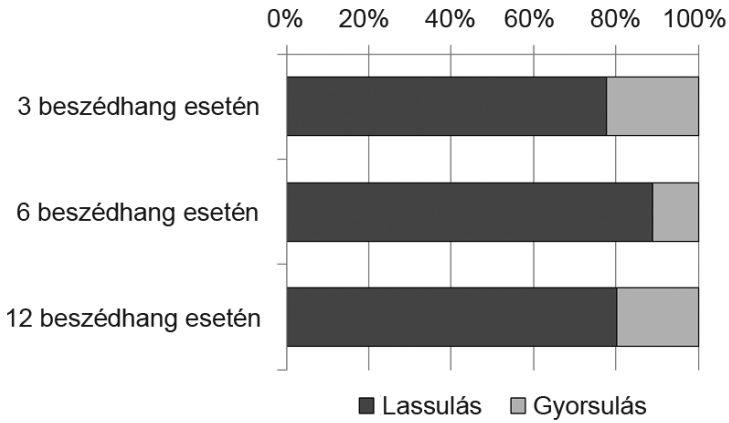
A fenti paraméterek, valamint a mondat- és a szövegfelolvasás közötti lehetséges eltérések elemzése a teljes vizsgálati korpusz tekintetében megvalósult. A statisztikai elemzések (χ^2 -próba, Mann–Whitney-próba¹⁰, Spearman-féle korrelációelemzés) az SPSS 20.0 program segítségével történtek. Az adatok nem normál eloszlása (Shapiro–Wilk-próba alapján) és több esetben a nem arányskála jellege tette indokolttá a nemparaméteres-eljárásokat.

4.3.2. Eredmények¹¹

A megnyilatkozások végét általánosságban csökkenő artikulációs tempó, azaz lassulás jellemzi. Az utolsó 3 beszédhang artikulációs tempóját viszonyítva az azt megelőző 3 hang artikulációs tempójához, az összes megnyilatkozás 77,8%-ában a megnyilatkozás utolsó szakaszának artikulációs tempója kisebb volt, mint a megelőző szakaszé. Az utolsó 6 beszédhang és az azt megelőző 6 hang artikulációs tempóját figyelembe véve, ez az arány valamelyest nagyobbak bizonyult (89,0%). A gyorsuló trendet mutató szakaszvégek aránya mindkét esetben egyértelműen alacsonynak mutatkozott (3 beszédhang esetében: 22,2%, 6 beszédhang esetében: 11,0%). A megnyilatkozások végének lassulása még az utolsó 12 beszédhang és az azt megelőző 12 hang összehasonlításakor is egyértelműen tetten érhető. Az utolsó szakaszok artikulációs tempója az esetek 80,3%-ában alacsonyabb a megelőző szakasz artikulációs tempójánál. A megnyilatkozás végi lassulás tehát a szakirodalommal megegyezően következetesen megjelenik a megnyilatkozások végén, ugyanakkor a lassulás nemcsak az utolsó egy-két szótagnyi nagyságú egységeket jellemzi, hanem nagyobb szövegrészekben is felfedezhető (42. ábra). A különböző nagyságú egységek összehasonlítása azonban csak az általános trendet jelzi, tehát nem a szakaszban lévő összes hang nyúlásáról van szó. A lassuló trendet mutató esetek artikulációs tempókülönbségének átlagos értéke kisebb is a 12-12 beszédhangot összehasonlító elemzésnél: 12 hang esetében 2,9 hang/s (szórás: 1,4 hang/s); 6 beszédhang esetében 3,9 hang/s (szórás: 2,0 hang/s); 3 beszédhang esetében 3,6 hang/s (szórás: 2,2 hang/s).

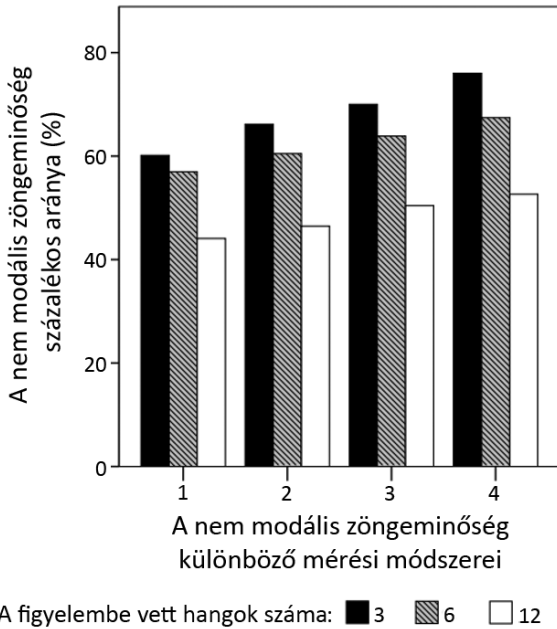
¹⁰ E kutatásban az adatok nem tették lehetővé párosított próba használatát, mivel a vizsgált tényezők (megnyilatkozás végi lassítás és gyorsítás előfordulása, valamint a glottalizáció) előfordulási aránya határozta meg a különböző csoportokban lévő esetek darabszámát.

¹¹ Ezen fejezet alapját egy korábbi kutatásunk (KOHÁRI–MARKÓ 2015) képezi.



42. ábra

A lassuló és gyorsuló szakaszvégek aránya különböző méretű vizsgált egységek esetén

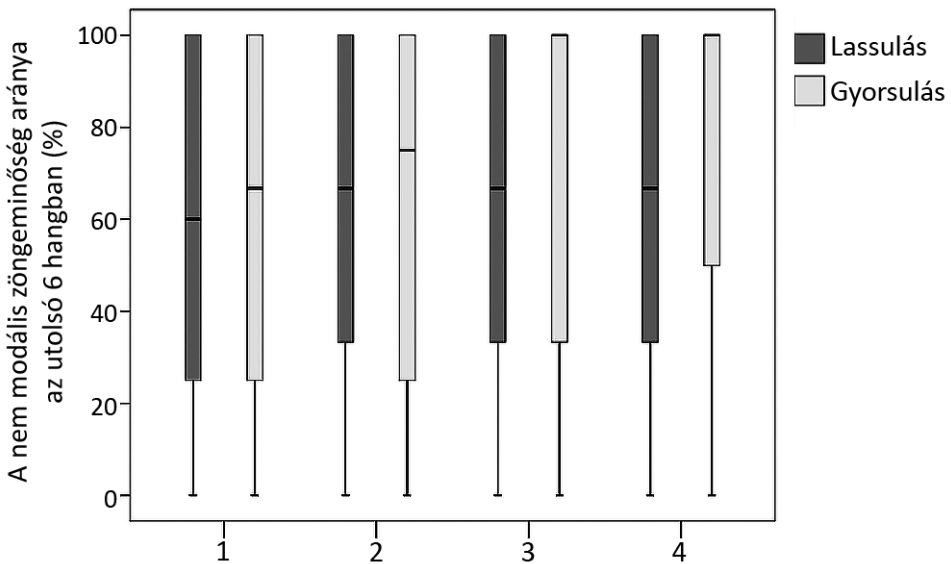


43. ábra

A jelöltség (irregularitás, illetve a modálistól való eltérés) mértéke a különböző nagyságú egységek és az elemzési tartományok függvényében (1 = az irregularis zöngével megvalósult szegmentumok aránya az összes $V + C_{szon}$ számához viszonyítva; 2 = a nem modális zöngével megvalósult szegmentumok aránya az összes $V + C_{szon}$ számához viszonyítva; 3 = az irregularis zöngével megvalósult V -k aránya az összes V számához viszonyítva; 4 = a nem modális zöngével megvalósult V -k aránya az összes V számához viszonyítva.)

Az irreguláris zöng megjelenése ugyancsak jellemző a megnyilatkozásvégeken, és a közlés végéhez közeledve egyre nagyobb a glottalizált beszédhangok aránya. A 43. ábra azt szemlélteti, hogy az összes megnyilatkozás különböző méretű szakaszvégeiben átlagosan hány százalékos a glottalizáció/irreguláris zöngképzés megjelenése. (A 100% minden esetben azoknak az adott típusú szegmentumoknak a számát jelenti, amelyek az adott egységben egyáltalán irreguláris zöngével valósulhattak meg).

Az ábrán egyértelműen látszik az a tendencia, hogy a megnyilatkozás végéhez közeledve mind önmagában az irregularitás, mind együttesen a modálistól eltérő zöngeminőség egyre nagyobb arányú. Továbbá ha a vizsgálatot a magánhangzókra korlátozzuk, ezek az arányok még nagyobbak, mint az összes szonor ($V + C_{szon}$) körében.



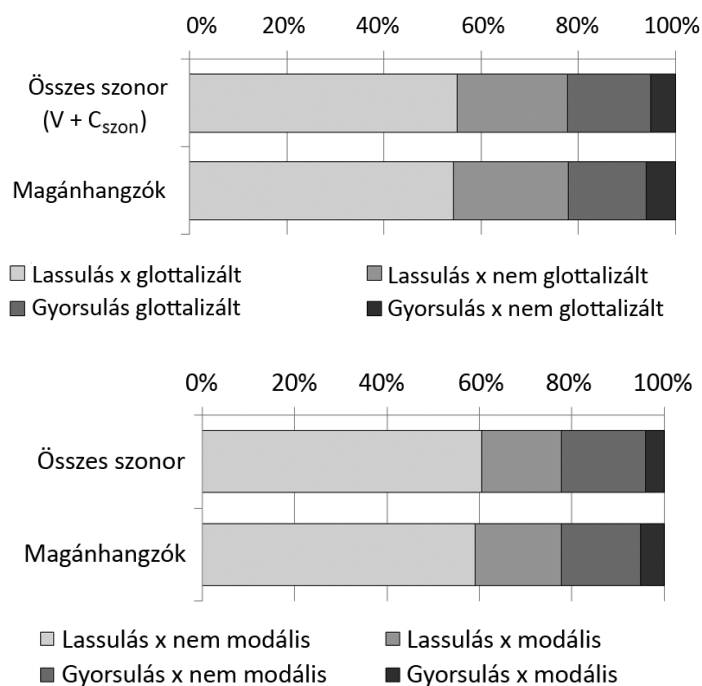
44. ábra

A jelöltség (irregularitás, illetve a modálistól való eltérés) mértéke a lassuló és a gyorsuló trendet mutató utolsó 12 hangban (fent) és 6 hangban (lent) (1 = az irreguláris zöngével megvalósult szegmentumok aránya az összes $V + C_{szon}$ számához viszonyítva; 2 = a nem modális zöngével megvalósult szegmentumok aránya az összes $V + C_{szon}$ számához viszonyítva; 3 = az irreguláris zöngével megvalósult V-k aránya az összes V számához viszonyítva; 4 = a nem modális zöngével megvalósult V-k aránya az összes V számához viszonyítva.)

A 12 beszédhangra terjedő elemzés eredménye szerint az összes szonor beszédhang közül a glottalizáltak aránya 44,1% (szórás: 31,0%), míg az összes nem modális realizáció aránya 46,5% (szórás: 31,1%). A magánhangzók között az irregulárisan ejtettek aránya 50,4% (szórás: 33,7%), a nem modális realizációké 52,6% (szórás: 33,9%). Ha az elemzést 6 beszédhangnyi terjedelempre korlátozzuk, az összes szonor ($V + C_{szon}$) 57,0%-a (szórás: 35,9%) irreguláris, és összesen 60,5%-a (szórás: 35,3%) nem modális. Ebben a körben a magánhangzók 63,9%-a (szórás: 38,2%) glottalizált, 67,4%-a (szórás: 37,5%) nem modális. Ha csak a megnyilatkozások utolsó 3 beszédhangját tekintjük, a szonorok ($V + C_{szon}$) 60,1%-a (szórás: 42,2%) irreguláris és 66,2%-a (szórás: 40,0%) nem modális; míg a magánhangzók 70,0%-a (szórás: 45,0%) glottalizált, és 76,0%-a (szórás: 42,0%) nem modális. Mindebből az is kiderül, hogy a lehetséges zöng a megnyilatkozás végén viszonylag ritka, míg a glottalizáció igencsak gyakori. Ugyanakkor a megnyilatkozás utolsó 6 beszédhangján (vagyis a 2-3 záró szótagon) mindkét jelenség gyakrabban fordul elő, mint korábban.

Ahogy vártuk, mind a glottalizáció (illetve tágabban a nem modális zöngeminőség), mind a lassulás gyakori jelenség a megnyilatkozás végén. Felmerül a kérdés, hogy csak a megnyilatkozás vége okozza a két paraméter együttes előfordulását, vagy összefüggés is kimutatható a megjelenésük között. Az elemzések ezért kiterjedtek a lassuló és a gyorsuló trendet mutató megnyilatkozásvégek összevetésére a glottalizált és a nem modális zöngeminőség megjelenési aránya szerint. Az utolsó 12 hang és az utolsó 6 hang vizsgálata alapján is a glottalizáltság vagy a modálistól eltérő zöngeminőség aránya a gyorsuló szakaszokban volt nagyobb (44. ábra). A szórások azonban minden esetben nagyak mondhatók, a statisztikai próbák nem is támasztották alá az eltérések szignifikáns voltát (12 beszédhang esetében a Mann–Whitney-próba: $Z \geq -1,685$, $p \geq 0,091$; 6 beszédhang esetében pedig: $Z = -1,011$, $p \geq 0,312$). Ez azt jelenti, hogy ezekben az egységekben sem a lassulás és a glottalizáció, sem a gyorsulás és a glottalizáció következetes együtt járása nem igazolható statisztikailag.

Az elemzések érintették a jelöltség mértékének és a temporális sajátosságoknak a kapcsolatát az utolsó 3 hang esetében is. Mivel az utolsó 3 hang vizsgálatakor a szonor mássalhangzókra számolt glottalizáltság vagy modálistól való eltérés aránya 5-féle kategóriát vehet fel, a magánhangzókra számolt pedig csak 3-féleképpen realizálódott, ezért a megjelenési arány vizsgálata elmaradt. Helyette az elemzések érintették azt a kérdéskört is, hogy az utolsó három beszédhang szonor mássalhangzói és/vagy magánhangzói eltérnek-e a modális zöngétől, ha lassuló vagy ha gyorsuló trendet mutató szakasról van szó (45. ábra). Akkor szerepelt egy szakasz glottalizáltként vagy nem modálisként, ha az utolsó három hang közül a magánhangzók vagy a szonor mássalhangzók valamelyike így realizálódott.



45. ábra

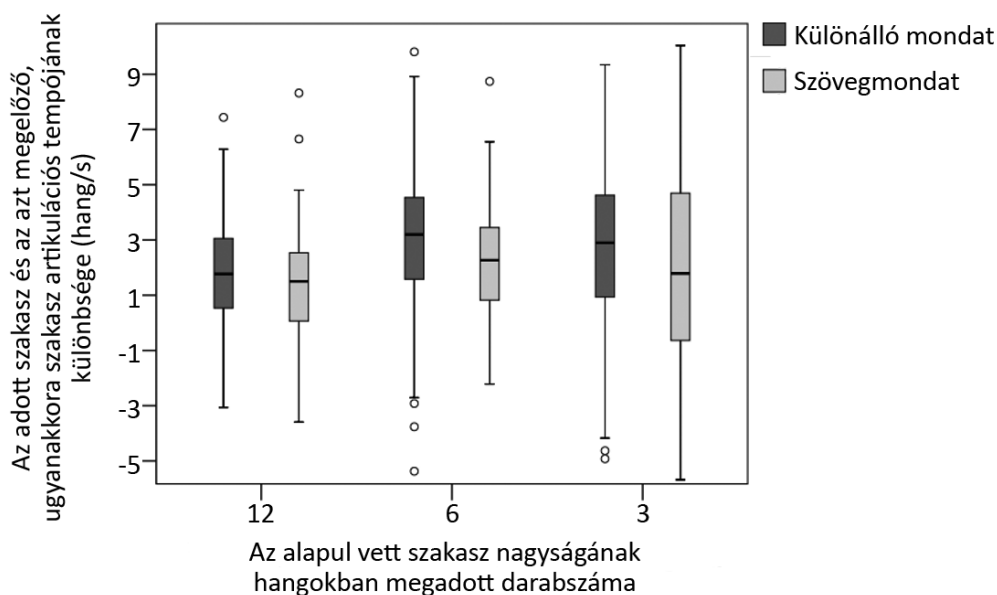
A temporális változások jellege és a zöngeminőség közötti összefüggések a megnyilatkozásokat záró 3 beszédhang esetén

Az összes szonort ($V + C_{szon}$) figyelembe véve az tapasztalható, hogy a glottalizált, lassuló trendet mutató szakaszvégek 55,1%-ban fordulnak elő, míg a nem glottalizált, lassuló szakaszok 22,7%-ban. A megnyilatkozásvégek 17,1%-a glottalizált, gyorsuló trendet mutató szakaszként realizálódott. Alig volt olyan eset (5,1%), amelyben sem lassulás, sem glottalizáltság nem volt megfigyelhető a megnyilatkozás végén. A többi mérési módszer eredményeként is hasonló arányok mutatkoztak. A megnyilatkozásvégek magánhangzói és szonor mássalhangzói 60,4%-ban nem modálisok voltak, 17,4%-ban modálisok lassuló trend esetében. A megnyilatkozások 18,2%-a volt nem modális és csak 4,0%-a modális a gyorsuló szakaszokban. Csupán a magánhangzókat figyelembe véve, a megnyilatkozásvégek 54,3%-a volt glottalizált a lassuló szakaszokban, 23,5%-uk volt nem glottalizált a lassuló szakaszokban, 16,0%-uk volt glottalizált a gyorsuló szakaszokban, és 6,1%-uk volt nem glottalizált a gyorsuló szakaszokban. A magánhangzóknál a leheletes zöngét is figyelembe véve, a megnyilatkozásvégek 59,1%-a volt nem modális a lassuló szakaszokban, 18,7%-uk volt modális a lassuló szakaszokban, 17,1%-uk volt nem modális a gyorsuló szakaszokban, végül 5,1%-uk volt modális a gyorsuló szakaszokban. Összességében tehát elmondható, hogy a megnyilatkozások végét vagy lassulás vagy nem modális zöngeminőség jelzi, igen ritkán fordul

elő, hogy egyik sem jelenik meg a megnyilatkozás végén. Ugyanakkor a két paraméter nem mutat összefüggést a χ^2 -próba alapján egyik, a zöngeminőség mérésére alkalmazott módszer esetében sem ($\chi^2(1) \leq 1,254, p > 0,257$), tehát a megnyilatkozás végi glottalizáció vagy modálislól eltérő zöngeminőség megjelenése függetlennek mutatkozik az artikulációs tempó változtatásától.

Az eddigiekben a megnyilatkozás végi temporális változást statikus jelenségként kezeltük, a lassulásnak és a gyorsulásnak azonban a mértéke is megragadható. A vizsgált szakaszok temporális változásának mértékét úgy határozhatjuk meg, hogy a szakasz nagyságával azonos nagyságú megelőző szakasz artikulációs tempójából kivonjuk a vizsgált szakasz artikulációs tempóját. Az így kapott különbség adja meg a szakaszra jellemző lassulás vagy gyorsulás jellemző mértékét. Az artikulációs tempó különbségének értéke nem mutatott összefüggést a glottalizált vagy a nem modális hangok arányával. Az utolsó 12 beszédhangot alapul véve a temporális változás mértéke nem mutatott szignifikáns összefüggést semelyik lehetséges zöngeminőséggel kapcsolatos aránnyal ($\rho \leq -0,066, p \geq 0,209$). Az utolsó 6 beszédhang esetében szintén nem mutatott szignifikáns összefüggést a statisztikai próba ($\rho \leq -0,090, p \geq 0,084$). (3 beszédhangra ez az összefüggés nem volt vizsgálható.) A megnyilatkozások végén tehát a gyorsítás vagy a lassítás mértéke nem függ össze az ott megjelenő hangok zöngeminőségével.

Az elemzés a mondat- és a szövegfelolvasás közötti lehetséges eltérésekre is kiterjedt. A megnyilatkozás végén lévő lassulás és gyorsulás mértéke függött attól, hogy a beszélők különálló mondatokat vagy szövegben szereplő mondatokat olvastak fel. A szövegolvasásnál kisebb az artikulációs tempóban lévő különbség (vagyis egyenletesebb az artikulációs tempó), mint különálló mondatok felolvasásakor, függetlenül az alapegység nagyságától. A megnyilatkozások utolsó 12 beszédhangjának és az azt megelőző 12 hang artikulációs tempójának különbsége a mondatfelolvasásokban átlagosan 1,8 hang/s (szórás: 1,9 hang/s), a szövegfelolvasásokban 1,3 hang/s (szórás: 2,0 hang/s). 6 beszédhangot alapul véve a mondatfelolvasásban mért artikulációs tempó különbsége átlagosan 3,2 hang/s (szórás: 2,6 hang/s) volt, a szövegfelolvasásban pedig 2,2 hang/s (szórás: 1,9 hang/s). 3-3 hang artikulációs tempóját összehasonlítva pedig a mondatfelolvasásokban átlagosan 2,8 hang/s (szórás: 2,7 hang/s), a szövegfelolvasásokban 1,7 hang/s (szórás: 3,5 hang/s) volt a különbség (46. ábra). A Mann-Whitney-próba mindhárom esetben szignifikánsnak mutatta az eredményeket (12 beszédhang esetében: $Z = -2,193, p = 0,028$; 6 beszédhang esetében: $Z = -4,264, p < 0,001$; 3 beszédhang esetében $Z = -2,948, p < 0,003$).



46. ábra

A gyorsulás és a lassulás mértéke a felolvasás típusától függően a különböző nagyságú szövegegységekben

Felmerül a kérdés, hogy vajon a szövegfelolvasás és a mondatfelolvasás jellemző artikulációs tempója befolyásolhatta-e az eredményeket, ezért az elemzés a mondatok teljes hosszára számolt artikulációs tempóra is kiterjedt. Az artikulációs tempó azonban mindkét típusban nagyon hasonló (mondatfelolvasás esetében az átlag 13,1, a szórás 1,3 hang/s; szövegfelolvasás esetében az átlag 13,2, a szórás 1,1 hang/s). A két típus mondatainak artikulációs tempója nem tér el szignifikánsan ($Z = -1,005$, $p = 0,315$), így a jellemző artikulációs tempó nem befolyásolhatta a lassulás vagy gyorsulás mértéke és a felolvasás típusa között talált összefüggést.

A felolvasandó anyag eltérő volta nemcsak a temporális változások mértékére, hanem a glottalizációs arányokra is befolyással lehet, ezért az elemzés a glottalizált vs. nem glottalizált, illetve a modális vs. nem modális arányokra is kiterjedt a felolvasás típusának függvényében. A jelöltségi arányok átlagosan magasabbnak mutatkoztak a szövegben, mint a felolvasásokban. Az utolsó 12 beszédhang esetében a szonorok ($V + C_{szon}$) átlagosan 43,0%-ban (szórás: 30,4%) voltak glottalizáltak különálló mondatok felolvasása esetében, a szövegfelolvasásban pedig 46,1%-ban (szórás: 31,9%). Szintén 12 beszédhangnyi egységek esetében a szonorok ($V + C_{szon}$) átlagosan 45,1%-a (szórás: 30,4%) realizálódott nem modális zöngképzéssel különálló mondatok felolvasásakor, a szövegfelolvasásban viszont 49,1%-uk (szórás: 32,4%). A magánhangzók glottalizáltsága szintén átlagosan nagyobb arányú volt szövegfelolvasásokban (átlag: 52,3%, szórás: 33,8%), mint különálló mondatokban (átlag: 49,4%, szórás: 33,8%). A magánhangzók modálistól eltérő volta is gyakrabban fordult elő szövegfelolvasásokban (55,2%, szórás: 34,3%),

mint különálló mondatokban (átlag: 51,3%, szórás: 33,6%). A statisztikai próbák azonban egyik esetben sem támasztották alá, hogy a különbség jelentős lenne ($Z = 16933,5$, $p \geq 0,230$). 6 beszédhang esetében a szonorok ($V + C_{szon}$) átlagosan 54,8%-a (szórás: 35,3%) realizálódott glottalizáltan különálló mondatok felolvasásakor, a szövegfelolvasásban viszont 61,0%-uk (szórás: 36,9%). Ugyanekkora szakaszhosszúságnál a magánhangzók átlagos glottalizációs aránya 62,0% (szórás: 38,1%) volt különálló mondatok felolvasásában, 67,6% (szórás: 38,3%) pedig szövegfelolvasásban. A statisztikai próba itt sem mutatott szignifikáns eltérést ($Z = -1,669$, $p \geq 0,095$). A modálisok és a nem modálisok arányát elemezve a felolvasás típusának függvényében, a statisztikai próbák igazolták az eltérések jelentőségét, amely eltérések a következőképpen alakultak: a szonorokra ($V + C_{szon}$) számolt nem modális arány kisebbnek mutatkozott különálló mondatokban (átlag: 58,0%, szórás: 34,6%), mint szövegfelolvasásokban (átlag: 65,2%, szórás: 36,3%), amely különbség szignifikáns ($Z = -2,080$, $p = 0,037$). A magánhangzókra számolt arány is átlagosan nagyobbnak mutatkozott a szövegfelolvasásokban (átlag: 71,9%, szórás: 37,9%), mint a különálló mondatokban (átlag: 65,1%, szórás: 37,2%), és az eltérés itt is szignifikáns a statisztikai próba alapján ($Z = -1,978$, $p = 0,048$). A nem modális zöngeminőség tehát statisztikailag is alátámaszthatóan gyakrabban jelent meg a szövegfelolvasásokban, mint a különálló mondatok felolvasásakor.

4.3.3. Következtetések

Ezen fejezetben a megnyilatkozások végének két lehetséges jelölési módját, a nem modális zöngeminőséget (ezen belül az irreguláris és a leheletes zöngét), valamint a temporális szerkezetet, illetve ezek összefüggéseit vizsgáltuk 10 magyar nyelvű beszélő felolvasásában.

Az eredmények szerint a lassulások nagy arányban jelennek meg a megnyilatkozásvégeken a korábbi kutatási eredményekhez hasonlóan, ugyanakkor a lassulás nemcsak az utolsó egy-két szótagnyi nagyságú egységeket jellemzi, hanem az artikulációs tempó csökkenése nagyobb szakaszokban (akár 6, 12 hangnyi távolságban) is felfedezhető. Az irreguláris zöng megjelenése szintén gyakori jelenség a megnyilatkozásvégeken (BÖHM–UJVÁRY 2008; MARKÓ 2010 eredményeivel is egybehangzóan), a megnyilatkozás végéhez közeledve mind önmagában az irregularitás, mind együttesen a modálistól eltérő zöngeminőség egyre nagyobb arányú. Habár az eredményekből nem dönthető el egyértelműen, hogy mekkora szakasz vizsgálata lenne alkalmasabb a szakaszvégi lassulások és a glottalizáció megragadására, a megnyilatkozás utolsó 6 beszédhangján (vagyis a 2-3 záró szótagon) mindkét jelenség gyakrabban fordul elő, mint ezt megelőzően.

Mind a lassulás, mind a zöng nem modálissá válása jelzi a megnyilatkozás végét, de a két tényező között nem volt lineáris korreláció, sem a gyakoriságuk, sem a mértékük tekintetében ebben a pozícióban. A lassulások és gyorsulások mérésére ugyanakkor nincs kialakult, bevált módszertan, tehát a kísérletben használt metódus csak egy lehetséges megközelítés. Ezért a kutatás jelen munkában közölt méréseken túl további lehetséges eljárásokra is kiterjedt.

Többek között a kiértékelések tartalmazták az artikulációs tempó különbsége helyett a nyúlás százalékos értékét is (hasonlóan: TURK–SHATTUCK–HUFNAGEL 2007), illetve finomítva a bináris kategorizáción, a lassulás és gyorsulás mellé egy stagnáló csoport is bekerült, amely csoportba azon megnyilatkozások végeinek adatai sorolódtak, ahol a változás értéke nem haladt meg az önkényesen meghatározott (10% vagy 20%) küszöbértéket. Az ekképpen módosított módszertannal lefuttatott elemzések azonban nem hoztak a fent tárgyaltaktól eltérő eredményt, ezért nem is részleteztem őket.

Összegezve: úgy tűnik, hogy pusztán a megnyilatkozás vége idézi elő a két paraméter együttes előfordulását, és alig fordul elő olyan eset, hogy egyik jelenség sem jelenik meg megnyilatkozásvégén. A megnyilatkozás vége tehát valamelyik vagy mindkét tényező által nagy valószínűséggel jelöltté válik. Ez a megállapítás nyilvánvalóan csak az adott anyagra, tehát a felolvasásra és ezen 10 beszélőre érvényes; hogy más beszédmódok és több beszélő esetén hogyan érvényesül a határjelölés, további kutatásokat igényelne. Ugyanakkor a vizsgált anyagban is előfordultak eltérések a felolvasott anyag jellegétől függően. A különálló mondatok felolvasása esetében a beszélők jobban lassítottak a megnyilatkozás végén, mint a szövegfelolvasásokban. A glottalizáció esetében ugyan nem mutatkozott ilyen szisztematikus, statisztikailag is kimutatható eltérés a két felolvasástípus között, de a szövegmondatok végi glottalizáció valamelyest nagyobb mértékű volt. Ezek az eltérések feltehetően két okra vezethetők vissza, ezek közül az egyik a beszélő attitűdje, szándéka a kétféle anyag felolvasása közben. A különálló mondatok végi lassítás mögött okként húzódhat az, hogy a különálló mondat lezárt, kompakt egység, ezért önmagában erősebben jelölt a megnyilatkozásvég. A nagyobb arányú glottalizáció a szövegmondatokban pedig talán éppen a folytatástól való markánsabb eltérés jelölésének szándéka miatt lehetséges. A másik lehetséges magyarázat a mondatok hosszában keresendő, mivel a szöveg mondatai átlagosan jelentősen hosszabbak, mint a különálló mondatok. A rövidebb különálló mondatok esetében talán emiatt nagyobb mértékben kontrollálható a tempóváltoztatás, a hosszú szövegmondatokban pedig esetleg a hangterjedelmi lehetőségek szabnak határt az f_0 -változásnak, és ezért csap át gyakrabban irregulárisba a zöngé.

Felvetődött, hogy a glottalizáció és a leheletes zöngé funkciókörében lehet átfedés, a leheletes zöngé kisszámú előfordulása miatt azonban ez a kérdéskör ebben az anyagban nem volt vizsgálható.

Összességében tehát a vizsgált anyagon nem sikerült közvetlen kapcsolatot kimutatni a nem modális zöngemínőség és a lassulás megjelenése között a megnyilatkozások végén, ennek ellenére az adatok arra utalnak, hogy a tempólassulás és a glottalizáció jellegzetes mintázatokat vesz fel, és ezek között vannak bizonyos összefüggések.

4.4. Összegzés és kitekintés

Kutatásaim elsődleges célja az volt, hogy néhány, a magyar beszédre jellemző, alapvető időzítésbeli sajátosságot feltárjam. A munka három fő kérdéskör köré szerveződött. A vizsgálatok első része arra irányult, hogy hogyan jellemezhető a magyar nyelv beszédritmusának időbeli dimenziója a nemzetközi szakirodalomból ismert beszédritmus-mérőszámokkal (lásd *4.1. fejezet*). Továbbgondolva ezen mérőszámokat elemeztem a magyar beszédben megjelenő tempóingadozásokat is (lásd *4.2. fejezet*). Továbbá kísérletet tettem a megnyilatkozás végén megjelenő tempóingadozások (lassulás/gyorsulás) és a nem modális (irreguláris/leheletes) zöngeminőség lehetséges összefüggéseinek feltárására is (lásd *4.3. fejezet*).

Eredményeim alapján arra a következtetésre jutottam, hogy a magyar beszéd beszédritmusa az általam vizsgált elméleti és módszertani keretben a szótag-időzítésű nyelvekéhez hasonló tulajdonságokat mutat (*4.1.3. fejezet*). A következő általános időzítésbeli sajátosságokat figyeltem meg az általam vizsgált magyar nyelvű spontán beszédben és felolvasásban: a magánhangzók és a mássalhangzók időtartamának aránya a beszédben közel azonos, továbbá a magánhangzós és a mássalhangzós szakaszok időtartamára nem jellemző olyan nagy variabilitás, mint ahogyan például a klasszikusan hangsúly-időzítésű angol vagy holland nyelvben mérhető (*4.1.3. fejezet*). Kevésbé jellemző rájuk a szélsőségesen rövid vagy szélsőségesen hosszú érték. Figyelembe véve sorrendjüket, ritkán található a beszéd folyamán olyan szakasz, ahol több egységen keresztül folyamatosan nőne vagy folyamatosan csökkenne a magánhangzós és mássalhangzós szakaszok vagy beszédhangok időtartama (*4.2.2.6. fejezet*). A tagmondatokat megvizsgálva mégis kimutatható, hogy tendenciaszerűen többször és nagyobb mértékben nő a mássalhangzós és magánhangzós szakaszok, valamint a beszédhangok időtartama, mint csökken (*4.2.2.1. fejezet*). Ha azonban a tagmondatoknál nagyobb egységet, például egy teljes szöveg felolvasását vizsgáljuk meg, nem tapasztaljuk ezt az általános lassuló trendet, hanem a lassulások és a gyorsulások egyfajta kiegyenlítetttsége jelenik meg. A tagmondatnál nagyobb egységet vizsgálva azt találtam, hogy az egymást követő magánhangzós és mássalhangzós szakaszok időtartam-különbségei általában nagyobbak voltak, ha egy rövidebb szakasz következett, mint amikor egy hosszabb. Továbbá az időtartamok gyakrabban növekedtek a beszéd folyamán, mint csökkentek. Ez a két összefüggés pedig a lassulások és a gyorsulások egyfajta kiegyensúlyozottságát jelenti az adatok alapján (*4.2.2.2. és 4.2.2.3. fejezet*). A lassulások egyik tipikus megjelenési helye a beszédben a megnyilatkozások vége, ami a vizsgált anyagon is igazolódott. Az irreguláris zöngé megjelenése szintén gyakori jelenség a megnyilatkozásvégeken, és az eredmények azt mutatták, hogy vagy az időszerkezetben megjelenő lassulás, vagy a zöngé nem modalitás volta, vagy mindkettő egyszerre egyfajta jelzéseként szolgál a megnyilatkozás végének. Alig fordult elő ugyanis olyan eset, amelyben egyik jelenség sem volt tetten érhető. A glottalizáció és a temporális változás közvetlenebb kapcsolatát viszont nem sikerült kimutatni (*4.3. fejezet*).

Az eddig összefoglalt eredmények felfoghatók a beszédben megjelenő szabályszerűségekként is, ily módon alkalmazhatók lennének a beszédtechnológia különböző területein. Ha nem is építhetők be feltétlen közvetlenül a jelenlegi eljárások statisztikai alapú modelljeibe, beszédfelolvasó rendszerek minőségellenőrzésére alkalmasak lehetnek (NORRENBROCK et al. 2012). A beszédritmus-mérőszámokkal mért eredmények alkalmazhatóságára ugyanakkor már több példát is láthattunk. Felhasználhatók például nyelvdetektálás során (LOUKINA et al. 2011) és érzelemfelismerésben (RINGEVAL–CHETOUANI–SCHULLER 2012; MEFIAH–ALOTAIBI–SELOUANI 2015) is. Az eredmények alapján a megnyilatkozás végén megjelenő irreguláris zöngé és/vagy a tempólassulás figyelembevétele hozzájárulhat a frázisvégek detektálásához. Természetesen nem csak a beszédtechnológia területén használhatók ezen eredmények. A beszédritmus-mérőszámokat alkalmazzák idegennyelv-elsajátítás kutatásában (WHITE–MATTYS 2007; MOK–DELLWO 2008), anyanyelv-elsajátítás kutatásában (PAYNE et al. 2011), sőt, a gyógypedagógia több területén is hasznosnak bizonyultak, például a dizartria diagnosztizálásában (LISS et al. 2009). A dadogás és a hadarás két további olyan beszédzavar, amelynek fő tünetei közé tartozik a sajátos időzítés, a ritmustalan beszéd és az átlagostól eltérő artikulációs és beszédtempó (BÓNA 2005). A legújabb kutatási eredmények alapján mind a dadogás, mind a hadarás esetében a beszédritmus-mérőszámok alkalmazásával jól megragadhatók a hadaró, illetve dadogó beszélők beszédprodukciónak időzítésbeli sajátosságai (MARUTHY–VENUGOPAL–PARAKH 2017; BÓNA–KOHÁRI 2018), aminek a gyakorlatban történő alkalmazása segítheti a személyre szabott diagnózis felállítását, és ezáltal hatékonyabb terápia kidolgozásához vezethet. A más tudományágból kölcsönzött és a beszédre alkalmazott lépéstatistikai módszer használata a beszédritmus-mérőszámokhoz hasonlóan feltehetően több területen ezek közül szintén hasznos eszköznek bizonyulna, hiszen ezen eljárás is a beszéd szupraszegmentális időbeli tulajdonságait próbálja megragadni, csak más aspektusból. Hangsúlyoznunk kell azonban, hogy a beszédritmus-mérőszámok más területen való alkalmazása során számtalan tényezőre tekintettel kell lenni, hiszen több faktor is befolyásolhatja az eredményeket nyelven belüli használat esetén is. A szakirodalom is tárgyalja a beszédmód hatását és a beszélők közti különbségeket különböző nyelvekben (ARVANITI 2012). Kutatásaim során azonban egy további összefüggést is találtam. A különböző beszélők globális artikulációs tempója összefüggést mutatott a magánhangzós szakaszok időtartamának arányával a teljes beszédidőtartamhoz képest (%V), annak ellenére, hogy ennek alapvetően függetlennek kellene lennie az artikulációs tempótól (4.1.2.5. fejezet).

Eredményeim alapján a beszédmód a magyar nyelvben a következőképpen volt hatással a beszéd időzítésbeli sajátosságaira: a spontán beszédben általánosságban nagyobb időtartambeli variabilitást tapasztaltam, azaz szélsőségesebb időtartamértékek jelentek meg ebben a beszédmódban, mint felolvasásban. Ezt az általános trendet többféle módszerrel is igazoltam. Beszélőnként azonban változó volt, hogy ez az eltérés valóban realizálódott-e (4.1.2.1. fejezet). További eltéréseket találtam a szövegfelolvasás és a spontán beszéd időzítésbeli struktúrájában is. Habár szövegfelolvasásban és spontán beszédben is gyakrabban nőtt az egymást követő beszédhangok időtartama, mint csökkent, szövegfelolvasásnál ez még inkább

jellemző volt. A csökkenő lépések nagysága jelentősebb volt, mint a növekvő lépéseké mindkét beszédmódban, de szövegfelolvasásban a növekvő és csökkenő lépések átlagos nagysága még nagyobb különbséget mutatott, mint spontán beszédben. Szövegfelolvasásban tehát az időtartamadatokban nagyobb fokú rendeződést találtam, mint spontán beszédben. A mondatfelolvasás és a szövegfelolvasás között is találtam egyértelmű eltérést (4.2.2.4. fejezet). A különálló mondatok felolvasása esetében a beszélők jobban lassítottak a megnyilatkozás végén, mint a szövegfelolvasásokban (4.3.2. fejezet). Nemcsak a különböző beszédmódok esetében találtam eltéréseket az időzítésben, hanem a nemek között is. A mássalhangzós szakaszokra építő és artikulációs tempótól független mérőszámok alapján a férfiak ejtésében átlagosan nagyobb időtartambeli variabilitás volt jellemző a nőkéhez képest (4.1.2.3. fejezet). Feltehetően a beszédritmus-mérőszám értékeire, illetve a lépésszisztematikai módszer eredményeire más tényezők is befolyással vannak, de ezen kutatási eredmények jó kiindulási alapot adhatnak további felhasználásuk anyagának és módszerének meghatározásakor.

- ABERCROMBIE, David 1967. *Elements of General Phonetics*. Edinburgh University Press, Edinburgh.
- ALLEN, Jonathan – HUNNICUTT, Sheron – KLATT, Dennis – ARMSTRONG, Robert C. – PISONI, David B. 1987. *From text to speech. The MITalk system*. Cambridge University Press, Cambridge.
- ANDÓ Éva 2002. *A történetmondás kommunikatív jellemzői*. PhD-értekezés. ELTE, Budapest.
- ARVANITI, Amalia 2009. Rhythm, timing, and the timing of rhythm. *Phonetica* 66. 46–63.
- ARVANITI, Amalia 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40. 351–373.
- BARBOSA, Plínio Almeida – VIANA, Céu – TRANCOSO, Isabel 2009. Cross-variety rhythm typology in portuguese. In *Tenth Annual Conference of the International Speech Communication Association*. Brighton, UK. 1011–1014.
- BARRY, William J. – ANDREEVA, Bistra 2001. Cross-language similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association*, 31/01. 51–66.
- BARRY, William J. – ANDREEVA, Bistra – RUSSO, Michela – DIMITROVA, Snezhina – KOSTADINOVA, Tanja 2003. Do rhythm measures tell us anything about language type? In SOLÉ, Maria-Josep – RECASENS, Daniel – ROMERO, Joaquín (eds): *Proceedings of 15th International Congress of Phonetic Sciences*. Barcelona, Spain. 2693–2696.
- BARRY, William J. – ANDREEVA, Bistra – STEINER, Ingmar 2007. The phonetic exponency of phrasal accentuation in French and German. In VAN HAMME, Hugo – VAN SON, Rob (eds): *Eighth Annual Conference of the International Speech Communication Association*. Antwerp, Belgium. 1010–1013.
- BARRY, William J. – ANDREEVA, Bistra – KOREMAN, Jacques 2009. Do rhythm measures reflect perceived rhythm? *Phonetica* 66. 78–94.
- BARTKOVA, Katarína – SORIN, Christel 1987. A model of segmental duration for speech synthesis in French. *Speech Communication* 6. 245–260.
- BEKE András – GÓSY Mária 2014. Phonetic analysis and automatic prediction of vowel duration in Hungarian spontaneous speech. *Intelligent Decision Technologies* 8/4. 301–314.
- BENTON, Matthew – DOCKENDORF, Liz – JIN, Wenhua – LIU, Yang – EDMONDSON, Jerry 2007. The continuum of speech rhythm: Computational testing of speech rhythm of large corpora from natural Chinese and English speech. In TROUVAIN, Jürgen – BARRY, William J. (eds): *Proceedings of 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany. 1269–1272.

- BERTINETTO, Pier Marco – BERTINI, Chiara 2008. On modeling the rhythm of natural languages. In BARBOSA, Plinio A. – MADUREIRA, Sandra – REIS, César (eds): *Proceedings of Speech Prosody 2008*. Campinas, Brazil. 427–430.
- BERTINETTO, Pier Marco – BERTINI, Chiara 2010. Towards a unified predictive model of Natural Language Rhythm. In RUSSO, Michela (ed.): *Prosodic Universals. Comparative studies in rhythmic modeling and rhythm typology*. Aracne, Napoli. 43–77.
- BOERSMA, Paul – WEENINK, David 2009. *Praat: doing phonetics by computer*. <http://www.praat.org> (Letöltés ideje: 2014. 08. 29.)
- BÓNA Judit 2005. A hadaró és a gyors beszéd temporális sajátosságai. *Magyar Nyelvőr* 129. 235–242.
- BÓNA Judit 2007. *A felgyorsult beszéd produkciós és percepciós sajátosságai*. PhD-értekezés. ELTE BTK, Budapest.
- BÓNA Judit 2009. *A gyors beszéd. Produkciós és percepciós sajátosságok*. MTA Könyvtára – Lexica Kiadó, Budapest.
- BÓNA Judit 2010. Beszédtervezési folyamatok az életkor és a beszédstílus függvényében. *Magyar Nyelvőr* 134. 332–341.
- BÓNA Judit 2011. A különböző beszédstílusok az akusztikai-fonetikai és percepciós vizsgálatok tükrében. *Alkalmazott Nyelvtudomány* 11. 39–48.
- BÓNA Judit 2013a. A beszédszünetek fonetikai sajátosságai a beszéd típus függvényében. *Beszédkutatás* 2013. 60–75.
- BÓNA Judit 2013b. *A spontán beszéd sajátosságai az időskorban*. ELTE Eötvös Kiadó, Budapest.
- BÓNA Judit 2015. Különböző beszéd típusok temporális sajátosságai az életkor és a nem függvényében. *Magyar Nyelvőr* 139. 201–213.
- BÓNA Judit – KOHÁRI Anna 2018. A hadaró beszéd ritmusa különböző beszédhelyzetekben. *Magyar Fonetikai, Foniátriai és Logopédiai Társaság Konferenciája (MFFLT)*. Budapest, 2018. június.
- BŐHM Tamás – UJVÁRY István 2008. Az irreguláris fonáció mint egyéni hangjellemző a magyar beszédben. *Beszédkutatás* 2008. 108–120.
- BYRD, Dani 1994. Relations of sex and dialect to reduction. *Speech Communication* 15. 39–54.
- BYRD, Dani – KRIVOKAPIC, Jelena – LEE, Sungbok 2006. How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America* 120/3. 1589–1599.
- CAMBIER-LANGEVELD, Tina 1997. The domain of final lengthening in the production of Dutch. In DE HOOP, Helen – COERTS, Jane A. (eds): *Linguistics in the Netherlands*. John Benjamins, Amsterdam. 13–24.
- CAMBIER-LANGEVELD, Tina – TURK, Alice E. 1999. A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics* 27/3. 255–280.

- CICHOCKI, Wladyslaw – SELOUANI, Sid-Ahmed – PERREAULT, Yves 2014. Measuring rhythm in dialects of New Brunswick French: Is there a role for intensity? *Canadian Acoustics* 42/3. 90–91.
- CHO, Taehong – KEATING, Patricia 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29. 155–90.
- CRYSTAL, Thomas H. – HOUSE, Arthur S. 1988. Segmental durations in connected-speech signals: Syllabic stress. *Journal of the Acoustical Society of America* 83/4. 1574–1585.
- CUMMING, Ruth E. 2010. The Interdependence of Tonal and Durational Cues in the Perception of Rhythmic Groups. *Phonetica* 67. 219–242.
- CUMMING, Ruth E. 2011. Perceptually Informed Quantification of Speech Rhythm in Pairwise Variability Indices. *Phonetica* 68. 256–277.
- DANKOVIČOVÁ, Jana 1997. The domain of articulation rate variation in Czech. *Journal of Phonetics* 25. 287–312.
- DANKOVIČOVÁ, Jana 1999. Articulation rate variation within the intonational phrase in Czech and English. In OHALA, J. John – HASEGAWA, Yoko – OHALA, Manjari – GRANVILLE, Daniel – BAILEY, Ashlee C. (eds): *Proceedings of the 14th International Congress of Phonetic Sciences*. San Francisco. 269–272.
- DAUER, Rebecca M. 1983. Stress-timing and syllable-timing re-analysed. *Journal of Phonetics* 11. 51–62.
- DAUER, Rebecca M. 1987. Phonetic and phonological components of language rhythm. GAMKRELIDZE, Tamaz V. (ed.): *Proceedings of the 11th International Congress of Phonetic Sciences*. Talin, Estonia. 447–450.
- DELLWO, Volker 2009. Choosing the right rate normalization method for measurements of speech rhythm. In SCHMID, Stephan – SCHWARZENBACH, Michael – STUDER-JOHO, Dieter (eds): *La dimensione temporale del parlato*. Torriana. 13–32.
- DELLWO, Volker 2010. *Influences of speech rate on the acoustic correlates of speech rhythm: An experimental phonetic study based on acoustic and perceptual evidence*. PhD thesis. Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany.
- DELLWO, Volker – LEEMANN, Adrian – KOLLY, Marie-José 2012. Speaker idiosyncratic rhythmic features in the speech signal. In *Thirteenth Annual Conference of the International Speech Communication Association*. 1584–1587.
- DE LOOZE, Celine 2010. *ADoReVA and ADoTeVA, two PRAAT plugins for the Automatic Detection of Register and Tempo Variations*. http://celinedelooze.com/2010DeLooze_Submission_SPSASSD.pdf (Letöltés ideje: 2015. 01. 20.)
- DEWAELE, Jean-Marc – FURNHAM, Adrian 2000. Personality and speech production: a pilot study of second language learners. *Personality and Individual Differences* 28. 355–365.
- DILLEY, Laura – SHATTUCK-HUFNAGEL, Stefanie – OSTENDORF, Mari 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24. 423–444.
- DUEZ, Danielle 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25. 11–25.

- É. KISS KATALIN – SIPTÁR Péter – KIEFER Ferenc 2003. *Új magyar nyelvtan*. Osiris Kiadó, Budapest.
- FANT, Gunnar – KRUCKENBERG, Anita 1989. Preliminaries to the study of Swedish prose reading and reading style. *Speech Transmission Laboratory Quarterly Progress and Status Report* 30/2. Royal Institute of Technology, Stockholm. 1–80.
- FLETCHER, Janet 2010. The prosody of speech: Timing and rhythm. In HARDCASTLE, William J. – LAVER, John – GIBBON, Fiona E. (eds): *The Handbook of Phonetic Sciences*. 2nd edition. Wiley-Blackwell, Oxford. 521–602.
- FÓNAGY Iván 1958. A hangsúlyról. *Nyelvtudományi Értekezések* 18.
- FÓNAGY, Iván – MAGDICS, Klára 1960. Speed of utterances in phrases of different length. *Language and Speech* 3. 179–193.
- FOUGERON, Cécile 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29/2. 109–36.
- FOUGERON, Cécile – KEATING, Patricia A. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101/6. 3728–3740.
- FOURAKIS, Marios – MONAHAN, Carolina B. 1988. Effects of metrical foot structure on syllable timing. *Language and Speech* 31/3. 283–306.
- FOX, Anthony 2000. *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford University Press, Oxford, USA.
- FROTA, Sónia – VIGÁRIO, Marina 2001. On the correlates of rhythmic distinctions: the European/Brazilian Portuguese case. *Probus* 13/2. 247–275.
- FUCHS, Robert 2014. Integrating variability in loudness and duration in a multidimensional model of speech rhythm: Evidence from indian english and british english. In CAMPBELL, Nick – GIBBON, Dafydd – HIRST, Daniel (eds): *Proceedings of Speech Prosody* 7. Dublin, 290–294.
- FUCHS, Robert 2016. *Speech Rhythm in Varieties of English: Evidence from Educated Indian English*. Springer, Germany.
- GÁBOR Ignác 1925. *A magyar ritmus problémája*. Lampel, Budapest.
- GALES, Mark – YOUNG, Steve 2008. The application of hidden Markov models in speech recognition. *Foundations and trends in signal processing* 1/3. 195–304.
- GIBBON, Dafydd – KLESSA, Katarzyna – BACHAN, Jolanta 2014. Duration and speed of speech events: a selection of methods. *Lingua posnaniensis* 56/1. 59–83.
- GOCSÁL Ákos 2000. A beszéd időviszonyai különböző életkorú személyeknél. *Beszédkutató 2000*. 39–50.
- GOCSÁL Ákos 2001. Gyorsabban beszélnek-e a nők, mint a férfiak? *Beszédkutató 2001*. 61–72.
- GOCSÁL Ákos 2010. *A beszédakusztiai paraméterek és a beszélő személyiségjegyei közötti összefüggések vizsgálata*. PhD-értekezés. ELTE BTK, Budapest.
- GOCSÁL Ákos 2014. Az artikulációs tempó vizsgálata régi magyar filmekben. *Beszédkutató 2014*. 144–154.
- GOCSÁL Ákos – HUSZÁR Ágnes 2003. Csábító hangok. *Beszédkutató 2003*. 9–18.

- GORDON, Matthew – LADEFOGED, Peter 2001. Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29. 383–406.
- GÓSY Mária 1997. A magyar beszéd tempója és a beszédmegértés. *Magyar Nyelvőr* 121. 129–139.
- GÓSY Mária 2000a. A beszédritmus elemzésének egy lehetséges megközelítése. *Magyar Nyelvőr* 124. 273–287.
- GÓSY Mária 2000b. A beszédszünetek kettős funkciója. *Beszédkutató* 2000. 1–15.
- GÓSY Mária 2003. A spontán beszédben előforduló megakadási jelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- GÓSY Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- GÓSY Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- GÓSY Mária 2008. Magyar spontánbeszéd-adatbázis – BEA. *Beszédkutató* 2008. 194–207.
- GÓSY Mária – BEKE András 2010. Magánhangzó-időtartamok a spontán beszédben. *Magyar Nyelvőr* 134. 140–165.
- GÓSY Mária – BEKE András – HORVÁTH Viktória 2011. Temporális variabilitás a spontán beszédben. *Beszédkutató* 2011. 5–30.
- GÓSY Mária – BÓNA Judit – BEKE András – HORVÁTH Viktória 2013. A kitöltött szünetek fonetikai sajátosságai az életkor függvényében. *Beszédkutató* 2013. 121–143.
- GÓSY Mária – GYARMATHY Dorottya – HORVÁTH Viktória – GRÁCSI Tekla Etelka – BEKE András – NEUBERGER Tilda – NIKLÉCZY Péter 2012. BEA: beszélt nyelvi adatbázis. In Gósy Mária (szerk.): *Beszéd, adatbázis, kutatások*. Akadémiai Kiadó, Budapest. 9–24.
- GÓSY Mária – KREPSZ Valéria 2016. A magánhangzó nyúlása a szótagszám és a mondatpozíció függvényében. *Beszédkutató* 2016. 59–88.
- GRABE, Esther – LOW, Ee Ling 2002. Durational variability in speech and the Rhythm Class Hypothesis. *Papers in Laboratory Phonology* 7. 515–546.
- GRÁCSI Tekla Etelka 2012. *Zörejhangok akusztikai fonetikai vizsgálata a zöngésségi oppozíció függvényében*. PhD-értekezés. ELTE, Budapest.
- GRAWUNDER, Sven – ENGERT, Hans – KAISER, Helen – UNGER, Stephanie – BOSE, Ines 2008. Pitch, speech rate and rhythm as speaking style characterizing parameters in German news presentation. *Workshop on Empirical Approaches to Speech Rhythm at UCL*. March 2008, London. <http://email.eva.mpg.de/~grawunde/files/LondonRhWS08poster.pdf> (Letöltés ideje: 2012. 06. 03.)
- GREGORY, Michelle L. – RAYMOND, William D. – BELL, Alan – FOSLER-LUSSIER, Eric – JURAFSKY, Daniel 1999. The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society* 35. 151–166.
- GYÜRE, Balázs – BARTOS, Imre – JÁNOSI, Imre Miklós 2007. Nonlinear statistics of daily temperature fluctuations reproduced in a laboratory experiment. *Physical Review E* 76. 037301.
- HARDCASTLE, William J. – LAVER, John – GIBBON, Fiona E. 2010. *The Handbook of Phonetic Sciences*. Blackwell, Oxford.

- HASELAGER, G. J. T. – SLIS, I. H. – RIETVELD, A. C. M. 1991. An alternative method of studying the development of speech rate. *Clinical Linguistics and Phonetics* 5/1. 53–63.
- HAYES, Bruce 1983. A grid-based theory of English meter. *Linguistic Inquiry* 14/3. 357–393.
- HELDNER, Mattias – STRANGERT, Eva 2001. Temporal effects of focus in Swedish. *Journal of Phonetics* 29. 329–361.
- HENTER, Gustav Eje – RONANKI, Srikanth – WATTS, Oliver – WESTER, Mirjam – WU, Zhizheng – KING, Simon. 2016. Robust TTS duration modelling using DNNs. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 5130–5134.
- HENTON, Caroline – BLADON, Anthony 1988. Creak as a sociophonetic marker. In HYMAN, Larry M. – LI, Charles N. (eds): *Language, speech and mind. Studies in honour of Victoria A. Fromkin*. Routledge, London – New York. 3–29.
- HINTERLEITNER, Florian – NORRENBROCK, Christoph R. – MÖLLER, Sebastian 2013. Is intelligibility still the main problem? A review of perceptual quality dimensions of synthetic speech. In BONAFONTE, Antonio (ed.): *Proceedings of ISCA Speech Synthesis Workshop*. Universitat Politècnica de Catalunya, Barcelona, Spain, 147–151.
- HOCKEY, Beth Ann – FAGYAL, Zsuzsanna 1999. Phonemic length and pre-boundary lengthening: an experimental investigation on the use of durational cues in Hungarian. In OHALA, John J. – HASEGAWA, Yoko – OHALA, Manjari – GRANVILLE, Daniel – BAILEY, Ashlee C. (eds): *14th International Congress of Phonetic Sciences*. University of California, Berkeley. 313–316.
- HOEQUIST, Charles 1983. Syllable duration in stress-, syllable and mora-timed languages. *Phonetica* 40. 203–237.
- HOFHUIS, E. M. F. J. – GUSSENHOVEN, C. – RIETVELD, T. 1995. Final lengthening at prosodic boundaries in Dutch. In ELENUS, Kjel – BRANDERUD, Peter (eds) *13th International Congress of Phonetic Sciences* 1. Stockholm University, Stockholm, Sweden. 154–158.
- HONBOLYGÓ Ferenc 2009. *A beszéd prozódiai jellemzőinek észlelése*. PhD-értekezés. ELTE PPK, Budapest.
- HONBOLYGÓ Ferenc – KOLOZSVÁRI Orsolya 2015. A hangsúly észlelésének akusztikai meghatározói. *Beszédkutatás* 2015. 21–34.
- HORNE, Merle – STRANGERT, Eva – HELDNER, Mattias 1995. Prosodic boundary strength in Swedish: Final lengthening and silent interval duration. In ELENUS, Kjel – BRANDERUD, Peter (eds): *13th International Congress of Phonetic Sciences* 1. Stockholm University, Stockholm, Sweden. 170–173.
- HORVÁTH, Viktória 2010. Filled pauses in Hungarian: Their phonetic form and function. *Acta Linguistica Hungarica* 57/2–3. 288–306.
- IMRE Angéla 2005. Különböző műfajú szövegek szupraszegmentális jellemzői. *Magyar Nyelvőr* 129. 510–520.
- ISHI, Carlos Toshinori – ISHIGURO, Hiroshi – HAGITA, Norihiro 2010. Analysis of the roles and the dynamics of breathy and whispery voice qualities in dialogue speech. *EURASIP Journal on Audio, Speech, and Music Processing* 2010. 528193. <http://asmp.erasipjournals.com/content/pdf/1687-4722-2010-528193.pdf> (Letöltés ideje: 2015. 01. 20.)

- JACEWICZ, Ewa – FOX, Robert Allen – O’NEILL, Caitlin – SALMONS, Joseph 2009. Articulation rate across dialect, age, and gender. *Language Variation and Change* 21. 233–256.
- JACEWICZ, Ewa – FOX, Robert Allen – WEI, Lai 2010. Between-speaker and withinspeaker variation in speech tempo of American English. *Journal of the Acoustical Society of America* 128. 839–850.
- JURGEC, Peter é. n. *Creaky voice in Slovene*. www.jurjec.net/publications/Creaky.pdf (Letöltés ideje: 2012. 10. 20.)
- KASSAI Ilona 1979. Időtartam és kvantitás a magyar nyelvben. *Nyelvtudományi Értekezések* 112. Akadémiai Kiadó, Budapest.
- KASSAI Ilona 1982. *A magyar beszédhangok időtartamviszonyai*. In BOLLA Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest.
- KECSKÉS András 1966. A komplex ritmuselemzés elvi kérdései. *Irodalomtörténeti Közlemények* 1–2. 106–140.
- KECSKÉS, András 1984. *A magyar vers hangzásszerkezete*. Akadémiai Kiadó, Budapest.
- KESZLER Borbála 2000. Az egyszerű és az összetett mondat határsávja. In KESZLER Borbála (szerk.): *Magyar Grammatika*. Nemzeti Tankönyvkiadó, Budapest. 461–471.
- KISLER, Thomas – SCHIEL, Florian – SLOETJES, Han 2012. Signal processing via web services: the use case WebMAUS. In *Proceedings Digital Humanities 2012*. University of Hamburg, Hamburg, Germany, 30–34.
- KLATT, Dennis H. 1973. Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54. 1102–1104.
- KLATT, Dennis H. 1975. Vowel lengthening in syntactically determined in a connected discourse. *Journal of Phonetics* 3. 129–140.
- KLATT, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59/5. 1208–1221.
- KOHÁRI Anna 2012. Rövid és hosszú magánhangzók a szóhosszúság függvényében. In VÁRADI Tamás (szerk.): *VI. Alknyelvdok Konferencia kötet*. MTA Nyelvtudományi Intézet, Budapest. 80–88.
- KOHÁRI Anna 2013a. A magyar nyelv beszédritmusának vizsgálata egy időtartamokon alapuló modell segítségével. In DRÁVUCZ Fanni – HAINDRICH Helga Anna – HORVÁTH Krisztina – KARÁCSONY Fanni (szerk.): *Félúton* 8. ELTE BTK Nyelvtudományi Doktori Iskola, Budapest. 155–172.
- KOHÁRI Anna 2013b. Az időzítés és a beszédritmus néhány összefüggése magyar spontán beszédben. In VÁRADI Tamás (szerk.): *VII. Alkalmazott Nyelvészeti Doktoranduszkonferencia kötet*. MTA Nyelvtudományi Intézet, Budapest. 93–103.
- KOHÁRI, Anna 2013c. Temporal patterns of segments and intervals in Hungarian language. In MERTENS, Piet– SIMON, Anne Catherine (eds): *Proceedings of the Prosody-Discourse Interface Conference (IDP)*. Leuven, 2013. 51–57.

- KOHÁRI Anna 2014a. A magánhangzók időtartama a beszédritmus tükrében. In GÉCSEG Zsuzsanna (szerk.): *Lingdok*. Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola, Szeged. 109–124.
- KOHÁRI Anna 2014b. Tempóváltozások a vizsgált szakasz nagyságának függvényében. *Beszédkutató* 22. 189–206.
- KOHÁRI Anna – MARKÓ Alexandra 2015. A megnyilatkozás végének jelzése: temporális szerkezet és zöngeminőség felolvasásban. *Beszédkutató* 23. 35–52.
- KOHLER, Klaus J. 1983. Prosodic boundary signals in German. *Phonetica* 40. 89–134.
- KOHLER, Klaus J. 1988. Zeistrukturierung in der Sprachsynthese. *ITG-Tagung Digitalc Sprachverarbeitung* 6. 165–170.
- KOHLER, Klaus J. 2000. Linguistic and paralinguistic functions of non-modal voice in connected speech. In *Proceedings of the 5th Seminar on Speech Production: Models and Data*. Kloster Seeon, Germany. 121–124.
- KOHLER, Klaus J. 2009. Whither speech rhythm research? *Phonetica* 66. 5–14.
- KOVÁCS Magdolna 2000. Beszédhangok kontextusfüggő időviszonyai. *Beszédkutató* 2000. 15–24.
- KOVÁCS Magdolna 2002. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen.
- KOVÁCS Magdolna 2009. *Tendenciák és szabályszerűségek a magánhangzó-időtartamok produkciójában és percepciójában*. PhD-értekezés. Debreceni Egyetem, Debrecen.
- KRIVOKAPIC, Jelena 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics* 35/2. 162–179.
- LADEFOGED, Peter 1975. *A Course in Phonetics*. Harcourt Brace Jovanovich, New York.
- LADEFOGED, Peter 2011. *A Course in Phonetics*. Harcourt Brace Jovanovich, New York.
- LAVER, John 1994. *Principles of Phonetics*. Cambridge University Press, Cambridge.
- LEHISTE, Ilse 1965. Juncture. In ZWIRNER, Eberhard – BETHGE, Wolfgang (eds): *Proceedings of the 5th International Congress of Phonetic Sciences, Münster 1964*. S. Karger, New York. 172–200.
- LEHISTE Ilse 1970. *Suprasegmentals*. MIT Press, Cambridge.
- LEHISTE, Ilse 1972. The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America* 51. 2018–2024.
- LIN, Hua – WANG, Qian 2007. Mandarin rhythm: An acoustic study. *Journal of chinese language and computing*. 17/3. 127–140.
- LINDBLOM, Björn 1968. *Temporal organisation of syllable production*. Quarterly progress and status report 2/3. Speech Transmission Laboratory, Stockholm.
- LINDBLOM, Björn 1990. Explaining phonetic variation: A sketch of the H&H theory. In HARDCASTLE, William J. – MARCHAL, Alain (eds): *Speech production and speech modelling*. Kluwer, Dordrecht. 403–439.

- LINDBLOM, Björn – RAPP, Karin 1975. Some temporal regularities of spoken Swedish. In FANT, Gunnar – TATHAM, Michel A. (eds): *Auditory analysis and perception of speech*. Academic Press, New York. 387–396.
- LISS, Julie M. – WHITE, Laurence – MATTYS, Sven L. – LANSFORD, Kaitlin – LOTTO, Andrew J. – SPITZER, Stephanie M. – CAVINESS, John N. 2009. Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research* 52/5. 1334–1352.
- LOUKINA, Anastassia – KOCHANSKI, Greg – ROSNER, Burton – KEANE, Elinor – SHIH, Chilin 2011. Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America* 129/5. 3258–3270.
- MACHAČ, Pavel – SKARNITZL, Radek 2009. *Principles of Phonetic Segmentation*. Nakladatelství Epocha, Praha.
- MADDIESON, Ian 1997. Phonetical universals. In HARDCASTLE, William J. – LAVER, John (eds): *The handbook of phonetic sciences*. Wiley-Blackwell, Oxford. 619–639.
- MAGDICS Klára 1964. A magyar érzelmi hanglejtésformák akusztikájához. *Nyelvtudományi Közlemények* 66/2. 417–420.
- MAGDICS Klára 1966. A magyar beszédhangok időtartama. *Nyelvtudományi Közlemények* 68. 125–139.
- MAIRANO, Paolo – ROMANO, Antonio 2007. A. Inter-subject agreement in rhythm evaluation for four languages (English, French, German, Italian) In TROUVAIN, Jürgen – BARRY, William J. (eds): *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken. 1149–1152.
- MAIRANO, Paolo – ROMANO, Antonio 2011. Rhythm metrics for 21 languages. In LEE, Wai-Sum – ZEE, Eric (eds): *Proceedings of the 17th International Conference on Phonetic Sciences*. City University of Hong Kong, Hong Kong. 1149–1152.
- MARKÓ Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége*. PhD-értekezés. ELTE, Budapest.
- MARKÓ Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutatás 2010*. 82–99.
- MARKÓ Alexandra 2012. A magyar hangsúly realizációinak és észlelésének összefüggése felolvasásban és spontán beszédben. In MARKÓ Alexandra (szerk.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngelkedési időig*. ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, Budapest. 277–303.
- MARKÓ Alexandra 2013. *Az irreguláris zöngé funkciói a magyar beszédben*. ELTE Eötvös Kiadó, Budapest.
- MARKÓ Alexandra 2014. Az irreguláris zöngeminőség gyakorisága és pozíciói különféle spontán beszédhelyzetekben. *Beszédkutatás* 22. 69–86.
- MARUTHY, Santosh – VENUGOPAL, Sahana – PARAKH, Priyanka 2017. Speech rhythm in Kannada speaking adults who stutter. *International journal of speech-language pathology* 19/5. 529–537.

- MEFIAH, Ali – ALOTAIBI, Yousef A. – SELOUANI, Sid-Ahmed 2015. Arabic speaker emotion classification using rhythm metrics and neural networks. *Signal Processing Conference (EUSIPCO), 2015 23rd European*. Nice, France. 1426–1430.
- MENYHÁRT Krisztina 2000. A beszéd temporális sajátosságai kétnyelvűeknél (kisiskoláskortól időskorig). *Beszédkutató 2000*. 51–62.
- MENYHÁRT Krisztina 2010. A beszédsebesség objektív mérései és szubjektív észlelési eredményeinek összefüggései mai és 60 évvel ezelőtti beszélőknél. *Beszédkutató 2010*. 110–124.
- MEYER, Ernst Alfred – GOMBOCZ, Zoltán 1909. *Zur Phonetik der ungarischen Sprache*. Berlins Buchdruckerei, Uppsala.
- MIHAJLIK Péter 2010. *Spontán magyar nyelvű beszéd gépi felismerése nyelvspecifikus szabályok nélkül*. PhD-értekezés. Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Médiainformatikai Tanszék, Budapest.
- MIHAJLIK, Péter – TUSKE, Zoltán – TARJÁN, Balázs – NÉMETH, Bottyán – FEGYÓ, Tibor 2010. Improved Recognition of Spontaneous Hungarian Speech: Morphological and Acoustic Modeling Techniques for a Less Resourced Task. *IEEE Transactions on Audio Speech and Language Processing* 18/6. 1588–1600.
- MOK, Peggy Pik Ki – DELLWO, Volker. 2008. Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, Beijing Mandarin and English. In BARBOSA, Plinio A. – MADUREIRA, Sandra – REIS, César (eds): *Proceedings of Speech Prosody 2008*. Campinas, Brazil. 423–426.
- NAGY, Péter – NÉMETH, Géza 2016. DNN-Based Duration Modeling for Synthesizing Short Sentences. In RONZHIN, Andrey – POTAPOVA, Rodmonga – NÉMETH, Géza (eds): *Speech and Computer. 18th International Conference. SPECOM 2016. Lecture Notes in Computer Science* 9811. Springer, Cham. 254–261.
- NAKAI, Satsuki – KUNNARI, Sari – TURK, Alice – SUOMI, Kari – YLITALO, Riikka 2009. Utterance-final lengthening and quantity in Northern Finnish. *Journal of Phonetics* 37. 29–45.
- NESPOR, Marina – VOGEL, Irene 1986. *Prosodic phonology*. Foris Publications, Dordrecht.
- NEUBERGER Tilda 2013. A spontán beszéd temporális sajátosságai 6–14 év közötti gyermekek-nél. *Anyanyelvpedagógia* 2013/2. <http://www.anyp.hu/cikkek.php?id=451> (Letöltés ideje: 2014. 08. 17.)
- NOLAN, Francis – JEON, Hae-Sung 2014. Speech rhythm: a metaphor? *Philosophical Transactions of Royal Society B: Biological Sciences* 369/1658. 20130396.
- NOOTEBOOM, Sieb G. 1972. *Production and perception of vowel duration: A study of durational properties in Dutch*. Ph.D. dissertation. University of Utrecht, Utrecht.
- NOOTEBOOM, Sieb G. 1997. The prosody of speech: melody and rhythm. In HARDCASTLE, William J. – LAVER, John (eds): *The handbook of phonetic sciences* 5. 640–673.
- NORRENBROCK, Christoph R. – HINTERLEITNER, Florian – HEUTE, Ulrich – MÖLLER, Sebastian 2012. Instrumental assessment of prosodic quality for text-to-speech signals. *IEEE Signal Processing Letters* 19/5. 255–258.

- O'SHAUGHNESSY, Douglas 1984. A multispeaker analysis of duration in read French paragraphs. *The Journal of the Acoustical Society of America* 75/S1. S84–S84.
- OLASZY Gábor 2000. Kísérlet a magyar beszédhangok specifikus időtartamainak meghatározására folyamatos beszédre. *Beszéd kutatás* 2000. 26–38.
- OLASZY Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Akadémiai Kiadó, Budapest.
- OLASZY 2010a. A beszéd szupraszegmentális szerkezete. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 171–208.
- OLASZY Gábor 2010b. A beszéd időszerkezetének szabály alapú modellezése. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 449–455.
- OLASZY Gábor 2010c. A beszéd szegmentális szerkezete. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest. 95–170.
- OLLER, D. Kimbrough 1973. The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America* 54/5. 1235–1247.
- ÖHMAN, Sven E. G. 1967. Numerical model of coarticulation. *Journal of the Acoustical Society of America* 41. 310–320.
- PAYNE, Elinor – POST, Brechtje – ASTRUC, Lluïsa – PRIETO, Pilar – VANRELL, Maria del Mar 2011. Measuring child rhythm. *Language and Speech* 54. 1–27.
- PETERSON, Gordon E. – LEHISTE, Ilse 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32. 693–703.
- PFITZINGER, Hartmut R. 1996. Two Approaches to Speech Rate Estimation. In McCORMACK, Paul – RUSSEL, Alison (eds): *Proceedings of the 6th Australian International Conference on Speech Science and Technology (SST '96)*. Adelaide, Australia. 421–426.
- PFITZINGER, Hartmut R. 1999. Local Speech Rate Perception in German Speech. In OHALA, J. John – HASEGAWA, Yoko – OHALA, Manjari – GRANVILLE, Daniel – BAILEY, Ashlee C. (eds): *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco. 893–896.
- PIKE, Kenneth 1945. *The intonation of American English*. University of Michigan Press, Ann-Arbor.
- PORTELE, Thomas – SENDLEMEIER, Walter – HESS, Wolfgang. 1990. A system for German speech synthesis based on demisyllables, diphones, and suffixes. In BAILLY, Gérard – BENOÎT, Christian (eds): *ESCA Workshop on Speech Synthesis Autrans*. Autrans, France. 161–164.
- PRESS, William H. – TEUKOLSKY, Saul A. – VETTERLING, William T. – FLANNERY, Brian P. 2002. *Numerical Recipes*. Cambridge University Press, Cambridge.
- PRIETO, Pilar – VANRELL, Maria del Mar – ASTRUC, Lluïsa – PAYNE, Elinor – POST, Brechtje 2010. Speech rhythm as durational marking of prosodic heads and edges. Evidence from Catalan, English, and Spanish. *Proceedings of Speech Prosody*. Chicago, USA. 951.

- PRIETO, Pilar – VANRELL, Maria del Mar – ASTRUC, Lluïsa – PAYNE, Elinor – POST, Brechtje 2012. Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication* 54/6. 681–702.
- PYLKKÖNEN, Janne – KURIMO, Mikko 2004. Duration modeling techniques for continuous speech recognition. *Proceedings of the 8th International Conference on Spoken Language Processing (Interspeech 2004)*. International Speech Communication Association (ISCA), Jeju Island, Korea. 385–388.
- QUENÉ, Hugo 2004. What is the Just Noticeable Difference for tempo in speech? In QUENÉ, Hugo – HEUVEN, Vincent J. (eds): *On speech and language: Studies for Sieb G. Nooteboom*. Netherlands Graduate School of Linguistics, Utrecht. 149–158.
- QUENÉ, Hugo 2007. On the just noticeable difference for tempo in speech. *Journal of Phonetics* 35. 353–362.
- QUENÉ, Hugo 2008. Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *Journal of the Acoustical Society of America* 123. 1104–1113.
- RAKERD, Brad – SENNET, William – FOWLER, Carol A. 1987. Domain-final lengthening and foot-level shortening in spoken English. *Phonetica* 44. 147–155.
- RAMUS, Franck – NESPOR, Marina – MEHLER, Jacques 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 72. 1–28.
- RAO, K. Sreenivasa 2012. *Predicting Prosody from Text for Text-to-Speech Synthesis*. Springer Science – Business Media.
- RIETVELD, Toni – KERKHOFF, Joop – GUSSENHOVEN, Carlos 2004. Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics* 32. 349–371.
- RINGEVAL, Fabian – CHETOUANI, Mohamed – SCHULLER, Björn W. 2012. Novel Metrics of Speech Rhythm for the Assessment of Emotion. In *Interspeech 2012. 13th Annual Conference of the International Speech Communication Association*. International Speech Communication Association (ISCA), Portland, USA. 346–349.
- ROACH, Peter 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In CRYSTAL, David (ed.): *Linguistic Controversies*. Edward Arnold, London. 73–79.
- ROACH, Peter 2002. *A little encyclopaedia of phonetics*. University College Dublin, Dublin. <http://www.linguistics.readline.ac.uk/staff/peter.Roach> (Letöltés ideje: 2014. 09. 01.)
- RUSSO, Michela – BARRY, William J. 2008. Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. In BARBOSA, Plínio A. – MADUREIRA, Sandra – REIS, César (eds): *Proceedings of Speech Prosody 2008*. Campinas, Brazil. 419–422.
- SAVY, Renata – VOGHERA, Miriam 2010. A corpus-based study on syntactic and phonetic prosodic phrasing boundaries in spontaneous Italian speech. In *Proceeding of Speech Prosody 2010*. Chicago. 1892–1896.
- SCHERER, Klaus R. 1995. How emotion is expressed in speech and singing. In ELENIUS, Kjell – BRANDERUD, Peter (eds): *Proceedings of the 13th International Congress of Phonetic Sciences*. Stockholm. 90–96.

- SCHERER, Klaus R. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40. 227–256.
- SCHIEL, Florian 1999. Automatic Phonetic Transcription of Non-Prompted Speech. In OHALA, J. John – HASEGAWA, Yoko – OHALA, Manjari – GRANVILLE, Daniel – BAILEY, Ashlee C. (eds): *Proceedings of the 14th International Congress of Phonetic Sciences*. University of California, San Francisco. 607–610.
- SILVERMAN, Kim – PIERREHUMBERT, Janet 1990. The timing of prenuclear high accents in English. *Papers in laboratory phonology* I. 72–106.
- SIPTÁR, Péter – TÖRKENCZY, Miklós 2000. *The phonology of Hungarian*. Oxford University Press, Oxford.
- SLIFKA, Janet 2006. Some physiological correlates to regular and irregular phonation at the end of an utterance. *Journal of Voice* 20/2. 171–186.
- SLIFKA, Janet 2007. Irregular phonation and its preferred role as a cue to silence in phonological systems. In TROUVAIN, Jürgen – BARRY, William J. (eds): *Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken 6–10 August 2007*. Pirrot GmbH., Dudweiler. 229–232.
- SUOMI, Kari 2007. On the tonal and temporal domains of accent in Finnish. *Journal of Phonetics* 35. 40–55.
- SZABÓ Eszter 2008. A szomorú és a vidám érzelmi állapot megjelenése a beszédben. *Magyar Pszichológiai Szemle* 63/4. 691–708.
- SZASZÁK György – MIHAJLIK Péter – FEGYÓ Tibor 2010. A beszéd gépi észlelése. In NÉMETH Géza – OLASZY Gábor (szerk.): *A magyar beszéd. Beszédkutatás, beszédtechnológia, beszéd-információs rendszerek*. Akadémiai Kiadó, Budapest. 333–410.
- TABAIN, Marija 2003. Effects of prosodic boundary on /aC/ sequences: Articulatory results. *Journal of the Acoustical Society of America* 113. 2834–2849.
- TARNÓCZY Tamás 1974. A magánhangzók akusztikai vizsgálatának problémái. *Általános Nyelvészeti Tanulmányok* 10. 153–180.
- TAYLOR, Paul 2009. *Text-to-speech synthesis*. Cambridge university press.
- TILSEN, Sam – JOHNSON, Keith 2008. Low-frequency Fourier analysis of speech rhythm. *Journal of the Acoustical Society of America* 124/2. 34–39.
- TORGERSEN, Eivind Nessa – SZAKAY, Anita 2012. An investigation of speech rhythm in London English. *Lingua* 122/7. 822–840.
- TÓTH Bálint 2013. *Rejtett Markov-modell alapú gépi beszéd-keltés*. PhD-értekezés. Budapesti Műszaki és Gazdaságtudományi Egyetem Távközlési és Média-informatikai Tanszék, Budapest.
- TÓTH, Bálint – NÉMETH, Géza 2010. Improvements of Hungarian Hidden Markov Model-based Text-to-Speech Synthesis. *Acta Cybernetica* 19/4. 715–731.
- TÓTH, László – KOCSOR, András 2005. Explicit duration modelling in HMM/ANN hybrids. *Proceeding of International Conference on Text, Speech and Dialogue*. Springer-Verlag, Berlin, Heidelberg. 310–317.

- TROUVAIN, Jürgen 2004. *Tempo variation in speech production: Implications for speech synthesis*. Doctoral dissertation. Saarland University, Saarbrücken.
- TURK, Alice E. – SAWUSCH, James R. 1997. The domain of accentual lengthening in American English. *Journal of Phonetics* 25. 25–41.
- TURK, Alice E. – SHATTUCK-HUFNAGEL, Stefanie 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28. 397–440.
- TURK, Alice E. – SHATTUCK-HUFNAGEL, Stefanie 2007. Multiple targets of phrase-final lengthening in American English. *Journal of Phonetics* 35. 445–472.
- TURK, Alice E. – SHATTUCK-HUFNAGEL, Stefanie 2014. Timing in talking: What is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society B: Biological Sciences* 369/1658. 20130395.
- TURK, Alice E. – WHITE, Laurence 1999. Structural influences on accentual lengthening in English. *Journal of Phonetics* 27. 171–206.
- VAN SANTEN, Jan 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.
- VÁRADI Viola 2009. Hallásalapú és vizuális alapú közlések. *Beszéd kutatás 2009*. 228–39.
- VÁRADI Viola 2010. A felolvasás és a spontán beszéd temporális sajátosságainak összehasonlítása. *Beszéd kutatás 2010*. 100–109.
- VÁRADI Viola 2011. A felolvasás és a spontán beszéd összevetésének pedagógiai vonatkozásai. *Anyanyelv-pedagógia* 2011/4. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=343> (Letöltés ideje: 2016. 03. 07.)
- VÁRADI Viola 2013. *A spontán beszéd szegmentálása produkciós és percepciósszempontról*. PhD-értekezés. Eötvös Loránd Tudományegyetem Bölcsészettudományi Kar, Budapest.
- VÁRADI Viola – BEKE András 2013. Az artikulációs tempó variabilitása felolvasásban. *Beszéd kutatás 2013*. 26–41.
- VARGA, László 1998. Rhythmical variation in Hungarian. *Phonology* 15. 227–266.
- VARGA, László 2002. *Intonation and stress: evidence from Hungarian*. Palgrave Macmillan, New York.
- VARGA László 2005. *Metrikus fonológia és ritmikai hangsúlyváltozás*. Tinta Könyvkiadó, Budapest.
- VERHOEVEN, Jo – DE PAUW, Guy – KLOOTS, Hanne 2004. Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech* 47/3. 297–308.
- VOLSKAYA, Nina B. 2003. Virtual and real pauses at clause and sentence boundaries. In SOLÉ, Maria-Josep – RECASENS, Daniel – ROMERO, Joaquín (eds): *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona. 499–502.
- WAGNER, Petra – WINDMANN, Andreas 2009. No Time to Lose? Time Shrinking Effects enhance the Impression of Rhythmic „Isochrony” and Fast Speech Rate. *10th Annual Conference of the International Speech Communication Association 2009. Interspeech*

2009. International Speech Communication Association (ISCA), Brighton, United Kingdom. 1523–1526.
- WAGNER, Petra – WINDMANN, Andreas 2011. The shrinking effects on speech tempo perception. In LEE, Wai-Sum – ZEE, Eric (eds): *Proceedings of the 17th International Conference on Phonetic Sciences*. City University of Hong Kong, Hong Kong. 2082–2085.
- WATTS, Oliver – HENTER, Gustav Eje – MERRITT, Thomas – WU, Zhizheng – KING, Simon 2016. From hmms to dnns: where do the improvements come from? *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 5505–5509.
- WHITE, Laurence 2002. *English speech timing: A domain and locus approach*. University of Edinburgh dissertation, Edinburgh.
- WHITE, Laurence – MÁDY, Katalin 2008. The long and the short and the final: phonological vowel length and prosodic timing in Hungarian. In BARBOSA, Plinio A. – MADUREIRA, Sandra – REIS, César (eds): *Proceedings 4th Speech Prosody Conference*. Campinas. 363–366.
- WHITE, Laurence – MATTYS, Sven L. 2007. Calibrating rhythm: first language and second language studies. *Journal of Phonetics* 35. 501–522.
- WHITE, Laurence – MATTYS, Sven L. – SERIES, Lucy – GAGE, Suzi 2007. Rhythm metrics predict rhythmic discrimination. In TROUVAIN, Jürgen – BARRY, William J. (eds): *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken. 1009–1112.
- WHITE, Laurence – MATTYS, Sven L. – WIGET, Lukas 2012. Language categorization by adults is based on sensitivity to durational. *Journal of Memory and Language* 66. 665–679.
- WHITESIDE, Sandra P. – HODGSON, Carolyn 2000. Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication* 32/4. 267–285.
- WIGHTMAN, Colin W. – SHATTUCK-HUFNAGEL, Stefani – OSTENDORF, Mari – PRICE, Patti J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91. 1707–1717.
- XU, Yi 2009. Timing and coordination in tone and intonation. An articulatory-functional perspective. *Lingua* 119/6. 906–927.
- ZEN, Heiga 2015. Acoustic Modeling in Statistical Parametric Speech Synthesis-From HMM to LSTM-RNN. *Proceedings of Machine Learning in Speech and Language Processing (MLSPL)*. University of Aizu, Aizu-Wakamatsu, Japan. <https://pdfs.semanticscholar.org/0495/69a2aa4f228626755b8979e0e6144feb5605.pdf> (Letöltés ideje: 2018. 07. 05.)

1. számú melléklet:

Szövegfelolvasás

Méreg vagy vitamin?

A zöldségeken és gyümölcsökön lévő növényvédő szerek egy része lemoshatatlan, ezért fogyasztás során a szervezetünkbe kerül, ahol immunrendszeri zavarokat, daganatos megbetegedést okozhat. Télen és kora tavasszal nehéz ellenállni a boltokban és a piacon kínált friss salátának, reteknek vagy éppen az epernek. Aki mégis beadja a derekát, kétszeresen is csaltkozhat. Egyrészt a várt ízvilág messze elmarad a szezonális termékekétől, viszont ennél talán súlyosabb problémát jelentenek a primőr zöldségekbe, gyümölcsökbe felszívódott lemoshatatlan növényvédő szerek, amelyeket gyanútlanul elfogyasztunk. Hazánkban évente közel 20 ezer tonna növényvédő szert használnak fel. Ez azt jelenti, hogy Magyarországon évente egy lakosra két kiló méreg jut. A legtöbb háztartásban télen és tél végén kedvelt vitaminforrás a primőr fejes saláta, amely a szakhatóságok elmúlt években végzett vizsgálatai alapján a legtöbb lemoshatatlan mérget tartalmazza. A kismamák sem sejtik, hogy a bébiételek 20 százalékában találtak a növényvédő szerekből visszamaradt vegyi anyagokat. A hazai hatóságok rendszeresen vizsgálják a boltokban és piacokon kapható zöldségek, gyümölcsök, illetve a feldolgozott és a bébiételek növényvédőszer-tartalmát is. A vizsgálatok szerint a hazai eredetű termékekben valamivel kisebb, az importált árukban pedig magasabb arányban mutathatóak ki a különféle növényvédő szerek. A zöldségekbe, a gyümölcsökbe olyan mérgek kerülhetnek, amelyek még mosás után is hormonális, immunrendszeri zavarokat, daganatos megbetegedéseket okozhatnak. A permetszerek hatóanyagainak jelentős része nem a kezelni kívánt növényekre és kártevőkre kerül, hanem a környezetbe jutva károsíthatja a környező területek élővilágát, szennyezheti a talajt és a vizeket – olvasható a Levegő Munkacsoport tanulmányában.

2. számú melléklet:

Mondatfelolvasás

A nagymama specialitása kétséget kizáróan a barackbefőtt.
Az ügyfeleknek kompromisszumot kellett kötniük.
A tegnapi esti repülőszerecsétséget senki sem élte túl.
A farsangi bálban mindenkinek szép jelmeze volt.
A tulipánágások gyomlálásával ütötte el az időt.
A forgalom miatt csak nehezen lehetett előrehaladni az autópályán.
A kormányrendelet alapján tilos lesz petárdázni szilveszterkor is.
Hétfvégén a gyerekek a Bakonyba mennek kirándulni.
A gyermekek bukfencezni is megtanulnak testnevelésórán.
Nem lehet eléggé hangsúlyozni a túlzásba vitt napozás veszélyét.
Betegség miatt maradt el a tegnapi estére tervezett színházi előadás.
A gyöngyhalászok hosszú ideig képesek a víz alatt maradni.
A kalauz szigorúan ellenőrzi a menetjegyeket és az igazolványokat.
A magyar úrturista kalandjait az egész országban figyelték.
Nem lehetett teljes bizonyossággal megítélni a vádlott elmeállapotát.
Külföldi festményeket raboltak el tegnap egy vidéki galériából.
Megéri biztosítást kötni minden külföldi utazás előtt.
Szerencsétlenül alakultak az események a tegnapi túszmntő akció során.
Nem lehet egyértelműen kijelenteni, hogy kinek volt igaza.
Több héten keresztül nem esett az eső, így a föld kiszáradt.
Nem lehetett megjósolni, mikorra várható a vulkánkitörés.
A minap önmagát kiáltotta ki a legnagyobb énekesnek a világon.
A hegyimentőknek sűrű köddel kellett megküzdeniük.
Nem kötött biztosítást, ezért kisebb vagyonba került a kórházi ellátás.
Még nem lehet tudni, milyen súlyos a buszbaleset sérültjeinek állapota.

*3. számú melléklet:***Minta egy beszélő spontán beszédéből**

hát az eltén vagyok magyar szakos és ő ezt elsősorban az irodalmas része miatt választottam a dolognak mert m hát amikor középiskolás voltam akkor még sokkal inkább az irodalom érdekelt mint a nyelvtan vagy a nyelvészet meg középiskolában nem is annyira volt rálátásunk erre tehát az hogy nyelvészet az nekünk tehát ilyen nyelvtanórában kimerült és ennyi és ő hát nagyon szerettem verseket olvasni meg verset elemezni meg ilyesmi és hát úgy gondoltam hogy akkor így meg kellene próbálkozni egy magyar szakkal bár már akkor is ilyen furcsán néztek rám emiatt hogy hát ugye ez nem annyira ilyen nem is tudom olyan szakma ahol így halálra keresi magát az ember meg ilyesmi meg hogy eleve nem tudják így elképzelni az emberek hogy mit kezd magával egy magyar szakos bölcsész aztán ő hát végül is nagy bátran idejöttem az eltére magyar szakra és akkor rájöttem hogy ugye nem az irodalom meg a magyar tanulás az nem merül ki az irodalomban az irodalom sem merül ki abban amit középiskolában tanulunk hát a nyelvészet meg pláne nem és ő hát nagyon jó volt mert egy csomó mindennel így megismerkedtem és akkor rájöttem hogy tulajdonképpen engem a nyelvész része sokkal jobban érdekel mint az irodalmas és hát ő jelenleg igazából így legjobban ez a fonetikai része érdekel a dolognak és hát azért is jöttem adatközlőnek mert nekem az tök érdekes hogy így miket csinálnak szóval hogy hogy abból mit lehet összehozni öm és hát még az egy jó kérdés hogy mit szeretnék majd magammal kezdeni de hát ő voltam a ezen az umnya kérdőívezésen dialektológiával kapcsolatban szóval hogy így a bárt jánosnál tanultam dialektológiát és ő mondta hogy lehet menni és akkor így valahogy nem tudom ez így nagy áttörés volt nekem hogy mi a különbség a között amikor így tudod tanuló tankönyvből a dolgokat meg amikor tényleg kint agy terepen és akkor találkozol egy jelenséggel és ez ilyen nagyon pozitív élmény volt úgyhogy azóta ez most már elégé biztos hogy hogy ezzel szeretnék foglalkozni csak még ugye akkora rálátom nincs a dolgokra mert ugye első éves vagyok

Temporal patterns in Hungarian speech

Understanding the general trends and variability in the temporal patterns of speech has attracted attention over many years. In the present volume traditional and novel techniques and mathematical tools are presented to quantify timing structures in Hungarian read and spontaneous speech.

The first part of the book provides an overview of current research on timing in speech, discusses the relevant terminological issues, and focuses on the temporal aspects of speech rhythm theories. The second part then presents my experiment-based acoustic phonetic investigations pertaining to temporal patterns in Hungarian speech. The main results can be summarized as follows:

1. By analyzing measured data (obtained from 20 speakers of standard Hungarian) I have confirmed the statement found in the literature that the characteristics of Hungarian speech rhythm are similar to languages traditionally classified as syllable-timed.

2. I found differences between the speech rhythm properties of different speech styles in Hungarian speech. Generally, the duration of both vocalic and consonantal intervals exhibited higher variability in spontaneous speech than in read speech. The values of vowel-based speech rhythm measures (ΔV , VarcoV , nPVI-V) obtained for the whole record were found to be consistently higher in spontaneous speech than in read speech production for all speakers. In terms of the other measures this difference between the speech styles did not appear manifestly in every realization.

3. I have disproved the statement found in the literature that the ratio between the combined duration of all vocalic intervals and the total duration of the speech (%V) is independent from articulation rate. In the data the global articulation rate of different speakers showed a certain dependency on %V. This observation implies that speaking slowly or fast is connected to the duration of consonantal and vocalic intervals relative to each other. However, variations within the articulation rate of a given speaker do not show any connection to this ratio.

4. I have confirmed (analyzing data from 20 speakers of standard Hungarian) that the segmental durations in utterances of read speech exhibit a generally decelerating temporal trend.

5. Concerning the read speech and spontaneous monologues of the speakers considered as a whole, no significant global (decelerating or accelerating) trends were found in terms of stepnumber-stepsize statistics. Instead, it appears that the local temporal patterns (decelerations and accelerations) balance out on such longer time scales.

6. When reading separate unrelated sentences aloud speakers tend to decelerate their speech at the end of the utterance more than in the case of reading aloud a coherent text.

7. I have demonstrated that in read speech the utterance-final positions are frequently marked by nonmodal phonation, decelerating speech tempo, or both. However, with the applied methodology, no direct connection could be revealed between the utterance-final appearance of the two factors.

A
BESZÉD • KUTATÁS • ALKALMAZÁS

című sorozat eddig megjelent kötetei:

MARKÓ ALEXANDRA:

Az irreguláris zöngé funkciói a magyar beszédben

ISBN 978-963-312-195-5

BÓNA JUDIT:

A spontán beszéd sajátosságai az időskorban

ISBN 978-963-312-199-3

HORVÁTH VIKTÓRIA:

Hezitációs jelenségek a magyar beszédben

ISBN 978-963-312-205-1

NEUBERGER TILDA:

A spontán beszéd sajátosságai gyermekkorban

ISBN 978-963-312-204-4

BEKE ANDRÁS:

Gépi beszélődetektálás magyar nyelvű spontán társalgásokban

ISBN 978-963-312-234-1

DEME ANDREA:

Magánhangzók ejtése és észlelése a szopránéneklésben

ISBN 978-963-312-261-7

TAR ÉVA:

Fonológiai fejlődés, variabilitás, beszédhanghibák

ISBN 978-963-312-289-1

KOHÁRI ANNA:

Időzítési mintázatok a magyar beszédben

ISBN 978-963-312-295-2

A magyar beszéd időzítésével kapcsolatban rengeteg leírás megjelent már. A beszédritmus eddigi megközelítései, az ezzel kapcsolatos korábbi állítások a ritmus sajátosságainak megragadhatatlanságát, a sok változóból adódó bizonytalanságot sugallták, nem véletlenül. A jelenség valóban soktényezős, az adatok gyakran ellentmondásosak, nehéz fogódzókat találni a megfelelő módszertan kialakításához. A könyv szerzője a nemzetközi szakirodalom széles körének ismeretében új, korábban a (magyar) beszédre még nem alkalmazott módszertanokat használva, különféle metódusokat ötvözve, szisztematikus elemző munkával jutott el azon eredményekig és megállapításokig, amelyek ebben a kötetben napvilágot látnak, és amelyek a legkorszerűbb ismereteinket foglalják össze a (magyar) beszéd ritmusának vonatkozásában.

A kötet nemcsak a szűkebb, fonetikusokból álló olvasótábor érdeklődésére tarthat számot, hanem olyan területek művelői is építhetnek a benne foglalt ismeretekre, akiknek kutatása a beszéd időzítéséhez bármilyen módon kapcsolódik, a logopédiától a beszédtechnológiáig. Azonban az a kutatómódszertani innováció, amelyre a kötet példát ad, még távolabbi, a beszédhez nem vagy kevésbé kapcsolódó, de az időzítés mintázatait magában rejtő jelenségek leírásában is haszonnal kecsegtet. Így a könyv bátran ajánlható a szélesebb érdeklődő közönség számára is.

ISBN 978-963-312-295-2



9 789633 122952