

A vizuális csatorna befolyása a beszédalkalmazkodásra – adatok az alapfrekvenciáról és a beszédtempóról

Maros Anna – Négyesi Regina Gréta – Németh Aletta – Varga Anna Enikő

Eötvös Loránd Tudományegyetem

Bölcsészettudományi Kar

Tartalom

1	Bevezetés	3
1.1	A téma ismertetése, kontextualizálása	3
1.2	A kutatás célja, a témaválasztás oka	3
2	Háttér.....	3
2.1	Elméleti háttér.....	3
2.2	A kutatás közvetlen előzményei.....	7
3	A kutatás célja, kérdése és hipotézise	9
4	Módszertan.....	10
4.1	A kísérletben résztvevő személyek	10
4.2	Anyag.....	11
4.3	Módszer.....	11
5	Eredmények	13
5.1	Alapfrekvencia	13
5.2	Hangterjedelem	16
5.3	Beszédtempó	17
6	Következtetések	20
6.1	Továbbvezető kérdések	22
7	Összefoglalás	22
8	Szakirodalom	24

1 Bevezetés

1.1 A téma ismertetése, kontextualizálása

A kutatásunk fonetikai értelmezésben és vonatkozásban vizsgálja a beszédalkalmazkodást. A beszédalkalmazkodás Goldinger (1998) megállapításából kiindulva Babel és munkatársai (2014) szerint fonetikai imitáció, amely az egyik beszélő által a másik beszélő beszédjellemzőinek spontán, nem szándékos átvételét jelenti.

Ebből kifolyólag a fonetikai alkalmazkodás egy olyan tendencia, amelynek értelmében a beszélők akusztikailag konvergálnak a beszélgetőtársukhoz kommunikációs helyzetekben. Az alkalmazkodás célja lehet a kommunikáció sikerességének elérése, a szimpátia elnyerése vagy növelése, esetleg pozitív benyomás keltése a hallgatóban (Barón-Birchenall 2023).

1.2 A kutatás célja, a témaválasztás oka

A jelen kutatás célja a beszédalkalmazkodás vizsgálata kvantitatív módszerekkel, fonetikai paraméterek mentén. A témaválasztást indokolja, hogy korábban kevés olyan kutatást végeztek, amely kitért volna arra, hogy milyen befolyása lehet a beszédalkalmazkodás folyamatára annak, ha a felek látják, és nem csak hallják egymást. A jelen vizsgálat erre a kérdésre keresi a választ.

2 Háttér

2.1 Elméleti háttér

Kutatásunk Roman Jakobson 1958-as modellje alapján értelmezi a kommunikációt, aki a feladó, a címzett, az üzenet, a kód, a kontextus és kontaktus tényezői figyelembevételével alkotta meg a modelljét. A feladó az, aki üzenetet küld a címzetthez. Közös kódra van szükség ahhoz, hogy mindkét fél számára érthető legyen az üzenet. A kód akkor működőképes, ha egy mindkét fél számára ismert valóság létezik. A kontaktus jelenti a feladó és a címzett közötti kapcsolatot, ami feltételezi a megfelelő fizikai csatorna (látási, hallási) meglétét (Fercsik–Raátz 2018).

A kommunikáció nemcsak információcserét jelent, hanem kapcsolódási pontot személy és csoport között, tehát a kommunikáció alakítja az emberi, csoportos kapcsolatokat (Gallois et al. 2005). A Howard Giles által kidolgozott kommunikációs alkalmazkodáselmélet (CAT – Communication Accommodation Theory) azzal foglalkozik, hogy az emberek viselkedése

hogyan változik a kommunikáció során (Giles 1991, Gallois et al. 2005). Hogyan hangolják a kommunikációjukat partnerükre, milyen kitételei vannak annak, hogy a kommunikációs partnerük úgy érezze, hogy a kommunikáció során a résztvevő a partnerre hangolódott. A CAT elmélete továbbá azt is számításba veszi, hogy milyen okok és célok vezérlik a kommunikációs viselkedést (Gallois et al. 2005).

A beszédalkalmazkodás-elmélet (SAT – Speech Accommodation Theory) azért jött létre, hogy magyarázatot adjon a motivációkra, amelyek a szociális helyzetekben az egyének beszédstílusának változása mögött állnak. Továbbá célja az elméletnek, hogy magyarázatot találjon a beszédalkalmazkodásból eredő társadalmi következményekre. A SAT alapja a kölcsönös kommunikációs hatások vizsgálata, ehhez figyelembe veszi a társadalmi és kognitív tényezőket, illetve a kommunikációs jellegzetességek a kommunikációs folyamatban tapasztalható változások mögött rejtőző társadalmi okokat, és az ezeket meghatározó tényezőket és motivációkat. A SAT elméletét a beszélők által produkált viselkedések széles körére lehet alkalmazni, például a beszédre, hangzó kommunikációra is (Gallois et al. 2005).

A beszédalkalmazkodás elmélete három fő szempont szerint vizsgálja a beszédalkalmazkodást. Ezek a beszédalkalmazkodás beszédprodukciós oldala (azaz annak a kérdése, hogy hogyan valósul meg a beszédalkalmazkodás a beszéd létrehozásában), a beszédalkalmazkodás mértéke és a beszédalkalmazkodás fogadtatása a hallgatóban, címzettben.

A beszédprodukció, azaz a beszédalkalmazkodás megvalósítása szempontjából háromféle stratégia-típus különíthető el a beszédalkalmazkodásban: a konvergencia, a divergencia és az egyéni beszédmintázatok fenntartása. A konvergencia az a stratégia, amely szerint az egyének úgy alkalmazkodnak kommunikációs viselkedésükben, hogy hasonlóvá váljanak a hallgatóhoz, a címzethez (Gallois et al. 2005). A konvergencia (közeledés) jelenthet tudatos alkalmazkodást (Bata 2009): a beszélők nagyobb eséllyel igazítják, konvergálják a beszédüket a címzethez, hallgatóhoz, ha a beszélők a hallgató elismerését, jóváhagyását, szimpátiáját szeretnék elnyerni. Fontos kitétele ennek, hogy az alkalmazkodás, a beszédben meghozott kompromisszum ne követeljen nagyobb erőfeszítést, mint az azért kapott hallgatói elismerés, jóváhagyás, az elnyert szimpátia (Gallois et al. 2005). Ezzel szemben a divergencia (eltávolodás) tartózkodást jelent a hallgatóhoz, címzethez való alkalmazkodástól, mely lehet tudatos távolodási szándék eredménye is (Bata 2009). A divergenciához hasonló a harmadik stratégia, az egyéni beszédmintázatok fenntartása, amely szerint a beszélő meg kívánja őrizni a beszéde stílusát, hangzásának sajátosságait függetlenül a kommunikációs helyzettől és a hallgatótól, címzettől (Gallois et al. 2005). A beszélő nagyobb eséllyel tartja fenn a saját

beszédmintázatait vagy tér el, távolodik el tudatosan a beszélőpartnere beszédétől, ha személyesen el szeretne szakadni tőle (Bata 2009, Gallois et al. 2005). A divergencia és a saját, egyéni beszédmintázatok fenntartása abban az esetben is bekövetkezhet, amikor az interakciót csoportközi viszonyként éli meg a beszélő, és pozitív csoportidentitást szeretne megtartani (Gallois et al. 2005).

A beszédalkalmazkodás elemezhető az alkalmazkodás mértéke szerint, ami a konvergenciára, divergenciára és az egyéni beszédmintázatok fenntartására is vonatkoztatható. A konvergencia során a közelítés mértéke függ attól, hogy milyen repertoárral rendelkezik a beszélő. Az is tényező lehet, hogy milyen személyiségbeli és környezeti tényezők növelik a beszélő igényét az elfogadásra a címzett részéről. A divergencia esetében szintén függ a beszélő repertoárjától a távolodás mértéke, továbbá olyan kontextuális tényezőktől is, amelyek képesek fokozni a beszélőben a csoportidentitás jelentőségét és a pozitív csoportidentitás iránti vágyat. A másik személy nem szimpatikus tulajdonságai is közrejátszhatnak a divergencia mértékének alakulásában az egyének közötti interakció alatt (Gallois et al. 2005).

Végül értelmezhető a beszédalkalmazkodás a fogadtatás felől is, tehát afelől, hogy a hallgató hogyan viszonyul a beszédprodukciónak, a beszélő által megvalósított beszédalkalmazkodáshoz. A hallgató értékelheti pozitívan a konvergenciát, ha az azt eredményező viselkedést pozitív szándéknak fogja fel, illetve akkor, ha a közeledés optimális szociolingvisztikai távolságot eredményez a felek között. Ez utóbbi arra utal, hogy a beszélgető partnerek között nem feltétlenül az egyenrangúság elérése a cél, és bizonyos helyzetekben (pl. hierarchizált viszonyokban) a csoportközi és interperszonális különbségek tükröztetése jelenti a beszédalkalmazkodás által elérendő optimumot. Az egyéni beszédmintázatok fenntartását és a divergenciát a hallgatók negatívnak tarthatják, ha negatív szándékot olvasnak ki belőle (Gallois et al. 2005).

A beszédalkalmazkodás érintheti a beszéd szegmentális és szuprasegmentális szintjét is (Bata 2009). Mivel a jelen dolgozatban a szuprasegmentális szint néhány jelenségét elemezzük a beszédalkalmazkodás kontextusában, ezek közül háromról részletesebben is szólnunk: a beszédnek a beszélő által használt átlagos alapfrekvenciájáról, az alapfrekvencia terjedelméről és a beszédtempóról.

A szuprasegmentális szerkezet, más néven prozódia a szegmentális szerkezeten (beszédhangokon, hangkapcsolatokon és hangsorokon) túli beszédjellegzetességeket jelenti, amelyek pusztán a szegmentumok összességéből nem levezethetők, nem következnek, de a szegmentális szerkezettel egyidőben jönnek létre a beszédalkeltés során, és ezeket is a szegmentumokat alkotó hangtest egyes részei, sajátosságai alakítják. A szuprasegmentumok

közé tartozik a beszéddallam, a hangsúly, a tempó, a szünettartás, a ritmus, a hangerő és a hangszínezet (Gósy 2004, Markó 2017).

A szupraszegmentális jellemzők jelentős többsége többféle akusztikai sajátosság, illetve ezek változásának következményeként áll elő (Markó 2017). Fizikailag minden rezgés a frekvencia, az intenzitás és az időtartam mennyiségeivel jellemezhető. A frekvencia jelzi, hogy hányszor ismétlődik az egyensúlyi helyzetből való kitérés egy adott időegység alatt a légnyomásban (Kassai 1998). A gégeszintű hangforrás és az artikulációs csatorna együttes tulajdonságaiból keletkezik a zöngés elemek frekvenciaszerkezete. A zöngé ciklusaira jellemző, hogy ismétlődésük nem tökéletesen, csupán körülbelül azonos időközönként következik be (Markó 1998). A zöngé összetett rezgés, tartalmazza az alaphangot, amely a legkisebb frekvenciájú és legnagyobb amplitúdójú összetevő, tehát az alapfrekvencia (f_0), és a felharmonikusokat, ennek egész számú többszöröseit (Markó 1998). A hangszalagok rezgése határozza meg a zöngé alapfrekvenciáját (Gósy 2004). Az alapfrekvencia, illetve ennek változása sok prozódiai jellemző megvalósításáért felel(het). Ilyennek tekinthető például a hangsúly, amelynek a kifejezése során az intenzitás mellett az alapfrekvencia is megváltozhat, de ilyen a beszéddallam is, amelynek bizonyos sajátosságait, a beszélő által használt átlagos alapfrekvenciát, illetve a hangterjedelmet a dolgozatban is elemezzük (Markó 2017).

A beszélő által használt átlagos alapfrekvencia, más szavakkal az átlagos hangmagasság, amely azzal függ össze, hogy a beszélő hangját milyen magasnak/mélynek érzékeljük. A hangterjedelem egy adott beszélő alaphangmagasság-változásainak a keretét jelenti, amit a legalacsonyabb és legmagasabb alapfrekvencia értékek jelölnek ki (Markó 2017). Beszédhangfekvésnek, másként hangterjedelemnek nevezzük azt a frekvenciatartományt, melyben megvalósítható a zöngéképzés hétköznapi körülmények között. Ez a beszédparaméter a beszéd dallamosságáról árulkodik, ugyanis minél nagyobb ez az érték, annál változatosabb a beszéd dallama. Ez és a beszéd közben használt átlagos alapfrekvencia egyénenként eltérő lehet, de hathatnak az alapfrekvenciára a közlési helyzet jellemzői, illetve a kulturális normák is. Részben ezek magyarázzák azt, hogy az f_0 -átlagról és az f_0 szórásának tartományáról eltérő adatokat találnak a kutatók a különböző vizsgálatokban (Gósy 2004).

A beszéd sebessége az artikulációs működések lassúságán vagy gyorsaságán alapul. A beszédhangok gyorsabb egymás utáni kiejtése gyorsabb beszédet eredményez, míg annak az eredménye, ha a beszédhangokat lassabban ejtjük, lassabb közlés lesz. A beszédtempót számos tényező meghatározza, például az életkor, foglalkozás, beszédhelyzet vagy az aktuális beszédtema. Ezen túl a beszédtempót befolyásolhatják még a beszélő egyéni adottságai is. (Gósy 2004).

A tempó mérésekor megkülönböztetendő az artikulációs tempó és a beszédtempó. Az artikulációs tempó a beszélő által adott időegység alatt kiejtett beszédegységeket, szavakat, szótagokat, beszédhangokat számszerűsíti úgy, hogy ezeknek a számát a beszédprodukcióban tartott szünetek (Markó 2017) és megakadásjelenségek (Gósy 2004) nélküli beszédidőre vetítjük. Ezzel szemben a beszédtempó mérésekor beleszámítjuk a beszédidőbe a szünetek időtartamát (Markó 2017) és a megakadásjelenségeket (Gósy 2004) is.

Mindkét tempóérték, a beszédtempó és az artikulációs tempó többféle mértékegységgel jellemezhető, mely függ attól, hogy milyen nyelvi egységeket mérünk, és ezeket mekkora időegységre levetítve szeretnénk megadni. Így a beszéd sebessége megadható például beszédhang/másodpercben vagy beszédhang/percben, illetve szótag/másodpercben, szótag/percben és akár szó/percben is (Gósy 2004). A tempó észlelését az időegység alatt elhangzott beszédegységek számán túl számos más paraméter is befolyásolja. A tempó nem tekinthető homogénnek egy-egy hangzásegységen belül, mert a beszédben lassulások és gyorsulások is megfigyelhetők. A tempó számítását befolyásolják a mérés során figyelembe vett alapegységek és időintervallumok. Ugyanakkor a mérési módszertantól függetlenül elkülönítendő az objektív, azaz mért tempó és a szubjektív, észlelt tempó. Kutatások bizonyítják, hogy az észlelt tempókülönbségekkel összefüggést mutatnak az artikuláció tempó és a beszédtempó objektív értékei (Markó 2017).

A jelen vizsgálatban a beszédnek a beszélő által használt átlagos alaphékvenciájában, az alaphékvencia terjedelmében és a beszédtempóban vizsgáljuk a beszédalkalmazkodást. A következőkben a beszédalkalmazkodás vizsgálatának szakirodalmából mutatunk be olyan eredményeket, melyek ezekre a beszédjellemzőkre vonatkoznak.

2.2 A kutatás közvetlen előzményei

Az alaphékvenciabeli alkalmazkodásra korábban több kutatás is irányult (Babel 2009, Garnier et al. 2013, Szabó 2019). Garnier és munkatársai (2013) többek között azt vizsgálták, hogy a kísérleti személyek alaphékvencia-értékeiben milyen különbségek fedezhetők fel akkor, ha a résztvevőket a beszélgetést megelőzően felszólítják arra, hogy alkalmazkodjanak a beszélgetőtársukhoz, illetve abban az esetben, ha nem instruálják erre őket. A kutatásban egyirányú kommunikációs helyzetet elemeztek, ahol a kísérleti személyek magánhangzókat hallgattak, és azokat kellett visszaismételniük. Eredményeik azt mutatták, hogy a kísérleti személyek akkor is imitálták a megnyilatkozó beszédjellemzőit (jelen esetben az alaphékvenciát), ha erre nem vonatkozott utasítás.

Ahogy az korábban is említettük, arra eddig kevés figyelmet fordítottak a kutatók, hogy miként hat az alkalmazkodásra az, ha a beszélő nem csak hallja, hanem látja is a beszélgetőtársát. Babel (2009) a vizsgálatában két csoportra osztotta a kísérleti személyeket. Mindkét csoport meghallgatott egy-egy előzetesen rögzített hangfelvételt két különböző férfi beszélőtől. Az egyik csoport csak a beszélő hangját hallotta, a másik csoport a hang mellett egy állóképet is látott az adott beszélőről. Az eredmények azt mutatták, hogy a vizuális inger, azaz az állókép hatással volt a kísérleti személyeknek a beszélőhöz való viszonyulására, valamint a beszélőhöz való alaphérvenciabeli alkalmazkodására: a férfiak nagyobb mértékben adaptálódtak a beszélőhöz akkor, ha arról képet is láttak, míg a nők akkor, ha a beszélőnek csak a hangját hallották.

Fontos kiemelni, hogy az idézett kísérletben (Babel 2009) a vizuális inger (azaz a beszélőt ábrázoló állókép) nem hordozott magában olyan, a beszédre és a beszélőre vonatkozó fonetikai információkat, mint például az artikulációs sajátosságok, így nem tudjuk meg azt, hogy ezek a mindennapi beszédhelyzetek nagy részében elérhető fonetikai természetű információk fokozzák-e a beszédalkalmazkodást. A jelen vizsgálatban erre a kérdésre fókuszáltunk, és állókép helyett előzetesen rögzített videófelvevételeket mutattunk a beszélőnek a másik megnyilatkozóról. Ezáltal nem csak a kísérleti személyekben keltett szubjektív benyomás alakulhatott ki, amelyet az állókép is képes volt biztosítani, hanem hozzáférhetővé váltak a beszéd és a beszélő további fonetikai, artikulációs sajátosságai is. Az idézett kísérlet eredményeit inkább úgy értelmezhetjük, hogy a beszédalkalmazkodást fokozta a kísérleti személyekben keltett szubjektív (társas) benyomás a beszélőről.

A hangterjedelmet érintő alkalmazkodással szintén több korábbi kutatás foglalkozott (pl. Bosshardt et al. 1997, Aubanel–Nguyen 2020). Aubanel és Nguyen (2020) egy olyan feladatban, ahol a beszélők két elszigetelt csendesített szobában ülve oldottak meg egy közös beszédfeladatot azt találták, hogy nem csak az alaphérvencia átlagában, de a maximum és a minimum értékeiben, azaz a hangterjedelemben is bekövetkezett alkalmazkodás. Eredményeik szerint a hangterjedelmet érintő konvergencia gyakrabban fordult elő akkor, ha a beszélgető felek előzetesen mért értékei eredetileg is közel álltak egymáshoz.

A beszédtempót illetően is számos kutatást végeztek korábban a beszédalkalmazkodás vizsgálatában (pl. Bosshardt et al. 1997 dadogókkal, ill. Casasanto et al. 2010, Tiede et al. 2010, Szabó 2019 egészséges beszélőkkel). Casasanto és munkatársai (2010) kísérletük során ugyancsak biztosítottak vizuális ingert a kísérletben részt vevők számára a beszélőről az audiofelvétel mellett: egy videót láttak róla, amelyet egy virtuális valóságot közvetítő eszközzel (VR szemüveggel) közvetítettek. Kutatási kérdésük elsősorban az volt, hogy a

kísérleti személyek alkalmazkodnak-e a beszédtempót illetően akkor, ha a megnyilatkozó nem egy valós személy (hanem egy fiktív karakter egy virtuális világban), és a résztvevők csak virtuálisan állnak vele kapcsolatban. Kutatásuk eredményei azt mutatták, hogy a kísérleti személyek alkalmazkodtak a beszélő beszédtempójához.

Szabó (2019) kétirányú párbeszédekben vizsgálta a beszédtempó alakulását a beszédalkalmazkodásban. Eredményei azt mutatták, hogy az alkalmazkodás a beszédtempóban jóval később jelentkezik, mint az alaphékvenciában, megközelítőleg 10 perc (600 s) elteltével.

Összefoglalóan: a fentiek nyomán azt állapíthattuk meg, hogy a beszélőalkalmazkodás nem tudatos imitáció, és hatása van rá annak, ha a beszélő látja a beszélgetőtársát egy állóképen, illetve a virtuális világban úgy, hogy a beszélgetőtárs nem valós személy. Arra a kérdésre ugyanakkor egyelőre nem kaptunk választ, hogy ehhez hasonló mértékben (vagy akár ennél jobban) befolyásolja-e az alkalmazkodást az, ha a beszélő videófelvételt lát a beszélgetőtársáról, aki valódi, hús-vér ember, tehát az, ha a beszélő hozzáfér olyan fonetikai jellegű információkhoz a beszélgetőtársról mint például annak artikulációs sajátosságai. Kérdés továbbá, hogy mely paraméterekben és milyen mértékben érhető tetten az a beszédalkalmazkodás, amit a beszélgetőtárs hallgatása, illetve együttes hallgatása és nézése eredményez.

Megállapíthattuk, hogy az alaphékvenciát érintő alkalmazkodás már az interakció kezdetekor, szinte azonnal bekövetkezik, míg a beszédtempóbeli ehhez képest később. Ezeket a szempontokat is figyelembe vettük az alább bemutatandó kísérletünk tervezésekor.

A jelen vizsgálatban építünk ezekre az eredményekre abban, hogy szintén virtuális és egyirányú kommunikációt mutató beszédhelyzetekben elemezzük az alkalmazkodást, ahol a vizsgált beszélő „beszélgetőtársa” egy videó- vagy hangfelvételen szerepel. Ennek a módszertani megoldásnak az volt az indoka, hogy az eredményeket befolyásolta volna, hogy mind a kísérleti személy, mind a “beszélgetőtárs” alkalmazkodott volna egymáshoz. Ezt a jelenséget nem kívántuk vizsgálni, így kizártuk ennek a lehetőségét.

3 A kutatás célja, kérdése és hipotézise

A fentiekben megfogalmazott kérdések nyomán a jelen kutatásban arra a kérdésre keressük a választ, hogy van-e különbség a beszédalkalmazkodásban akkor, ha a beszélő nem csak auditív ingerekre hagyatkozik a beszédpercepció során, hanem vizuálisakra is. Más szavakkal a kérdésünk az, hogy befolyásolhatja-e a beszédalkalmazkodás mértékét az a tény, hogy a beszélők látják is azt a személyt, akinek a beszédét hallgatják.

A kutatásunk hipotézise szerint az audiovizuális ingert kapó kísérleti személyek nagyobb mértékben alkalmazkodnak a másik félhez a kommunikációs helyzetben, mint a csak hangingert kapó kísérleti személyek, és ez az alkalmazkodás az alaphékvenciában és beszédtempóban is megjelenik.

4 Módszertan

A kutatás megvalósítására az ELTE Bölcsészettudományi Kar Kari Kutatástámogató Bizottsága által kiadott etikai engedély birtokában került sor.

4.1 A kísérletben résztvevő személyek

Kutatásunkhoz férfi kísérleti személyeket és egy férfi mintabeszélőt (azaz *model talkert*; Babel–Bulatov 2011) választottunk azért, hogy minél jobban facilitáljuk a beszédalkalmazkodás folyamatát. A korábbiakban említett Babel (2009) által végzett kísérlet eredményei alapján elmondható, hogy a férfiak nagyobb mértékben alkalmazkodtak a mintabeszélőhöz akkor, amikor nem csak a hangját hallották a beszédfeladatok közben, de az arcát is látták álló fényképfelvételen. Továbbá Szabó (2019) azonos nemű személyek között folytatott beszélgetésekben azt találta, hogy a férfiak nagyobb mértékben alkalmazkodnak az alaphékvenciát illetően más férfiakhoz, mint a nők más nőkhöz.

A kísérletben húsz, a mintabeszélőt nem ismerő személy vett részt, akik magyar anyanyelvű, heteroszexuális felnőtt férfiak (19-35). Saját bevallásuk szerint ép hallásúak és nem rendelkeznek beszéd- és/vagy nyelvi zavarral, beszédhanghibával.

A heteroszexualitás kérdésének azért van jelentősége, mert korábbi eredmények azt mutatták (Namy et al. 2002), hogy a beszédalkalmazkodás egyik kiváltó oka lehet a szimpátia elnyerése, a másinak való megfelelés, vagy épp a bizalom keltése a beszélgetőtársban. Ezek alapján felmerül annak a lehetősége is, hogy a szexuális vonzalom ugyancsak hatással van az akkomodációra úgy, hogy fokozza azt. Ezt az aspektust jelen kutatásunkban nem kívántuk vizsgálni, és ennek kiküszöbölése érdekében választottunk a mintabeszélővel azonos nemű, heteroszexuális beszélőket.

A kísérlet megkezdése előtt a kísérleti személyeket tájékoztattuk a kísérlet menetéről, a kezelt adatok köréről, és minden kísérleti személy beleegyező nyilatkozatot írt alá.

4.2 Anyag

A résztvevők a kísérletben a *szál*, *szed*, *szín*, *szúr* monoszillabikus, CVC-szerkezetű magyar szavakat ejtették ki önálló mondatokként. A szavak kiválasztásának szempontjai a következők voltak. A szavakban szereplő magánhangzók a magyar magánhangzótér szélső hangjai, a mássalhangzók pedig szókezdő helyzetben egységesen *sz*, szózáró helyzetben egységesen zöngések és alveoláris képzési helyűek voltak. Ezeket a szempontokat azért érvényesítettük az anyag megtervezésénél, mert a kutatás folytatásában a magánhangzók minőségét is elemezni kívánjuk, amelyhez a hangsorok ilyen fokú illesztése szükséges, miközben a jelen vizsgálatban ezek a szavak önálló megnyilatkozásokként elemezhetők.

4.3 Módszer

A kísérlet során a résztvevők vagy egy hangfelvételen vagy egy videófelvételen rögzített hanganyagot hallgattak meg, amelyen a mintabeszélő szerepelt. Ezt előzetesen rögzítettük. Ezzel a kísérleti résztvevők számára egy egyirányú kommunikációs helyzetet hoztunk létre, ahol kiküszöböltük annak a lehetőségét, hogy a mintabeszélő a kísérleti személyekhez alkalmazkodjon, így biztosítva azt, hogy minden kísérleti személy ugyanazt az ingert hallja és/vagy lássa.

A mintabeszélőről tehát videó- és hangfelvétel is készült. A mintabeszélő a célszavakat ötször ismételte el véletlenszerű sorrendben, ez szolgált a beszélők utánmondási feladatának alapjául, illetve felolvasott egy 8 perc hosszúságú mesét, amelyben egy-egy alkalommal szerepelt az összes célszó. Ez utóbbit a beszélőknek végig kellett hallgatnia a kísérlet második részében. A meséről készült felvétel azért volt éppen 8 perc hosszúságú, mert a korábbi korpuszalapú tanulmányok azt mutatták ki, hogy megközelítőleg ennyi idejű kitettségre van szükség ahhoz, hogy a beszélő a beszélgetőtársához a beszédtempójában alkalmazkodjon (Szabó 2019), és a jelen vizsgálatban ezt a beszédjellemzőt is elemeztük. Mivel azonban a jelen vizsgálatban egyoldalú kommunikációs helyzetet teremtettünk, ami hosszútávon fárasztó, és magas fokú kognitív terheléssel jár, ezért a kapott inger időtartamát lecsökkentettük a korábban említett 10 percről (Szabó 2019) 8 percre. A felvételek külső hangkártyával és fejmikrofonnal készültek, illetve egy laptop kamerájával rögzítettük a videófelvételt a mintabeszélőről. Ezt az anyagot hallgattattuk vagy néztünk meg az adatfelvétel során a kísérleti személyekkel.

A kísérleti személyeket két csoportra osztottuk. Az első, „auditív” csoportban lévő résztvevők (10 fő) csak auditív ingert kaptak: hallották a mintabeszélőt, de a felvett videót nem

látták. A második, „audiovizuális” csoportban (10 fő) a résztvevők auditív és vizuális ingert is kaptak: hallották és látták is a mintabeszélőről készült felvételeket.

A kísérleti személyekkel egyesével rögzítettük a vizsgálandó hangfelvételeket ugyancsak külső hangkártya és fejmikrofon segítségével. A kísérlet során a kísérleti személyek mellett tartózkodott egy kísérletvezető is, aki a hangfelvételt készítette. A felvételek egy hangszigetelt szobában készültek. A kísérletet vezető személy a hangfelvétel készítése során nem kommunikált a kísérleti személlyel, minden instrukció és felolvasandó szöveg a monitoron jelent meg, ezzel kerültük el, hogy a kísérletvezető személye befolyásolja a beszédalkalmazkodást, így például azt is, hogy a kísérleti személy beszéde a mintabeszélő helyett a kísérletvezetőéhez alkalmazkodjon.

A résztvevők első feladata mindkét csoportban azonos volt: fel kellett olvasniuk a célszavakat. A szavakat mondszerűen olvasták fel, melynek oka az egységes prozódia megteremtése volt, így a kísérleti személyek minden célszó felolvasását kijelentő mondatok formájában valósították meg. Minden célszó esetében 5-5, így összesen beszélőnként húsz felvételt készítettünk a felolvasásról. Ez azért volt fontos, hogy a beszélők (általunk vizsgálni kívánt) beszédjellemzőiről (az alapfrekvenciáról és a beszédtempóról) reprezentatívabb képet kaphassunk. A célszavakat a beszélők kvázi-random sorrendben olvasták fel, ahol két azonos szó nem követhette egymást. Így a felvétel elején a kezdeti megilletődés, a helyzet természetellenessége, valamint a kísérlet végén a túlzott monotonitás, unalom, fáradtság stb. okozta hatásokat kiegyenlítettük az anyagban az egyes beszélők, és így a célszavak között is. Ezután a második feladatban a résztvevők meghallgatták a mintabeszélő által felolvasott mesét: ebben a feladatban, mint fentebb írtuk, az auditív csoport csak a beszélő hangját hallotta, míg az audiovizuális csoport látta is a beszélő arcát. Azért, hogy a szöveg felolvasása alatt biztosítsuk azt, hogy a résztvevők figyeljenek a mintabeszélő hangjára, azt a feladatot adtuk nekik a felvételkészítés kezdetekor, hogy figyelmesen hallgassák végig a mesét, mert ahhoz kapcsolódóan a kísérlet végén kérdésekre kell majd válaszolniuk. A mese meghallgatásának a célja az volt, hogy biztosítsa a kísérleti személyek számára az alkalmazkodáshoz elegendő auditív vagy audiovizuális ingert, azaz a huzamosabb kitettséget. Erre azért volt szükség, hogy elősegítsük a beszédtempóbeli alkalmazkodást tekintve, hogy a beszédtempó szerint a beszédalkalmazkodás később jelentkezik, mint az f_0 szerint (Szabó 2019).

A harmadik feladat az utánmondás volt: itt a korábban már elhangzott, az első fázisban felolvasott mondatokat ismételték vissza a résztvevők (szintén 5 alkalommal, az előzőtől eltérő sorrendben) úgy, hogy minden egyes mondat előtt először meghallgatták az adott mondatot a

mintabeszélő ejtésében. Itt az egyik csoportnak ismét csak a mintabeszélő hangját, a másiknak pedig a róla készült videófelvételt is levetítettük.

A hanganyagokat a Praat (Boersma–Weenink 2023) programban címkéztük és szegmentáltuk manuálisan. Automatikusan mértük az elhangzott mondatok időtartamát, ezzel a beszédtempót számszerűsítettük: minél nagyobb időtartamú egy szó, annál gyorsabb a beszéd sebessége. Ugyancsak automatikusan mértük az egyes mondatokban a beszélők átlagos alaphangfrekvenciáját, melyet az egyes szavakban mért f_0 -érték mediánjaként nyertünk ki. Emellett a használt alaphangfrekvencia hangterjedelmét is elemeztük a beszédhang kihasználásának mérésére. A hangterjedelem vizsgálatához kinyertük az egyes mondatokban a legmagasabb és legalacsonyabb f_0 -értékeket, és kiszámítottuk ezek különbségét.

A beszédtempóbeli, illetve az alaphangfrekvenciabeli alkalmazkodást háromféle összevetésben elemeztük. Először összehasonlítottuk az auditív és az audiovizuális csoportban a felolvasásban és utánmondásban kapott átlagértékeket a mintabeszélő átlagértékeivel. Ezek után megnéztük, hogy a két csoportban mennyire térnek el a felolvasásban mért értékek az utánmondásban mért értékektől kísérleti személyenként. Végül pedig azt is elemeztük, hogy az egyes beszélők felolvasásban és utánmondásban mért értékei mennyire térnek el a mintabeszélő beszédében mértéktől. A hangterjedelemben tapasztalható alkalmazkodás esetében csak a csoportátlagokat vetettük össze a mintabeszélő beszédében mértékekkel.

5 Eredmények

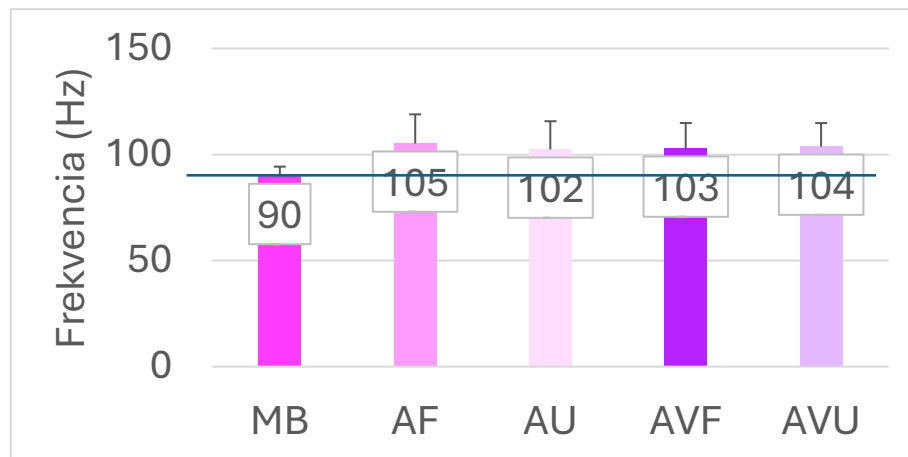
A továbbiakban részletesen ismertetjük a kísérleti személyek átlagos alaphangfrekvencia-, hangterjedelem-, valamint beszédtempó-értékeit.

5.1 Alaphangfrekvencia

A mintabeszélő átlagos alaphangfrekvenciája (f_0) 90 Hz volt, ehhez viszonyítottuk a kísérleti személyek két csoportjában mért értékeket. A kísérleti személyek beszédének alaphangfrekvenciájában a következőket figyelhettük meg (1. ábra). Az auditív és audiovizuális csoport felolvasásában (auditív csoport: 105 Hz, audiovizuális csoport: 103 Hz) és utánmondásában (auditív csoport: 102 Hz, audiovizuális csoport: 104 Hz) mért alaphangfrekvencia-értékek átlagai egyöntetűen magasabbak voltak, mint a mintabeszélő átlagos alaphangfrekvenciája. Ez azt jelenti, hogy a mese meghallgatását követően az utánmondásban nem történt számottevő változás a kísérleti személyek beszédében. Azonban megfigyelhettük azt is, hogy jelentős volt a szórás az egyes csoportokban (a mintabeszélő esetében 4 Hz, az auditív csoportban a

felolvasás esetében 14 Hz, az utánmondás esetében 13 Hz, az audiovizuális csoportban a felolvasás esetében 12 Hz, az utánmondás esetében 11 Hz). (A szórást a diagramon a hibaszáv jelöli fekete függőleges vonallal.)

Az eredmények csoportszinten, az alaphfrekvencia-átlagot vizsgálva azt mutatják, hogy a kísérleti személyek f_0 -ja megközelítőleg megegyezett a felolvasás és az utánmondás esetében.

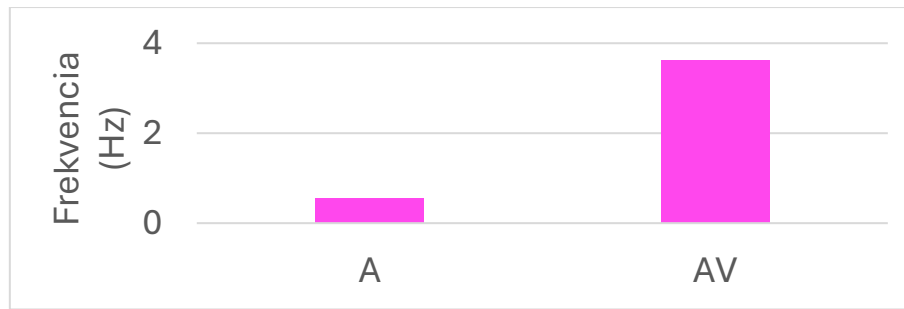


1. ábra. Átlagos alaphfrekvencia (MB = mintabeszélő, AF = auditív csoport, felolvasás, AU = auditív csoport, utánmondás, AVF = audiovizuális csoport, felolvasás, AVU = audiovizuális csoport, utánmondás; a vízszintes vonal a mintabeszélő átlagát mutatja)

Az adatelemzés következő lépéseként megvizsgáltuk a felolvasás és az utánmondás során mért alaphfrekvencia értékek különbségét kísérleti személyenként. Ebben a lépésben azt elemeztük, hogy az alaphfrekvencia-értékek távolodnak-e (divergálnak) a mintabeszélő átlagos alaphfrekvenciájától vagy közelítenek (konvergálnak) hozzá.

Az elemzés ezen szakaszában a beszédalkalmazkodást az alaphfrekvencia tekintetében a következőképpen számoltuk ki: átlagoltuk az alaphfrekvencia-értékeket kísérleti személyenként, majd kivontuk egymásból az utánmondásban és felolvasásban kapott átlagértékeket. Akkor, ha a kísérleti személy konvergált a mintabeszélőhöz, ehhez az értékhez pozitív előjelet tettünk, ha divergált tőle, akkor negatív.

Ezeket az egyéni tendenciákat leképező adatokat átlagolva azt állapíthattuk meg, hogy az auditív csoport tagjainak f_0 -ja átlagosan 4 Hz-cel közeledett a mintabeszélő alaphfrekvenciájához, míg az audiovizuális csoport tagjaié átlagosan 1 Hz-cel (2. ábra).

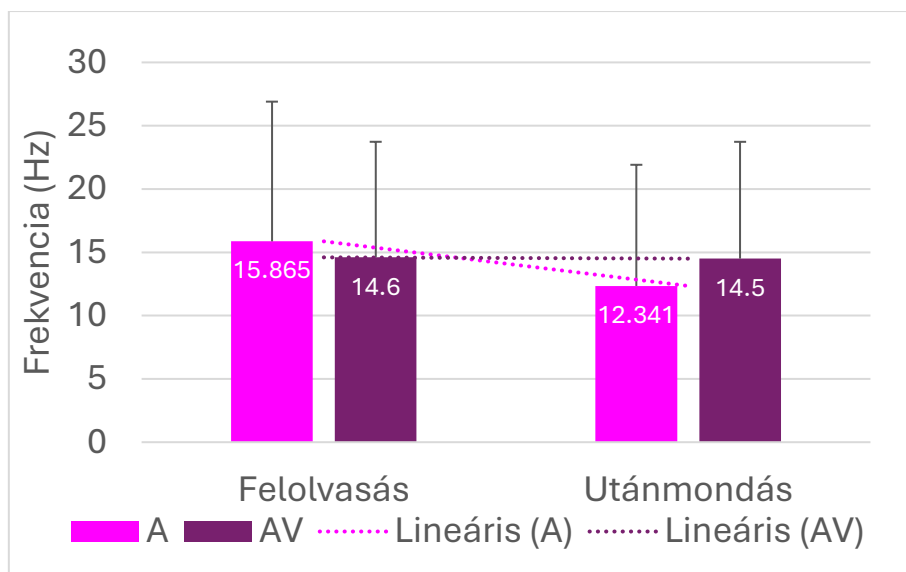


2. ábra. Az átlagos alapfrekvencia-értékek változása a felolvasás és az utánmondás között csoportonként (A = auditív csoport, AV = audiovizuális csoport)

Mivel fennállhat annak a lehetősége, hogy egy erősen konvergáló (alkalmazkodó) és egy erősen divergáló (távolító) résztvevő értékei a csoportátlagban kiegyenlítik egymást, ezért szükségesnek tartottuk, hogy az alkalmazkodás mértékét egyénenként is megvizsgáljuk mindkét csoport esetében.

Megvizsgáltuk, hogy hogyan változott a kísérleti személyek alapfrekvenciájának különbsége a felolvasás és az utánmondás feladatok között. Megfigyelhetjük, hogy mindössze két beszélő esetén nem volt nagyobb az eltérés 1 Hz-nél (m04, m17).

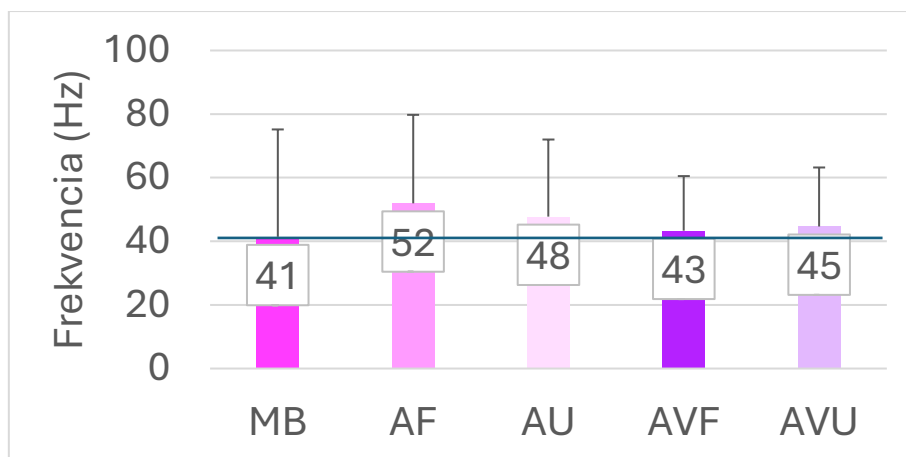
Ebben az elemzésben a kapott különbségek abszolút értékét átlagoltuk (3. ábra). Eszerint az elemzés szerint a kísérleti személyek és a mintabeszélő beszédalapfrekvencia-értékeinek különbsége átlagosan 16 Hz volt az auditív, 15 Hz az audiovizuális csoport esetében a felolvasásban. A mese meghallgatása után, az utánmondás alatt a különbség rendre 12 Hz és 15 Hz volt. Így tehát az auditív csoport esetében a különbség csökkent 4 Hz-cel, vagyis a résztvevők az f_0 értékében közeledtek a mintabeszélő átlagos alapfrekvenciájához, míg az audiovizuális csoport esetében a beszélők és a mintabeszélő közötti átlagos különbség nem változott.



3. ábra. Az alapfrekvencia-értékek eltérése a mintabeszélő megfelelő értékétől az auditív és az audiovizuális csoportban (A = auditív csoport, AV = audiovizuális csoport)

5.2 Hangterjedelem

A hangterjedelem számszerűsítéséhez a beszélőknek a megnyilatkozások alatt mért alapfrekvencia-értékeinek a maximumából kivontuk ugyanezen értékek minimumát. A hangterjedelmet Hz-ben mértük, ami minden esetben pozitív szám (illetve nulla akkor, ha a beszéddallamban nem történik változás). A hangterjedelmet csoportonként átlagoltuk. Ezek az átlagértékek láthatók az 4. ábrán a szórást jelölő hibasávokkal. A mintabeszélő esetében a hangterjedelem átlaga 41 Hz volt, a szórása pedig (a csoportok összevetésében kiemelkedően magas) 34 Hz. Az auditív csoport esetében alkalmazkodást figyelhettünk meg, ugyanis 52 Hz-ről 48 Hz-re (4 Hz-cel) csökkent a hangterjedelem, illetve a hangterjedelem szórása is kisebb lett (ez 28 Hz-ről 24 Hz-re csökkent). Az audiovizuális csoport értékei ezzel szemben távolodtak a mintabeszélő esetében mért értékektől, míg a felolvasásban a mintabeszélőével közel azonos értéket mértünk (43 Hz, szórás: 17 Hz), addig a mese meghallgatása után, az utánmondás folyamatában ez az érték nőtt, azaz távolodott a mintabeszélő hangterjedelmétől (45 Hz), miközben az adatok szórása is nagyobb lett (19 Hz).

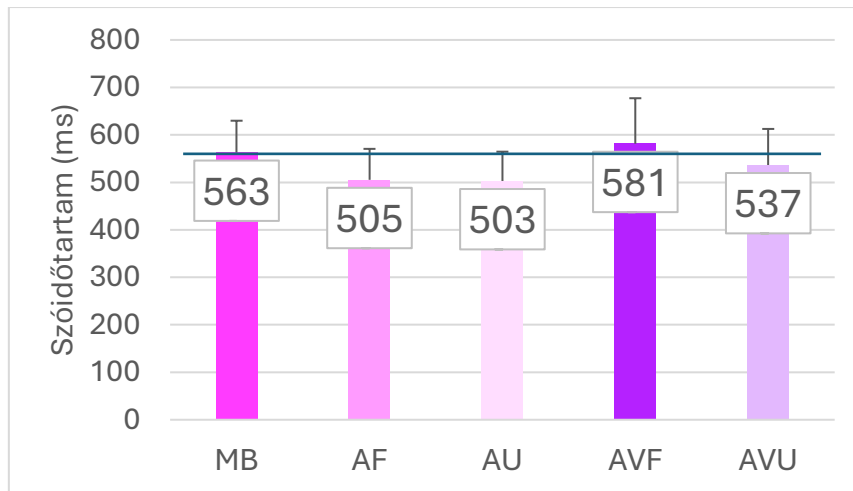


4. ábra. Átlagos hangterjedelem (MB = mintabeszélő, AF = auditív csoport, felolvasás, AU = auditív csoport, utánmondás, AVF = audiovizuális csoport, felolvasás, AVU = audiovizuális csoport, utánmondás; a vízszintes vonal a mintabeszélő átlagát mutatja)

5.3 Beszédtempó

A beszédtempó elemzésének esetében hasonlóan jártunk el, mint az alaphékvencia esetében. A felolvasás és az utánmondás során kapott adatokat csoportonként összesítve, valamint egyénenként is megvizsgáltuk, hogy átfogóbb képet kaphassunk az auditív és az audiovizuális csoportban találtokról.

Az alaphékvenciánál látottakkal egyezően elsőként itt is csoportokra bontva vizsgáltuk meg, hogy hogyan változik a mintabeszélő és a kísérleti személyek átlagos beszédtempója a felolvasás és az utánmondás esetében, majd kiszámítottuk a kísérleti személyek értékeinek saját magukhoz viszonyított különbségét, végül pedig a beszélők és a mintabeszélő értékeinek különbségét is.



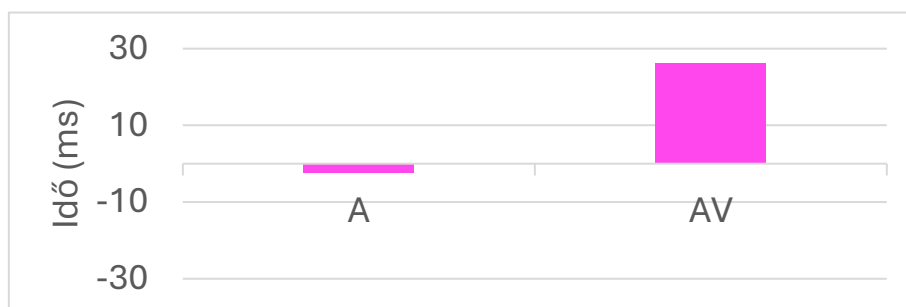
5. ábra. Átlagos szóidőtartam (MB = mintabeszélő, AF = auditív csoport, felolvasás, AU = auditív csoport, utánmondás, AVF = audiovizuális csoport, felolvasás, AVU = audiovizuális csoport, utánmondás; a vízszintes vonal a mintabeszélő átlagát mutatja)

Az 5. ábrán látható, hogy az auditív csoport átlagos szóidőtartama 58 ms-mal alacsonyabb, így beszédtempója gyorsabb volt a felolvasásban, mint a mintabeszélőé. A mese meghallgatását követően ez a csoportszintű átlag közel ugyanannyi maradt, mindössze 2 ms különbség volt megfigyelhető a felolvasás és az utánmondás között az auditív csoportban. A csoport átlagos beszédtempója tehát nem mozdult el a mintabeszélő beszédtempójának irányába.

Az audiovizuális csoport felolvasásban mért szóidőtartamai magasabbak voltak, mint a mintabeszélő átlagos szóidőtartamai, ám ebben az esetben a csoport és a mintabeszélő közti különbség kisebb volt, mindössze 18 ms (5. ábra). Megfigyelhetjük azonban azt is, hogy az utánmondás esetében az audiovizuális csoport átlagos szóidőtartamai nem csak a saját, a felolvasáskor mért átlagukhoz képest, de a mintabeszélőhöz képest is alacsonyabbak lettek. A tempó növekedése hozzájárult ahhoz, hogy a mintabeszélő és az audiovizuális csoport átlagos szóhosszúsága (így beszédtempója) közti különbség megnövekedjen: a felolvasás esetében ez a differencia a korábban említett 18 ms-ról 26 ms-ra nőtt.

Ahogy az f_0 esetében is, a kísérleti személyek egyes szavakban mért beszédtempóját második lépésben (abszolút szóidőtartamait) átlagoltuk a két feladatban, és ezekből számoltuk ki az alkalmazkodást. Ahogy az előzőekben is, konvergálás (a mintabeszélő értéke irányába történő elmozdulás) esetén ez az érték pozitív előjelet kapott, divergálás (a mintabeszélő értékétől történő távolodás) esetén negatív. Ezeket az egyéni tendenciákat leképező adatokat átlagolva azt állapíthattuk meg, hogy míg az auditív csoport átlaga -2 ms volt, addig az audiovizuális csoporté 26 ms (6. ábra). Ennek értelmében az auditív csoport -2 ms-ot változva

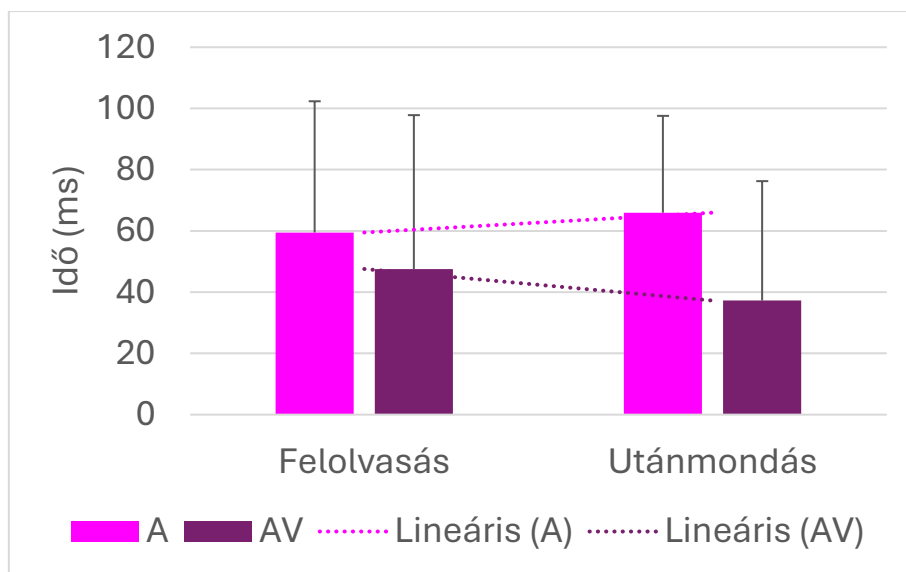
mozdult el a mintabeszélő értékeinek irányából, míg az audiovizuális csoport esetében 26 ms-nyi elmozdulással közeledést láttunk.



6. ábra. Az átlagos beszédtempó-értékek változása a felolvasás és az utánmondás között csoportonként (A = auditív csoport, AV = audiovizuális csoport)

Az egyének szintjén történő eltérések vizsgálatakor megfigyelhetjük, hogy mindössze három kísérleti személy esetében nem következett be mérhető változás (m05, m11, m16), míg a fennmaradó tizenhét esetben vagy konvergálás vagy divergálás volt megfigyelhető. Tehát ebben az esetben is elmondható, hogy a csoportátlag többségében reprezentálja a beszélők egyéni beszédtempó-értékeit, ezáltal valós képet adva mind az auditív, mind az audiovizuális csoport beszédjellemzőiről.

A 7. ábrán láthatjuk a kísérleti személyek átlagos szóidőtartam-értékeinek (azaz beszédtempójának) eltérését a mintabeszélő megfelelő értékétől az auditív és az audiovizuális csoportban. Megállapíthatjuk, hogy az auditív csoport eltérése a mintabeszélőhöz képest a mese meghallgatása előtt (a felolvasásban) 60 ms, a meghallgatása után (az utánmondásban) pedig 66 ms volt, míg ugyanezek az értékek az audiovizuális csoport esetében rendre 48 ms és 37 ms voltak. Ennek értelmében az auditív csoport és a mintabeszélő értékei közti különbség nőtt, míg az audiovizuális csoport és a mintabeszélő értékei között csökkent.



7. ábra. A beszédtempó-értékek eltérése a mintabeszélő megfelelő értékétől az auditív és az audiovizuális csoportban (A = auditív csoport, AV = audiovizuális csoport)

6 Következtetések

A dolgozatban bemutatott kutatás a beszédalkalmazkodás megjelenését vizsgálta fonetikai paraméterek mentén. A hipotézisünk az volt, hogy a beszédalkalmazkodás nagyobb mértékben megy végbe akkor, ha a hallgató nem csak auditív, hanem egyszerre auditív és vizuális ingert is kap a megnyilatkozó részéről. Ezt az alapfrekvenciában, a hangterjedelemben, valamint a beszédtempóban vizsgáltuk.

A hipotézisünk az alapfrekvenciát illetően nem nyert alátámasztást. Az eredmények azt mutatták, hogy az átlagos alapfrekvencia esetében mindössze néhány hertz eltérés volt a felolvasás és az utánmondás értékei között mind az auditív, mind az audiovizuális csoportban, valamint a csoport részvevőinél egyenként is. Az auditív csoportban ez a differencia valamivel nagyobb (3 Hz) volt, mint az audiovizuális csoportban (1 Hz). Ezek azonban csekély eltérésnek számítanak, hiszen ezen az alapfrekvencia-tartományon az emberi észlelés számára feldolgozható félhang ennél nagyobb (kb. 6-7 Hz) különbséget jelent (Weisstein 2019).

A hangterjedelem esetében hasonló eltéréseket figyelhettünk meg mindkét csoportban a két feladat, a felolvasás és az utánmondás között. Ahogyan az átlagos alapfrekvencia-értékekben, úgy a hangterjedelem értékeit illetően is az auditív csoportban mutatkozott nagyobb változás a beszélők által használt hangterjedelem értékeiben (4 Hz), mint az audiovizuális csoportban (2 Hz). És bár a szórások értékei magasnak számítottak, a hangterjedelmet illetően összességében alacsony volt a differencia a két beszédfeladat során mért értékek között.

A fenti eredmények tehát összefoglalóan azt mutatják, hogy a két csoport átlagos alaphfrekvenciája, valamint hangterjedelme nem változott a mintabeszélő beszédének hatására. Érdeemes azonban kiemelni, hogy divergenciát, azaz távolodást sem az auditív, sem az audiovizuális csoport ejtésében nem tapasztaltunk. Éppen ezért arra következtethetünk, hogy a mintabeszélőnek ezen beszédjellemezője nem volt hatással a kísérleti személyek beszédére. Ennek oka lehetett az is, hogy a mintabeszélő értékei hasonlóak voltak a felolvasásban mért csoportos átlagokhoz, illetve az is, hogy a mintabeszélő nem váltotta ki a kísérleti személyekből az alkalmazkodásra való szándékot.

A beszédtempó értékeiben hasonló tendenciát tapasztalhattunk az auditív csoportban, mint az alaphfrekvenciát és a hangterjedelmet illetően. A csoport átlagos beszédtempója jelentősen alacsonyabb volt a felolvasáskor, mint a mintabeszélőé, az utánmondásban mért érték megközelítőleg egyezett a felolvasásban mértékkel, tehát nem változott a kísérlet során. Az audiovizuális csoportban azonban megfigyelhettünk látványos különbségeket a két feladat között. A huzamosabb kitettség, tehát az, hogy a kísérleti személyek 8 percen keresztül látták és hallották a mintabeszélőt, hatással volt a kísérleti személyek beszédtempójára. A csoportnak a korábban, a felolvasásban mért átlagos beszédtempó-értéke magasabb volt a mintabeszélőénél, az utánmondásban pedig a mintabeszélő beszédében mért értéknél alacsonyabb lett. Felmerül azonban a kérdés, hogy ebben az esetben alkalmazkodásról, azaz konvergenciáról, vagy inkább távolodásról, tehát divergenciáról van-e szó. Az audiovizuális csoport felolvasás során mért beszédtempó-értéke (581 ms) közelebb állt a mintabeszélőéhez (563 ms), mint az utánmondás során mért értéke (537 ms). Ebből kifolyólag következtethetnénk arra is, hogy a csoport tagjai távolodtak a beszélőtől. Azonban a változás a mintabeszélő beszédtempójának irányába történt, tehát a kísérleti személyek csökkentették saját beszédtempójukat, csak épp az alkalmazkodás (konvergencia) esetében a vártnál nagyobb mértékben.

Az alaphfrekvenciát és a beszédtempót illetően a két csoport eredményei nem egyeztek meg. Az átlagos alaphfrekvenciában és a hangterjedelemben sem az auditív, sem az audiovizuális csoportban nem mutatkozott jelentős alkalmazkodást. Azonban a beszédtempót illetően az audiovizuális csoport az auditív csoporthoz képest nagyobb mértékben alkalmazkodott a mintabeszélőhöz.

Elmondható tehát, hogy a beszédalkalmazkodásra a vizuális csatorna jelentős befolyással bír elsősorban a beszédtempót illetően: az auditív csoport beszédtempó-, alaphfrekvencia- és a hangterjedelem-értékeiben nem következett be jelentősebb változás a mese meghallgatását

követően, míg az audiovizuális csoport tagjai számottevő mértékben változtattak az átlagos beszédtempójukon.

6.1 Továbbvezető kérdések

Zárásul arról is teszünk említést, hogy a kutatás milyen további lehetőségeket tartogat magában, illetve milyen egyéb paraméterek vizsgálatát tervezzük a jövőben.

A továbbiakban érdemes lehet a vizsgálatot nagyobb számú kísérleti személlyel megismételni. Ezzel reprezentatívabb képet kaphatnánk arról, hogy a vizuális csatorna milyen és mekkora szerepet tölt be a beszédalkalmazkodás folyamatában. A vizsgálatot ki lehetne egészíteni azzal, hogy a kísérleti személyek ne csak virtuálisan, egy egyirányú kommunikációs helyzetben hallják, vagy hallják és lássák a mintabeszélőt: az élő interakcióval elősegíthető a személyes kapcsolat megteremtése, mely fontos szerepet játszhat a beszédalkalmazkodásban. Azonban ebben az esetben egy olyan módszer kidolgozása szükséges, mellyel biztosíthatjuk, hogy a mintabeszélő ne alkalmazkodjon a kísérleti személyekhez. Az ehhez szükséges körülmények megteremtése rendkívül nehéz, így a vizsgálat ily módon történő lebonyolítása kérdéses.

Mint arra már a 4. fejezetben (Módszertan) is utaltunk, szándékosan úgy választottuk ki a kísérletben megadott célszavakat, hogy ugyanezeket a hangfelvételeket felhasználva egy későbbi kutatásban a magánhangók minősége is elemezhető legyen. A későbbiekben a magánhangzók időtartamát, valamint hangszínét is tervezzük megvizsgálni, utóbbit az első és a második formáns értékével számszerűsítve. Ezen kívül a felvételek tanúsága szerint a mintabeszélő egyik gyakran megjelenő beszédjellemzője volt az irreguláris zöngé, tehát a szabálytalan hangszalagrezgés, „rekedtes hang” megjelenése. Ez a magyar beszédben leginkább a megnyilatkozások végére esik, és – ami a jelen vizsgálat szempontjából különösen érdekes –, nagy mértékben egyénfüggő, egyénre jellemző beszéd-sajátosság is (Markó 2017). A továbbiakban az is vizsgálható lehet, hogy az irreguláris zöngét érinti-e a beszédalkalmazkodás folyamata, azaz például megnő-e az irreguláris zöngé megjelenésének gyakorisága a résztvevők beszédében a mintabeszélő beszédének hatására.

7 Összefoglalás

A jelen kutatásban a vizuális ingerek befolyásoló hatását vizsgáltunk a beszédalkalmazkodás folyamatában. Az alaphérfkvenciában nem tapasztaltunk alkalmazkodást sem az auditív, sem az audiovizuális csoportban, azaz függetlenül attól, hogy a kísérletben részt vevők látták-e a

mintabeszélőt vagy sem. A beszédtempót illetően azonban elmondható, hogy a kísérletben részt vevő két csoport közül az audiovizuális csoport részvevői alkalmazkodtak nagyobb mértékben a mintabeszélőhöz.

Levonhatjuk tehát azt a következtetést, hogy fokozza a beszédalkalmazkodás mértékét az, ha a hallgató a beszélőt nem csak hallja, de egyúttal látja is a megnyilatkozó arcát, valamint a szabad szemmel megfigyelhető artikulációs mozgásait. Mivel ezek az artikulációról szóló információk ismert módon segítik a beszédfeldolgozást (Gósy 2005), ezért az eredményeink végeredményben arra is mutatnak, hogy az a tény, ha a beszélő hatékonyabban észleli/érti meg a beszélgetőtársa beszédét, a beszédalkalmazkodást is facilitálhatja.

8 Szakirodalom

- Aubanel, Vincent – Nguyen, Noël (2020). Speaking to a common tune: Between-speaker convergence in voice fundamental frequency in a joint speech production task. *PLoS ONE*, 15 (5). <https://doi.org/10.5281/zenodo.3630439>
- Babel Molly – Bulatov D. (2011). The role of fundamental frequency in phonetic accommodation. *Language and Speech*, 55(2), 231–248.
- Babel Molly (2009). *Phonetic and Social Selectivity in Speech Accommodation*. Doktori értekezés. University of California, Berkeley.
- Babel, Molly – McGuire, Grant – Walters, Sophia – Nicholls Alice (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology* 5(1), 123–150.
- Barón-Birchenall, Leonardo (2023). Phonetic accommodation during conversational interactions: An overview. *Revista Guillermo de Ockham*, 21(2), 493–517.
- Boersma, Paul – Weenink, David (2023). Praat: doing phonetics by computer [Computer program]. Version 6.3.10. <http://www.praat.org/>
- Bosshardt, Hans-Georg – Sappok, Christian – Knipschild, M. – Hölscher, Christoph (1996). Spontaneous imitation of fundamental frequency and speech rate by nonstutterers and stutterers. *Journal of Psycholinguistic Research*, 26/4, 425–448.
- Casasanto, Laura Statum – Jasmin, Kyle – Casasanto Daniel (2010). Virtually accommodating: Speech rate accommodation to a virtual interlocutor. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 32(32), 127–132.
- Fercsik Erzsébet – Raátz Judit (2018). *Kommunikáció és nyelvhasználat*. Műszaki Könyvkiadó, Budapest.
- Gallois, Cindy, – Ogay, Tatyana – Giles Howard (2005). Communication accommodation theory. A look back and a look ahead. In W. B. Gudykunst (Ed.) *Theorizing about intercultural communication*. (pp. 121-148). Thousand Oaks: Sage.
- Garnier, Maëva – Lamalle, Laurent – Sato Marc (2013). Neural correlates of phonetic convergence and speech imitation. *Frontiers in Psychology*.
- Giles, Howard – Coupland, Nikolas – Coupland, Justine (1991). Accommodation theory: Communication, context, and consequence. In: Giles H, Coupland J, Coupland N, eds. *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Studies in Emotion and Social Interaction. Cambridge University Press; 1991:1-68.
- Gósy Mária (2004). *Fonetika, a beszéd tudománya*. Osiris, Budapest.

- Gósy Mária (2005). *Pszicholingvisztika*. Osiris, Budapest.
- Markó Alexandra (2017). Hangtan. In: Tolcsvai Nagy Gábor szerk., *Nyelvtan*. Osiris, Budapest.
- Namy Laura L. – Nygaard Lynne C. – Sauerteig, Denise (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*, 21(4), 422–432. <https://doi.org/10.1177/026192702237958>
- Szabó Ildikó Emese (2019). Phonetic selectivity in accomodation: The effect of chronological age. In Sasha, C., Escudero, P., Tabain, M., Warren, P. (szerk.): *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS) 2019*, Australasian Speech Science and Technology Association Inc., 3195–3199.
- Tiede, Mark – Bundgaard-Nielsen, Rikke – Kroos, Christian – Gibert, Guillaume – Attina, Virginie – Kasisopa, Benjawan – Vatikiotis-Bateson, Eric – Best, Catherine (2010). Speech articulator movements recorded from facing talkers using two electromagnetic articulometer systems simultaneously. *Acoustical Society of America, Proceedings of Meetings on Acoustics 11*. [DOI: 10.1121/1.4772388]
- Weisstein, Eric. Equal Temperament. In from Eric Weisstein's Treasure Trove of Music. <https://web.archive.org/web/20190614131343/> Letöltés: 2025. 01. 13.