

AZ ARTIKULÁCIÓS TEMPÓ ÉS AZ ÁTLAGOS ALAPHANG SZEREPE A BESZÉLŐ ÉLETKORÁNAK MEGBECSLÉSÉBEN

Gocsál Ákos

Bevezetés

Számos hazai és külföldi tanulmány foglalkozik azzal a kérdéssel, hogy az emberi hang a beszélő személy mely tulajdonságait tükrözi, milyen akusztikai paraméterek formájában, illetőleg a hallgató a beszélőt a hangja alapján milyen tulajdonságúnak véli. A beszélő hangjában tükröződő emberi tulajdonságokat Traunmüller (1994) személyes minőségnek (personal quality), a közölt üzenettől és a beszélő szándékától független extranyelvi „üzenetnek” nevezi.

A beszédparaméterek, illetve a beszélő személyes tulajdonságai közötti összefüggések vizsgálata három fő kérdés köré irányul. Egyrészt: összefüggéseket keresünk az objektív módon mérhető, meghatározható tulajdonságok (a beszélő neme, életkora, testalkatának jellemzői stb.) és beszédének akusztikai paraméterei között. Másrészt: vizsgáljuk, hogy az objektív módon mért tulajdonságra a hallott hangfelvételek alapján vissza tudnak-e következtetni a lehallgatók, azaz mennyire pontosak az ilyen becslések, ítéletek (életkorra, testalkatra pl. Gósy 2001). A harmadik fő terület pedig annak vizsgálata, hogy mely beszédakusztikai paraméterek korrelálnak a becslésekkel, azaz melyek azok a paraméterek, amelyeket a lehallgatók a beszélőkkel kapcsolatos benyomásaik kialakításakor felhasználnak.

E kutatási terület egyik gyakran vizsgált részterülete a beszélő személy életkorának hangja alapján történő megbecslésével foglalkozik. Az életkorbecslés kutatásakor a lehallgatók vagy éves pontossággal határozzák meg a beszélő életkorát, vagy korcsoportba sorolják őket. Az éves pontossággal megbecsült életkori adatok átlaga, illetve a beszélők valós életkorának korrelációja segítségével meghatározható, hogy a lehallgatók mennyire adtak következetes becsléseket. Ezzel az eljárással több kutató is szignifikáns kapcsolatot talált. Az elsők között vizsgálta ezt a jelenséget Shipp és Hollien, akik 1969-es tanulmányukban $r = 0,88$ értékű korrelációs együtthatót közöltek (idézi: Huntley et al. 1987). Winkler, Brückl és Sendlmeier (2003) izolált hangokkal, felolvasott szöveggel és spontán beszéddel is végeztek életkorbecslési kísérletet. Az észlelt és a valós életkor minden esetben szignifikáns korrelációt mutatott, de az r értéke spontán beszédnél (0,864) és felolvasásnál (0,862) volt a legmagasabb. Az izolált hangoknál 0,344 és 0,738 közötti, de

DOI: 10.15775/Beszkut.2017.151-167

szintén szignifikáns r értékek adódtak. Cerrato és munkatársai (2000) kutatásában pedig telefonbeszélgetésekből rögzített hangfelvételek alapján $r = 0,77$ adódott. Braun és Cerrato (1999) német és olasz anyanyelvű beszélőkkel és hallgatókkal végeztek kísérleteket. Az olasz beszélők életkorát mindkét hallgatói csoport 0,7-nél magasabb korrelációs együtthatóval tudta megítélni (olasz hallgatók: $r = 0,751$, német hallgatók: $r = 0,745$), míg az német beszélőknél alacsonyabb korrelációs együtthatók adódtak (olasz hallgatók: $r = 0,300$, német hallgatók: $r = 0,578$). Stölten–Engstrand (2002), hat beszélőre (három fiatal, három idős) adott életkorbecslések alapján $r = 0,96$ -es korrelációs együtthatót kaptak a természetes (nem manipulált) bemondásoknál. Saját, korábbi kutatásunk is hasonló eredményt mutatott, az ott nem közölt, de az akkori adatokkal utólag kiszámolt korrelációs együttható értéke is szignifikáns, $r = 0,875$, $p < 0,001$, felolvasott szöveg, mint akusztikai inger használatával (Gocsál 1998).

A szignifikáns összefüggés léte azonban csak azt igazolja, hogy a különböző életkorú beszélők esetén a hallgatók következetes becsléseket adtak, de az a következetesség nem jelenti a becslések pontosságát. Már Shipp és Hollien fent idézett, 1969-es eredményei is rámutattak a fiatal beszélők életkorának felülbecslésére, illetve az idősek „fiatalítására”. A becslések pontatlanságának megállapításához a becslült és a valós életkorok közötti különbségeket határozza meg a kutató, azaz az abszolút hibát számítja ki. Hughes és Rhodes (2010) eredményei a felnőtt beszélők életkorának szignifikáns alulbecslését jelzik, különösen a 46–55 éves korosztályban, ahol a férfiak életkorát közel 15 évvel becslték alul a hallgatók. Egy másik kutatás különböző kísérleti körülményeknél hasonlította össze a becslült és a valós életkorok különbségeit, azonban itt az idősebb (60–70 éves) korosztálynál adódott ilyen mértékű alulbecslés (Skoog Waller et al. 2015). 5–18 éves gyerekek, tizenévesek életkorbecslésénél alkalmazta az észlelt és a valós életkorok különbségét Assmann és munkatársai (2013). Fiú beszélőknél 15 éves kortól, lányoknál már 12 éves kortól figyelte meg az alulbecslést, amikor a beszélő neve ismert volt.

A becslések pontatlanságának meghatározása más módszerrel is történhet. A valós és a becslült életkor hányadosaként megállapított érték – a kettő közül a kisebbet elosztva a nagyobbal – szükségképpen 1-nél kisebb szám, illetve a két adat egyenlősége esetén értelemszerűen 1 a hányados értéke (Cerrato et al. 2000). Ezzel a megoldással Cerrato és munkatársai a kutatásukban szereplő hét korcsoport közül a legfiatalabb és a legidősebb korcsoportban 68%-os, illetve 78%-os pontosságot találtak, míg a középső korcsoportban 96% adódott. Krepsz és Gósy (2016) ugyanezzel a számítási módszerrel jellemzően 80% fölötti értékeket kaptak a különböző beszélőkre, 60% alatti érték csak egy idős férfi beszélő esetén adódott.

Az életkorbecslési kutatások másik módszere nem pontos évszámokat, hanem életkori tartományokat, kategóriákat alkalmaz. A legegyszerűbb számí-

tási módszer ilyenkor az, hogy a kutató megvizsgálja, a becslések hány százaléka esik abba a kategóriába, amelyikbe a beszélő tartozik. Tatár (2013) eredményei szerint a 20–29 éves beszélők életkorára adott becslések mindössze 40%-a volt helyes, míg a legpontosabb becsléseket az 50–65 éves beszélők kapták, 80%-os pontossággal. Amir és munkatársai (2012) 8–18 éves gyerekek, tizenévesek hangjával végeztek életkorbecslési kísérletet. A beszélőket 8, 10, 12, 14, 16 és 18 éves korcsoportokba osztották, és a kísérleti személyeknek is ennek megfelelően kellett becsléseiket megadniuk. A fiúknál a helyes becslések aránya 33,03–50,53 % volt (a legalacsonyabb a 14 éves, a legmagasabb a 16 éves fiúknál), míg a lányoknál 8,16–59,87% (legalacsonyabb a 18, a legmagasabb a 8 éveseknél). Amikor a kutató helyes becslésként kezelte azt is, ha a kísérleti személy a valós életkor melletti kategóriába eső becslést adott, a helyes becslések aránya a fiúknál 80,13–97,89%, a lányoknál pedig 30,26–97,71% lett (Amir et al 2012).

Számos tanulmány foglalkozott azzal, hogy a becslések pontatlanságának okát felderítse. Pontatlanságot okozhat például a lehallgatók életkora. Eppley és Mueller (2001) kutatásában 71 és 82 év közötti beszélők életkorát kellett 18–22, illetve 61–84 éves lehallgatóknak megbecsülniük. A beszélők átlagos életkorát (76,44 év) mindkét hallgatói csoport jelentősen alulbecsülte. A kutatók feltételezése szerint a fiatalok pontatlanabb becsléseket adnak, ami tendenciaszerűen ugyan igazolódott (átlagos becsült életkor: 57,27 év), de az idős hallgatók által adott, átlagosan 61,56 éves becsült életkor nem volt szignifikánsan magasabb. Egy másik kutatás viszont igazolta a hallgató életkorának szignifikáns hatását. Négy korcsoportba osztott hallgatók közül a legidősebbek csoportjába tartozók pontosabban ítélték meg az legidősebb beszélők életkorát, mint a fiatalabb hallgatók. (Hughes–Rhodes 2010). Egy régebbi kutatás pedig azt találta, hogy a 40 év alatti beszélők életkorát sokkal jelentősebben felülbecsülték az idős és a serdülőkorú lehallgatók, mint a fiatal és a középkorú felnőttek (Huntley et al. 1987).

Pontatlanság eredhet abból is, hogy bizonyos életkori csoportokhoz tartozó személyek életkora esetlegesen jobban, míg másoké kevésbé jól ítéltető meg. Emellett a beszélő neme is befolyásolja a becslés pontosságát. Hummert és munkatársai (1999) felolvasott szöveggel végzett kísérletükben az 60 év fölötti beszélőket több csoportra bontották (60–69, 70–79, 80–89 évesek), majd a lehallgatóknak szintén ezeknek megfelelő kategóriákba kellett a beszélőket besorolniuk, kiegészítve a „60 éves vagy az alatti”, illetve „90 éves vagy az fölötti” kategóriákkal. A hallgatók 25,5%-ban adtak helyes becslést a 70–79 éves női beszélők életkorára, míg 46,5 %-ban becsülték helyesen a 60–69 éves női beszélők életkorát. A többi kategóriába eső beszélőkre adott helyes becslések aránya e két érték közé esett. A kísérleti személyek jellemzően alulbecsülték a beszélők életkorát. Egy másik kutatás pedig arra mutatott rá, hogy a 23–34, illetve 46–54 éves női beszélők életkorát pontosabban becsülték a kísérleti személyek, mint az ugyanilyen életkorú férfiakét. Emellett pe-

dig a gyerekek és serdülőkorúak életkorának becslése bizonyult a legpontosabbnak (Hughes–Rhodes 2010). Különböző arányú pontosságértékek adódtak Reubold és munkatársai (2010) kutatásában is, amelyben két beszélő különböző időpontokban rögzített hangjával végezték életkorbecslési kísérletüket. A női beszélőt 46 és 57 éves korában rögzített hangja alapján a kísérleti személyek jellemzően 5-10 évvel fiatalabbnak vélték, míg a férfi beszélő életkorát 39 éves korában rögzített hangja alapján – az átlagos becslt életkort tekintve – jól becsülték meg, a 62 évesen rögzített hang alapján kissé alulbecsülték, 82 évesen rögzített hangja alapján viszont nagyon jelentősen alulbecsülték.

További érdekes kérdéseket vet fel az, hogy a hallgatók vajon milyen akusztikai kulcsokat használnak fel az életkorbecsléshez.

Hummert és munkatársai (1999) azt találták, hogy a férfi beszélőknél észlelt életkorral sem az alaphangmagasság, sem annak szórása (SD), szórási tartománya, és még a jitter sem korrelált, egyedül a hangerő és annak szóródása mutatott tendenciaszerű, de nem szignifikáns összefüggést (a hangosabban beszélőt tendenciaszerűen idősebbnek gondolták a hallgatók). A női beszélőknél pedig a nagyobb f_0 -szóródással beszélőket szignifikánsan idősebbnek gondolták, de az alaphang magassága nem mutatott szignifikáns összefüggést a becslt életkorral. Tendenciaszerű korreláció ($p < 0,1$) a jitter és a hangerő szóródásával mutatkozott. Winkler–Sendlmeier (2006) a hangrés nyitott állapotának idejét jellemző OQ (open quotient) és az elképzelt beszélő életkora között találtak pozitív korrelációt. Kitarított [a] hangokkal végzett kísérletükben a nagyobb OQ-értékekkel kiejtett hangzók idősebb beszélő benyomását keltették, de ez csak a férfi beszélőknél volt jellemző.

Reubold és munkatársainak (2010) kísérlete az f_0 és az F_1 szerepét vizsgálta. A korábbiakban már említett két beszélő hangján manipulációkat végeztek, mesterségesen növelték, illetve csökkentették az f_0 és az F_1 értékeit, és az így előállított hangmintákkal végeztek további életkorbecslési kísérleteket. Az eredmények azt mutatták, hogy az alaphang emelése és mélyítése szignifikáns hatással volt a becslt életkorra, a magasabb alaphang fiatalabb beszélő benyomását keltette és fordítva. A jelenség a női beszélő hangjaival erőteljesebben jelentkezett. A szerzők egyes esetekben az F_1 hatását is kimutatták, bár egyértelműen meghatározható szabályszerűséget nem tudtak megfogalmazni a különböző kísérleti összeállításokban kapott eredmények alapján. A tempó manipulációja hasonlóképpen szignifikáns hatással volt a becslt életkorra egy másik kísérletben is (Skoog Waller et al. 2015). A természetes hangbemondások tempójának 10%-os növelésével fiatalabbnak, 10%-os lassításával viszont idősebbnek tűnt a beszélő. Mivel ebben a kísérletben is tapasztalható volt a felülbecslés a fiataloknál, illetve alulbecslés az időseknél, a fiatalok életkorának becslése a gyorsítással, az időseké pedig a lassítással valamivel pontosabbá vált.

A tempó és az alaphang kombinált manipulációi azonban egyértelműen csak a tempó szignifikáns hatását igazolták, az alaphang manipulációi önmagukban már nem eredményeztek változást a becslésekben (Harnsberger et al. 2008). Stölten és Engstrand (2003) mesterségesen lassították a beszéd tempóját, illetve megemelték a beszélő alaphangját. A valós és az észlelt életkor között szignifikáns összefüggés adódott az eredeti ($r = 0,96$) és a manipulált ($r = 0,86$) hangfelvételek lejátszásakor, de a manipulált hangok alapján az elképzelt beszélőt mintegy 10 évvel idősebbnek vélték a hallgatók, mint az eredeti hang alapján. A szerzők a jelenség okának a lassabb tempót jelölték meg, de mélyebb statisztikai elemzéseket, különösen az f_0 szerepével, esetleges hatásával kapcsolatban nem végeztek.

Természetes bemondások akusztikai paraméterein alapuló szintézist használt Winkler (2007) annak megállapítására, hogy szavak különböző tempó- és alaphang-kombinációival előállított változatai eltérő életkorbenyomásokat alakítanak-e ki a hallgatókban. A gyorsabb tempók a férfi és a női változatban előállított szavaknál is fiatalabb beszélő benyomását keltették. Az alaphang emelése a női változatú szavaknál lassú és a normál tempó mellett nem eredményezett szignifikánsan elérő becsült életkorokat, csak a gyors tempókánál, itt fiatalabbnak vélték a kísérleti személyek a beszélőt. A férfi változatban szintetizált szavaknál viszont az alaphang változtatása kevésbé befolyásolta az életkori ítéleteket a gyors és normál tempóban előállított szavaknál, a leglassabb tempóval előállított szavaknál viszont a legmagasabb alaphang szignifikánsan idősebb beszélő benyomását keltette.

A hazai szakirodalomban a mai napig hat tanulmányról tudunk, amelyek az életkorbecslés problémakörét vizsgálták. Saját kutatásunkban a beszélők életkorát év szerinti pontossággal becsülték meg a hallgatók, azonban a valós és a becsült életkorok összefüggésének, illetve az alul- és felülbecslések statisztikai alátámasztása már nem történt meg (Gocsál 1998). Egy újabb kutatásban Bóna (2013: 126, 128) a valós és az átlagos becsült életkori adatok között szignifikáns kapcsolatot talált 60 év fölötti férfi ($r = 0,907$, $p < 0,001$) és női ($r = 0,809$, $p = 0,003$) beszélőknél. A kísérleti személyek által adott helyes becslések arányát életkori kategóriákba soroltan közölte Gósy (2001), aki azt találta, hogy a 40 év fölötti beszélők életkorát mintegy 65%-ban, a 20 és 30 év közöttieket pedig közel 75%-ban becsülték helyesen a hallgatók. A fiatalok életkorának alulbecslése, az idősök felülbecslése egyértelműen kirajzolódik Krepsz és Gósy (2016) eredményei alapján. Tatár (2013) korábban említett adatai azonban az idősebb beszélők életkorának pontosabb becslésére utalnak. Ismeretesek még gyermekek életkorának becsléséről közölt adatok. Tóth (2014) azt találta, hogy óvodás- illetve iskoláskorúak 90%-os, illetve 86%-os pontossággal tudták a hallgatók besorolni a beszélő gyermekeket.

Arra a kérdésre, hogy milyen beszédparaméterek befolyásolhatják az életkori ítéleteket, több hazai kutatás is kereste a választ, részben a kísérleti személyek megkérdezésével, részben mérésekkel. Bóna (2013: 128) a lehallga-

tóktól jellemzően azt a választ kapta, hogy a tempó, az alaphang és a hangszínezet alapján becsülték meg a beszélő életkorát. Krepesz és Gósy (2016) a megakadások hatását keresték. Eredményük szerint a megakadások inkább megzavarták a becsléseket, az idősebb beszélők esetében pontatlanabb becslések adódtak, amikor a beszédminta megakadásokat tartalmazott, mint amikor nem. Saját kutatásunk pedig a formánsszerkezet esetleges hatását vizsgálta három magánhangzó különböző realizációival. A női változatban előállított magánhangzók nem idéztek elő különböző életkorbenyomásokat, míg a férfi változatú hangoknál az F_1 alacsonyabb értékei idősebb beszélő benyomását keltették (Gocsál 2016).

A fentiek alapján az alábbi kérdéseket fogalmazzuk meg. 1. Magyar anyanyelvű beszélők és hallgatók esetében mennyire pontosak az életkorbecslések? Milyen erős összefüggés mutatkozik a valós és a spontán beszéd alapján megbecsült életkor között? A szakirodalomból ismert adatok alapján hogyan jellemezhetőek a pontatlanságok? 2. A szakirodalomban felmerül a tempó, illetve esetlegesen az f_0 szerepe az életkorbecslésben. Kérdés tehát, hogy a magyar anyanyelvű beszélők esetén, magyar spontán beszéd alapján a becsült életkort mennyiben magyarázza a beszélő artikulációs tempója és alaphangja? Kísérletünkben az alaphanghoz kapcsolódó más paraméterek esetleges hatását is keressük.

Kísérleti személyek, anyag, módszer

Akusztikai ingerként a BEA adatbázisból (Gósy et al. 2012) 24 férfi beszélő hangját használtuk fel, életkoruk 20–72 év. A beszélőket úgy választottuk ki, hogy életkoruk eloszlása megközelítőleg egyenletes volt az említett tartományban. Két, életkorukat tekintve egymást követő beszélő között átlagosan 2,16 év volt a korkülönbség, ténylegesen 0–4 éves lépésköz volt közöttük.

Mindegyik beszélő felsőfokú végzettségű és nem dohányzik. A lejátszáshoz a BEA protokoll szerinti „interjú”, illetve „érvelés” című részekből 20-30 másodperc hosszúságú részleteket választottuk ki, melyeken az interjúalanyok a foglalkozásukról, szabadidős tevékenységeikről, vagy más, általános témákról, például a városi közlekedésről beszéltek. Ezeket a részleteket úgy választottuk ki, hogy tartalmuk szerint bármilyen életkorú személytől származhattak volna. A kísérletben 49 lehallgató (38 nő, 11 férfi) szerepelt, életkoruk: 19–37 év, átlagos életkoruk 22 év. Mindannyian egyetemi hallgatók, fonetikai szempontból képzetlenek.

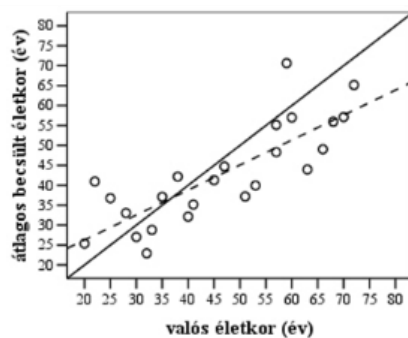
A hangminták lejátszására korábbi kísérletünkhöz hasonló körülmények között, a PTE Művészeti Karának előadótermében került sor (Gocsál 2016). A lejátszást megelőzően a kísérlet vezetője ismertette a kísérlet menetét, majd próbaként lejátszott néhány hangmintát, hogy a kísérleti személyek megismerkedjenek a feladattal. Minden hangminta egy 1 másodperces, 440 Hz-es figyelemfelhívó hangjelzéssel kezdődött, amelyet egy 2 másodperces szünet követett. A kutató akkor indította a következő felvétel lejátszását,

amikor már minden kísérleti személy rögzítette a becslését az adatlapon. A próbát követően minden hangminta csak egy alkalommal hangzott el.

Az 1. kísérlet eredményei

A kutatás során a kísérleti személyeknek évre pontosan kellett megbecsnüük a beszélő személy életkorát. Az így kapott becslésértékekből minden egyes beszélőre vonatkozóan kiszámítottuk a becslt életkori adatok átlagértékét, majd a korrelációs számítás módszerével meghatároztuk az átlagértékek és a valós életkor közötti összefüggés szorosságát. Második lépésként lineáris regressziós modellt állítottunk fel, amelyben a valós életkor mint magyarázó változó szerepelt. Harmadik lépésként a pontatlanság mértékét számítottuk ki mindkét korábban ismertetett módszerrel, egyrészt a valós életkortól való abszolút eltéréseket, másrészt a Cerrato és munkatársai (2000) illetve Krepsz és Gósy (2016) által ismertetett arányszám (valós és becslt életkor összefüggése, VBÖ) segítségével.

Az 1. ábra a beszélők valós életkora és az átlagos becslt életkorok közötti összefüggést mutatja. Az ábrán minden pont egy-egy beszélőre utal. A két adatsor közötti korrelációs együttható: $r = 0,808$, $p < 0,001$, ami erős, szignifikáns összefüggést jelez. Ha a kísérleti személyek a beszélők életkorát pontosan eltalálták volna, akkor a diagram pontjai az ábrán folytonos vonallal jelölt $y = x$ egyenesre illeszkedtek volna. Az ábráról azonban leolvasható, hogy a pontok a szaggatott vonallal jelzett regressziós egyenesre illeszkednek, amelynek a meredeksége kisebb, mint az $y = x$ egyenesé. Megállapítható, hogy a regressziós egyenes a megközelítőleg 35 évnél fiatalabb beszélőknél az $y = x$ egyenes fölött helyezkedik el. Ez arra utal, hogy a kísérleti személyek a fiatal beszélők életkorát jellemzően felülbecsülték, míg az annál idősebbeket jellemzően fiatalabbnak vélték.



1. ábra

A valós életkorok és az átlagos becslt életkorok összefüggése

A kapott adatok alapján a regressziós egyenes egyenlete is meghatározható. A regressziós modell felállítása során magyarázó változónak a valós életkort, függő változónak az átlagos becsült életkort tekintjük. A számítás eredménye egy szignifikáns magyarázóerejű lineáris regressziós modell, melynek determinációs együtthatója: $R^2 = 0,652$. Ez az érték arra utal, hogy az átlagos becsült életkorok varianciáját 65,2%-ban a beszélő valós életkora magyarázza. A regressziós egyenes egyenlete pedig a következő:

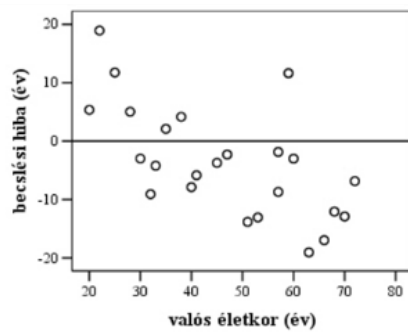
$$\text{átlagos becsült életkor} = 0,627 \text{ valós életkor} + 13,74$$

A regressziós egyenlet segítségével kiszámítható a pontos becslések helye. Pontos a becslés, ha az átlagos becsült életkor megegyezik a valós életkorról, azaz ahol az 1. ábrán a regressziós egyenes és az $y = x$ egyenes metszi egymást. Megállapítható, hogy a modell 36,83 éves valós életkorra jósolja a pontos becslést. Ettől távolodva, a két egyenes szétartása az alul- és a felülbecsléseket mutatja. A valós életkor pozitív, de 1-nél kisebb együtthatója – a regressziós egyenes meredeksége – magyarázza az idősebb beszélőknél megfigyelhető alulbecslést: a valós életkor egységnyi, pl. egy évvel való növelésével a becsült életkornak csak 0,627 éves növekedése jár együtt. Egy egyszerű példával illusztrálva az alulbecslést: a modell azt jósolja, hogy a pontos becslések helyétől számítva a 30 évvel idősebb személyt (azaz 66,83 éves beszélőt) a modell szerint csak $30 \times 0,627 = 18,81$ évvel gondolják idősebbnek a hallgatók, azaz átlagosan $36,83 + 18,81 = 55,64$ évesnek fogják becsülni.

A következő lépés a becslések pontatlanságát meghatározása. A 2. ábrán az idézett szakirodalomban közöltekhez hasonló elrendezésű pontthalmazt kaptunk. A legnagyobb mértékű felülbecslést a 22 éves beszélőnél találtuk, őt közel 20 évvel vélték idősebbnek, míg a 63 éves beszélő volt az, akinek a korát a leginkább alulbecsülték. Nála szintén közel 20 év volt a különbség, őt átlagosan megközelítőleg 43 évesnek vélték.

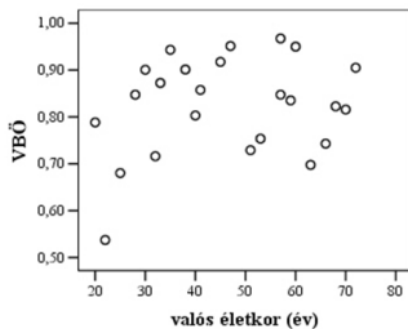
A 3. ábra a VBÖ-értékeket mutatja az életkor függvényében. Krepsz és Gósy (2016) felhívja a figyelmet arra, hogy a VBÖ kizárólag a pontatlanság mértékét mutatja, az irányát nem. A kapott adatok egy kivételével 0,6 fölöttiek, átlagértékük 0,82. A grafikon nem mutat semmilyen összefüggést arra vonatkozóan, hogy az életkor függvényében a VBÖ-érték szignifikánsan eltérne (3. ábra).

Az eddig elvégzett számítások során az egyes beszélőkre kapott átlagos becsült életkori értékeket használtuk. Az átlagszámítás, bár használata kényelmes, információvesztéssel jár együtt, mivel az adatsor egyetlen értékkel való helyettesítésével elvesz az adatsor szóródása. Ennek ellensúlyozására külön számítást végzünk, azaz az egyes beszélőknél kapott becsült életkorok szóródási (SD) értékeinek alakulását vizsgáljuk meg a valós életkor függvényében (4. ábra).



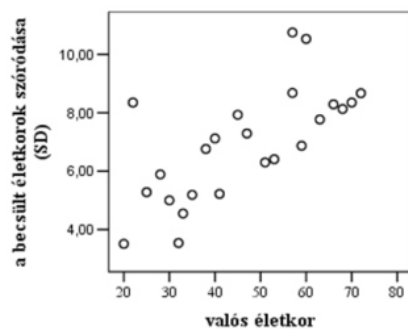
2. ábra

A valós és az átlagos becslült életkorok különbségeinek összefüggése



3. ábra

A VBÖ-érték és a valós életkorok összefüggése



4. ábra

A becslésértékek szóródása és a valós életkorok összefüggése

A 4. ábráról leolvasható, hogy az idősebb beszélők esetében lényegesen magasabb SD-értékek adódtak a becslések során, mint a fiatalabbaknál. Az összefüggés szignifikáns ($r = 0,692$, $p < 0,001$). Ez az eredmény azt mutatja, hogy a lehallgatók között lényegesen nagyobb volt az egyetértés a fiatalabb beszélők életkorának megítélésében, mint amikor idősebb beszélők életkorát ítélték meg.

A 2. kísérlet eredményei

A 2. kísérletünkben azt kerestük, hogy a becsült életkorok átlagértékeit mennyiben magyarázzák a lejátszott hangminták egyes akusztikai paramétereit. A kísérlethez az alábbi paramétereket mértük meg, illetve számítottuk ki:

- artikulációs tempó
- az f_0 mediánja
- az f_0 szórási tartománya
- az f_0 eloszlásának csúcsossága
- az f_0 eloszlásának ferdesége
- az irregulárisan képzett f_0 -értékek százalékos aránya

Az alaphang mérése során a következő eljárást alkalmaztuk. Minden egyes hangmintára vonatkozóan Praat program *Pitch listing* funkciójával megjelenítettük az összes mért f_0 -értéket, majd ezeket táblázatkezelő programba importáltuk (Boersma–Weenink 2014). Oszlopdiaagramos ábrázolás segítségével megvizsgáltuk az f_0 -értékek eloszlását. Esetenként hibás mérési adatokat találtunk, ezek jellemzően zöngétlen réshangokhoz tévesen rendelt 500–600 Hz-es értékek voltak. A számítások során ezeket a téves mérési adatokat figyelmen kívül hagytuk. Az oszlopdiaagramok alapján megállapítottuk azt is, hogy mely beszélőknél fordultak elő irreguláris zöngék. Az irregulárisan képzett zöngék frekvenciája a szakirodalmi adatoknak megfelelően (vö. Markó 2013: 19) igen mélyen, jellemzően 30–70 Hz között fordult elő. Az f_0 -t jellemző, fentiekben felsorolt paraméterek megállapításához csak a modális zöngét használtuk fel. Ezzel azonban elvesz az adathalmazból az irreguláris zöngékkel kapcsolatos információtartalom, márpedig ennek sajátos hangzása miatt – esetlegesen – szerepe lehet az életkor megbecslésében. Így, 100 százaléknak tekintve az összes helyesen mért f_0 -értéket, megállapítottuk minden egyes beszélőre vonatkozóan az irregulárisan képzett zöngék f_0 -értékeinek százalékos arányát is, amelyet szintén felhasználtunk a további számításokhoz.

Az f_0 középértékének jellemzéséhez a mediánt használtuk, mivel az eloszlási grafikonok aszimmetrikus jelleget mutattak (vö. Gocsál 2012). Emiatt az eloszlást is az általában használatos átlagos eltérés vagy szórás (SD) helyett a csúcsosság, ferdeség és a szórási tartomány segítségével jellemeztük.

Az említett módszerekkel az alábbi adatokat kaptuk. Artikulációs tempó: 10,29–16,14 hang/s (átlag: 13,59 hang/s), f_0 : 86–121 Hz (átlag: 102,97 Hz), f_0 szórási tartománya: 36,11–122,86 Hz (átlag: 77,85 Hz), ferdeség: 0,46–3,81

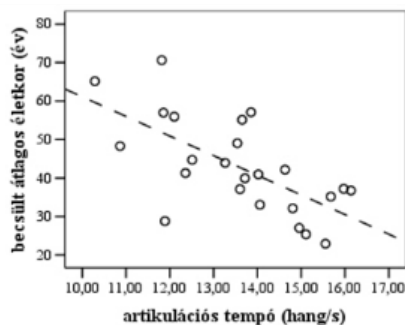
(átlag: 1,61), csúcsosság: 0,54–19,4 (átlag: 4,6), irreguláris zöngék 0–27% (6,38%, 6 esetben nagyobb 10%-nál). Az adatok rögzítését, illetve kiszámítását követően lineáris regressziós modellt állítottunk fel, melyben függő változóként az átlagos becsült életkor, magyarázó változóként pedig a mért, illetve kiszámított akusztikai paraméterek szerepeltek.

A modell felállításához a FORWARD módszert használtuk. A számos lehetséges magyarázó változó ellenére csak egyetlen szignifikáns magyarázó-erejű modell adódott, melynek determinációs együtthatója: $R^2 = 0,433$, azaz a modell 43,3%-ban magyarázza az átlagos becsült életkorok varianciáját. A feltételezett magyarázó változók közül egyedül az artikulációs tempóról bizonyosodott be, hogy szignifikáns hatással van a függő változó varianciájára. Ennek megfelelően a regressziós egyenes egyenlete az alábbiak szerint alakult:

$$\text{átlagos becsült életkor} = -5,104 \text{ artikulációs tempó} + 112,16$$

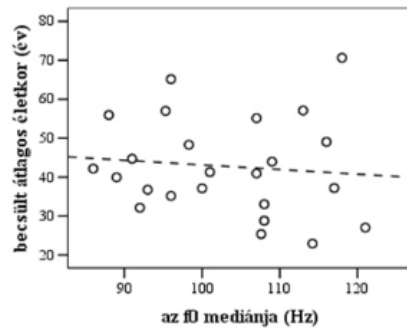
Mivel csak egy független változó szignifikáns hatása igazolódott, az összefüggés pontdiagramon ábrázolható (5. ábra). A két mennyiség közötti korrelációs együttható értéke: $r = 0,658$, $p < 0,001$. Az artikulációs tempó együtthatója negatív előjelű, ami arra utal, hogy az artikulációs tempó emelkedésével a modell által megmagyarázott átlagos becsült életkor csökken, mégpedig 1 hang/másodperces tempónövekedés 5,104 éves csökkenést jelent az átlagos becsült életkorban. Másképp fogalmazva, a gyorsabb tempók fiatalabb beszélő benyomását keltik, és fordítva.

Kutatási céljaink között szerepelt az f_0 és a becsült életkorok közötti összefüggések vizsgálata is. A 6. ábrán ugyan az látható, hogy a pontokra illesztett egyenes enyhén lejt, de a korrelációs együttható nullához közeli ($r = -0,1$), így még tendenciaszerű kapcsolat sem mutatható ki a két mennyiség között.



5. ábra

Az artikulációs tempó és a becsült átlagos életkor összefüggése



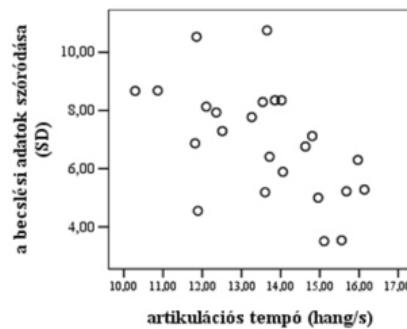
6. ábra

Az f_0 mediánja és a becsült átlagos életkor összefüggése

Az 1. kísérlet során bebizonyosodott, hogy a becslési értékek szóródása függ a beszélő életkorától, azaz az idősebb beszélők esetében kevésbé voltak egységesek a kísérleti személyek becslései, mint a fiatalabbaknál. Felmerül a kérdés, hogy vajon – a becslésértékekhez hasonlóan – a becslésértékek szóródását hogyan befolyásolják az akusztikai paraméterek.

Az előzőekhez hasonlóan lineáris regressziós modellt állítunk fel, melyben függő változó az egyes beszélőknél mért becslési adatok szóródása (SD), magyarázó változóknak pedig a mért akusztikai paramétereket tekintjük.

A számítás a korábbihoz hasonló eredményt adott. Szignifikáns magyarázóerejű modellt kaptunk ($R^2 = 0,302$), de ebben az esetben is csak az artikulációs tempó magyarázóereje igazolódott. A lassabban beszélők esetében a becslési adatok nagyobb mértékben szóródnak (7. ábra).



7. ábra

Az artikulációs tempó és a becsült átlagos életkor szóródásának összefüggése

A 7. ábrán mutatkozó összefüggés megerősíti az eddigieket, de új információval nem szolgál. A lassabb artikuláció ugyanis szoros összefüggésben van a magasabb életkorral. A 4. ábra és a hozzá kapcsolódó számítás alapján pedig már beláttuk, hogy az idősebbeknél az életkorbecslések szóródása nagyobb. A regressziószámítást azonban mégis célszerű volt elvégezni, mivel – elviekben – más akusztikai paraméterek hatása is jelentkezhett volna.

Következtetések

Első kísérletünkben a becsült és a valós életkorok összefüggését vizsgáltuk. A jelen kísérleti összeállításban – magyar anyanyelvű beszélők és hallgatók alkalmazásával, spontán beszéd alapján – a korábbi kutatásokban közöltékhez hasonló mértékű összefüggést ($r = 0,808$) sikerült igazolni. A kutatás megerősítette a más kísérletekben szintén tapasztalat alul- és felülbecslést a fiatal, illetve idős beszélőknél, ennek magyarázatára regressziós modellt állítottunk fel. Bóna (2013: 127) 63–90 éves férfi beszélők esetében átlagosan 19,8 éves alulbecslést mért úgy, hogy átlagosan legpontosabb becslést a 70 éves beszélő esetében kapta, 15,1 éves alulbecsléssel. Regressziós modellünk a 70 éves beszélőre 57,63 éves megbecsült életkort jelez előre, azaz ebben az esetben nem egészen 3 éves eltéréssel ad jóslást a becsült életkorra. Egy másik, Bóna tanulmányában szereplő, 90 éves férfi beszélő életkorát pedig – az ott közölt 7.2. ábráról leolvastva – a lehallgatók megközelítőleg 70 évesre becsülték, a modell szerint az megjósolt becsült életkor 70,17 év.

A modell azonban Krepisz és Gósy (2016) tanulmányában szereplő adatokkal már kevésbé működik jól. A szerzők közlése szerint a becslések átlaga az 55 éves beszélők esetében közelítette meg a legjobban a valós életkort, míg a kísérleti személyek a 35 évesek életkorát is jelentősen alulbecsülték. Ennek oka lehet az, hogy modellünk kizárólag férfi beszélőkre vonatkozik, míg Krepisz és Gósy (2016) kutatásában férfi és női beszélők egyaránt szerepeltek. Feltehető – és a szakirodalomban is találunk ilyen utalásokat, pl. Hummert és munkatársai (1999) –, hogy eltérő módon, eltérő pontossággal működik a férfi és női beszélők esetében az életkorbecslés, emiatt célszerűnek tűnik női beszélőkkel is megismételni a kísérletet, és külön elvégezni a regressziós számítást.

A VBÖ-értékek a szakirodalmi adatokhoz hasonlóan alakultak, a beszélő életkorával való összefüggés nem mutatkozott. A 3. ábráról azonban az olvasható le, hogy a lehallgatók az egyik, életkorban hozzájuk közel álló beszélő esetében lényegesen alacsonyabb VBÖ-t mutattak. Ez az eredmény összhangban van azzal, hogy ugyanennél a beszélőnél volt a legjelentősebb a felülbecslés, illetve a becsült életkorok szóródása is. A kutatásunkban mért paraméterek nem adnak magyarázatot arra, hogy miért kaptunk ennyire eltérő adatokat az ő esetében.

A szóródási értékek magasabb életkorú beszélő esetében nagyobbak voltak, mint a fiatalabbaknál (4. ábra), azaz, minél távolabb esett a beszélők

életkora a lehallgatókétól, annál bizonytalanabbak voltak a becsült életkori értékek. Ilyen összefüggés nem mutatkozott Krepsz és Gósy (2016) kutatásában, azonban ott kevesebb beszélő vett részt a kutatásban, 10 éves különbségekkel, korcsoportonként csak egy férfi és egy nő. Az azonban saját kutatásunkból, Krepsz és Gósy (2016), illetve Bóna (2013: 127) adataiból is világosan látszik, hogy egy-egy beszélő életkorának megítélésében igen nagy különbségek lehetnek a szóródási adatokban. Még az azonos korcsoporthoz tartozó beszélőknél is jelentős különbség mutatkozhat az életkor megítélésében. Feltételezhetően ennek az az oka, hogy egyes beszélők hangja jól megfelel a vélt életkornak megfelelő hangprototípusnak, így nagyobb a lehallgatók közötti egyetértés, míg mások hangjában a különböző paraméterek különböző életkoroknak megfelelő hangprototípusoknak felelhetnek meg. Elképzelhető, hogy ilyen „ellentmondásos” esetekben az egyes lehallgatók más-más stratégiával próbálják a beszélő életkorát megbecsülni, emiatt adódik a nagyobb szóródás. Az sem zárható ki, hogy az életkori hangprototípusok kialakulása egyénileg eltér, és a saját életkortól távolabb eső korcsoportokhoz tartozó hangprototípusok egyéneknél is nagyobb különbségeket mutatnak.

Kutatásunk második része az akusztikai paraméterek életkorbecslésben betöltött szerepére vonatkozott. A szakirodalommal megegyezően kimutattuk a tempó szignifikáns hatását, regressziós modellünk szerint 43,3%-ban magyarázza az átlagos becsült életkort. Más paraméterek, főleg az f_0 hatása nem igazolódott. Ez összhangban van néhány korábbi kutatás eredményeivel (Hummert et al. 1999; Harnsberger et al. 2008), más kutatók azonban az f_0 hatását is kimutatták (Reubold et al. 2010). Ennek a látszólagos ellentmondásnak lehetséges oka a kísérleti összeállítás: ahol manipulált hangokat használtak, ott a megváltoztatott paraméteren, azaz a megemelt vagy mélyített alaphangon kívül minden más paraméter változatlan maradt. Így, miután a különböző hangmintákat hallva a tempó nem adott támpontot a lehallgatóknak az életkorbecsléshez, az alaphangot vették alapul. Elképzelhető tehát, hogy léteznek „elsődleges”, illetve „másodlagos” akusztikai kulcsok az életkorbecslés során. A tempó ilyen „elsődleges” kulcs, önmagában szignifikáns magyarázóerőt ad a regressziós modellben, ami nem is meglepő, hiszen kutatásunkban a valós életkor és az artikulációs tempó között szignifikáns az összefüggés ($r = -0,58$, $p < 0,01$), az f_0 és a valós életkor között viszont nem áll fenn ilyen kapcsolat ($r = -0,052$, nem szignifikáns). Ez alapján tehát „helyesen” ítélték kutatásunkban a lehallgatók, amikor becsléseikben nem használták fel az alaphangot.

Az alaphang akusztikai kulcsként viszont akkor játszhatott az életkorbecslésben meghatározó szerepet, amikor ugyanazon személy hangjával történtek manipulációk (Reubold et al. 2010), miközben az alaphangon kívül más paraméterek változatlanul maradtak. A kísérleti személyek döntése ebben az esetben is „helyes”, mivel úgy tűnik, „tudják”, hogy ugyanazon személy

alaphangja az életkor előrehaladtával hogyan változik (vö. pl. Balázs 1993; Russell et al. 1995). Jelen kutatásunkat azonban nem ugyanazon személyek különböző időpontokban rögzített felvételeivel, hanem különböző személyek hangmintáival végeztük, így az alaphang mellett minden más paraméter is különbözött. Ez lehetett az oka annak, hogy jelen kísérletünkben a kísérleti személyek az alaphangot nem használhatták akusztikai kulcsként.

Az azonban, hogy az artikulációs tempó 43,3%-ban magyarázta a becsült életkort, azt is jelenti, hogy közel 57%-ban itt nem vizsgált tényezők érvényesültek. További akusztikai paraméterek mellett akár az elmondott szöveg jellege, stílusa, vagy akár egy-egy, fiatalok vagy idősek nyelvhasználatára jellemző szó is befolyásolhatta. Bár a szövegek kiválasztásakor törekedtünk arra, hogy azok tartalmukban ne tartalmazzanak semmilyen kulcsot, ami befolyásolhatja a lehallgatók döntéseit, nem akusztikai – például szociolingvisztikai – tényezők hatása nem zárható ki teljes bizonyossággal. A megmagyarázatlan 57%-nyi varianciához magyarázó tényezőinek felderítése tehát további feladat.

Kutatásunkban korlátozó tényező volt, hogy akusztikai ingerként csak férfiak hangját használtuk, továbbá a lehallgatók egyetemisták voltak, és nemi megoszlásuk aránytalansága miatt nem különítettük el a férfiak és a nők válaszait. További kutatások során női beszélők és idősebb hallgatók alkalmazásával, továbbá a férfi és női válaszadók becsléseinek különválasztásával árnyaltabb képet kaphatunk az életkorbecslés mechanizmusairól. Különösen ígéretesnek látjuk a beszélőknél és a lehallgatóknál megfigyelhető egyéni különbségek további, még részletesebb vizsgálatát is.

Irodalom

- Amir, Ofer – Enger, Merav – Shabtai, Esther – Amir, Noam 2012. Identification of children's gender and age by listeners. *Journal of Voice* 26/3. 313–321.
- Assmann, Peter – Barreda, Santiago – Nearey, Terrance 2013. Perception of speaker age in children's voices. In: *Proceedings of Meetings on Acoustics*. 2–7 June 2013. Montreal. http://www.ica2013montreal.org/Proceedings/mss/060059_1.pdf (A letöltés ideje: 2016. november 29.)
- Balázs Boglárka 1993. Az időskori hangképzés jellemzői. *Beszédkutatás '93*. 156–165.
- Boersma, Paul – Weenik, David 2014. *Praat: doing phonetics by computer*. [Computer program]. Version 5.3.82. <http://www.praat.org/> (A letöltés ideje: 2014. június 12.)
- Bóna Judit 2013. *A spontán beszéd sajátosságai az időskorban*. ELTE Eötvös Kiadó, Budapest.
- Braun, Angelika – Cerrato, Loredana 1999. Estimating speaker age across languages. In: *Proceedings of the 14th International Congress of Phonetic Sciences*. University of California, Berkeley, San Francisco. 1369–1372. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/papers/p14_1369.pdf (A letöltés ideje: 2016. november 26.)
- Cerrato, Loredana – Falcone, Mauro – Paolini, Andrea 2000. Subjective age estimation of telephonic voices. *Speech Communication* 31. 107–112.

- Eppley, Brock D. – Mueller, Peter B. 2001. Chronological age judgments of elderly speakers: The effects of listeners' age. *Contemporary Issues in Communication Science and Disorders* 28. 5–8.
- Gocsál Ákos 1998. Életkorbecslés a beszélő hangja alapján. *Beszédkutatás '98*. 122–134.
- Gocsál Ákos 2012. A beszéd alaphangmagasságának mérése spontán beszédből és izolált [ə]-hangokból, férfi és női beszélőknél. In Markó Alexandra (szerk.): *Beszédtudomány*. MTA Nyelvtudományi Intézete – ELTE Bölcsészettudományi Kar, Budapest. 316–331.
- Gocsál Ákos 2016. Szintetizált [ɔ], [e] és [a:] hangok által keltett benyomások és életkorbecslések vizsgálata. *Beszédkutatás 2016*. 103–138.
- Gósy Mária 2001. A testalkat és az életkor becslése a beszéd alapján. *Magyar Nyelvőr* 125. 137–148.
- Gósy Mária – Gyarmathy Dorottya – Horváth Viktória – Gráczki 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.
- Harnsberger, James D. – Shrivastav, Rahul – Brown Jr., W.S. – Rothman, Howard – Hollien, Harry 2008. Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice* 22/1. 58–69.
- Hughes, Susan M. – Rhodes, Bradley C. 2010. Making age assessments based on voice: the impact of the reproductive viability of the speaker. *Journal of Social, Evolutionary, and Cultural Psychology* 4/4. 290–304.
- Hummert, Mary Lee – Mazloff, Debra – Henry, Clark 1999. Vocal characteristics of older adults and stereotyping. *Journal of Nonverbal Behavior* 23/2. 111–132.
- Huntley, Ruth – Hollien, Harry – Shipp, Thomas 1987. Influences of listener characteristics on perceived age estimations. *Journal of Voice* 1/1. 49–52.
- Krepsz Valéria – Gósy Mária 2016. A hangzásidő és a megakadásjelenségek hatása az életkorbecslésre. In Balázs Géza – Veszelszki Ágnes (szerk.): *Generációk nyelve*. ELTE BTK Mai Magyar Nyelvi Tanszék, Inter Nonprofit Kft. – MSZT, Budapest. 49–62.
- Markó Alexandra 2013. *Az irreguláris zöngé funkciói a magyar nyelvben*. ELTE Eötvös Kiadó, Budapest.
- Reubold, Ulrich – Harrington, Jonathan – Kleber, Felicitas 2010. Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52. 638–651.
- Russell, Alison – Penny Lynda – Pemberton, Cecilia 1995. Speaking fundamental frequency changes over time in women: a longitudinal study. *Journal of Speech and Hearing Research* 38/1. 101–110.
- Skoog Waller, Sara – Eriksson, Márten – Sörquist, Patrik 2015. Can you hear my age? Influences of speech rate and speech spontaneity on estimation of speaker age. *Frontiers in Psychology* 6. 978.
- Stölten, Katrin – Engstrand, Olle 2002. Effects of sex and age in the Arjeplog dialect: a listening test and measurements of preaspiration and VOT. In: *Proceedings of Fonetik* 44/1. 029–032. http://www.ling.gu.se/~anders/SWEDIA/papers/Stoelten_%26_Engstrand_Fonetik2002.pdf (A letöltés ideje: 2016. november 12.)

- Stölten, Katrin – Engstrand, Olle 2003. Effects of perceived age on perceived dialect strength: A listening test using manipulations of speaking rate and F0. *PHONUM* 9. 29–32. http://www.ling.gu.se/~anders/SWEDIA/papers/Stoelten_&_Engstrand_Fonetik_2003.pdf (A letöltés ideje: 2016. november 12.)
- Tatár Zoltán 2013. Beszélőprofil-alkotás lehetőségei a kriminalisztikai fonetikában. *Alkalmazott Nyelvtudomány* XIII/1–2. 121–130.
- Tóth Andrea 2014. Gyermekek nemének és életkorának meghatározása a beszédük alapján. *Beszédkutató* 2014. 98–111.
- Traunmüller, Hardmut 1994. Conventional, biological and environmental factors in speech communication: a modulation theory. *Phonetica* 51. 170–183.
- Winkler, Ralf – Brückl, Markus – Sendlmeier, Walter F. 2003. The aging voice: an acoustic, electroglottographic and perceptive analysis of male and female voices In: *Proceedings of the 15th International Congress of Phonetic Sciences*. 3–9 August 2003. Barcelona. 2869–2872.
- Winkler, Ralf – Sendlmeier, Walter F. 2006. EGG open quotient in aging voices – Changes with increasing chronological age and its perception. *Logopedics Phoniatrics Vocology* 31/2. 51–56.
- Winkler, Ralf 2007. Influences of pitch and speech rate on the perception of age from voice. In: *Proceedings of the XVI International Congress of Phonetic Sciences*. 6–10 August 2007. Saarbrücken. 1849–1852.

A jelen tudományos közleményt a szerző a Pécsi Tudományegyetem alapításának 650. évfordulója emlékének szenteli.

The role of articulation rate and mean f_0 in speaker age estimation

The purpose of this study was to determine the degree of correlation between mean estimated age and calendar age of Hungarian speaking men, based on spontaneous speech, and to find acoustic correlates of age estimations. 49 Hungarian students were involved as listeners. In line with previous research, correlation at $r = 0.808$ was found between calendar age and mean perceived age. A linear regression model proved the significant role of articulation rate in age estimations but f_0 was not proved to be significant a predictor in the model, neither were the other parameters, such as f_0 range, its skewness and kurtosis, and the percentage of irregular phonation.