

Beszéd alapján történő személyazonosítás új kihívásai a kriminalisztikában

FEJES Attila¹

A beszéd alapján történő személyazonosítás módszertana jelentős mértékű változáson esett át az elmúlt több mint egy évtizedben a hangbiometria technológiájának megjelenésével. A biometrikus beszélőazonosítás megjelenése fontos mérföldkő volt a szakterület életében, egyben új kihívások sokasága társult hozzá. A hangbiometriai rendszerek számadatokat szolgáltatnak, amelyeket különböző szempontok szerint értékelni kell, továbbá biztosítani kell az egyes rendszerek egységes sztenderd szerinti összehasonlítását, validálását.

Kulcsszavak: személyazonosítás, beszélőazonosítás, hangbiometria, valószínűségi arány, valószínűségi skála, minőségbiztosítás

A kriminalisztikai célú beszélőazonosítás – különösen a klasszikus azonosítási módszerekhez képest – viszonylag rövid múltra tekint vissza Magyarországon. Alkalmazása számos elméleti és gyakorlati kérdést felvet,² az általános (nem magyar nyelvspecifikus) elmélete megannyi kutatásra ad lehetőséget és követel meg egyben az anyanyelvünk vonatkozásában, hiszen nem létezik univerzális, nyelvfüggetlen komplex azonosítási módszer (percepció-akusztikai-fonetikai-biometrikus), minden nyelvnek egyéni jellemzői vannak, amelyek befolyásolják a végrehajtható metódust.

A beszélőazonosítást vizsgálva felmerül az első kérdés: személy- vagy hangazonosításról beszélhetünk? A kérdés egyszerűen eldönthető, hiszen egy tárgy csak önmagával lehet azonos,³ így ez utóbbi lehetőség kiesik. Azt is mondhatjuk, hogy az emberi beszéd(hang) nagyfokú variabilitása egyértelműen meghatározza az alaptételt: egy hangot nem lehetséges pontosan ugyanúgy kiejteni eltérő időpontokban – nem beszélve egy komplex beszédprodukcióról. A személyazonosítás tűnik a helytállóbb meghatározásnak, lévén, hogy a beszéde alapján – megfelelő körülmények esetén – a személy azonosítható, így meghatározható, mi annak a valószínűsége, hogy az ismeretlen személytől származó beszéd az ismert személytől származik (ha a leggyakoribb példát vesszük

¹ FEJES Attila nb. őrnagy, hangtechnikai szakértő, Nemzetbiztonsági Szakszolgálat Szakértői Intézet, PhD-hallgató, NKE Rendészettudományi Doktori Iskola
Attila FEJES national security major, audio expert, Special Service for National Security Institute of Expert Services, PhD student, National University of Public Services, Doctoral School of Law Enforcement
orcid.org/0000-0003-4139-5718, fejes.attila@uni-nke.hu

² Nogel (2017a) 110.

³ Katona (2004) 71.

alapul). Itt kell meghatározni, hogy a hangtechnikai szakértő a személyazonosítás elvégzése során nemcsak a hangot, mint akusztikai rezgést vizsgálja, hanem analizálja a beszédet, annak tartalmát, jellegzetességeit is, a kiejtett hang akusztikai tulajdonságaitól függetlenül.

Összességében meghatározható, hogy a hangazonosítás kifejezés nem fedti le pontosan a mögötte álló módszertant és célt, helytállóbb, ha beszéd alapján történő személyazonosítást, röviden beszélőazonosítás kifejezést alkalmazunk a módszertan megnevezése során. Egy részletekbe menő megközelítés, ha arról beszélünk, hogy létezik azonosítás,⁴ felismerés és hitelesítés.⁵

Itt már módszertani kérdésről van szó.⁶ Míg az azonosítás általában azt fejezi ki, hogy beszéde alapján egy konkrét személyt (be)azonosítunk, a felismerés a kriminalisztikán belül pontosabb kifejezésnek tűnhet azokban az esetekben, amikor több személy közül kell az azonosítási eljárás során kiválasztani, hogy az inkriminált beszéd melyik személytől származik. Ezt nevezzük 1:N módszertannak, míg a hitelesítés annak „igazolása”, hogy a beszéd egy konkrét személytől származik. A nemzetközi szakirodalomban ennek megfelelően alkalmazzák a *speaker identification*, *speaker verification* és a *speaker recognition* fogalmakat. Manapság egyre inkább a *Forensic Speaker Recognition* (FSR) és a *Forensic Automatic Speaker Recognition* (FASR) fogalmakat használják.⁷

A beszéd

A beszéd akusztikai rezgések sorozata, amelyben a prozódia biztosítja az értelmi, érzelmi elkülönítést, jelentésbeli tagolást. A beszéd szegmentális és szupraszegmentális (prozódia) szerkezetre osztható, amelyek tulajdonságai különböző típusú vizsgálati módszereket követelnek meg. A szegmentális szerkezetre jellemzőbb az egyező hangok, hangkapcsolatok paramétereinek összehasonlítása a két beszédprodukciónál, míg a szupraszegmentumok analízise sokszor eltérő szövegtartalmú beszéden valósul meg. Ezenkívül jellemző még, hogy a prozódia vizsgálható perцепciós eszközökkel is, igaz, itt a megállapítások objektivitása (értve ezalatt az értékítéletek nélkülözését), a vizsgálatok más szakértő által történő megismételhetősége és ugyanazon eredmény szavatolása nem mondható el minden esetben, hiszen az észleléses vizsgálatok mindig magukban foglalják a szakértő egyéni tapasztalatát, elgondolásait, benyomásait.

A beszéd⁸ a hangképzés egyediségének, a beszédképzés sajátosságainak és a nyelvhasználat egyéni jellegzetességeinek következtében alkalmas személyazonosítás elvégzésére. Habár soha nem vizsgálta senki, de mégis elfogadott állítás, hogy nincs két ember a Földön, akiknek beszédhangja, nyelvhasználata, beszédprodukciónja teljesen megegyező lenne. Ezt azért lehetséges kijelenteni, mert szintén nincs két, fizikai fel-

⁴ A kriminalisztikai azonosítás elméletéről bővebben: Czebe (2017) 65–73.

⁵ Kovács–Nagy–Nogel (2015) 56.

⁶ Nogel (2017b) 116.

⁷ Drygajlo (2017) 221.

⁸ Gósy (2004) 11.

építésében, a beszédképzésben részt vevő szervek biológiai felépítésében teljesen megegyező ember. Így azt mondhatjuk, hogy beszéde alapján a személy azonosítható, az eljárásban az azonosság valószínűsége meghatározható. Ez az állítás a – remélhetőleg minél távolabbi – jövőben látszólag megcáfolható, amennyiben a humán klónozás megtörténik, de addig is egyrészt azt mondhatjuk, hogy még az azonos szociokulturális közegben felnőtt egypetűjű ikreknek is jól mérhetően különbözik a beszéde.⁹ Másrészt – csak elméleti síkon vizsgálódva – ha feltételezzük is a 100%-ban egyező anatómiai szerkezetet, a beszédet meghatározó további sajátosságok nem lehetnek ugyanazok: például általános műveltség, érdeklődési kör, szókincs, EQ és IQ stb., tehát csupa olyan dolog, amit a beszélő szociokulturális közege és saját személyisége, képességei határoznak meg.

A hangtechnikai szakértő számára a beszéd egyrészt akusztikai jelek sorozatát jelenti, másrészt szöveget, amely a nyelvhasználat megjelenési formájaként is jellemezhető. Mivel a beszéd az emberek közötti kommunikáció egyik alapvető eszköze, ezért vizsgálhatóak a beszélő személy társas kommunikációban felvett jellemzői is. Ilyenek lehetnek: visszacsatolások megléte, egyedi tulajdonságai (például ismétlődő töltelék-szavak formájában), ezek átlagos időtartama, a másik félhez viszonyított verbális helyzet felvétele (ha egyértelműen megfigyelhető, például: alá-, fölé-, mellérendelt szituációk), beszédhang intenzitása kérdés vagy válasz elhangzásakor stb.

Ugyanakkor elmondható, hogy a beszéd alapján történő azonosítás megbízható módon történő elvégzéséhez mindkét elem (a hang és a szöveg) vizsgálata szükséges a legmegbízhatóbb eredmény elérése érdekében. Itt felvetődik az első kérdés: milyen módon azonosítható a szakértő számára ismeretlen nyelven beszélő személy? Erre tudományosan megalapozott választ¹⁰ csak a hangbiometria ad, hiszen ebben az esetben a szöveg nem vizsgálható, az akusztikai mérések döntő többsége nem végezhető el (amelyek egyező hangok, hangkapcsolatok összehasonlításán alapul), a kisebb rész csak csekély támpontot jelent a szakértő számára, a percepció elemzés pedig önmagában nem kellő megalapozottságot lehetővé tévő módszertan. A biometrikus beszélőazonosítás létrejötté és egyre fejlettebb algoritmusokkal, módszerekkel (GMM-UBM, i-Vektor, neurális hálózatok) történő alkalmazása teszi lehetővé, hogy az azonosítás olyan körülmények között is elvégezhető legyen, ahol a nyelv, a rögzítési csatorna nem meghatározó elem.

Akusztikai-fonetikai (félautomatikus) beszélőazonosítás

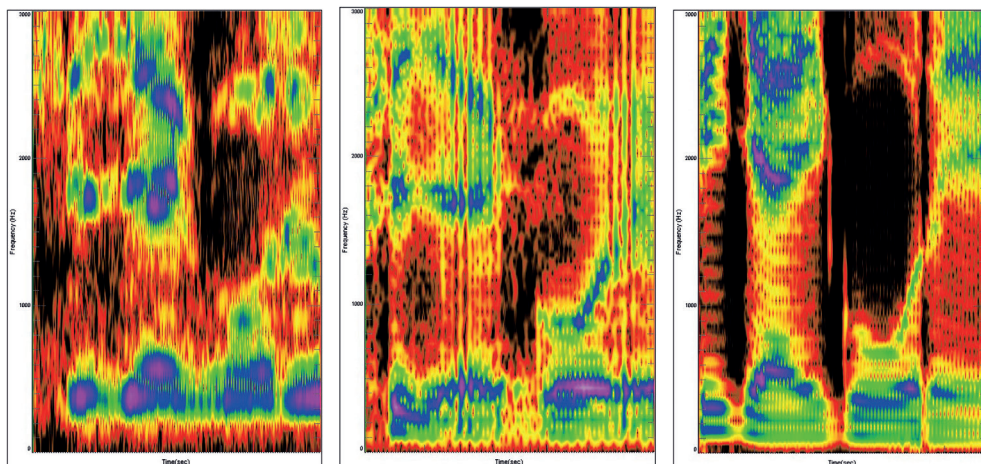
Az eljárás percepció elemzésből és akusztikai mérések sorozatából áll. Az első vizsgálati módszer során észlelés útján tárja fel a szakértő a vizsgált személy beszédének egyedi jellegzetességeit. Ezek lehetnek objektív módon meghatározható elemek, például tájnyelvi dialektus jelenléte, ismétlődő szavak, szókapcsolatok, kiejtésbeli hibák,

⁹ Künzel (2010) 251–277.

¹⁰ Kovács–Nogel (2017) 257.

valamint az emberi fül és az agy hallásért és dekódolásért felelős részeinek működése következtében előállt megállapítások. Ehhez kapcsolódik a neurális spektrogram¹¹ fogalma, amely a beszéd akusztikai jellemzőinek az agy számára kódolt megfelelője és amelynek „előhívásával” a beszédkommunikációban részt vevő fél beazonosítja a beszélőt – amennyiben korábban azt ismerte vagy a hangját hallotta. Ezt a szakértő oly módon „használja ki”, hogy az ismeretlen beszédet többször, egymás után lejátszva, tulajdonképpen „rögzíti” a neurális spektrogramot, így a később hallott összehasonlító hanganyag alapján, ideális esetben képes megállapítást tenni a beszélők azonosságára vonatkozóan.

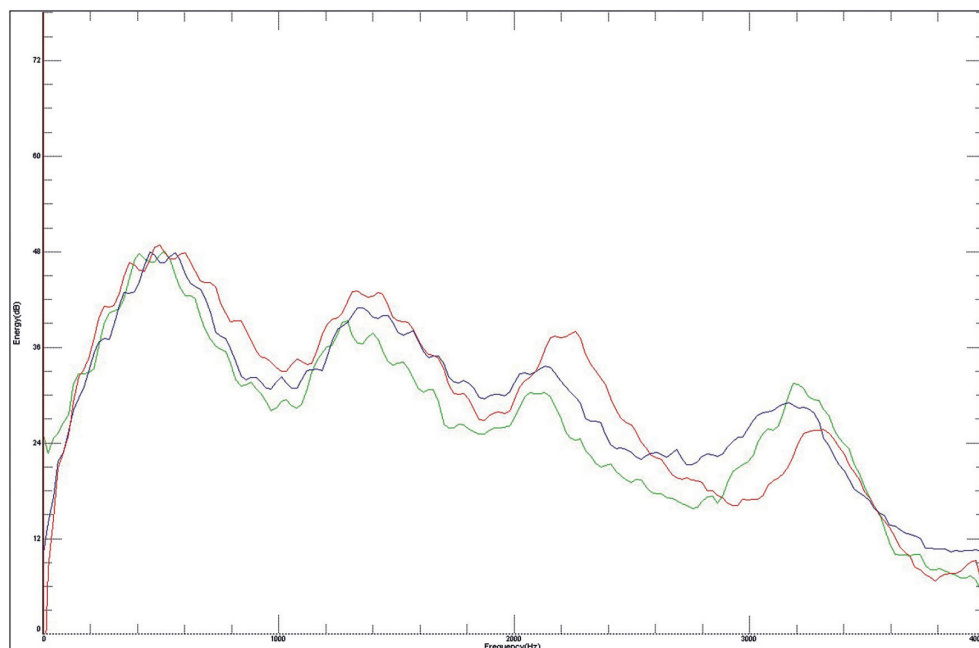
Az akusztikai jellemzők mérése talán még nagyobb dilemmát vet fel, amelynek illusztrálására nézzük meg a következő három képet.



1. ábra: Ismeretlen beszélő (1.); gyanúsított (2.); tesztszemély (3.) spektrogramja. Forrás: a szerző

Mindhárom képen ugyanazon hangsornak („góre”) a spektrografikus (SPG) képe látható. Az ismeretlen beszélő és a gyanúsított megegyezik, míg a tesztszemély tőle különbözik. Jól megfigyelhető az első két ábra hasonlósága, míg a 3. kép tőlük való relatív eltérése. A sötétkéssel jelzett formánsok (a zöngé rezonátorüregben felerősödött felharmonikusai) frekvenciaértékei és alakzata az első két kép esetében sokkal hasonlóbbak egymáshoz, mint a 3. kép esetében az előzőekhez viszonyítva. A hangtechnikai szakértő számára az SPG-kép az igazi „térkép” a beszédhang vonatkozásában, hiszen az a hangot három dimenzióban ábrázolja – idő, frekvencia, intenzitás – míg a rezgés-kép csak az amplitúdó (intenzitás) változását mutatja meg az idő függvényében. Vagy nézzük a következő ábrát, amelyen egy konkrét beszélő különböző telefonhívásaiból kiemelt „ö” hangjainak FFT-leképezése (*Fast Fourier Transformation* – a kiejtett hang intenzitásának változása a frekvencia függvényében) látható.

¹¹ Gósy (2004) 275.



2. ábra: FFT-görbék. Forrás: a szerző

Ami elsőre feltűnik, az a görbék szemmel látható hasonlósága. Ugyanakkor – mindkét példát vizsgálva – felmerül a kérdés, hogyan tudnánk objektív módon meghatározni a hasonlóság vagy a különbözőség szempontjait? A helyes válasz: sehogya. Ugyanis a beszéd variabilitása következtében a vizsgált alakzatok sohasem lesznek egybevágóak, így viszont az összevetéskor támaszkodni kell a szakértő értékítéletére, tapasztalatára is. Ennek megfelelően a módszer nem mondható objektívnek, hiszen, az előbb említettekén kívül, a vizsgálatok megismétlése során nem garantálható ugyanaz az eredmény minden alkalommal, így a megismételhetőség követelménye száz százalékosan nem teljesíthető.

A biometrikus (automatikus) beszélőazonosítás

A módszertan alapja, hogy számítógépes szoftverrel a beszédhangból nagyszámú jellemző nyerhető ki (*feature extraction*), amelyek matematikai-statisztikai módszerek alkalmazásával olyan rendszerré formálhatóak, amellyel meghatározható, hogy az ismeretlen beszédhang inkább az ismert személyhez tartozik vagy valaki máséhoz. Ez a klasszikus összehasonlítási példa, amikor a hatóság birtokába egy ismeretlen személytől származó hangfelvétel kerül, majd látókörbe vonnak egy ismert személyt (gyanúsított, tanú), akitől hangmintát vesznek és a szakértőnek azt kell megállapítania, hogy az ismeretlen beszédhang származhat-e az ismert személytől. Természetesen

ez csak egyetlen példa a gyakorlatból, a szakértő sok esetben ismeretlen beszédproduktókat hasonlít össze az egyező beszélő(k) megállapítása céljából, vagy pedig az ismeretlen hanganyagot több ismert beszélő hangmintájával veti össze.

A hangbiometria a Bayes-tételből ismert Likelihood Ratio (LR) aránypárt alkalmazza annak érdekében, hogy a feltételes valószínűség tétele a beszélőazonosításon belül értelmet nyerhessen. Ez oly módon történik, hogy meghatározunk axiómákat. Egy per során – amennyiben feltételezzük, hogy már az igazságszolgáltatás keretein belül kell döntenie egy terhelt bűnösségéről – mindig van hipotézis és ellenhipotézis. Az elsőt a vád, a másodikat a védelem képviseli. Ebben a megközelítésben a hangszakértő nem tudja megmondani (nem mondhatja meg), hogy mekkora a valószínűsége a priori és/vagy a posteriori valószínűségnek. A szakértő nem ismerheti az a priori állítás alátámasztását jelentő adatokat, csakúgy, mint a posteriori ismeretek sem lehetnek számára nyilvánvalóak. A szakértő csakis két dolgot láthat (pontosabban hallhat): a két összehasonlítandó beszédproduktót. Minden további adat, információ hat az értékítéletére, a módszertan azon elemeire, amelyek az emberi közreműködés következtében szubjektív elemeket hordoznak magukban. Csak a bizonyíték bizonyítási erejét (nem pedig a hipotézis valószínűségét) ismertetheti a beszéd elemzésén keresztül.

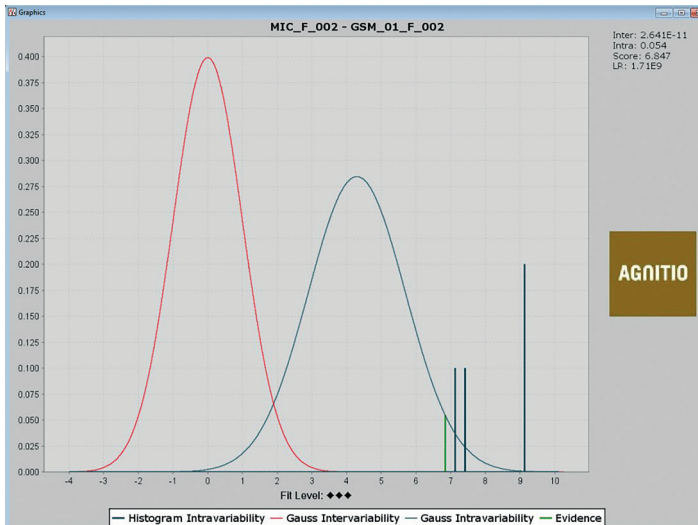
A biometrikus azonosítórendszerek nélkülözik a szakértői közreműködést a hasonlóság valószínűségének megállapítása során – leszámítva azt, hogy a hangszakértő készíti el a bemenetre továbbított hangmintákat. Azonban ezen módszertan során is szerepet kap a szakértő szaktudása. A következő ábrán egy biometrikus beszélőazonosító rendszer eredménytáblájának részletét láthatjuk. A teszt során 10 személy különböző csatornán (mobiltelefonos beszélgetés, stúdiómikrofonnal felvett spontán interjú) rögzített hangmintája összehasonlításának eredményei láthatóak. Az első oszlopban az ismeretlen beszélőként jelzett felvételek láthatóak, a „Model Order 1–5” oszlopokban ugyanezen beszélők stúdiómikrofonos hangmintái vannak jelölve. Az első score oszlopban láthatók az azonos beszélők összehasonlításának LR-eredményei, a többi score oszlopban a különböző beszélőké. Így tehát összesen 10 azonos és 90 különböző beszélő eredményeit kapjuk meg a teljes mátrix formájában.

A 2. oszlop „Score” eredményei jól szemléltetik a biometrikus beszélőazonosítás problematikáját, ugyanis a számértékek jelentős nagyságrendi különbséget vesznek fel annak ellenére, hogy biztosan tudható, hogy ebben az oszlopban azonos beszélők hangmintájának összehasonlítási eredményei láthatóak.

Model Order 1				Model Order 2			Model Order 3			Model Order 4			Model Order 5		
Name	Score	Fit Level	Name	Score	Fit Level	Name	Score	Fit Level	Name	Score	Fit Level	Name	Score	Fit Level	
GSM_01_F_001	MIC_F_001	1.9092E8 Error	MIC_F_009	0.49486	Ok	MIC_F_007	0.12508	Ok	MIC_F_006	0.09285	Ok	MIC_F_003	0.07102	Ok	
GSM_01_F_002	MIC_F_002	1.7098E9	MIC_F_006	0.08656	Ok	MIC_F_004	0.03925	Ok	MIC_F_008	0.01246	Ok	MIC_F_010	0.01236	Ok	
GSM_01_F_003	MIC_F_003	3.28981E7	MIC_F_006	2.33732	Ok	MIC_F_007	0.59647	Ok	MIC_F_010	0.48832	Ok	MIC_F_005	0.00685	Ok	
GSM_01_F_004	MIC_F_004	2.32173E4	MIC_F_005	0.03395	Ok	MIC_F_010	0.01922	Ok	MIC_F_002	0.0101	Ok	MIC_F_007	0.00777	Ok	
GSM_01_F_005	MIC_F_005	5006.33783	MIC_F_009	0.5296	Ok	MIC_F_007	0.15626	Ok	MIC_F_010	0.09037	Ok	MIC_F_002	0.04079	Ok	
GSM_01_F_006	MIC_F_006	5.73095E4	MIC_F_007	1.43699	Ok	MIC_F_002	0.09752	Ok	MIC_F_003	0.05146	Ok	MIC_F_004	0.04969	Ok	
GSM_01_F_007	MIC_F_007	3.61656E6	MIC_F_004	0.07732	Ok	MIC_F_005	0.02483	Ok	MIC_F_002	0.02243	Ok	MIC_F_006	0.01102	Ok	
GSM_01_F_008	MIC_F_008	2.7147E9	MIC_F_002	0.21021	Ok	MIC_F_006	0.10845	Ok	MIC_F_004	0.02542	Ok	MIC_F_007	0.01433	Ok	
GSM_01_F_009	MIC_F_009	2.50578E8	MIC_F_005	0.11148	Ok	MIC_F_006	0.0506	Ok	MIC_F_007	0.04722	Ok	MIC_F_010	0.00824	Ok	
GSM_01_F_010	MIC_F_010	2.08215E6	MIC_F_005	4.72374	Ok	MIC_F_007	3.14648	Ok	MIC_F_006	0.07247	Ok	MIC_F_003	0.00686	Ok	

3. ábra: Automatikus beszélőazonosító rendszer eredmény-táblája (LR) (részlet). Forrás: a szerző

A második sor „Score” eredménye azt mutatja, hogy $17\,098 \times 10^9$ -szer (1 709 800 000) valószínűbb, hogy a GSM_01_F_002 felvételen hallható beszéd a MIC_F_002 felvételen hallható személytől származik, mint valaki mástól. Grafikusan ez az eredmény a 4. ábrán látható módon ábrázolható.



4. ábra: Személyazonosság ábrázolás grafikonon. Forrás: a szerző

Itt a piros Gauss-görbe a referencia populációs adatbázis jelöli, kék színű a MIC_F_002 felvétel alapján képzett haranggörbe, zöld színű a GSM_01_F_002 hangfelvétel alapján

számított egyenes. A zöld színű egyenes, a Gauss-görbékhez képest elfoglalt helyzete, a jobb felső sarokban látható intravariabilitás és intervariabilitás értékek fejezik ki, hogy mennyi a valószínűsége annak, hogy a két összehasonlított felvételen hallható beszédhang ugyanazon személytől származik, vagy sem.

Visszatérve a 3. ábrához, látható, hogy az első eredményoszlop, a 005 jelzésű beszélő kivételével nagy számértékeket mutat (a maximum a 10^{10} érték, nem utolsósorban azért, mert hozzávetőlegesen 7 Mrd ember él a Földön), különösen a további oszlopok adataihoz képest. A 005 jelzésű személy esetében az LR azonban „csak” ~5006. Egyes esetekben ez lecsökkenhet pár százas nagyságrendre. Ugyanakkor előfordul, hogy különböző beszélők esetén is százas nagyságrendű eredmény születik. Itt adódik a kérdés, hol a hibahatár? A válasz a jól ismert: sehol, ugyanis nincs kemény küszöb, csak gradiens.

A biometrikus beszélőazonosító rendszerek hibaaránya háromféle módon fejezhető ki. Téves elfogadás (*false accept rate* – FAR), téves elutasítás (*false reject rate* – FRR), egyenlő hibaarány (*equal error rate* – EER), amely fogalomrendszerben az FAR különböző személyek téves beazonosítását, az FRR azonos személyek tévesen különbözőnek történő megállapítását mutatja meg. Az EER az a pont, ahol az FAR és az FRR előfordulásának száma megegyezik. Az, hogy mennyi a tényleges Threshold értéke, úgy határozható meg, hogy vesszük azt a pontot, ahol az FRR és az FAR megegyezik. Ez egy hatékony módja a rendszer teljesítőképességnek, robusztusságának meghatározására, azonban nagyon sok tesztet kell végrehajtani annak érdekében, hogy a nyelvspecifikus eredmények létrejöhessenek. A fejlesztők 5% körül szokták az rendszereik EER-értékeit meghatározni, ez a gyakorlatban, „éles” hangfelvételekkel inkább 10% körülire tehető.

Összefoglalás

A biometrikus beszélőazonosítás módszertana a szakértő szubjektív ismerete, jellemzői nélkül állapítja meg, hogy mekkora a valószínűsége annak, hogy két vizsgált beszédhang azonos vagy különböző személytől származik. Ennek óriási előnye, hogy nincs torzító effektus a szakértői megállapításban.¹² Hatalmas hátrány, hogy bitek sorozatára bizzuk a büntetőeljárásban részt vevő személy sorsát. Kihívás, hogy mi a legjobb módszer az azonosító rendszer megbízhatóságának meghatározására¹³ (magyar nyelvű beszélők esetében), vagy mi a követendő eljárás abban az esetben, ha a szoftver téves elfogadás/téves elutasítás intervallumokhoz közel álló eredményt szolgáltat.

További probléma a kategorikus szakvélemény kérdésköre.¹⁴ A 2000-es évek elején az ENFSI (*European Network of Forensic Science Institutes*) javasolta a verbális skála alkalmazását a szakvéleményben és ezen belül a kategorikus megállapítások alkalmazását.¹⁵ Több mint tíz év elteltével (és számos kutatás tapasztalatai alapján) ugyanezen

¹² Kovács–Nogel (2016) 598.

¹³ Pádár–Nogel–Kovács (2015) 412.

¹⁴ Kovács (2010) 310.

¹⁵ Katona (2004) 75.

szervezet ajánlásban fogalmazta meg, hogy a kategorikus véleményalkotás a beszélő-azonosítás terén kerülendő, helyette a valószínűsítő verbális skála, vagy a numerikus (LR, százalékpont stb.) skála alkalmazása elfogadható.¹⁶ Léteznek továbbá mechanikus LR értékelési skálák,¹⁷ azonban ezek univerzálisak, nincs bennük a nyelvre vonatkozó specifikusság, amelynek figyelembevételéhez szükség van magyar nyelven beszélő személyek hanganyagán történő, széles körű kutatási tevékenységre, végső soron a hangszakértői vélemény megbízhatóságának, objektivitásának emelésére a kriminalisztikai hiba elkerülése érdekében.¹⁸ Látható, hogy csak a fenti néhány gondolat alapján, számos kérdés merül fel, amelyek csak nagyon jól átgondolt, részletes kutatómunka alapján válaszolhatóak meg, így lehetővé válik tudományos igényű szakmai vita lefolytatása és a szakismeretek, az alkalmazott kutatási eredmények széles körben történő megosztása.

IRODALOMJEGYZÉK

- Czebe András (2017): A kriminalisztikai azonosítás törvényszerűségei. *Magyar Bűnüldöző*, 1-2. sz. 65–73.
- Drygajlo, Andrzej – Haraksim Rudolf (2017): Biometric Evidence in Forensic Automatic Speaker Recognition. In Massimo, Tistarelli – Christophe, Champod (2017): *Handbook of Biometrics for Forensic Science*. Cham, Springer International Publishing AG.
- Drygajlo, Andrzej – Jessen, Michael – Gfroerer, Stefan – Wagner, Isolde – Vermeulen, Jos – Niemi, Tuija (2015): *Methodological Guidelines for Best Practice in Forensic Semiautomatic and Automatic Speaker Recognition*. Forrás: http://enfsi.eu/wp-content/uploads/2016/09/guidelines_fasr_and_fsasr_0.pdf (2018. 01. 18.)
- Fenyvesi Csaba (2015): A kriminalisztika veszélyei. In Gaál Gyula – Hautzinger Zoltán: *Modernkori veszélyek rendészeti aspektusai. Pécsi Határőr Tudományos Közlemények, XVI.*
- Finszter Géza (2003): *A rendészet elmélete*. Budapest, KJK-KERSZÖV Jogi és Üzleti Kiadó Kft.
- Gósy Mária (2014): *Fonetika a beszéd tudománya*. Budapest, Osiris Kiadó.
- Katona Géza (2004): Kriminalisztikai elméletek. In Bócz Endre szerk.: *Kriminalisztika*. Budapest, BM Kiadó.
- Kovács Gábor (2010): Igazságügyi szakértők átértékelődő szerepe a büntetőeljárásban. In Bihari Mihály – Patyi András szerk.: *Ünnepi kötet Szalay Gyula tiszteletére, 65. születésnapjára*. Győr, Universitas Győr. 310–315.
- Kovács Gábor – Nagy Klára – Nogel Mónika (2015): Accreditation of forensic laboratories as a part of the „European Forensic Science 2020“ concept. In Keller Éva szerk.: *AAP 2015 Programme and Abstract Book: 24th International Meeting on Forensic Medicine Alpe-Adria-Pannonia*. Budapest, Expert Quality Congress.
- Kovács Gábor – Nogel Mónika (2016) Forensic experts and expert testimony in criminal procedure. *IAML Intersocietal Symposium, P5 Medicine & Justice: Abstract Book*. Velence.
- Kovács Gábor – Nogel Mónika (2017): Comparative Analysis of the Legal Regulation of Forensic Experts in Europe. *Forensic Science International*, No. 277. 81–82.
- Künzel, Hermann J. (2010): Automatic speaker recognition of identical twins. *The International Journal of Speech, Language and the Law*, Vol. 17. No. 2. 251–277.

¹⁶ Drygajlo et al. (2015) 46.

¹⁷ Pápay (2007) 111.

¹⁸ Fenyvesi (2015) 125.

Nogel Mónika (2017a): A szakértőjelöltekre vonatkozó szabályozás minőségbiztosítási szempontú vizsgálata. *Belügyi Szemle*, 65. évf. 9. sz. 105–114.

Nogel Mónika: (2017b): A szakértői minőségbiztosítási rendszer kialakításának előkérdései és alapfogalmi. *Jog – Állam – Politika*, 9. évf. 1. sz. 115–133.

Pádár Zsolt – Nogel Mónika – Kovács Gábor (2015): Accreditation of forensic laboratories as a part of „European Forensic Science 2020” concept in countries of Visegrad Group. *Forensic Science International Genetics Supplement Series*, Vol. 5. 412–413.

Pápay Kinga (2007): Valószínűségi skálák az igazságügyi nyelvészetben. In Váradi Tamás szerk.: *I. Alkalmazott Nyelvészeti Doktorandusz Konferencia*. Budapest, MTA Nyelvtudományi Intézet.

ABSTRACT

New Challenges of Speaker Recognition in Forensic Science

FEJES Attila

The method of speaker recognition has changed significantly in the past decade due to the appearance of voice biometrics. The emergence of voice biometrics was an important landmark in forensic science, at the same time a number of challenges were associated with it. Voice biometrics systems provide data which must be evaluated from different aspects in addition to the comparison and validation of certain systems that must be provided according to a uniform standard.

Keywords: *identification, speaker recognition, voice biometrics, Likelihood Ratio, probability scale, accreditation*