Доц. др Јадранка ОТАШЕВИЋ[1]
Академија за хумани развој, Београд

ДОИ: 10.5937/bezbednost2601071O
УДК: 343.982.323:81'342.1
343.983:81'342.1
Прегледни научни рад
Примљен: 5. 3. 2026. године
Датум прихватања: 13. 3. 2026. године

Форензичко-фонетска анализа гласа у криминалистичкој идентификацији говорника

Апстракт: *Форензичко-фонетска анализа гласа представља значајан поступак у криминалистичкој идентификацији говорника, нарочито у истрагама које се врло често ослањају на аудио-снимке настале у реалним, неконтролисаним условима. Циљ рада је да прикаже улогу, могућности и ограничења форензичко-фонетске анализе у идентификацији говорника, са нагласком на криминалистичком и доказном контексту. Рад је конципиран као прегледни рад заснован на анализи релевантне домаће и међународне литературе из области форензичког препознавања говорника. Разматрају се аудитивно-перцептивни, инструментално-акустички и комбиновани приступи, као и савремени евалуативни модели извештавања у оквиру односа вероватноћа (likelihood ratio – LR). Посебно се истиче да „неусклађеност услова" између спорних и референтних снимака (session mismatch) и некритичко прихватање концепта гласовног отиска (voiceprint) могу довести до погрешне процене поузданости налаза.*

Анализа указује да највећу практичну поузданост даје комбиновани приступ који интегрише акустичке, фонетске и лингвистичке показатеље, уз незаменљиву улогу експерта у интерпретацији налаза са јасним навођењем ограничења. Закључак је да форензичко-фонетски налаз не представља категоричку идентификацију, већ евалуативну процену јачине доказа у складу са европским форензичким стандардима.

[1] jadrankastevovic@yahoo.com ORCID: 0000-0001-7051-3897

Кључне речи: *форензичко препознавање говорника, криминалистичка идентификација говорника, варијабилност говора, однос вероватноћа (likelihood ratio – LR), евалуативно извештавање.*

Увод

Глас представља специфичан криминалистички траг нематеријалне природе. Његов идентификациони потенцијал произлази из комбинације физиолошких и бихевиоралних особина говорника, али и из карактеристика говора и језика (фонетских, прозодијских и лингвистичких елемената). У том смислу, глас се у форензичком поступку може посматрати као биометријско обележје, док се у криминалистичкој пракси третира као траг и потенцијално доказно средство, када је потребно идентификовати осумњичено лице или искључити лице које није учинилац дела (Оташевић и сар., 2024). Савремена дигитална ера довела је до пораста броја аудио-снимака који се појављују у криминалистичким истрагама (пресретнути разговори, тајно прибављени снимци, снимци из неконтролисаних услова). То је додатно повећало значај стручне обраде гласовних и звучних записа у кривичним предметима, јер налази могу имати непосредне процесне последице и захтевају методолошки транспарентан приступ (Carić & Širić, 2023).

Циљ овог рада је да прикаже улогу и значај форензичко-фонетске анализе гласа у криминалистичкој идентификацији говорника, са фокусом на практичним условима истрага и интерпретацији налаза у доказном контексту. Рад је конципиран као прегледни (теоријски) рад заснован на анализи релевантне домаће и међународне научне литературе из области форензичке фонетике и форензичког препознавања говорника, укључујући савремене евалуативне моделе извештаја и смернице добрих пракси у судском контексту.

Глас као криминалистички траг

Глас (говорни сигнал) може представљати криминалистички траг када је забележен у оквиру извршења или припреме кривичног дела (нпр. пресретнути позиви, претње, уцене, договарање испорука, гласовне поруке, снимци са надзора). У таквим ситуацијама говор улази у домен дигиталних доказа и постаје предмет форензичке об-

раде. Тада се пореди познати узорак (неспорни) са трагом (спорни узорак) ради процене да ли потичу од истог говорника (Kruthika et al., 2025). У криминалистичко-форензичком смислу, типична ситуација је да постоји снимак са лица места (нпр. „глас починиоца" из прислушкивања или са места догађаја) и референтни снимак (глас осумњиченог прикупљен у поступку). Форензички задатак је поређење та два материјала, уз свест да су услови настанка траговног снимка често неконтролисани, а последице закључка правне (Јовичић & Кашић, 2009). Треба имати на уму да се, за разлику од контролисаних биометријских система, форензичко препознавање говорника одвија у реалним, неконтролисаним условима. Говорни материјал који се користи у криминалистичким истрагама најчешће је јединствен, непоновљив и често оптерећен различитим сметњама као што су бука, реверберација, редуковани спектар телефонског канала, компресија аудио-записа или преклапање говора. Поред техничких ограничења, на говор могу утицати и бројни индивидуални фактори, као што су намерно маскирање гласа, болест, стрес или дејство алкохола и психоактивних супстанци. Ти снимци могу бити кратки и са присуством шума, због чега је поуздано разликовање говорника захтевно и за човека и за примену аутоматских метода (Campbell et al., 2009).

Специфичност гласа као трага огледа се и у томе што он носи и акустичке и језичко-лингвистичке информације. У пракси се разматра широк спектар обележја: изговор, језичка компетенција, дисфлуентност и акустички параметри (Sergidou et al., 2024: 490). Форензичко-лингвистички докази обухватају све типове текста – говорни, писани или потписани, који може бити коришћен у кривичној истрази или као доказ пред судом. Тиме су обухваћени телефонски позиви, анонимна писма, поруке и изјаве осумњичених (Correa, 2013: 6). У том смислу, глас као траг не представља само акустички сигнал, већ и дискурзивни и прагматички материјал чија анализа може допринети расветљавању околности кривичног дела. У форензичком поређењу се често комбинују аудитивна-експертска анализа и инструментална (рачунарска) анализа, уз идеју да се субјективност смањи, а објективност мерења повећа (Нешић и сар., 2011). Да би глас као траг имао доказну вредност у поступку, налаз се не своди на „поклапа се / не поклапа се", већ се у савременој пракси тежи извештавању кроз *likelihood* (*LR*) – однос вероватноћа доказа по две супротне хипотезе (један говорник наспрам другог говорника) (Sergidou et al., 2024). У

литератури се наводи и постојање различитих оквира односа вероватноћа (UK framework, и др.), као и потреба емпиријске валидације система и опрез у судској употреби (Kruthika et al., 2025). У том контексту, није довољно процењивати само „сличност" између узорака, већ и „типичност" у односу на релевантну популацију, јер јачина доказа зависи од тога колико је уочени образац истовремено карактеристичан за посматраног говорника и неуобичајен у популацији. Стога аутоматски системи могу помоћи у процени односа вероватноћа, али не уклањају потребу за стручним тумачењем и испитивањем претпоставки (Campbell et al., 2009). За разлику од многих физичких трагова, глас има изражене интраговорничке варијације, што отежава стабилно издвајање маркера и захтева опрез у интерпретацији (Јовичић & Кашић, 2009). Посебно су савремени изазови и манипулације релевантни у дигиталном контексту те се стога у литератури истичу пораст значаја форензичке анализе говора у дигиталним истрагама и потреба за напреднијим алатима (Kruthika et al., 2025).

У криминалистичкој обради, глас као траг најчешће служи:

1. *за усмеравање истраге и сужавање круга осумњичених* (оперативно: индиције, повезивање догађаја и снимака), (Оташевић и сар., 2024) *и за форензичко вештачење* (доказно: структурисана процена јачине доказа кроз однос вероватноћа или вербалну скалу), (Sergidou et al., 2024);

2. *за комбиновање са другим доказима* (јер је сам глас често подложан варијацијама и техничким деградацијама) (Оташевић и сар., 2024).

Форензичко-фонетска анализа гласа

Форензичка фонетика је примењена интердисциплинарна научна област која се бави анализом гласа и говора у форензичком и криминалистичком контексту, с циљем поређења спорних узорака и неспорног говорног узорка, идентификације или елиминације говорника, као и стручног тумачења говорних карактеристика релевантних за кривични поступак (Нешић и сар., 2011, Оташевић и сар., 2024). За разлику од говорних технологија и биометријских система, у којима се препознавање говорника одвија у контролисаним условима, домен форензичке фонетике обухвата анализу реалних говорних материјала насталих у различитим комуникационим и технич-

ким околностима. У том смислу, форензичка фонетика интегрише знања из фонетике, лингвистике, акустике, обраде говорног сигнала, психологије и криминалистике, при чему основни задаци вештака јесу стручно поређење говорних и језичких обележја и њихова интерпретација у доказном поступку (Јовичић, 2001, Нешић и сар., 2011). Савремени приступи проширују домен форензичке фонетике на примену рачунарских и статистичких метода у евалуацији говорног доказа, уз задржавање кључне улоге експерта у селекцији параметара и интерпретацији резултата (Sergidou et al., 2024; Carroll, 2024).

У форензичко-фонетском вештачењу не примењује се једна изолована метода, већ се користи скуп комплементарних аналитичких поступака чија је сврха поуздано поређење спорних и неспорног говорног узорка. Избор метода зависи од квалитета и количине расположивог говорног материјала, као и од захтева конкретног криминалистичког или судског предмета (Атанасов, 2016; Оташевић и сар., 2019). Аудитивно-перцептивна анализа заснива се на стручној процени гласа и говора од стране обученог форензичког експерта. Овим приступом идентификују се карактеристике као што су висина и боја гласа, интонација, ритам, темпо говора, артикулационе особине, као и индивидуални језички обрасци који могу имати идентификациони значај. Ова метода омогућава сагледавање говорних особина у њиховом природном контексту и представља основни сегмент форензичко-фонетске експертизе (Јовичић & Кашић, 2009; Оташевић и сар., 2019).

Акустичка анализа подразумева инструментално испитивање говорног сигнала применом специјализованих софтверских алата. Анализирају се квантитативни параметри као што су фундаментална фреквенција, форманти, интензитет, трајање сегмената, спектралне карактеристике и микроваријације гласа (Heđever, 2010). Овај приступ омогућава објективизацију одређених говорних обележја и њихову статистичку обраду, што је нарочито значајно у поткрепљивању налаза аудитивне анализе (Оташевић и сар., 2019). Комбиновани приступ представља стандард у савременој форензичко-фонетској пракси и подразумева интеграцију аудитивно-перцептивне и акустичке анализе. Циљ овог приступа јесте да се субјективна процена експерта допуни и провери инструменталним подацима, као и да се нумерички резултати правилно интерпретирају у контексту конкретних криминалистичких околности (Јовичић & Кашић, 2009;

Оташевић и сар., 2024). У савременим форензичким системима резултати комбиноване анализе све чешће се изражавају кроз структурисане евалуационе оквире, укључујући моделе вероватноће и однос вероватноћа, чиме се обезбеђује транспарентније и разумљивије изношење налаза пред судом (Sergidou et al., 2024; Carroll, 2024).

Говорне карактеристике у идентификацији говорника

Идентификација говорника у форензичко-фонетској анализи заснива се на систематском испитивању говорних карактеристика које носе индивидуалне, релативно стабилне и поновљиве информације о говорнику. Те карактеристике могу се груписати на сегментне, супрасегментне и индивидуалне говорне обрасце, при чему се њихов значај процењује у оквиру конкретног криминалистичког случаја (Јовичић & Кашић, 2009).

Сегментне карактеристике

Сегментне карактеристике односе се на реализацију појединачних гласова у говору, односно на фонетске и акустичке особине самогласника и сугласника. У форензичкој фонетици, ове карактеристике имају значај јер могу указивати на индивидуалне обрасце артикулације, карактеристике дијалекта или патолошка одступања у говору (Јовичић & Кашић, 2009). Код самогласника, анализа се најчешће усмерава на формантску структуру (F1, F2, F3), стабилност реализације и контекстуалне варијације, које могу бити под снажним утицајем темпа говора, емоционалног стања и коартикулације. У спонтаном говору, нарочито при убрзаном темпу, вредности форманата често одступају од номиналних, што представља један од главних извора интраговорничке варијабилности (Јовичић & Кашић, 2009). Код сугласника, нарочито фрикатива, плозива и назала, анализирају се спектралне карактеристике, трајање, енергија шумних компоненти и стабилност артикулације. Ове особине могу имати идентификациону вредност када се уочавају доследна одступања или специфични обрасци реализације код истог говорника (Оташевић и сар., 2019). Артикулационе особине говора рефлектују се кроз акустичке параметре говорног сигнала и представљају везу између физиолошких карактеристика вокалног тракта и продукованог звука. У

форензичкој анализи, стабилност ових особина посматра се у оквиру њиховог варијационог поља, при чему се процењује однос интраговорничких и интерговорничких варијација (Јовичић & Кашић, 2009).

Супрасегментне карактеристике

Супрасегментне карактеристике односе се на особине говора које су саставни део појединачних гласова и обухватају прозодијске елементе као што су висина гласа, интонација, темпо, ритам и интензитет. Ове карактеристике често носе значајну индивидуалну компоненту, али су истовремено подложне утицају емоционалних и ситуационих фактора (Јовичић & Кашић, 2009). Фундаментална фреквенција (Ф0) представља један од основних акустичких параметара у форензичкој анализи гласа. Иако Ф0 зависи од биолошких карактеристика гласница, она је изразито променљива у зависности од психоемотивног стања, комуникационог контекста и амбијенталних услова, што захтева опрез у њеној интерпретацији (Јовичић & Кашић, 2009). Интонацијске варијације и темпо и ритам говора представљају супрасегментне обрасце који могу бити карактеристични за одређеног говорника. Ипак, у форензичком контексту они се увек анализирају у корелацији са другим параметрима, јер су подложни променама услед стреса, узбуђења или комуникативне ситуације (Јовичић & Кашић, 2009). Интензитет говора може указивати на индивидуални стил говора, али умногоме зависи од амбијенталних услова, нарочито у бучним окружењима, где долази до појаве Ломбардовог рефлекса. Због тога се интензитет не користи као самосталан идентификациони маркер, већ искључиво у оквиру комбиноване анализе (Јовичић & Кашић, 2009).

Индивидуални говорни обрасци обухватају специфичне, често несвесне карактеристике говора које могу имати висок идентификациони потенцијал, посебно у комбинацији са сегментним и супрасегментним обележјима (Оташевић и сар., 2019). Дисфлуентности, попут понављања, продужавања гласова, блокада или невољних пауза, могу представљати релативно стабилне обрасце у говору одређеног говорника. Њихова идентификациона вредност је нарочито изражена када су присутне доследно у различитим говорним ситуацијама (Јовичић & Кашић, 2009). Анализа пауза, као и испуњених пауза (нпр. „ух", „хм"), омогућава увид у когнитивне и стилистичке

обрасце говорника. У форензичкој пракси, ове карактеристике могу допринети разликовању говорника када се уоче стални обрасци њиховог појављивања (Оташевић и сар., 2019). Говорни манири обухватају поштапалице, специфичне лексичке изборе, устаљене фразе и индивидуалне начине започињања или завршавања исказа. Ти елементи имају посебан значај у аудитивно-перцептивној анализи и често представљају снажне индикаторе индивидуалности говорника (Јовичић & Кашић, 2009).

Криминалистичка идентификација говорника

Криминалистичка идентификација говорника представља сложен процес у коме се резултати форензичко-фонетске анализе интегришу у доказни поступак с циљем процене да ли спорни и неспорни аудио-узорци потичу од исте особе. За разлику од чисто лингвистичког или акустичког приступа, криминалистичка идентификација захтева вредновање налаза у условима реалних ограничења, варијабилности говора и процесних захтева кривичног поступка (Атанасов, 2016; Оташевић и сар., 2019). Основ криминалистичке идентификације говорника чини упоредна анализа спорних и референтних (неспорних) узорака, при чему се испитује степен подударности говорних карактеристика у оквиру њиховог природног варијационог опсега. Анализа се спроводи на више нивоа: на сегментном, супрасегментном и на нивоу индивидуалних говорних образаца, уз комбиновање аудитивно-перцептивне и акустичке методе (Јовичић & Кашић, 2009). Посебан изазов у криминалистичкој идентификацији говорника представља варијабилност говора. Она се испољава како унутар говора исте особе (интраговорничка варијабилност), тако и између различитих говорника (интерговорничка варијабилност). Интраговорничке промене могу бити последица емоционалног стања, стреса, здравствених фактора, дејства алкохола или психоактивних супстанци, као и промена у темпу и ритму говора или комуникационој ситуацији. С друге стране, интерговорничке разлике односе се на варијације између различитих говорника и представљају основу за дискриминацију у поступку идентификације. Управо однос интраговорничких и интерговорничких варијација одређује степен поузданости закључивања и ширину такозване „зоне несигурности" у форензичком поређењу говорника. Због тога се у савременој форензичкој

пракси идентификација гласа не третира као детерминистички поступак, већ као евалуативни процес у коме се степен сличности и разлике тумачи у оквиру природног варијационог поља говора (Јовичић & Кашић, 2009; Оташевић и сар., 2019). У криминалистичкој пракси, посебно се наглашава потреба да референтни узорци буду упоредиви са спорним снимком, по садржају, говорној ситуацији и техничким условима, јер разлике у акустичким условима могу значајно утицати на поузданост закључака (Оташевић и сар., 2019). Акустички параметри који се анализирају (фреквенцијски опсег, интензитет, спектралне карактеристике) имају физичку основу у законитостима простирања и перцепције звука, што је посебно важно приликом тумачења резултата у условима деградираних снимака (Дринчић и сар.. 2018).

Акустички посматрано, промене у интензитету, фреквенцији и спектру говора могу бити последица физичких фактора, као што су амбијентална бука, рефлексије звука и ограничења микрофона, што захтева опрез у тумачењу измерених параметара (Дринчић и сар., 2018). У криминалистичком контексту, варијабилност се не посматра као сметња, већ као аналитичка категорија, јер правилно сагледавање граница варијација омогућава реалнију процену значаја уочених сличности и разлика (Оташевић и сар., 2019). Криминалистичка идентификација говорника не подразумева апсолутну сигурност, већ процену вероватноће у оквиру јасно дефинисаних методолошких и доказних ограничења. Форензичко-фонетски налаз представља експертско мишљење засновано на доступним подацима, а не категоријску тврдњу о идентитету говорника (Sergidou et al., 2024). Могућности идентификације зависе од: квалитета и трајања аудио-снимка, степена упоредивости узорака, броја и стабилности анализираних говорних карактеристика, примењене методологије и комбинованог приступа. Границе идентификације посебно долазе до изражаја код кратких, шумом деградираних или емоционално обојених снимака, где се поузданост закључака значајно смањује (Јовичић & Кашић, 2009; Оташевић и сар., 2019). Физичка ограничења звучног сигнала, укључујући слабљење, изобличења и промене спектра при простирању, додатно утичу на интерпретацију форензичких налаза, што захтева да се акустички резултати увек тумаче у контексту реалних услова снимања (Дринчић и сар., 2018).

Тумачење форензичко-фонетског налаза

Форензичко-фонетски налаз представља стручно мишљење вештака о степену сличности или различитости између спорних и референтних (неспорних) говорних узорака, али не и одлуку о идентитету учиниоца, која је искључиво у надлежности суда. Улога вештака огледа се у стручној анализи и евалуацији гласовних и говорних карактеристика, као и у јасном образлагању методологије и ограничења примењене анализе (Јовичић & Кашић, 2009). У форензичком препознавању говорника вештак делује у условима ограниченог и често технички компромитованог материјала, због чега се закључци не могу формулисати категорички, већ искључиво као вероватне процене засноване на целокупном миљеу опсервација (Posse, 2002; Јовичић, 2009). Овакав приступ у потпуности је у складу са ставом да је форензичко препознавање говорника евалуативни, а не детерминистички поступак.

Језик форензичко-фонетског налаза мора бити прецизан, неутралан и методолошки транспарентан, како би суд могао правилно да разуме домет и тежину изнетог мишљења. Будући да се право остварује кроз језик, јер без језика не би било ни закона ни судског поступка, свака интерпретација гласовног доказа неизбежно подразумева анализу начина на који се значење конструише у правном контексту (Correa, 2013: 5). Стога форензички експерт не анализира само акустичке параметре, већ и комуникативну функцију и могуће интерпретације исказа у институционалном дискурсу. У савременој пракси избегавају се формулације које имплицирају апсолутну сигурност (нпр. „глас несумњиво припада…"), а предност се даје вербалним скалама процене (нпр. „значајан степен вероватноће сличности") које омогућавају адекватно тумачење резултата без довођења суда у заблуду (Јовичић & Кашић, 2009; Association of Forensic Sciences Providers, 2009). Управо зато што је судница облик „институционалног дискурса", у којем су улоге учесника и „дозвољени" доприноси разговору унапред ограничени правилима и процедурама, комуникација је често асиметрична (нпр. једна страна управља темама и редом излагања), што директно утиче на начин на који се стручни исказ тумачи и вреднује (Coulthard & Johnson, 2007: 16). У том оквиру, питања и формулације унакрсног испитивања могу бити конструисани тако да сугеришу очекивани одговор или да „усмеравају"

закључак, па је тешко да вештак јасно раздвоји налаз (опсервације) од интерпретације и експлицитно наведе ограничења да би се смањио ризик доношења погрешних закључака (Coulthard & Johnson, 2007: 8). Овакав начин извештавања посебно је важан у случајевима када се користе статистички модели и мере вероватноће (нпр. *likelihood ratio* – *LR*), јер нумерички подаци без адекватног контекста могу бити погрешно интерпретирани у криминалистичкој и судској пракси (Росе, 2002; Јовичић & Кашић, 2009). Један од кључних проблема у примени форензичко-фонетских налаза јесте ризик њихове погрешне интерпретације. До тога може доћи услед: прецењивања појединачних говорних параметара, занемаривања интраспикерских варијација, некритичког ослањања на аутоматске системе за препознавање говорника и неадекватног разумевања вероватноће могућих закључака. Истраживања указују да интраспикерске варијације, условљене емоционалним стањем, контекстом комуникације, здравственим факторима или техничким условима снимања, могу бити толико изражене да значајно утичу на крајњу процену сличности (Јовичић & Кашић, 2009; Оташевић и сар., 2019). Због тога је неопходно да вештак у налазу експлицитно наведе све релевантне ограничавајуће факторе.

У савременој европској пракси тумачење форензичко-фонетског налаза све више се усклађује са смерницама које прописује *ENFSI* (*European Network of Forensic Science Institutes*). *ENFSI* стандарди наглашавају неколико кључних принципа: независност вештака, транспарентност методологије, јасно разграничење стручног и правног закључивања, као и обавезно навођење ограничења анализе (ENFSI, 2016). У складу са овим смерницама, вештак: не износи категоричке тврдње о идентитету, користи прихватљиву терминологију и вербалне скале процене, јасно раздваја налаз (чињенице и мерења) од мишљења (интерпретације), документује квалитет снимка, присуство шума и дисторзија. Овакав приступ у потпуности је у складу са ставовима домаћих аутора да је комбиновани аудитивно-инструментални приступ уз доминантну улогу експерта тренутно најпоузданији модел форензичког препознавања говорника (Јовичић & Кашић, 2009; Оташевић и сар., 2019). Примена *ENFSI* стандарда у домаћој форензичко-фонетској пракси доприноси повећању кредибилитета вештачења, уједначавању судске праксе и смањењу ризика од судских заблуда заснованих на погрешном тумачењу гласовних и говорних доказа. На тај начин, форензичко-фонетски налаз постаје процесно

одржив и међународно упоредив доказ, што је од посебног значаја у сложеним кривичним предметима са аудио-доказима.

Дискусија

Практично, један од највећих извора ризика у закључивању није само разлика између говорника, већ „неусклађеност услова" између спорних и референтних снимака (различит канал, уређај, окружење, психофизичко стање, језички садржај и сл.), која може изменити мерене особине и произвести привидне сличности или разлике које нису карактеристике идентитета говорника (Campbell et al., 2009). Због тога се у форензичком контексту посебно наглашава потреба за опрезом и избегавање наратива о гласовном отиску, који погрешно сугерише да спектрограм обезбеђује поузданост сличну отиску прста или ДНК. У реалности, говор је врло варијабилан и закључивање мора бити евалуативно (Campbell et al., 2009). Додатно, и када системи постижу добре резултате у стандардизованим евалуацијама, сама стопа грешке није довољна гаранција форензичке поузданости у конкретном предмету, посебно због питања калибрације скорова у различитим условима, што директно утиче на интерпретацију односа вероватноћа у пракси (Campbell et al., 2009). Одатле следи да аутоматизација може бити значајна подршка, али не може заменити експертску процену варијабилности, квалитета и упоредивости узорака, уз обавезно навођење ограничења и алтернативних објашњења (Carroll, 2024; Sergidou et al., 2024; AFSP, 2009; ENFSI, 2016).

Доступна литература указује да се у реалним случајевима најпоузданије закључивање постиже комбиновањем аудитивно-перцептивне и инструментално-акустичке анализе, уз укључивање лингвистичких индикатора када су они доступни. Такав „хибридни" приступ смањује ризик да налаз зависи од једног параметра и омогућава да се узме у обзир целокупно говорно понашање (Јовичић & Кашић, 2009; Оташевић и сар., 2019). Посебно је важно да се подударности и разлике процењују у оквиру варијационог поља, јер и код истог говорника постоје велике промене настале услед емоција, стреса, темпа говора, говора у окружењу са буком (Ломбардов рефлекс), здравствених фактора или дејства супстанци, док се интерговорничка варијабилност користи као основа дискриминације (Јовичић & Кашић, 2009; Оташевић и сар., 2019). Инструментална ана-

лиза се често ослања на физиолошке и фонетске параметре, као што су фундаментална фреквенција (*Fo*), форманти (*F1–F3*) и просечни дугорочни говорни спектар (*LTASS*), јер су повезани са карактеристикама вокалног тракта и говорним навикама. Ипак, њихова употреба у форензичком контексту зависи од квалитета снимка и упоредивости узорака – деградације (шум, канал, компресија, преклапање говора) – који могу променити или прикрити мерене вредности, па се закључци морају заснивати на статистички стабилним обрасцима и експлицитном навођењу ограничења (Јовичић, 2001; Kruthika et al., 2025). У том смислу, кратки и узорци са присуством шума шире „зону несигурности" и повећавају потребу за опрезним, евалуативним закључивањем (Нешић и сар., 2011).

Међународна литература наглашава пораст улоге аутоматизованих приступа и машинског учења у форензичком поређењу говорника. Међу тим приступима предњачи коришћење мелфреквенцијских кепстралних коефицијената (*MFCC*) и сродних спектралних карактеристика, као и генеративни модели (приступи засновани на генеративним моделима). Њихова највећа вредност је у стандардизацији обраде и могућности квантификације резултата, али поузданост мора бити поткрепљена емпиријском валидацијом на подацима који личе на реалне форензичке услове, јер деградације сигнала и кратко трајање узорака могу значајно нарушити перформансе (Kruthika et al., 2025). Форензичка лингвистика се данас у великој мери препознаје као засебна дисциплина која се глобално проширила, при чему се њен домет и примена шире у различитим правним системима и контекстима (Alduais et al., 2023). Развој говорних и текстуалних технологија додатно утиче на овај раст: технологија помаже у аутоматизацији анализе текста, али истовремено омогућава да се дигитални подаци све више испитују и кроз призму класичних лингвистичких и дискурзивних метода (Alduais et al., 2023). Посебно је релевантан налаз да је јединство лингвистичких и акустичких информација корисно управо у условима у којима је чисто акустичко поређење отежано (нпр. пресретнути телефонски разговори у бучном окружењу и ограниченим фреквенцијским опсегом). У таквим случајевима, лингвистички елементи (нпр. избор речи, граматички обрасци, дисфлуентности, говорни манири) могу допунити акустичке показатеље и учврстити укупну процену, док се квалитет система и његовог извештавања проверава стандардним мерама као што су *Cllr* (*cost of*

log-likelihood ratio) и *EER* (*equal error rate*), што доприноси транспарентности и контроли ризика од погрешне интерпретације (Sergidou et al., 2024).

Улога експерта остаје централна. Аутоматизација може бити значајна подршка, али не може заменити стручно тумачење варијабилности говора и процену квалитета и упоредивости узорака, уз обавезно навођење ограничења и алтернативних објашњења. То је посебно важно у судском контексту, где се налаз мора представити јасно, неутрално и евалуативно, уз избегавање категоричких формулација и уз објашњење шта резултат заиста значи у односу на постављене хипотезе (Carroll, 2024; Sergidou et al., 2024; Association of Forensic Science Providers [AFSP], 2009; Јовичић, 2001).

Закључак

Рад указује да форензичко-фонетска анализа гласа представља релевантан и применљив метод у криминалистичкој идентификацији говорника, али само уколико се њени резултати тумаче у евалуативном, а не категоричком оквиру. Савремена литература потврђује да се поузданост закључивања повећава применом комбинованог аудитивно-перцептивног и инструментално-акустичког приступа, уз интеграцију лингвистичких показатеља и извештавање у оквиру односа вероватноћа. Такав приступ омогућава транспарентније вредновање форензичког налаза и његову већу процесну прихватљивост у доказном поступку. Истовремено, резултати анализе указују да аутоматизовани системи и методе засноване на машинском учењу могу допринети стандардизацији обраде и квантификацији резултата. Њихова улога у форензичком контексту мора остати на нивоу помоћног средства, а примена захтева обавезну емпиријску валидацију на репрезентативним, реалним снимцима и интерпретацију резултата у складу са форензичким евалуативним оквиром. То значи да се не може посматрати као самостални доказ идентитета говорника.

С аспекта криминалистичке и безбедносне праксе, резултати рада указују на потребу за унапређењем процедура у две кључне области. Прво, неопходна је стандардизација прикупљања референтних говорних узорака, уз контролу трајања говорне ситуације и техничких услова снимања. На тај начин би се обезбедила већа упоредивост са спорним материјалом. Друго, неопходно је систематско јачање ком-

петенција криминалистичких и правосудних актера, посебно у разумевању евалуативне природе форензичко-фонетских налаза и правилне интерпретације односа вероватноћа (*LR*). Тако би се избегле погрешне или претеране интерпретације говорних доказа у поступку. Примена ових препорука, уз усклађивање са савременим европским форензичким стандардима (*ENFSI*), може допринети повећању поузданости, уједначавању праксе и јачању доказне вредности форензичко-фонетске анализе гласа у криминалистичком и судском контексту.

Литература

1. Alduais, A., Al-Khulaidi, M. A., Allegretta, S., & Abdulkhalek, M. M. (2023). Forensic linguistics: A scientometric review. *Cogent Arts & Humanities*, 10(1), 2214387. https://doi.org/10.1080/23311983.2023.2214387
2. Association of Forensic Science Providers. (2009). Standards for the formulation of evaluative forensic science expert opinion. *Science & Justice, 49*(3), 161–164. https://doi.org/10.1016/j.scijus.2009.07.004
3. Атанасов, С. (2016). *Улога сведока у откривању и доказивању кривичног дела* (Докторска дисертација). Универзитет у Приштини са привременим седиштем у Косовској Митровици, Правни факултет.
4. Дринчић, Д., Правица, П., & Новковић, Д. (2018). *Основи акустике*. Београд: Висока школа електротехнике и рачунарства, струковних студија.
5. European Network of Forensic Science Institutes. (2016). *Best practice guidelines for ENF analysis in forensic authentication of digital evidence: Forensic speech and audio analysis working group*. https://enfsi.eu/wp-content/uploads/2016/09/forensic_speech_and_audio_analysis_wg_best_practice_guidelines_for_enf_analysis_in_forensic_authentication_of_digital_evidence_0.pdf
6. Јовичић, С. Т. (2001). Форензички аспекти препознавања говорника. *Наука, техника и безбедност,* 1, 41–60.
7. Jovičić, S. T., & Kašić, Z. (2009). *Intra-spikerske varijacije u govoru: Forenzičke implikacije.* U International Conference on Applied Linguistics. Filozofski fakultet, Novi Sad, 123–132.

8. Kruthika, K., Raghavendra, R., Busch, C., & Ramachandra, R. (2025). Voice comparison using acoustic analysis and generative adversarial networks for forensics. *International Journal of Information and Graphics Processing, 17*(2), Article 2550007. https://doi.org/10.1142/S021946782550007X
9. Нешић, Л., Ковачевић, Ј., & Стевовић-Оташевић, Ј. (2011). Форензичка анализа говорног сигнала. *Правни живот*, 60(5–6), 87–102.
10. Otašević, J. R., Atanasov, S., & Čolić, G. (2019). Utilization of voice analysis method in criminal investigations. *Archibald Reiss Days*, *9*(2): 261–270.
11. Otašević, J., Otašević, B., & Atanasov, S. (2024). *Forensic voice analysis in criminal investigation.* U O. Czenczer, G. Kovács, & B. Mészáros (U.), Conference proceedings (p. 238–246). Hungarian Association of Police Science.
12. Rose, P. (2002). *Forensic speaker recognition*. Taylor & Francis.
13. Sergidou, E. K., Ypma, R., Rohdin, J., Worring, M., Geradts, Z., & Bosma, W. (2024). Fusing linguistic and acoustic information for automated forensic speaker comparison. *Science & Justice, 64*(5), 485–497. https://doi.org/10.1016/j.scijus.2024.07.001
14. Heđever, M. *Osnove fiziološke i govorne akustike*. Zagreb: Edukacijsko-rehabilitacijski fakultet Sveučilišta u Zagrebu, 2010.
15. Carroll, J. (2024). *Rhythm in forensic voice comparison: Perception, metrics and practice* (Doctoral dissertation). University of York.
16. Carić, M., & Širić, L. (2023). Primjena forenzičke akustike i fonetike u kaznenom postupku s posebnim osvrtom na vještačenje glasovnih zapisa. *Zbornik radova Pravnog fakulteta u Splitu*, *60*(1), 189–217.
17. Campbell, J. P., Shen, W., Campbell, W. M., Schwartz, R., Bonastre, J.-F., & Matrouf, D. (2009). *Forensic speaker recognition*. *IEEE Signal Processing Magazine, 26*(2), 95–103. https://doi.org/10.1109/MSP.2008.931100
18. Correa, M. (2013). Forensic linguistics: An overview of the intersection and interaction of language and law. *Studies About Languages*, (23), 5–13. https://doi.org/10.5755/j01.sal.0.23.5020
19. Coulthard, M., & Johnson, A. (2007). *An introduction to forensic linguistics: Language in evidence*. Routledge.

Forensic Phonetic Analysis of Voice in the Criminalistic Identification of Speakers

Abstract: *This paper examines the role and significance of forensic-phonetic voice analysis in the criminalistic identification of speakers, with a particular focus on its application under real investigative and judicial conditions. Voice is considered a specific criminalistic trace of intangible nature, whose identification value arises from a combination of physiological, behavioural, phonetic, and linguistic characteristics of the speaker. Unlike controlled biometric systems, forensic speaker recognition is performed in uncontrolled and often degraded conditions, which requires a cautious and evaluative approach to interpretation. The paper is based on a review and comparative analysis of relevant domestic and international literature in the fields of forensic phonetics, forensic speaker comparison, and criminalistics. Special attention is given to methodological approaches used in practice, including auditory-perceptual, acoustic-instrumental, and combined methods, as well as to the problem of intra- and inter-speaker variability and its impact on the reliability of conclusions. The paper highlights the importance of an evaluative framework in forensic voice comparison, particularly the use of the likelihood ratio approach, which allows the strength of the evidence to be assessed relative to competing hypotheses rather than through categorical identification. The growing role of automated and machine-learning-based systems is also discussed, emphasizing that such tools should be empirically validated and used as support to expert judgment, not as a substitute for it.*

The findings indicate that a combined expert-based and instrumental approach, aligned with international standards such as those of the European Network of Forensic Science Institutes (ENFSI), currently represents the most reliable model for forensic speaker identification. The paper concludes with recommendations aimed at improving forensic and criminalistic practice, including standardization of reference samples, transparent reporting of limitations, and continuous education of criminal justice professionals on the interpretation of forensic voice evidence.

Keywords: *forensic speaker recognition, criminalistics identification of speakers, speech variability, likelihood ration (LR), evaluation report*