

DOI: <https://doi.org/10.32839/2304-5809/2019-4-68-40>
УДК 343.9

Шелепкова С.М.

Харківський науково-дослідний експертно-криміналістичний центр
МВС України

АУДИТИВНИЙ МЕТОД ОЦІНКИ РОЗПІЗНАВАННЯ ГОЛОСУ І МОВЛЕННЯ ЛЮДИНИ У РІЗНИХ УМОВАХ ПОДАВАННЯ МОВЛЕННЕВОГО СИГНАЛУ

Анотація. На сьогодні криміналісти все частіше зіштовхуються у своїй практиці із необхідністю експертного вивчення об'єктів дослідження різного походження. Криміналістичне дослідження усного мовлення, зафіксованого на матеріальному носії, є одним із найцікавіших та перспективніших, хоча і мало вивчених завдань, здатних посісти пріоритетне місце у слідчому і судовому процесах. У статті розглянуто історію формування і розвитку судової експертизи відео-, звукозапису; вивчено перші спроби візуалізації мовленнєвих сигналів; проаналізовано аудиторний метод оцінки розпізнавання голосу і мовлення людини у різних умовах подання мовленнєвого сигналу; наведено відомі прийоми оцінювання природності мовлення; досліджено мовлення із погляду акустики, мовлення за своїми фізичними показниками, а також показана варіативність мовлення; оглянуто питання, пов'язані із природою мовленнєвого сигналу, зафіксованого на матеріальному носії; доведено, що голос в умовах мовленнєвого процесу характеризується своєю неповторною індивідуальністю.

Ключові слова: судова експертиза відео-, звукозапису, голос і мовлення людини, аудиторний метод, мовленнєвий сигнал, акустика, варіативність мовлення, запис, фонограма, файл.

Shelepko Svitlana

Kharkiv Scientific Research Forensic Center of MIA of Ukraine

AUDITING METHOD FOR EVALUATION OF VOTING AND LIBRARY RECOGNITION OF HUMAN IN VARIOUS CONDITIONS OF LINGUISTIC SIGNALING

Summary. Effective prevention and investigation of crimes requires continuous improvement of the activities of law enforcement agencies. Today criminologists increasingly encounter their practice with the need for expert study of objects of study of different origins. The forensic study of oral speech, recorded on a material medium, is one of the most interesting and perspective, though not well-studied tasks, capable of occupying a priority place in investigative and judicial processes. One of the most natural and at the same time effective methods for evaluating legibility is the method based on the auditory perception of speech by a person. A human hearing is not able to give an exact quantitative assessment of one or another physical characteristics of the perceptive speech signal, nevertheless, can make a global solution to the particular qualities of speech and voice, and to indicate the severity of its features. The article deals with the history of the formation and development of forensic examination of video and audio; the first attempts of visualization of speech signals were studied; the audit method of evaluation of speech recognition and speech recognition in different conditions of presentation of a speech signal is analyzed; famous methods of evaluation of naturalness of speech are given; Speech is investigated in terms of acoustics, speech by its physical characteristics, and also the variation of speech is shown; examined the issues connected with the nature of the speech signal, fixed on the material carrier; it is proved that the voice in the speech process is characterized by its unique personality; the issues related to counterfeiting of phonograms, which are carried out as purely technical techniques (gluing, editing, overlaying of different sources of sound, etc.), and directly forgery of the content and voice of the speaker through imitation, or vice versa, by changing their voice from purpose "to withdraw" from themselves suspicion, as a result of video and audio expertise, you can get information that is interesting for criminology, indicating the individual features of the speaker, as well as the "diagnosis" of phonogram rigging.

Keywords: forensic examination of video and audio, voice and speech of a person, auditing method, speech signal, acoustics, variation of speech, record, phonogram, file.

Постановка проблеми. Щоденно ми зіштовхуємося із фактом розпізнавання людини тільки за її голосом. Наприклад, подзвонивши у двері, ми чуємо питання: «Хто там?», на яке мимовільно даємо відповідь: «Я», не замислюючись над тим, що займенник «я» не несе інформації про конкретну особу. Ми підсвідомо посилаємо свій голос тому, хто слухає, як об'єкт для розпізнавання.

У чому ж полягає проблема розпізнавання (ідентифікації) голосу? Це питання не є риторичним, а далі у роботі ми намагатимемося відшукати відповідь на це питання.

Аналіз останніх досліджень і публікацій. Історія формування і розвитку судової експертизи відео-, звукозапису безпосередньо пов'язана з науково-технічним прогресом, який проходив

у західних країнах набагато раніше, ніж в Україні. Першу спробу наукового опису акустичної моделі людського голосу та мовлення здійснив у 1779 році Християн Готліб Кратценштейн, що винайшов прототип синтезатора мовлення, який вимовляв п'ять голосних звуків (а, е, і, о, у) і був представлений на конкурс Санкт-Петербурзької Імператорської академії наук [4].

Перші спроби візуалізації мовленнєвих сигналів, що дозволяють побачити мовленнєвий сигнал у формі, що відображає артикуляцію і голос людини, були зроблені ще у 20-30-х рр. минулого століття, але винахід такого пристрою вдалося здійснити тільки до кінця 1945 року. За допомогою цього приладу можна було здійснювати аудиторний та візуальний аналіз звукового мовлення.

У 1964 році польський кримінолог Анжей Шварц (Andrzej Witold Szwarz) у результаті узагальнення десятирічного досвіду теоретично обґрунтував методику дослідження і запропонував назвати вид експертизи, що розглядає записи усного мовлення, «фоноскопія» [1].

Цей термін надалі був запозичений іншими фахівцями і продовжує застосовуватися у деяких літературних джерелах і експертних установах.

Залучення фонограм із звукозаписною інформацією у кримінальний процес на різних стадіях розкриття та розслідування злочинів об'єктивно зумовило виникнення у 1964-1969 рр. самостійного роду експертної діяльності – криміналістичної експертизи фонограм.

На той час були розроблені і сформульовані загальні принципи проведення криміналістичної експертизи, проаналізовано основні стадії процесу криміналістичного дослідження, фактично сформовані технічні, наукові та правові основи виникнення криміналістичної експертизи фонограм.

Такого напрямку питання вирішували, зокрема Р.С. Белкін, В.Є. Бергер, А.І. Вінберг, Ю.Ф. Жаріков, М.В. Салтевський, С.І. Тихенко, А.О. Фокіна.

Процеси, що відбулися за останні роки, зробили актуальним проведення і лінгвістичних експертиз. Висвітлення цієї тематики знаходимо у роботах таких дослідників, як: Р.І. Аванесов, В.М. Алпанов, О.І. Галяшина, В.В. Іванов, В.В. Потапов та Р.К. Потапова, Д.М. Ушаков.

Докладний опис звукової системи української мови знаходимо у праці авторського колективу за редакцією І.К. Білодіда, де на підставі нових експериментальних досліджень, проведених у лабораторіях експериментальної фонетики Інституту мовознавства ім. О.О. Потебні та Київського державного університету імені Т.Г. Шевченка, подається анатомо-фізіологічна й акустична характеристика звуків. У написанні «Сучасної української мови. Вступ. Фонетика» брали участь такі дослідники, як: В.М. Брахнов, М.А. Жовтобрюх, П.П. Коструба, М.Ф. Наконечний, В.С. Перейбийніс, Л.І. Прокопова, Н.І. Тоцька.

Проте опубліковані роботи не вичерпують сучасних потреб науки та запитів практичної діяльності правоохоронних органів, оскільки відбулись зміни у криміналістиці, зокрема, з'явилися нові напрями розвитку криміналістичної техніки, які своєю появою значно розширюють її можливості у дослідженні судової експертизи відео-, звукозапису і вимагають серйозних теоретичних висновків.

Виділення невирішених раніше частин загальної проблеми. Між тим на сьогодні криміналісти усе частіше зіштовхуються у своїй практиці із необхідністю експертного вивчення об'єктів дослідження різного походження.

Криміналістичне дослідження усного мовлення, зафіксованого на матеріальному носії, є одним із найцікавіших та перспективніших, хоча і мало вивчених завдань, здатних посісти пріоритетне місце у слідчому і судовому процесах.

Сьогодні матеріали відео-, звукозапису додаються до протоколу слідчих дій не тільки як дані фіксації показань (свідчень) під час дізнання та попереднього слідства, але і як речові докази на суді. Дійсно, якщо відео-, звукозаписи використовуються у процесі слідства для протокольного оформлення

потрібного матеріалу, суттєво доповнюючи при цьому спосіб стенографування даних, було б помилковим не використати і відмовитися від цілої низки переваг, які притаманні фонограмі, у порівнянні із письмовим методом фіксації даних дошиту.

Такими перевагами, у першу чергу, є достовірність фактичного матеріалу, що виявляється у відео-, звукозаписах. У випадку, коли різняться письмові дані протоколу і дані відео-, звукозапису, то за основу, як показує практика, береться останній спосіб, якщо автентичність фонограми не викликає сумніву.

Крім достовірності смислової інформації, фонограма містить ще багато відомостей, які не можна описати стенографічно. Це і мовленнєві навички, етнографічне походження тощо, тобто усі відомості, які притаманні усному мовленню у порівнянні з письмовим записом.

Метою статті є аналіз аудиторного методу оцінки розпізнавання голосу і мовлення людини у різних умовах подання мовленнєвого сигналу. Незважаючи на складний процес «породження» фонограми мовлення, який характеризується великим різновидом чинників, що заважають та спотворюють нормальне мовлення, людське вухо досить добре розпізнає як зміст мовлення, зафіксований на матеріальному носії зі значним шумом, так і індивідуальні особливості мовлення.

Усе це свідчить про те, що у результаті експертизи відео-, звукозапису можна отримати цікаву для криміналістики інформацію, що свідчить про індивідуальні риси того, хто говорить, а також «діагноз» фальсифікації фонограм.

Виклад основного матеріалу. Поглянемо тепер на мовлення із погляду акустики і проаналізуємо аудиторний метод оцінки розпізнавання голосу і мовлення людини у різних умовах подання мовленнєвого сигналу. Під час мовлення утворюються звукові хвилі, або коливання, які, з одного боку, характеризуються амплітудою, а з іншого – частотою, тобто кількістю коливань за одиницю часу, наприклад, за секунду.

Максимальна амплітуда – це максимальне підвищення або зниження тиску повітря, що його створює звукова хвиля. Тиск повітря вимірюється у паскалях (Па). За оптимальних умов людське вухо здатне вловити звукові хвилі з надзвичайно малою амплітудою, яка викликає тиск усього 10-5Па.

Частота вимірюється у герцах (Гц), тобто у коливаннях за секунду, або у децибелах. За сприятливих умов людина здатна сприймати акустичні вібрації до 20 кілогерц (кГц). Найслабший звук, який ми чуємо, дорівнює 5дБ. Дуже тихий шепіт на відстані неповних 2 метрів дорівнює 20 дБ. Межа, на якій сприймання звуку викликає больові відчуття, – 140 дБ.

Діапазон людського голосу, від найтихішого шепоту до найсильнішого крику, дорівнює 60 дБ. Вухо людини найчутливіше до звуків, що мають 2-4 кГц. Настільки тонким є людський слух, показує той факт, що барабанна перетинка у вусі рухається у діапазоні однієї десятимільйонної частки метра. Якби вухо було ще чутливішим, воно вловлювало б випадковий рух молекул повітря.

Існують мільярди різних складних звуків, і вивчити їх усі неможливо. Учені встановили, що вловлюваний вухом звуковий діапазон охоплює приблизно 340 000 звуків, які теоретично

можна відрізнити один від одного. Вухом може розрізнити звуки лише тоді, коли вони сприймаються не одночасно, а послідовно – один за одним. Вухом найчутливіше сприймає звуки, що за енергією знаходяться довкола 50 дБ, а за частотою – у діапазоні між 500 і 4000 Гц.

У людського вуха є властивість, якої немає, наприклад, у зору. Воно може чути два звуки водночас, з яких один, скажімо, має 800 коливань, а інший – 1200. Натомість людське око не може бачити водночас червоне і жовте світло, що випромінюється із одного джерела. Обидва кольори зливаються в один – жовтогарячий.

Ця виняткова здатність людського слуху дає можливість чути голос людини, незважаючи на навколишні шуми. Ми можемо виділити цей голос, відігнати його. Постійні навколишні шуми є причиною того, чому мова має обмежену кількість звукових символів, які чітко відрізняються один від одного. Завдяки цьому ми не плутаємо звуки в умовах навколишніх шумів.

Одним із найбільш природних і разом із тим ефективних методів оцінки розбірливості мовлення є метод, заснований на адитивному (слуховому) сприйнятті мовлення людиною.

В основі аудитивних дослідів розпізнання голосів знаходяться відомі прийоми оцінювання природності мовлення [5], які були засновані на відборі групи дикторів і групи слухачів (до речі, голос та мовлення дикторів слухачам мав бути знайомий). Метою цих дослідів було вирішення найрізноманітніших завдань, таких як оцінювання різновидів між голосами дикторів, установа ролі окремих звуків щодо інформації про того, хто говорить.

Серед перших робіт щодо вивчення можливостей ідентифікації голосів аудитивним методом за різних умов демонстрації мовленнєвого сигналу виділено статтю Полока, Пікета та Самбі [6], що стосується проблеми розпізнання тих, хто говорить. Був покладений початок дослідження надійності розпізнавання, що залежить від таких чинників: кількість дикторів, тривалість наданого на дослідження мовленнєвого матеріалу тощо.

У статті Брикера та Пружанського [2] проводиться аналіз досліджень не тільки ізольованих фонем, але і мовленнєві відрізки більшої тривалості: склади, односкладові та двоскладові слова, а також речення. Результати правильного розпізнавання залежать від типу повідомлення і кількості фонем, які є у цьому повідомленні.

Виявилось, що мовленнєвий відрізок із багатшим змістом фонем більш інформативний. Не викликає сумніву і той факт, що головну роль у передаванні інформації про індивідуум відіграють саме голосні [8]. Голосним звукам найбільш властива інформативність, дзвінкі ж звуки краще розпізнаються, ніж глухі.

Мовлення за своїми фізичними показниками змінюється від вимови до вимови, якщо навіть у семантичному, синтаксичному й іншому аспектах вимовляються тотожні відрізки мовлення. Варіативність мовлення залежить також і від віку мовця. Як уже зазначалося, протягом усього життя голос людини змінюється як за висотою, так і особливо за тембром.

Людське мовлення підпадає під вплив залежно від географічного, етнографічного та соціального середовища, до якого належить диктор і під

вплив якого він потрапляє (набуття діалектних рис мовлення). Також на якості голосу можуть позначитися різні патологічні процеси, пов'язані із загальним фізичним станом (наприклад, захворювання горла, яке впливає на вимову).

Крім цього, людина може змінювати свій голос і навмисно. Бездоганне імітування голосів та вимови, що здійснюється професійними імітаторами, можливе лише у тому випадку, якщо «імітованому суб'єкту» притаманні дуже виразні особливості вимови (характерна інтонація, акцент, темп мовлення) або тембр (гугнявість, гаркавість) і імітатор чітко відчуває і сприймає їх.

До речі саме цим пояснюється той факт, що навіть професійні імітатори не можуть наслідувати звичайне мовлення, яке нічим не виділяється серед мовлення пересічних людей.

Висока мобільність мовленнєвого апарату, з одного боку, і великі можливості людини варіювати свою артикуляцію, з іншого, створюють умови не тільки для якісної зміни голосу, але і для наслідування артикуляції інших людей. Серед різних параметрів мовленнєвого сигналу можуть бути такі, що легко імітуються, а також такі, що важко імітуються.

Майже усі без винятку просодичні характеристики мовлення: висота голосу та його мелодика, інтенсивність, темп та ритміка мовлення, тривалість пауз тощо відносно легко імітуються [3]. Слід також зазначити високий рівень наслідування мелодики основного тону. Найбільш простим способом наслідування параметрів основного тону голосу є спів [7].

Наслідувати також можна деякі дефекти мовлення, а також ті особливості, які пов'язані зі свідомо спрямованою і при цьому контрольованою активністю людини.

Висновки і пропозиції. Питання, пов'язані із природою мовленнєвого сигналу, зафіксованого на матеріальному носії, привертають значну увагу, серед яких пріоритетними є ті, що пов'язані із індивідуальною варіативністю людського мовлення.

Насправді, якщо б голос людини за всіма своїми фізичними показниками дуже змінювався, тоді не тільки у різних умовах мовленнєвої активності, але і взагалі неможливо б було у повсякденному житті впізнати людину за її голосом та манерою вимови.

Оскільки кожний голос в умовах мовленнєвого процесу характеризується своєю неповторною індивідуальністю, яка є безпомилковою розпізнавальною ознакою для оточуючих, то необхідно вивчити і межі, і види його можливої змінності.

Серед питань, що вирішуються експертизою відео-, звукозапису, особливого значення набувають питання, пов'язані із гарантованою перевіркою їх автентичності. Підробка фонограм може здійснюватися як суто технічними прийомами (монтаж, накладання різних джерел звуку тощо), так і безпосередньо підробкою змісту і голосу мовця шляхом імітації, або ж навпаки, шляхом зміни свого голосу із метою «відвести» від себе підозру.

Незважаючи на те, що проблема імітації дуже важлива та актуальна як із практичної точки зору, так із пізнавальної, але на сьогодні вона ще далека від розв'язання і повинна стати у майбутньому предметом спеціального дослідження.

Список літератури:

1. Andrzej Witold Szwarc [w:] Doktorzy honoris causa Uniwersytetu Wrocławskiego : 1948-2002 / Józef Rozynek, Józef J. Ziółkowski. Wrocław : Wydaw. Uniwersytetu Wrocławskiego, 2002. S. 297–299.
2. Bricker P.D., Pruzansky S. Effect of stimulus content and duration on talker identification. *JASA*, 1966, v. 40, № 6.
3. Капанадзе Г.У. К вопросу имитации некоторых просодических характеристик произнесения. *Теория систем автоматического управления*. Тбилиси, 1977.
4. Копелевич Ю.Х., Цверева Г.К. Христиан Готлиб Кратценштейн. Л. : Наука, 1989. 131 с.
5. Покровский М.Б. Расчет и измерение разборчивости речи. Москва : Связьиздат, 1962.
6. Pollack L., Pickett J.M., Sumbly W.H. On the identification of speakers by voice. *JASA*, 1954, v. 25, № 3.
7. Федорова И.А. Имитация длительности, интенсивности и частоты основного тона изолированного гласного. *Управление движением*. Л., 1970.
8. Шейкин Р.Л. Оценка голоса диктора по изолированным гласным звукам. *Анализ речевых сигналов человеком*. Л. : Наука, 1974.

References:

1. Andrzej Witold Szwarc [w:] (2002). Doktorzy honoris causa Uniwersytetu Wrocławskiego : 1948-2002 / Józef Rozynek, Józef J. Ziółkowski. Wrocław : Wydaw. Uniwersytetu Wrocławskiego, s. 297–299.
2. Bricker P.D., Pruzansky S. (1966). Effect of stimulus content and duration on talker identification. *JASA*, v. 40, № 6.
3. Kapanadze G.U. (1977). K voprosu imitatsii nekotorykh prosodicheskikh kharakteristik proizneseniya. *Teoriya sistem avtomaticheskogo upravleniya* [To the question of imitation of some prosodic characteristics of pronunciation. *Theory of automatic control systems*]. Tbilisi.
4. Kopelevich Yu.Kh., Tsverava G.K. (1989). Khristian Gotlib Krattsenshteyn [Christian Gottlieb Kratzenstein]. L. : Nauka, 131 s.
5. Pokrovskiy M.B. (1962). Raschet i izmereniye razborchivosti rechi [Calculation and measurement of speech intelligibility]. Moskva : Svyaz'izdat.
6. Pollack L., Pickett J.M., Sumbly W.H. (1954). On the identification of speakers by voice. *JASA*, v. 25, № 3.
7. Fedorova I.A. (1970). Imitatsiya dlitel'nosti, intensivnosti i chastoty osnovnogo tona izolirovannogo glasnogo. *Upravleniye dvizheniyem* [Simulate the duration, intensity and frequency of the pitch of an isolated vowel. *Motion control*].
8. Sheykin R.L. (1974). Otsenka golosa diktora po izolirovannym glasnym zvukam. *Analiz rechevykh signalov chelovekom* [Evaluation of the speaker's voice on isolated vowel sounds. *Human speech analysis*]. L. : Nauka.