

# 50 lat polskiej fonoskopii

## Wstęp

Rok 2010 to czas szczególny dla polskich ekspertów zajmujących się analizą nagrań dźwięku i badaniem mowy. Mija bowiem 50 lat od pierwszego orzeczenia polskiego sądu, dokumentującego pojawienie się nowego typu dowodów – nagrań dźwięku. Efektem tego orzeczenia było powstanie w naszym kraju nowego działu techniki kryminalistycznej, który nazwano fonoskopią.

W ciągu tego półwiecza fonoskopia szybko się rozwijała, co można wiązać nie tylko z postępem technicznym w zakresie metod i nośników utrwalania dźwięku, ale także z opracowywaniem nowych kryminalistycznych technik badania nagrań dźwiękowych i identyfikacji mówców. Nie bez znaczenia były także coraz szersze możliwości wykorzystania wyników ekspertyz fonoskopijnych tak w pracy wykrywczej organów ścigania, jak też dla potrzeb postępowania dowodowego przed sądem. Polska fonoskopia ma się czym pochwalić, przy czym na podkreślenie zasługuje fakt, że nie tylko adaptowała rozwiązania techniczne i badawcze stosowane w innych krajach, ale jako jeden ze światowych pionierów fonoskopii proponowała nowe podejście do tych badań i wypracowała własne algorytmy i techniki badań. Dlatego obchodzona właśnie rocznica jest dobrą okazją do przedstawienia krótkiego rysu historycznego i współczesnych możliwości badawczych fonoskopii.

## Początki fonoskopii



Ryc.1. Wiesław Bednarczyk, ekspert Zakładu Kryminalistyki KG MO, pracujący przy stanowisku z magnetofonem szpulowym  
Fig. 1. Expert of Militia Headquarters Forensic Institute  
źródło (ryc. 1–2): Archiwum CLK KGP

Na przełomie lat 50. i 60. XX wieku taśma magnetofonowa nie była jeszcze powszechnie używanym nośnikiem dźwięku. Magnetofonów było niewiele, wykorzystywano jedynie ciężkie i duże szpulowe urządzenia analogowe (ryc. 1).

Z uwagi na niepoddające się miniaturyzacji gabaryty, magnetofony te nie nadawały się do zastosowania jako odtwarzacze przenośne. Zaczęto je jednak wykorzystywać do rejestracji zjawisk akustycznych (rozmów), które mogły mieć znaczenie w procesie karnym. Ponieważ nagrania dźwięku stanowiły nową rzeczywistość, w ich sprawie głos zabrali sędziowie Sądu Najwyższego. W sprawie sygn. akt I K 196/60 z 23 XII 1960 roku SN orzekł:

„(...) taśma magnetofonowa z zarejestrowanym na niej przebiegiem całości lub poszczególnych fragmentów postępowania przygotowawczego czy też rozprawy ma charakter dokumentu (...) znaczenie prawne dowodu z taśmy magnetofonowej nie zostało przez ustawodawcę dotychczas wyraźnie określone, niemniej jednak nie istnieją żadne przeszkody do traktowania tej zdobyczy w sensie dowodu z dokumentów”<sup>1</sup>.

Terminu fonoskopia po raz pierwszy w literaturze kryminalistycznej użył A. Szwarz, łącząc greckie *phone* i *skoepo* (dźwięk i patrzeć) oraz tworząc teoretyczne podstawy tego działu techniki kryminalistycznej<sup>2</sup>. Po latach okazało się, że określenie to zostało powszechnie przyjęte w Polsce<sup>3</sup>, mimo że w innych krajach stosowany jest termin „akustyka sądowa” (ang. *Forensic Acoustics*, niem. *Forensische Akustik*). Warto dodać, że obecnie w Niemczech funkcjonuje kilka nazw na określenie różnych rodzajów badań związanych z fonoskopią, np. rozpoznawanie mowy (*forensische Spracherkennung*), fonetyka sądowa (*forensische Phonetik*) czy też rozpoznawanie mówcy i ocena nośników (*Sprechererkennung und Tonträgerauswertung*), przy czym to ostatnie określenie jest oficjalną specjalnością biegłych<sup>4</sup>.

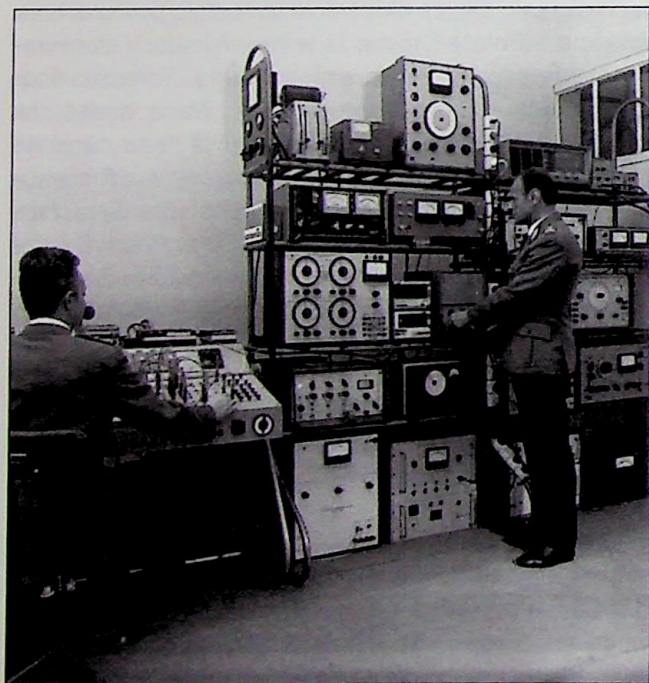
Należało jeszcze doprecyzować, czym fonoskopia ma się zajmować. Wykładnią dla sądów definiujących zakres badań fonoskopijnych stało się orzeczenie Sądu Najwyższego, który w sprawie sygn. akt III K 49/61 z 10 III 1961 orzekł:

„(...) dowód z taśmy magnetofonowej stanowiącej dowód rzeczowy wymaga przeprowadzenia z kolei dowodu na okoliczność identyczności zarówno utworzonych głosów jak i samej taśmy, a także braku w niej zmian”<sup>5</sup>.

Zakładano, że nowa dziedzina kryminalistyki powinna zajmować się nagraniami w szerokim zakresie: analizą mowy, treści rozmów oraz parametrów technicz-

nych zapisu magnetofonowego. Ze względu na zakres badań miała też reprezentować podejście interdyscyplinarne, dlatego powinna być uprawiana przez zespoły eksperckie, a nie – jak to jest praktykowane w wielu innych badaniach kryminalistycznych, np. w ekspertyzach pismoznawczych – przez indywidualnych biegłych. Pierwsza w Polsce komórka badań fonoskopijnych została utworzona na początku lat 60. XX wieku w Zakładzie Kryminalistyki Komendy Głównej MO. Na jej czele stanął późniejszy wieloletni naczelnik Wydziału Fonoskopii – Stanisław Błasikiewicz (ryc. 2). W trakcie wieloletniej pracy inspirował on szereg programów badawczych, których efektem były liczne, pionierskie publikacje. Do najważniejszych należy zaliczyć:

- S. Błasikiewicz, A. Miściuk, W. Wójcik: Podstawowy zakres badań fonoskopijnych prowadzonych w Zakładzie Kryminalistyki KG MO, „Problemy Kryminalistyki” 1967, nr 67–68, s. 303–327.
- W. Wójcik, S. Błasikiewicz: Podstawowe zasady pobierania materiału porównawczego do badań identyfikacyjnych mowy, „Problemy Kryminalistyki” 1969, nr 80–81, s. 577–589.
- S. Błasikiewicz: Metoda odsłuchu szeptu i mowy intensywnie zakłóconej, „Problemy Kryminalistyki” 1971, nr 90, s. 159–183.
- S. Błasikiewicz, W. Bednarczyk: Metoda badań autentyczności zapisu magnetofonowego, „Problemy Kryminalistyki” 1978, nr 131, s. 34–50.



Ryc. 2. Ekspersi Zakładu Kryminalistyki Komendy Głównej MO – Stanisław Błasikiewicz (stoi) i Wiesław Bednarczyk  
Fig. 2. Experts of Forensic Institute of Militia Headquarters: Stanisław Błasikiewicz (standing) and Wiesław Bednarczyk

Owocny etap pracy S. Błasikiewicza w charakterze biegłego nie zakończył się w 1991 roku, wraz z momentem jego przejścia na emeryturę. Nadal wydawał on opinie fonoskopijne jako ekspert Polskiego Towarzystwa Kryminalistycznego, wykorzystując swoje ogromne doświadczenie, za które wielokrotnie był honorowany. Zmarł w 2001 roku, zostawiając pokolenie wyszkolonych ekspertów, precedensowe opinie fonoskopijne oraz opracowania teoretyczne i praktyczne<sup>6</sup>.

Bez wątpienia Stanisławowi Błasikiewiczowi należy się miejsce wśród innych pionierów fonoskopii na świecie, do których zaliczają się: L.G. Kersta z USA – autor określenia *voiceprint*<sup>7</sup>, W. Gierasimow<sup>8</sup> z ZSRR i Ch. Koristka z NRD<sup>9</sup>.

Spadkobiercą dorobku pierwszej polskiej pracowni fonoskopii jest obecnie Centralne Laboratorium Kryminalistyczne Komendy Głównej Policji. Kolejnymi były pracownia fonoskopii Urzędu Ochrony Państwa oraz pracownia Instytutu Ekspertyz Sądowych. Obydwie ukształtowały metodykę badawczą pod wpływem dorobku CLK KGP.

Okres inicjacji prac nad metodami badań fonoskopijnych przypadł na lata, kiedy tworzono podwaliny teoretyczne współczesnych metod analizy sygnałów cyfrowych. Wystarczy wspomnieć prace B.P. Bogerta, M.J.R. Healya i J.W. Tukeya, które doprowadziły do analizy cepstralnej (rok 1963)<sup>10</sup>, czy F. Itakury i S. Saito, które stały się podstawą liniowego kodowania predykcyjnego (rok 1968)<sup>11</sup>.

Fonoskopia, będąc nauką interdyscyplinarną, od zawsze czerpała z osiągnięć nauk pokrewnych. Dobrym przykładem jest spektrograf. Urządzenie to zostało opracowane przez ośrodek pracujący na potrzeby wojskowe i dopiero po zakończeniu II wojny światowej zostało zaprezentowane użytkownikom cywilnym<sup>12</sup>. Szybko stało się podstawowym narzędziem pracy eksperta fonoskopii.

#### Wpływ ośrodków akademickich na rozwój fonoskopii w Polsce

Przedmiot badań fonoskopijnych w naszym kraju od lat pozostał ten sam – nagranie dźwiękowe, natomiast stopniowej modyfikacji ulegały metody badań. Jest to skutek ogromnego postępu technicznego, jaki dokonał się przez ostatnie lata. Przejawia się on m.in. w miniaturyzacji urządzeń rejestrujących, wprowadzeniu zapisu cyfrowego, rozpowszechnieniu telefonii komórkowej, poszerzeniu możliwości powszechnie dziś dostępnych komputerowych edytorów dźwięku. Wymaga to dostosowania metodyki do zmienności pola badań.

Materiałem dowodowym stały się zarówno nagrania rozmów bezpośrednich, jak i przeprowadzonych za pośrednictwem transmisji telefonicznej. Liczba wypowiedzi

dzi, odległość mówców od mikrofonu, warunki techniczne rejestracji – to tylko niektóre z czynników indywidualizujących nagrania. Nagrania dowodowe – nadsyłane do analiz fonoskopijnych przez prokuratury i sądy cechuje z reguły niska jakość. W wyniku niejawnej rejestracji, stosowania amatorskich urządzeń czy zużytych nośników na wypowiedzi dowodowe nakłada się szereg zakłóceń i zniekształceń, pogarszających percepcję słuchową.

Dlatego jednym z ważnych kierunków badań naukowych, kładących podwaliny pod rozwój ekspertyzy fonoskopijnej, była analiza sygnału mowy. Na tym polu wyraźnie zaznaczyła się działalność polskich naukowców; wśród nich pionierem był związany z Uniwersyteciem im. A. Mickiewicza w Poznaniu Wiktor Jassem, który już od 1952 r. prowadził badania m.in. nad stopniem dyskryminacji mówców na podstawie częstotliwości formantowych<sup>13</sup>. Prace W. Jassem'a stały się przez następne lata podstawą badań dla specjalistów z całego świata z dziedziny przetwarzania sygnału mowy, w tym ekspertów fonoskopii (prace te cytowane są na przykład przez H. Holliena i Ph. Rose'a)<sup>14</sup>. W nieco późniejszym okresie pojawiają się prace Wojciecha Majewskiego związanego z Politechniką Wrocławską, współpracującego z Oscarem Tosim – pionierem kryminalistycznej identyfikacji mówców w USA<sup>15</sup>. Pole badań W. Majewskiego było szerokie, koncentrował się on bowiem na analizie parametrów widmowych (innych niż częstotliwości formantowe, a dających się łatwiej wyekstrahować, co ma istotne znaczenie dla systemów automatycznych), takich jak widmo długoterminowe, częstotliwość przejść przez zero, częstotliwość tonu kraniowego<sup>16</sup>.

### Klasyczna ekspertyza fonoskopijna

Z dorobku naukowców wyrastała metodyka badań fonoskopijnych. Podstawowy zakres prowadzonych w Polsce badań fonoskopijnych składa się z trzech zagadnień, które zdefiniował Sąd Najwyższy. Taka trójetapowa opinia nazywana jest dzisiaj klasyczną ekspertyzą fonoskopijną. Składają na nią:

- spisanie treści nagrania,
- badanie autentyczności zapisu magnetofonowego,
- identyfikacja osób biorących udział w zdarzeniu.

Każdy z tych etapów badań zostanie pokrótce omówiony poniżej.

### Spisanie treści rozmów

Treść utrwalonych rozmów jest ustalana na bazie wrażeń słuchowych, biegły słucha nagrania i spisuje zasłyszane frazy. Metoda audytywna wykorzystuje pre-

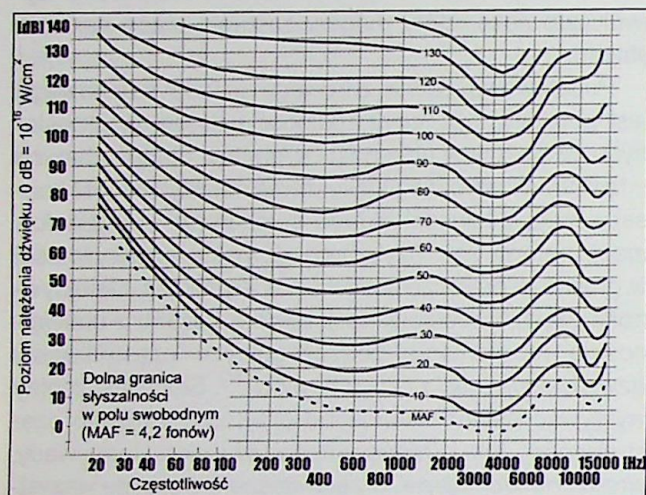
dyspozycje narządu słuchu, którego doskonałość, związana z czułością ucha ludzkiego jako systemu mechanicznego, zdecydowanie przewyższa precyzję najlepszych współczesnych mikroukładów<sup>17</sup>. Podręcznik biologii podaje: „Ucho ludzkie jest narządem słuchu tak dalece efektywnym, że dalsza ewolucja jego wrażliwości na dźwięki byłaby bezcelowa, gdyż umożliwiałaby rejestrowanie dźwięków towarzyszących ruchom cząsteczek powietrza, a to z kolei dostarczałoby męczących wrażeń w postaci stałego szumu, syku i brzęczenia”<sup>18</sup>. Dotychczas nie powstał żaden techniczny wytwór człowieka, który mógłby zastąpić pracę narządu słuchu.

Podstawową cechą zdrowego układu słuchowego jest selektywność, czyli zdolność usłyszenia określonych dźwięków w obecności sygnałów niepożądanych – tzw. tła akustycznego. Słyszenie selektywne jest konsekwencją zdolności układu słuchowego do transformacji parametrów fizycznych dźwięku: intensywności w głośność oraz częstotliwości w wysokość. Sелеktywność częstotliwościowa jest ściśle związana z umiejętnością rozróżniania poszczególnych częstotliwości (tzw. dyskryminacją częstotliwości)<sup>19</sup>. Słuch jest zdolny wykrywać pewne skończone przyrosty, zarówno częstotliwości, jak i intensywności dźwięku. Jego wady, wrodzone lub nabyte, przejawiające się przede wszystkim przesuniętą dolną granicą słyszalności, są elementem ograniczającym efektywność pracy eksperta.

Ryciny 3–6 opisują zdolność układu słuchowego człowieka do rozróżniania częstotliwości oraz intensywności dźwięku. Rycina 3 przedstawia krzywe jednakowego poziomu głośności. Charakterystyczne jest, że subiektywnie oceniana intensywność dźwięku dla dwóch tonów o tym samym natężeniu, lecz różnych częstotliwościach, może być zupełnie różna. Rycina 4 przedstawia zależność między względną zmianą częstotliwości a częstotliwością dla różnych poziomów bodźca. Wyniki te są prawdziwe tylko dla tonów (dźwięków sinusoidalnych), dźwięki złożone są rozróżniane również na podstawie różnic między harmonicznymi (na przykład tony 60 Hz i 62 Hz są nierozróżnialne, natomiast dźwięki złożone o takich częstotliwościach podstawowych mogą być rozróżniane dzięki różnicy między ich składowymi harmonicznymi)<sup>20</sup>. Na rycinie 5 pokazano tzw. powierzchnię słyszenia (wyznaczoną przez dolną granicę słyszalności – MAF oraz próg bólu) wraz z naniesionym zakresem sygnału mowy oraz muzyki. Rycina 5 uwidacznia fakt, że w zakresie sygnału mowy czułość słuchu jest największa. Rycina 6 obrazuje przykładowy audiogram, na którym wartość ubytku słuchu podawana jest dla wybranych częstotliwości w układach względnych. Na tego typu wykresach granica słyszenia normalnego, uzyskana w wyniku pomiarów statystycznych odpowiednio licznej grupy, re-

prezentowana jest na poziomie 0 dB. Oś pionowa na audiogramie to tzw. poziom HL (*Hearing Level*). Przykładowy audiogram obrazuje rycina 6. Widoczny na rycinie 6 niedosłuch typu odbiorczego ucha prawego dyskwalifikuje osobę charakteryzującą się taką ułomnością jako eksperta fonoskopii.

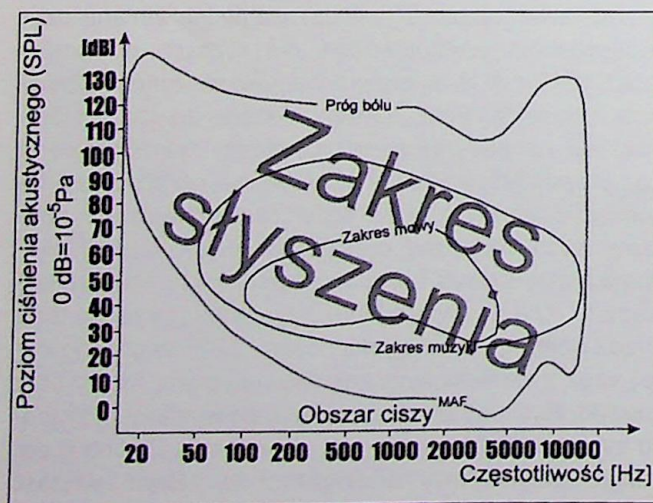
Oprócz indywidualnych uwarunkowań, w pracy eksperta bardzo istotny jest czynnik środowiskowy. Profesjonalne pracownie fonoskopijne powinny być tak projektowane, aby przy ich tworzeniu uwzględniano dwie



Ryc. 3. Krzywe jednakowego poziomu głośności (izofony) dla tonów wg Robinsona i Dadsona. Liczby nad krzywymi wskazują poziom głośności w fonach. Krzywe te zostały przyjęte przez Międzynarodowy Komitet Standaryzacyjny za charakterystyki normalnego słyszenia.

Fig. 3. Robinson-Dadson curves – equal loudness contours for pure tones. Numbers below curves indicate levels of loudness within phons. These curves are approved by the International Standards Committee as characteristics of normal hearing.

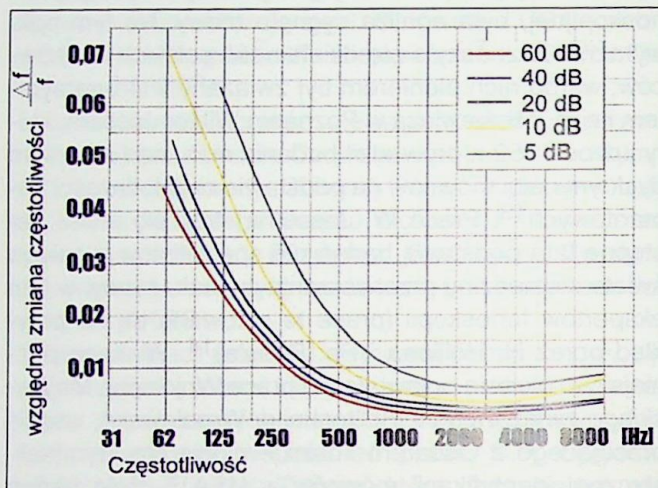
źródło (ryc. 3, 6): oprac. W. Maciejko na podstawie: J. Renowski „Akustyka psychofizjologiczna”, Wrocław 1974



Ryc. 5. Izofona MAF (dolna granica słyszenia) oraz próg bólu wyznaczają tzw. powierzchnię słyszenia

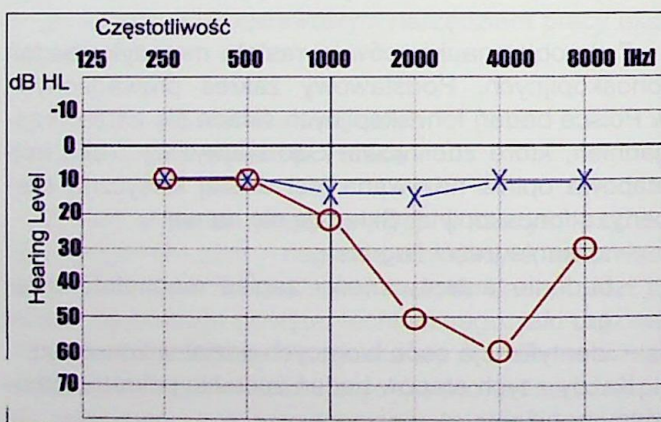
Fig. 5. MAF (threshold of hearing) and threshold of pain isophons are boundaries of so-called audible field

podstawowe właściwości ścian: pochłanianie przez nie dźwięku oraz izolacyjność akustyczną, zwłaszcza że zwykle materiały dobrze pochłaniające słabo izolują dźwięk i odwrotnie: ściana stanowiąca bardzo dobry izolator może niemal zupełnie nie pochłaniać dźwięku<sup>21</sup>. Izolację akustyczną stosuje się w celu zabezpieczenia wnętrza przed przenikaniem niepożądanych zakłóceń z zewnątrz, przy czym warunki akustyczne należy tak kształtować, aby poziom dźwięku tła akustycznego był minimalny. Kryteria oceny warunków



Ryc. 4. Najmniejsza wyczuwalna przez człowieka zmiana częstotliwości dla kilku poziomów ciśnienia akustycznego. Oś pionowa reprezentuje stosunek zmiany częstotliwości  $\Delta f$  do częstotliwości średniej  $f$ .

Fig. 4. Lowest change of frequency sensed by human ear. Vertical axis represents relation of frequency change (DELTA  $f$ ) to mean frequency ( $f$ ).  
źródło (ryc. 4–5): oprac. W. Maciejko na podstawie Z. Żyszkowski „Podstawy elektroakustyki”, Warszawa 1984



Ryc. 6. Krzywa niebieska prezentuje słyszenie w zdrowym uchu lewym, natomiast czerwona uwidacznia niedosłuch ucha prawego w zakresie wyższych częstotliwości

Fig. 6. Blue curve represents hearing in healthy left ear while the red one demonstrates hearing deficiency in right ear in the range of higher frequencies

akustycznych pomieszczenia zawarto w Polskiej Normie<sup>22</sup>, zgodnie z którą w pomieszczeniach do pracy umysłowej wymagającej silnej koncentracji uwagi ekwiwalentny poziom dźwięku A nie powinien przekraczać 35 dB. Drugim czynnikiem istotnie wpływającym na efektywność pracy eksperta jest średnie pochłanianie dźwięku przez ściany i wyposażenie pomieszczenia. Poprzez odpowiedni dobór materiałów zawieszanych na ścianie i suficie, kształtowana jest charakterystyka czasu pogłosu w zależności od częstotliwości. Czas pogłosu pomieszczeń ma ogromny wpływ na zrozumiałość mowy. Zbyt duży czas pogłosu powoduje, że dźwięki bezpośrednie i odbite interferują i zacierana jest ich wyrazistość. Z drugiej strony za krótki czas pogłosu, spowodowany zbyt dużym pochłanianiem, jest również niepożądany, gdyż skutkuje m.in. koniecznością prowadzenia odsłuchu przy większym „wysterowaniu sygnału”.

Kolejnym składnikiem laboratorium fonoskopijnego wpływającym na jakość ekspertyzy fonoskopijnej jest zestaw odsłuchowy, który powinien możliwie wiernie reprodukować zapis dźwiękowy (gwarantuje to wyrównana charakterystyka częstotliwościowa) oraz przenosić sygnał w pełnym zakresie. Podczas rozmieszczania urządzeń głośnikowych należy kierować się ogólnymi zasadami: ustawienie powinno być takie, aby nie powstawały cienie akustyczne, głośniki nie powinny być kierowane na przegrody twarde, moc głośnika powinna być dobrana do kubatury oraz warunków akustycznych pomieszczenia.

Stereotypowo uważa się, że odsłuch nagrań jest czynnością na tyle prostą, że może być prowadzony wszędzie i przez każdego. Praktyka ekspertów fonoskopii pokazuje jednak, że gorsze parametry stanowiska odsłuchowego wpływają na większą liczbę błędnych skojarzeń w stenogramie. Stąd spisujący treść rozmów dowodowych stosują odpowiednie techniki pozwalające na uzyskanie możliwie najwierniejszego przekładu rozmowy.

Treść spisana w trakcie pierwszego odsłuchania nagrania dowodowego zwykle zawiera błędy, zmieniające sens poszczególnych wypowiedzi i rolę konkretnych osób w zarejestrowanym zdarzeniu. Całkowite uniknięcie błędów odsłuchu wydaje się wprawdzie niemożliwe, jednak istnieją dwa podstawowe sposoby ich skorygowania w trakcie badań. Są to: kolejny odsłuch i odsłuch przez kolejnego eksperta.

Technika kolejnych odsłuchów to podstawowa metoda zwiększania dokładności odtwarzania treści nagrania dowodowego. W polskiej praktyce fonoskopijnej wypracowano następujący schemat: pierwsze odsłuchanie całego materiału dowodowego jest prowadzone pod kątem wstępnego spisania treści nagrania; następnie biegły słucha ponownie całości materiału

i uszczegóławia zwroty wątpliwe. Liczba powtórzeń odsłuchu jest kwestią indywidualną, zależną od predyspozycji psychofizycznej eksperta, jego doświadczenia w pracy z nagraniami, jak również od specyfiki nagrania. Wydaje się jednak, że możliwość ujawniania błędów przez stosowanie kolejnego odsłuchu jest ograniczona; nie istnieje bowiem prosta zależność: im więcej razy powtórzono odsłuch, tym treść jest wiernej spisana. Z każdym następnym odsłuchem rośnie rola sugestii spisanej treścią i maleje spostrzegawczość osoby odsłuchującej nagranie. Błędne skojarzenia są wynikiem zdolności rekonstrukcji szczątkowej informacji lingwistycznej, dokonywanej przez słuchacza posługującego się w tym celu zasobami pamięciowymi.

Biegły prowadzący spisanie treści na podstawie wielokrotnego odsłuchu wytwarza sobie pamięciowy obraz zdarzenia, przez co zwiększa swą podatność na autosugestie. Eliminacji błędnych skojarzeń służy weryfikacja odsłuchu przez innego eksperta, który dokonuje sprawdzenia spisanej treści nagrania dopiero w końcowym etapie tych badań, gdy pierwszy ekspert już przeprowadził kilkukrotny odsłuch. Na zagadnienie współpracy biegłych zwracał już uwagę S. Błasikiewicz: „(...) najlepsze rezultaty w rekonstrukcji treści szeptu i mowy intensywnie zakłóconej uzyskuje się w odsłuchu zespołowym”<sup>23</sup>. Metoda zaproponowana przez Błasikiewicza odznaczała się dużą pracowitością, gdyż była opracowana dla biegłych dokonujących ręcznego zapisu spisanej treści: „Poszczególne osoby zapisują, niezależnie od siebie, odsłuchaną treść. Po zakończeniu analizy kolejnego fragmentu następuje konfrontacja poszczególnych wyników. W przypadku uzyskania różnych wyników odsłuchu, powtarza się analizę spornych fragmentów”<sup>24</sup>. Obecnie weryfikacja wyników odsłuchu przebiega znacznie sprawniej, dzięki wykorzystaniu komputerowego edytora tekstu.

Konieczność przełożenia na spisaną treść subtelnej gry intonacją, akcentowaniem, głośnością i tempem wypowiedzi, stosowaniem celowych pauz między pytaniem i odpowiedzią itp. znacznie zubaża obraz odsłuchiwanej rzeczywistości. Środki takie, jak wykrzyknik, znak zapytania czy komentarz eksperta nie wystarczają do obiektywnego oddania złożoności komunikacji słownej. Element subiektywnej oceny przez biegłego posłyszanych słów jest nieuchronny. Spisana treść jest więc tylko przybliżonym zapisem przebiegu wydarzeń. Dla uzyskania pełnego obrazu zarejestrowanego zdarzenia konieczne jest odsłuchanie nagrania dowodowego – nawet mimo dysponowania spisanej treścią – przez prokuratora czy sędziego.

Spisana treść rozmów jest logicznie powiązana z oceną autentyczności zapisu i identyfikacją rozmówców. Może być polem do wskazania prób ewentualne-

go montażu, umożliwia precyzyjne przypisanie konkretnym osobom poszczególnych wypowiedzi. Spełnia funkcję swoistego szkieletu, z którym nierozzerwalnie łączy się wszelkie inne wyniki pracy biegłego z zakresu fonoskopii.

Często zdarza się, że do badań fonoskopijnych trafiają nagrania zakłócone i zniekształcone. Poprawa ich jakości, przejawiająca się polepszeniem wyrazistości i zrozumiałości wypowiedzi dowodowych, może być niekiedy uzyskana w wyniku korekcji, polegającej na tłumieniu bądź wzmacnianiu poszczególnych częstotliwości<sup>25</sup>.

Jednak oczekiwania organu procesowego – dysponenta nagrań, który liczy na spektakularne efekty takich zabiegów, nie zawsze możliwe są do spełnienia. Powszechne mniemanie o nieograniczonych możliwościach usuwania dźwięków niepożądanych klóci się bowiem z podstawową zasadą fizyki, że dwie lub więcej fal może przebiegać ten sam obszar niezależnie od siebie (zasada superpozycji)<sup>26</sup>. Konsekwencją jest niemożność usunięcia sygnału zakłócającego bez utraty jakości w konkretnych częstotliwościach, w których została dokonana ingerencja. Występuje tu analogia do fotografii: obrazy z dalszego tła przy niewłaściwym ujęciu przesłonięte bliższymi obiektami są bezpowrotnie tracone. Dlatego jeszcze przed podjęciem rejestracji należy zadbać o jak najlepszą jakość rejestrowanego sygnału i odrzucić stereotypowy pogląd, że wyposażenie pracowni fonoskopijnej jest tak zaawansowane technicznie, że pozwoli oczyścić zapis z zakłóceń.

Nie należy jednak w ogóle odrzucać korekcji jako procesu, chociaż jej wartość w warunkach fonoskopijnych jest na ogół niewielka. Ponieważ o jej powodzeniu decyduje wrażenie słuchowe eksperta, tak więc zasadność przeprowadzenia korekcji należy pozostawić indywidualnej ocenie odtwarzającego nagranie. Warto przy tym przytoczyć pogląd S. Błasikiewicza, który podchodząc z ostrożnością do tego zagadnienia, tak pisał o zakłóceniach szumowych: „(...) całkowite wyeliminowanie zakłóceń szumowych – jednakże kosztem częściowego ubytku informacji – jest możliwe tylko wtedy, gdy występują one w wąskim paśmie na określonej częstotliwości. Dokonując korekcji częstotliwościowej, należy uwzględnić wyrazistość i zrozumiałość głosek polskich w różnych pasmach przenoszenia. W przeciwnym przypadku wystąpią nie tylko trudności w odsłuchu, lecz także może to doprowadzić do fałszywych skojarzeń (...)”<sup>27</sup>.

Należy zatem podsumować, że stopień efektywności korekcji to cecha indywidualna nagrania, a jej efekty nie mogą być uogólniane na inne zapisy, nawet jeżeli zostały one utrwalone z wykorzystaniem analogicznego toru transmisji sygnału audio.

### Badanie autentyczności nagrania

W sentencji cytowanego już orzeczenia SN z 1961 roku pojęcie autentyczność zostało zdefiniowane jako „brak zmian w nagraniu”. Określenie to wymaga szerszego komentarza<sup>28</sup>.

S. Błasikiewicz i W. Bednarczyk proponują definicję badań autentyczności nagrań dźwięku, zgodnie z którą główny nacisk kładzie się na ciągłość zapisu całej rozmowy<sup>29</sup>. Stanowisko takie jest bliskie ogólnej definicji autentyczności dokumentów (nagranie rozmowy bywa nazywane audiodokumentem<sup>30</sup>) przytoczonej w pracy „Technika kryminalistyczna”, przeciwstawiającej pojęciu autentyczny słowo sfalszowany<sup>31</sup>. I choć Słownik języka polskiego podaje, że termin autentyczny może być używany zamiennie z terminem oryginalny<sup>32</sup>, to jednak – w kontekście badań fonoskopijnych – znaczenie obu terminów jest odmienne.

Nagranie jest autentyczne, gdy wiernie oddaje przebieg całego utrwalonego zdarzenia, w szczególności rozmowy<sup>33</sup>. Zastrzeżenie to dotyczy nagrania jako czynności utrwalenia, gdyż pojęcie autentyczność nagrania rozmowy ma inne znaczenie niż autentyczność rozmowy. Okoliczności przeprowadzenia rozmowy i treść wypowiedzi mogą zostać sfingowane (rozmówcy odgrywają ustalone role), ale jej zapis będzie autentyczny, tzn. magnetofon zostanie włączony na nagrywanie przed rozpoczęciem rozmowy i wyłączony po jej zakończeniu. Biegły fonoskopii, badając autentyczność nagrania rozmowy, wskazuje wszelkie nieciągłości zapisu i ewentualnie określa przyczyny ich wystąpienia (np. koniec taśmy w kasie skutkujący niemożnością zarejestrowania zakończenia rozmowy, wyczerpanie baterii zasilających magnetofon). Jeżeli jednak ma wątpliwości co do autentyczności nagrania, nimi również powinien podzielić się ze zlecniodawcą badań.

Nagranie dowodowe może być wielokrotnie kopiowane, a mimo to – choć nieoryginalne – może być autentyczne. Oczywiście nie każda kopia wiernie odzwierciedla przebieg zdarzenia utrwalonego w zapisie oryginalnym. W praktyce bowiem ujawniano niekiedy, że kopiowanie – poprzez pogorszenie jakości, spowodowane np. zaszumieniem – miało być środkiem do ukrycia śladów montażu.

Praktyka fonoskopijna dostarcza przykładów, które nie pozwalają literalnie stosować wytycznych Sądu Najwyższego w kwestii „braku zmian”. Nie ulegnie zaburzeniu autentyczność dowodu fonoskopijnego w powyżej podanym rozumieniu, jeżeli oryginalny zapis na analogowej taśmie magnetofonowej zostanie wiernie przekopiowany na cyfrowy dysk CDR. Nagranie nie traci przymiotu autentyczności, jeżeli zostanie poddane np. korekcji częstotliwościowej, mającej na celu usunięcie szumów i poprawę zrozumiałości wypowiedzi do-

wodowych. Tu jednak należy postawić pytanie, czy ingerencja w zapis nie była na tyle silna, że w wyniku korekcji usunięto jakiegokolwiek sygnały mogące mieć wpływ na interpretację dowodu z nagrania.

W polskiej praktyce fonoskopijnej badania autentyczności nagrań są prowadzone od 1965 roku. Początkowo wykonywano je nawet wtedy, gdy nie zostały zakreślone w postanowieniu o powołaniu biegłego<sup>34</sup>. Praktyka ta miała służyć wykryciu fałszywych dowodów, przedstawianych w celu postawienia w złym świetle drugiej strony sporu sądowego. Obecnie coraz częściej wspomniana zasada nie znajduje zastosowania, ponieważ rutynowe uzupełnianie badań ponad zleczone postanowieniem klóci się z obowiązkiem wykonywania przez biegłego badań wyłącznie w zleconym zakresie. Biegły, który dostrzega potrzebę przeprowadzenia dodatkowych badań, powinien ten fakt uzgodnić ze zlecającą.

W badaniach autentyczności nagrań analogowych stosowane są trzy podstawowe metody:

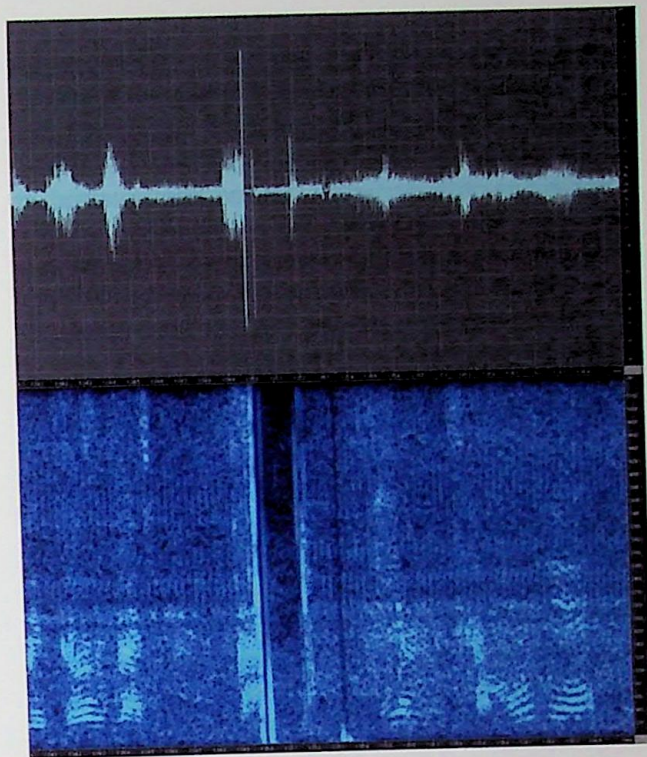
- wizualizacyjna – wymaga specjalnej aparatury optycznej, badaniu podlega taśma ferromagnetyczna, na której zarejestrowano kwestionowane nagranie. Badania opierają się na zjawisku Faradaya, które polega na zdolności niektórych przezroczystych materiałów do skręcania płaszczyzny polaryzacji przechodzącego przez nie światła pod wpływem pola magnetycznego<sup>35</sup>,
- analiza czasowo-częstotliwościowa z wykorzystaniem spektrografii (wskazująca ślady nieciągłości na obrazie widmowym sygnału akustycznego)<sup>36</sup>,
- audytywna (analiza logiczności prowadzenia nagranej rozmowy)<sup>37</sup>.

W przypadku nagrań analogowych możliwe są następujące techniki ingerencji w ciągłość zapisu zdarzenia:

- zatrzymanie rejestracji poprzez wciskanie przełącznika Start/Stop, wciskanie przełącznika Pause lub odłączenie magnetofonu od zasilania,
- montaż zapisu za pomocą cięcia i sklejania,
- kasowanie,
- dogrywanie rozmowy.

Ryciny 7–8 przedstawiają dwa przykłady wykorzystania analizy czasowo-częstotliwościowej, zatrzymanie rejestracji i ponowne uruchomienie Start/Stop (ryc. 7), wcisknięcie Pause (ryc. 8).

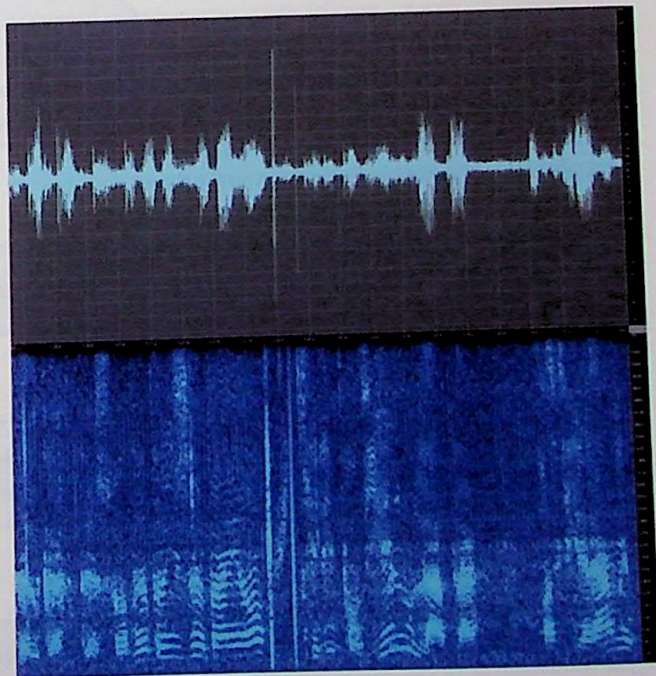
W XXI wieku na coraz większą skalę zaczęto przekazywać do badań fonoskopijnych zapisy cyfrowo utrwalonych rozmów, natomiast coraz rzadziej przedmiotem takich badań są nagrania analogowe. Wpływają z tego istotne konsekwencje dla wyboru metody badań i oceny ich wyników. W szczególności wobec zupełnie innej techniki zapisu nagrań cyfrowych nie ma zastosowania metoda wizualizacyjna, zaś w bardzo



Ryc. 7. Zatrzymanie zapisu poprzez wciśnięcie klawisza Stop oraz ponowne uruchomienie (Start)

Fig. 7. Interruption in recording by pressing „Stop” button and restarting by „Start” button

źródło (ryc. 7–8, 10–13): W. Maciejko



Ryc. 8. Nieciągłość zapisu, spowodowana najprawdopodobniej przez wciśnięcie klawisza Pause

Fig. 8. Discontinuity of recording most probably caused by pressing „Pause” button

ograniczonym zakresie może być wykorzystywana metoda spektrograficzna oraz audytywna. Jest to skutek tego, że ewentualna próba montażu nagrania cyfrowego zwykle nie powoduje powstania tak widocznych znamion ingerencji w zapis, jak na przykład znaki głowic magnetycznych w przypadku zapisu analogowego na taśmie ferromagnetycznej.

Obecnie najbardziej skuteczną metodą badania autentyczności nagrań cyfrowych jest badanie przebiegu tzw. przydźwięku sieciowego. Analiza taka opiera się na dwóch podstawowych założeniach:

1. stała częstotliwość prądu zmiennego w sieciach elektroenergetycznych wynosząca 50 Hz w rzeczywistości ulega niewielkim wahaniom  $f_{\text{prądu\_zmiennego}} = 50 \pm \Delta f [\text{Hz}]$ ,
2. chwilowe wahania wartości częstotliwości są przypadkowe, a ich sekwencje w kolejnych chwilach czasu niepowtarzalne<sup>38</sup>.

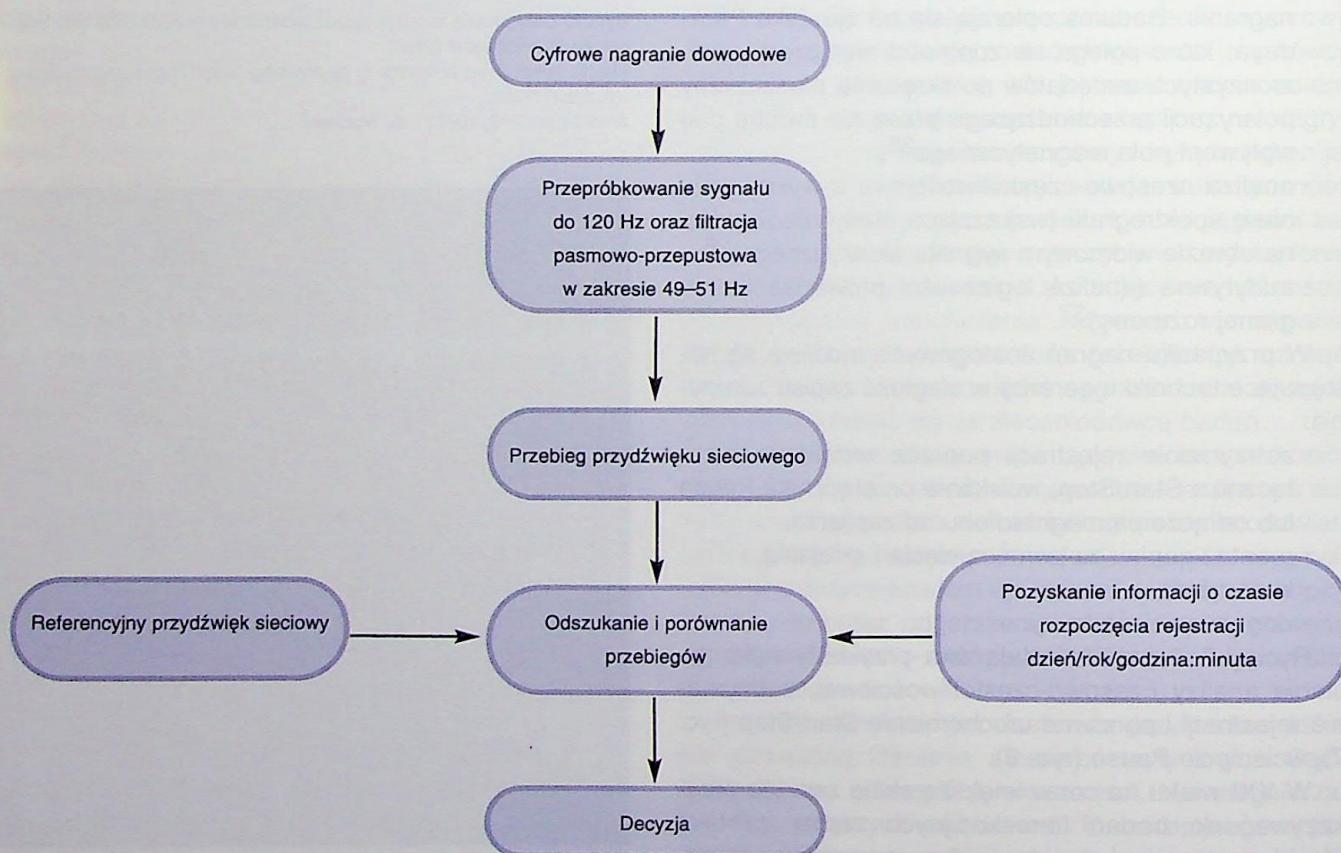
Wykorzystanie wskazanej metody jest możliwe pod warunkiem zarejestrowania się przydźwięku w trakcie zapisu rozmowy dowodowej oraz dysponowania prze-

biegiem referencyjnym. Odchylenia częstotliwości są w sposób ciągły monitorowane oraz archiwizowane przez operatora; w Polsce są nim Polskie Sieci Elektroenergetyczne Operator S.A. Schemat postępowania w procesie badawczym przedstawiono na rycinie 9.

Ryciny 10–11 przedstawiają przydźwięk sieciowy zarejestrowany w dwóch nagraniach, różnymi dyktafonami cyfrowymi, w tym samym czasie. Na rycinie 10 przedstawiono przebieg referencyjny, do którego odnieszono przebieg wyekstrahowany w drugim nagraniu, który ilustruje rycina 11.

Strzałki wskazują punkty charakterystyczne, za pomocą których można w przedmiotowym przypadku stwierdzić zgodność przebiegów.

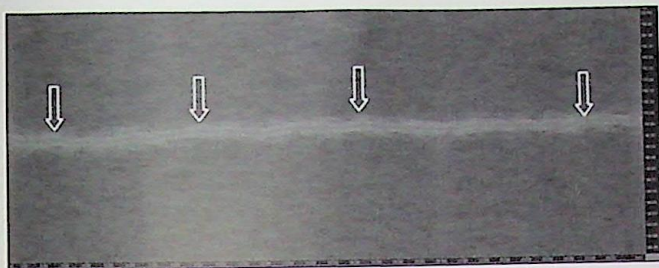
W realiach cyfrowych nabiera znaczenia wątpliwość, czy niewykrycie śladów zaburzenia autentyczności zapisu może być podstawą zawarcia w opinii sformułowania o autentyczności nagrania dowodowego. Czy raczej należy – domniemając, że istnieją techniki fałszerstw, których być może eksperci nie potrafią wykrywać – wstrzymać się od kategori-  
cznych sfor-



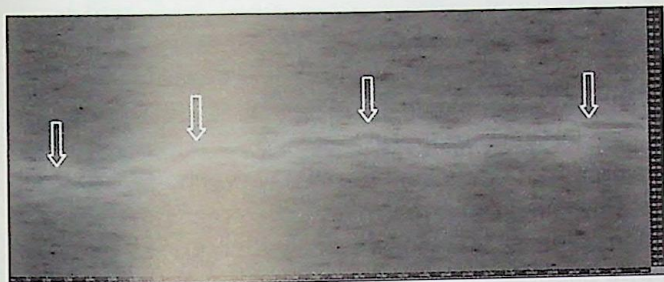
Ryc. 9. Algorytm postępowania w procesie porównania przydźwięku w zapisie referencyjnym oraz w nagraniu dowodowym

Fig. 9. Examination algorithm in process of comparing hum in referential and evidential recordings

źródło: oprac. W. Maciejko na podstawie Grigoras C.: Digital Audio Recording Analysis. The Electric Network Frequency Criterion. Catalin Grigoras, Ph. D, IAFPA 2003



Ryc. 10. Przebieg referencyjny  
Fig. 10. Referential curve



Ryc. 11. Przebieg w nagraniu porównywanym z zapisem referencyjnym  
Fig. 11. Curve in recording compared to referential curve

mułowań? Warto pamiętać, że tego typu wątpliwości towarzyszą przedstawicielom innych zawodów czy specjalności; postępując zgodnie z tym rozumowaniem, onkolog nie powinien podawać leków pacjentom, wiedząc, że w określonym procencie przypadków są one nieskuteczne. Zdaniem autorów dopuszczalne jest kategoryczne wypowiadanie się o autentyczności nagrań, jednakże pod warunkiem, że specyfika nagrania pozwala na zastosowanie kilku wiarygodnych metod, które nie wskażą miejsc mogących świadczyć o nieciągłości zapisu.

Stosując mianownictwo używane przez badaczy dokumentów, można stwierdzić, że niemal wszystkie próby fałszowania nagrań dowodowych są dokonywane przez przerabianie<sup>39</sup>. Sporadycznie zdarzają się przypadki, kiedy dokonuje się podrobienia, czyli rejestracji zdarzenia, którego okoliczności zostały sfingowane. Praktyka kryminalistyczna pokazuje bowiem, że trudność montowania (przerabiania) zapisów dowodowych wynika z braku możliwości dowolnego usuwania (względnie dodawania) fragmentów nagrania tak, żeby zachować logiczną spójność wypowiedzi uczestników dyskursu, a jednocześnie wszelkich zdarzeń akustycznych współtowarzyszących – w tle akustycznym – rozmowie dowodowej. Niemniej, szczególnie w odniesieniu do zapisów cyfrowych, warunkiem poprawności wnioskowania na podstawie badań autentyczności zapisu jest adekwatność zastosowanych metod do specyfiki nagrania. Dlatego na przykład izolowane użycie metody audytywnej w badaniu profesjonalnych zapisów cyfrowych z założenia jest błędem w sztuce.

## Identyfikacja mówców

Wprowadzenie do zagadnienia identyfikacji mówców wymaga przedstawienia podstawowych informacji o mówieniu.

Zdolność mówienia to jeden z najistotniejszych atrybutów człowieczeństwa. I. Styczek definiuje mowę jako „zespół czynności wykonywanych w podobny sposób przez członków danej społeczności językowej w celu komunikowania się”<sup>40</sup>. Z kolei K. Dejna określa mowę jako „procesy i wytwory porozumiewania się ludzi między sobą za pomocą wyrazów, stanowiących formę dźwiękowego przedstawienia pojęć przez nie oznaczanych”<sup>41</sup>. Na biologiczne i kulturowe uwarunkowania posługiwania się mową zwraca uwagę A. Prusiewicz: „jest [mowa – przyp. aut.] zjawiskiem intelektualnym, komponentem czynności poznawczej, ale też czynnikiem zachowań społecznych”<sup>42</sup>. Związek procesu mówienia z czynnością ośrodkowego układu nerwowego potwierdzają wyniki badań z zakresu neuropsychologii. Posługując się metodą kliniczną, wykazano, że poszczególne uszkodzenia mózgu znajdują odzwierciedlenie w konkretnych i powtarzalnych (u innych pacjentów z taką samą dysfunkcją) zaburzeniach mowy<sup>43</sup>.

Badania jednoznacznie wskazują na wiodącą rolę naśladownictwa w procesie kształtowania się mowy u dzieci<sup>44</sup>. Dzieci rodzą się z predyspozycjami do mówienia, takimi jak np. sprawny słuch i narządy zaangażowane w mówienie, prawidłowa czynność układu nerwowego, odpowiedni poziom intelektualny. Na pewnym etapie rozwoju dziecko dostrzega, że rodzice i osoby z najbliższego otoczenia porozumiewają się za pomocą mowy, również do niego zwracają się słowami. Jest to kolejny niezbędny czynnik właściwego rozwoju mowy – stymulacja socjopsychologiczna<sup>45</sup>.

Do dziecka dociera mieszanina wypowiedzi, te zgodne z regułami gramatycznymi przeplatają się z mową niespójną czy celowymi zdrobnieniami<sup>46</sup>. Zawsze jednak są to wypowiedzi przedstawicieli określonego kręgu kulturowego, środowiska, w którym dziecko się wychowuje. Przedstawiciele tej społeczności reprezentują określoną grupę zawodową, cechuje ich określony status materialny, przynależność do grupy społecznej władającej określonym dialektem. Jak podaje Dejna: „dialekt formuje się w obrębie obszaru etniczno-językowego dzięki wytwarzaniu się na części jego terytorium cech dialektalnych w odizolowanej geograficznie, politycznie, gospodarczo czy kulturowo podgrupie etnicznej. Dialekty więc są językami regionalnych grup etnicznych i różnią się od pozostałych dialektów szeregiem sobie właściwych innowacji o ograniczonym geograficznie zasięgu”<sup>47</sup>.

Według badań populacyjnych moment przyswojenia sobie języka i reguł gramatycznych w stopniu pozwala-

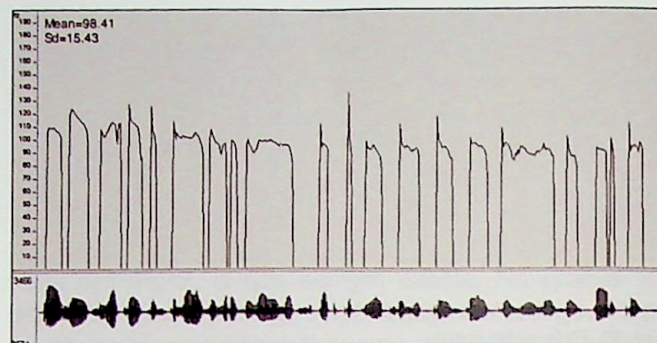
jącym na komunikację słowną i umiejętność artykulacji wszystkich głosek przypada na 5–6 rok życia<sup>48</sup>. Język znany dziecku w tym wieku został ukształtowany w przeważającej mierze przez opiekunów. W dalszych etapach życia jego znajomość ulegać będzie przemianom m.in. pod wpływem grupy, z którą osobnik się utożsamia, środków społecznego przekazu, miejsc kształtowania światopoglądu, takich jak szkoła, miejsce pracy itd. Wraz ze zdobywanym wykształceniem i migracją do większych ośrodków miejskich, język będzie nabierał coraz więcej cech języka ogólnonarodowego, mającego w szczególności cechy dialektu grupy przodującej w życiu społecznym i politycznym<sup>49</sup>. Rozszerzy się również zbiór używanych pojęć, bez których niemożliwe jest komunikowanie np. niektórych idei abstrakcyjnych czy dyskusja nad zagadnieniami branżowymi<sup>50</sup>.

W kontekście wprowadzenia do identyfikacji mówców należy także wspomnieć wybrane zagadnienia z zakresu lingwistyki. Kompetencja językowa to zdolność posługiwania się językiem<sup>51</sup>. Ulega modyfikacjom w ontogenezie, np. poprzez rozwój słownika pojęć fachowych, przyswajanych w związku z wykonywaną pracą czy nabycie biegłości w stosowaniu złożonych konstrukcji składniowych. Poziom kompetencji językowej indywidualizuje mówców, pozwala – pod warunkiem dysponowania obszernymi wypowiedziami – na ustalenie poziomu inteligencji i wykształcenia, rodzaju wykonywanej pracy itp.

Reasumując – indywidualności mowy, których wydzielenie jest istotą w kryminalistycznej identyfikacji mówców – mogą mieć swoje źródło w uwarunkowaniach biologicznych (w tym w czynnikach związanych zarówno z anatomią, jak również układem nerwowym), społecznych oraz kulturowych. Ich zbiór wydaje się nieskończony, choćby dlatego że poszczególne cechy mogą występować w różnym nasileniu<sup>52</sup>. Metoda identyfikacji mówcy, której celem jest wyszczególnienie cech mowy, to tzw. metoda językowa.

Dopełnieniem analizy językowej w procesie kryminalistycznej identyfikacji mówcy jest metoda pomiarowa. Najogólniej, opiera się ona na ilościowym określeniu cech głosu. Obecnie ilościowe cechy głosu otrzymywane są wyłącznie na drodze przekształceń matematycznych charakterystyki widmowej sygnału. Istnieje kilka podstawowych parametrów takiej analizy, wśród których do dzisiaj najchętniej wykorzystywanymi przez polskich ekspertów są ton kraniowy oraz częstotliwości formantowe.

Model wytwarzania sygnału mowy jest prezentowany w postaci układu źródło – filtr. Dla głosek dźwięcznych pobudzenie filtru następuje sygnałem okresowym, którego źródłem jest krtań, częstotliwość tego sygnału nazywana jest tonem kraniowym. Na rycinie 12 przedstawiono przykładowy przebieg tonu kraniowego.



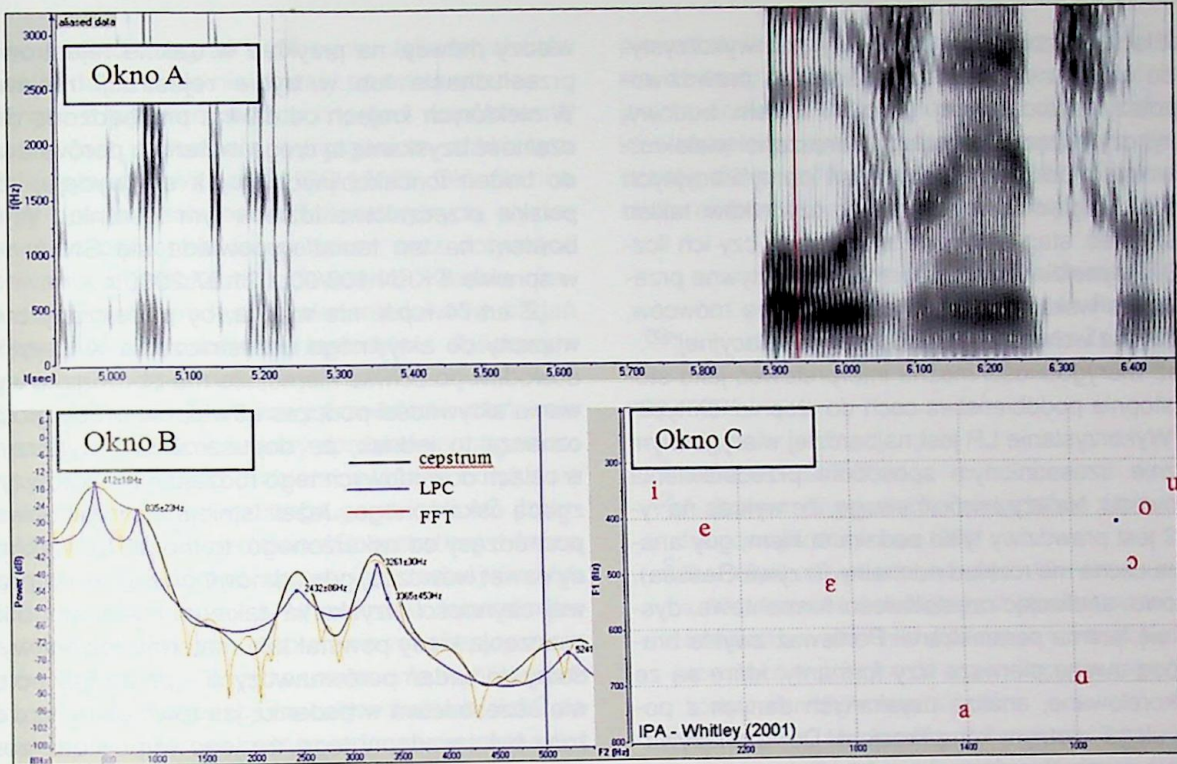
Ryc. 12. Przebieg tonu kraniowego oraz przebieg czasowy sygnału  
Fig. 12. Laryngeal tone and signal in time curves

Porównywanie częstotliwości tonu kraniowego – wobec licznych fluktuacji – wymaga analogicznych wypowiedzi dowodowych i porównawczych, ze szczególnym uwzględnieniem chwilowego stanu psychofizycznego, intonacji zdaniowej i siły wydechu<sup>53</sup>. Wielomilionowe populacje mówców cechują zbliżone wartości tonu kraniowego ( $F_0$ ), przeciętne ramy dla mężczyzn wynoszą 100–200 Hz, dla kobiet 200–400 Hz<sup>54</sup>. Dlatego  $F_0$  nie należy do głównych wyznaczników indywidualności mówienia, ma natomiast określoną wartość różniczącą. Zgoda odmienny wynik pomiaru tonu kraniowego w przebiegu dwóch zdań wygłoszonych w analogicznych warunkach (np. w zbliżonym stanie emocjonalnym) pozwala sądzić, że zostały one wypowiedziane przez różnych mówców. Parametryczne różnicowanie mówców opiera się na analizie formantowej.

Formanty samogłoskowe pod względem częstotliwości są ściśle zależne od budowy fizjologicznej toru głosowego i jako takie stanowią cechy osobnicze<sup>55</sup>. Reprezentacją graficzną formantów są lokalne wartości maksymalne w widmie danej głoski.

Zmienność budowy anatomicznej skutkuje różnorodnością barwy głosu w populacji ludzkiej<sup>56</sup>. Może być ona prezentowana w postaci wewnątrzpopulacyjnych różnic rozkładu struktur formantowych. Z punktu widzenia fizyka, mowa to przebiegi akustyczne zmienne w czasie w dziedzinie częstotliwości i natężenia<sup>57</sup>. Częstotliwościowy rozkład energii w sygnale mowy ilustruje spektrogram<sup>58</sup>. Rycina 13 przedstawia jedną z metod pomiaru częstotliwości formantowych. Głoski dźwięczne lokalizowane są audytywnie oraz za pomocą spektrogramu (okno A). Precyzyjne odczytanie częstotliwości następuje za pomocą charakterystyki częstotliwościowej (okno B). Za pomocą okna C dokonywana jest weryfikacja poprawności odczytanych wartości.

Na różnice w położeniu i kształcie formantów wpływa również kontekst fonetyczny i intonacja, ekspresja oraz impresja wypowiedzi. Fluktuacje parametrów sygnału mowy – warunkowane wieloczynnikowo – bywają bardzo znaczne, dlatego dla wykazania indywidual-



Ryc. 13. Metoda ekstrakcji częstotliwości formantowych (okno A – spektrogram, okno B – chwilowe widmo sygnału mowy, okno C – rozkład formantów samogłosek)

Fig. 13. Formant frequencies extraction method (window A – spectrogram, window B – momentary spectrum of speech signal, window C – distribution of vowel formants)

ności budowy anatomicznej aparatu mowy konieczne jest porównywanie struktur formantowych w analogicznych ciągach alofonów, wypowiedzianych w możliwie zbliżonym stanie emocjonalnym. Kolejnym warunkiem wykorzystania korelacyjnych pomiarów struktur formantowych jest relatywnie dobra jakość nagrania, gwarantująca odpowiednio szerokie pasmo częstotliwości zarejestrowanego sygnału (ujawniające kolejne formanty) i taki odstęp sygnał-szum, że formanty nie ulegają zatarciu w tle akustycznym.

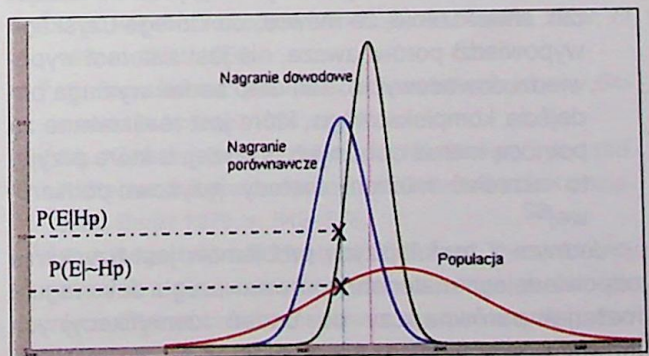
Kluczowym zagadnieniem jest szacowanie wartości dowodu z badań. Jeden ze sposobów wnioskowania opiera się na skrótowo opisanym poniżej modelu, uwzględniającym iloraz wiarygodności (*Likelihood Ratio* LR). Jeżeli założono alternatywne hipotezy:

1.  $H_p$  – hipoteza oskarżenia,
2.  $\sim H_p$  – hipoteza obrony,

to iloraz wiarygodności obliczany jest jako stosunek prawdopodobieństwa obserwacji dowodu (E), gdy prawdziwa jest hipoteza  $H_p$ , do prawdopodobieństwa obserwacji dowodu, gdy prawdziwa jest hipoteza alternatywna  $\sim H_p$ .

$$LR = \frac{P(E|H_p)}{P(E|\sim H_p)}$$

Jeżeli  $LR > 1$ , to wspierana jest wersja oskarżenia, natomiast w przypadku przeciwnym  $LR < 1$  – wspierana jest hipoteza obrony<sup>59</sup>. Rycina 14 przedstawia interpretację graficzną ilorazu wiarygodności w odniesieniu do analizy formantowej.



Ryc. 14. Rozkłady gęstości formantu F2 dla mowy w nagraniu dowodowym, porównawczym oraz w populacji

Fig. 14. Distributions of F2 densities in evidential and comparative recordings, and in population

źródło: W. Maciejko na podstawie P. Rose: Technical forensic speaker recognition: Evaluation, types and testing of evidence, „Computer Speech and Language” 2006, nr 20

Populacja to zbiór mówców, który jest wykorzystywany do określenia prawdopodobieństwa prawdziwości hipotezy alternatywnej ( $\sim H_p$ ). Problem budowy, struktury czy licznosci populacji poruszano wielokrotnie, również w odniesieniu do badań identyfikacyjnych mówców. Uwzględnić należy szereg czynników, takich jak: wiek, płeć, stan emocjonalny mówców czy ich licznosc. D. Reynolds pisze: „nie istnieją obiektywne przesłanki, które wskazywałyby właściwą liczbę mówców, jaka powinna wchodzić w skład bazy populacyjnej”<sup>60</sup>.

Iloraz wiarygodności można interpretować jako stosunek stopnia podobieństwa cech do stopnia ich typowości. Wykorzystanie LR jest najbardziej wiarygodnym i logicznie uzasadnionym sposobem przedstawienia wagi dowodu. Należy zwrócić uwagę, że wykres na rycinie 12 jest prawdziwy tylko pod warunkiem, gdy analizowana cecha ma rozkład normalny (krzywa Gaussa), dodatkowo, analizując częstotliwości formantowe, dysponuje się trzema parametrami. Ponieważ zwykle brane są pod uwagę pierwsze trzy formanty, które są ze sobą skorelowane, analiza uzyskanych danych z pomiarów ulega istotnym komplikacjom. Do rozwiązywania tego typu problemów wykorzystywane są na przykład estymatory jądrowe gęstości<sup>61</sup>.

Fonoskopijne badania identyfikacyjne są prowadzone w następującym zakresie:

- identyfikacja w obrębie materiału dowodowego – polegająca na przypisaniu poszczególnym mówcom konkretnych wypowiedzi z materiału dowodowego,
- identyfikacja z wykorzystaniem materiału porównawczego – przypisanie konkretnych wypowiedzi z materiału dowodowego osobie, od której został pobrany materiał porównawczy w postaci mowy (odwróceniem takiej identyfikacji jest eliminacja, tzn. stwierdzenie, że mówca, od którego uzyskano wypowiedzi porównawcze, nie jest autorem wypowiedzi dowodowych). Ten etap badań wymaga podejścia kompleksowego, które jest realizowane za pomocą metod opisanych powyżej, a które przyjąto określać mianem metody językowo-pomiarowej<sup>62</sup>.

Jednym z trudniejszych problemów jest uzyskanie odpowiedniego materiału porównawczego. Jako reguła materiał porównawczy do badań identyfikacyjnych mówcy powinien być pobierany przez eksperta w trakcie odrębnej czynności, z zachowaniem tak reguł metodologicznych, uwzględniających rodzaj i treść dowodowego nagrania, jak też prawnych. O ile stosunkowo łatwo jest uzyskać taki materiał wtedy, gdy osoba, od której pobierane są próbki, zgadza się na ich złożenie i współdziała z biegłym, o tyle może budzić wątpliwości natury procesowej zmuszanie takiej osoby do złożenia materiału porównawczego bądź jego uzyskanie bez

wiedzy mówcy, na przykład w trakcie rejestrowanego przesłuchania lub w trybie rejestracji niejawnej<sup>63</sup>. W niektórych krajach od dawna przesądzono dopuszczalność uzyskania tą drogą materiału porównawczego do badań fonoskopijnych<sup>64</sup>; jak się wydaje, również polskie orzecznictwo idzie w tym kierunku. Wyraźnie bowiem na ten temat wypowiada się SN w wyroku w sprawie II KKN 108/00 z 21.07.2000 r.:

„Z art.74 k.p.k. nie wynika, by podejrzany był obowiązany do aktywnego uczestniczenia w uzyskiwaniu dowodów przeciwko niemu, nie ma obowiązku wykazywania aktywności podczas utrwalania próby głosu. Nie oznacza to jednak, że dopuszcza się wykorzystanie w celach dowodowych tego rodzaju materiałów tylko za zgodą oskarżonego. Jeżeli istnieje materiał dowodowy pochodzący od oskarżonego, to może być wykorzystany nawet wówczas, gdy odmówił on udziału w procesowej czynności uzyskania takiego materiału. Nie ma znaczenia kiedy powstał taki materiał dowodowy przydatny do badań porównawczych – przed czy po odmowie uczestnictwa w badaniu, jak również to, czy oskarżony był świadomy tego, że jego wypowiedź zostanie utrwalona”<sup>65</sup>.

### Perspektywy fonoskopii w Polsce

Fonoskopia jest dyscypliną, która stale się rozwija, czego przejawem jest udoskonalanie metodyki, adaptowanie osiągnięć nauk pokrewnych i wdrażanie nowych rozwiązań.

Najbardziej dynamicznie rozwijającym się obszarem fonoskopii wydaje się identyfikacja mówców, gdyż przez ostatnie kilka lat dokonał się istotny postęp na tym polu. Współczesne systemy automatycznej identyfikacji mówców istotnie obiektywizują badania porównawcze. Działają one w oparciu o znacznie bardziej skomplikowane cechy widmowe niż częstotliwości formantowe (MFCC, PLP), wykorzystują parametryczne i nieparametryczne metody klasyfikacji (GMM, SVM), adaptacyjnie kompensują zakłócenia wprowadzane przez kanał transmisji<sup>66</sup>.

Należy prowadzić badania nad udoskonaleniem metodyki spisywania treści rozmów. Warto zastanowić się nad celowością wdrożenia metody, która została opisana w publikacji H. Holliena<sup>67</sup>. Zgodnie z tą metodą, ekspert słuchając w słuchawkach nagrania niskiej jakości, wykorzystuje nagranie zaszumione (oryginalne) i poddane korekcji częstotliwościowej (w zakresie nie pogarszającym zrozumiałości wypowiedzi). Obydwa nagrania są analizowane jednocześnie, podczas odsłuchiwania zapisu dwuścieżkowego.

W odniesieniu do metod badania autentyczności nagrań należy przeprowadzić eksperymenty, które precyzyjnie określą wartość diagnostyczną poszczególnych

metod w realiach nagrań cyfrowych. Jak dotychczas w Polsce badania takie nie zostały przeprowadzone.

Kolejnym z wyzwań stojących przed współczesną fonoskopią jest sprostanie wymaganiom stawianym przez międzynarodową normę ISO 17025. Określa ona standardy prowadzenia badań laboratoryjnych, wymagania wobec personelu laboratorium oraz dotyczące postępowania z próbkami badawczymi. Kładzie również nacisk na informowanie zlecniodawców badań o wiarygodności stosowanej metodyki i ilościowym przedstawianiu niepewności wyników<sup>68</sup>. Fonoskopia – z uwagi na ilościowo-jakościowy przedmiot badań – nie może w obecnym stanie wiedzy sprostać wymaganiom tej normy. Ale nie jest to problem, z którym boryka się wyłącznie fonoskopia. Dostrzega go m.in. ILAC (International Laboratory Accreditation Cooperation) – forum międzynarodowe, którego celem jest rozwój praktyk i procedur akredytacyjnych w zakresie laboratoriów i jednostek inspekcyjnych, wspieranie rozwijających się systemów akredytacji, a także uznawanie wyników badań i świadectw z inspekcji na całym świecie. ILAC za cel stawia sobie również pomoc jednostkom akredytującym z krajów będących w trakcie opracowywania własnych systemów akredytacji laboratoriów i/lub systemów akredytacji jednostek inspekcyjnych. Jednostki akredytujące mają zagwarantowane prawo przystąpienia do ILAC na zasadach członkostwa stowarzyszonego oraz mogą korzystać z wiedzy i doświadczenia członków ILAC o bardziej rozwiniętych i zaawansowanych systemach akredytacji. Korzystając z doświadczeń wszystkich stowarzyszonych członków, ILAC tak pisze o szacowaniu niepewności badań w odniesieniu do badań jakościowych: „Uważa się, że kwestia szacowania niepewności pomiaru w odniesieniu do wyników jakościowych jest obszarem, gdzie potrzebne są dalsze wytyczne. ILAC skoncentruje się, w pierwszym rzędzie, na wprowadzaniu niepewności pomiaru w odniesieniu do wyników badań ilościowych”<sup>69</sup>.

Należy zatem poszukiwać adekwatnych metod walidacji metod badawczych. Norma ISO 17025 dopuszcza w tym zakresie:

- wzorcowanie z wykorzystaniem wzorców odniesienia lub materiałów odniesienia,
- porównanie wyników uzyskanych innymi metodami,
- porównania międzylaboratoryjne,
- systematyczną ocenę czynników wpływających na wynik,
- ocenę niepewności wyników opartą na naukowym rozumieniu teoretycznych podstaw metody i praktycznym doświadczeniu.

Specyfika metod jakościowych wykorzystywanych w fonoskopii pozwala prowadzić walidację tylko w oparciu o badania międzylaboratoryjne. Problematyka ich

przewodzenia została już częściowo opisana<sup>70</sup>. Wydaje się, że testy międzylaboratoryjne są obecnie nie tylko jedyną naukowo uzasadnioną metodą walidacji wyników kompleksowych badań fonoskopijskich, ale także szansą dla samej fonoskopii na zobiektywizowanie wyników ekspertyz i podniesienie ich poziomu.

**Waldemar Maciejko**  
**Jacek Rzeszotarski**  
**Tadeusz Tomaszewski**

#### PRZYPISY

- <sup>1</sup> OSNKW 1963/7-8/131.
- <sup>2</sup> **A. Szwarc:** Kryminalistyczna ekspertyza zapisu magnetofonowego, Warszawa 1964.
- <sup>3</sup> **B. Hołyst:** Kryminalistyka, wyd. 12, Warszawa 2010, s. 904. Autor ten wskazuje, że obecnie termin ten nie jest adekwatny w stosunku do zakresu i rodzaju badań stosowanych w praktyce do analizy zapisów nagrań dźwiękowych i środków rejestrujących.
- <sup>4</sup> **M. Jessen, M. Jessen:** Forensische Sprechererkennung und Tonträgerauswertung in Praxis und Forschung, „Die Kriminalpolizei”, cz. 1, Dez. 2008 ([http://www.kriminalpolizei.de/articles/forensische\\_sprechererkennung\\_und\\_tontraegerauswertung\\_in\\_praxis\\_und\\_forschung\\_1,223.htm](http://www.kriminalpolizei.de/articles/forensische_sprechererkennung_und_tontraegerauswertung_in_praxis_und_forschung_1,223.htm)), cz. 2 März 2009 ([http://www.kriminalpolizei.de/articles/forensische\\_sprechererkennung\\_und\\_tontraegerauswertung\\_in\\_praxis\\_und\\_forschung\\_teil\\_2,1\\_236.htm](http://www.kriminalpolizei.de/articles/forensische_sprechererkennung_und_tontraegerauswertung_in_praxis_und_forschung_teil_2,1_236.htm)).
- <sup>5</sup> OSNKW 1962/1/8.
- <sup>6</sup> **A. Filewicz:** Słowo o płk. Stanisławie Błasikiewiczu, twórcy polskiej szkoły fonoskopii, „Problemy Współczesnej Kryminalistyki” 2003, t. VII, pod red. **E. Gruzy, T. Tomaszewskiego, M. Goca**, s. 13–14.
- <sup>7</sup> **L.G. Kersta:** Voiceprint Identification, „Nature” 1962, nr 196, s. 1253–1257.
- <sup>8</sup> **W. Gierasimow:** Identyfikacja liczności po głosie, „Sowietskaja Justicja” 1971, nr 20.
- <sup>9</sup> **Ch. Koristka**, [w:] „Sozialistische Kriminalistik”, pod red. **F. Stelzer**, t. 2, Naturwissenschaftlich-technische Kriminalistik, Berlin 1979, s. 549–576.
- <sup>10</sup> **B.P. Bogert, M.J.R. Healy, J.W. Tukey:** The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking, Proceedings of the Symposium on Time Series Analysis (**M. Rosenblatt**, Ed.) Chapter 15, 209–243, New York: Wiley, 1963.
- <sup>11</sup> **F. Itakura, S. Saito:** Analysis synthesis telephony based on the maximum likelihood method, Proc. 6th Intern. Congr. Acoust. C-5-5, Tokyo, Japan, August 1968.
- <sup>12</sup> **R.J. Baken, R.G. Daniloff:** Readings in Clinical Spectrography of Speech, KAY Elemetrics, New Jersey 1991, Preface.

- 13 **W. Jassem**: Formant Frequencies as Cues to Speaker Discrimination, *Speech Analysis and Synthesis* 1, 9–42, Państwowe Wydawnictwo Naukowe, Warszawa 1968.
- 14 **H. Hollien**: Forensic voice identification, New York 2002, s. 11; **Rose Ph.** Forensic speaker identification, London–New York 2002, s. 262.
- 15 **H. Hollien**: Forensic voice identification, op.cit., s. 147.
- 16 **W. Barycki, C. Basztura, W. Majewski**: Effectiveness of selected voice features in speaker identification procedures, Open Seminar in Acoustics OSA-89; **C. Basztura, W. Majewski**: „The application of long-term analysis of zero – crossing of speech signal in automatic speaker identification”, *Archive Acoustics* 1978; **W. Majewski, H. Hollien**: Euclidean distance between long-term speech spectra as a criterion for speaker identification, *Proc. Speech Comm Seminar* 1974; **W. Majewski, H. Hollien**: Speaker identification by long term spectra under normal and distorted speech conditions, *JASA* 1977; **W. Majewski, C.S. Basztura, H. Hollien**: Analiza przejść przez zero sygnału mowy jako metoda ekstrakcji parametrów w krótkoterminowym modelu, OSA 1977; **W. Majewski, C.S. Basztura**: Integrated approach to speaker recognition in forensic applications, *Forensic Linguistics* 1996.
- 17 **A. Jaroszewski**: Głuchniemy, „Audio Video” 2001, nr 7, s. 18–20.
- 18 **E.P. Solomon, L.R. Berg, D.W. Martin, C.A. Ville**: Biologia, Warszawa 1996, s. 873.
- 19 **U. Jorasz**: Selektowność układu słuchowego, Wydawnictwo Naukowe UAM, Poznań 1999, s. 48.
- 20 **Z. Żyszkowski**: Podstawy elektroakustyki, WNT, Warszawa 1984, s. 244.
- 21 Ibidem, s. 526.
- 22 PN-87/B-02151/02 Akustyka budowlana. Ochrona przed hałasem pomieszczeń w budynkach. Dopuszczalne wartości poziomu dźwięku w pomieszczeniach.
- 23 **S. Błasikiewicz**: Metoda odsłuchu szeptu i mowy intensywnie zakłóconej, „Problemy Kryminalistyki” 1971, nr 90, s. 159–183.
- 24 Ibidem.
- 25 **R. Haines**: Cyfrowe przetwarzanie dźwięku, Warszawa 2002, s. 169.
- 26 **R. Resnick, D. Halliday**: Fizyka dla studentów nauk przyrodniczych i technicznych, Warszawa 1975, s. 568.
- 27 **S. Błasikiewicz**, op.cit.
- 28 Obszerne uwagi na temat przedmiotu badań autentyczności nagrania zostały przedstawione przez **A. Jedy-nak i J. Rzeszotarskiego**: Definicja autentyczności zapisu dźwięku, w czasopiśmie „Iustitia” 2007, nr 2, s. 10–14 oraz „Problemach Kryminalistyki” 2007, nr 257, s. 23–27.
- 29 **S. Błasikiewicz, W. Bednarczyk**: Metoda badań autentyczności zapisu magnetofonowego, „Problemy Kryminalistyki” 1978, nr 131, s. 34–50.
- 30 Ibidem.
- 31 Technika kryminalistyczna, pod red. **W. Kędzierskiego**, Szczytno 1995, t. 2, s. 256.
- 32 Słownik języka polskiego, red. nauk. **M. Szymczak**, Warszawa 1992, t. 1, s. 98.
- 33 **B.E. Koenig**: Authentication of forensic audio recordings, *Journal of the Audio Engineering Society* 1990, nr 38 (1/2).
- 34 **S. Błasikiewicz, W. Bednarczyk**, op.cit.
- 35 **J.S. Bouten**: Applying an image enhancement technique to imaging polarimetry used in magneto optical investigations of audio tapes in authenticity investigations, *Forensic Science International*, vol. 136, suppl. 1, 2003; **D. Boss, S. Gfroerer, N. Neoustroev**: A new tool for the visualization of magnetic features on audio-tapes, *International Journal of Speech Language and the Law*, Vol. 10, No 2 (2003).
- 36 Szerzej **S. Błasikiewicz, W. Bednarczyk**, op.cit.
- 37 **J. Rzeszotarski**: Audytywna ocena autentyczności nagrania, „Problemy Kryminalistyki” 2004, nr 245, s. 50–53.
- 38 **C. Grigoros**: Digital Audio Recording Analysis, The Electric Network Frequency Criterion, Catalin Grigoros, Ph. D, IAFPA 2003.
- 39 Technika kryminalistyczna, op.cit., t. 2, s. 257.
- 40 **I. Styczek**: Zarys logopedii, Warszawa 1970, s. 13.
- 41 **K. Dejna**: Atlas polskich innowacji dialektalnych, Warszawa–Łódź 1994, s. 5.
- 42 Foniatria kliniczna, pod red. **A. Prusze-wicza**, Warszawa 1992, s. 11.
- 43 **M. Maruszewski**: Mowa a mózg, Warszawa 1970, s. 77 i nast.
- 44 **B. Wierzchowska**: Wymowa polska, PZWS, Warszawa 1971, s. 97.
- 45 Foniatria kliniczna, op.cit., s. 222.
- 46 **P.H. Lindsay, D.A. Norman**: Procesy przetwarzania informacji u człowieka, Państwowe Wydawnictwo Naukowe, Warszawa 1984, s. 459.
- 47 **K. Dejna**, op.cit., s. 8.
- 48 Foniatria kliniczna, op.cit., s. 224.
- 49 **K. Dejna**, op.cit., s. 7.
- 50 **P.G. Zimbardo**: Psychologia i życie, Warszawa 2001, s. 165 i nast.
- 51 **J. Strelau**: Psychologia. Podręcznik akademicki, Gdańsk 2002, s. 246.
- 52 **F. Nolan**: Auditory and acoustic analysis in speaker recognition, [w:] *Language and the Law*, pod red. **J. Gibbonsa**, London 1993, s. 326–345.
- 53 **L.R. Rabiner, M.J. Cheng, A.E. Rosenberg, C.A. McGonegal**: A comparative study of several pitch detection algorithms, *IEEE Transactions on Audio*,

- „Speech and Signal Processing” 1976, nr 24, s. 399–413.
- 54 **B. Wierzchowska**, op.cit., s. 27.
- 55 **H. Hollien**: The acoustic of crime, New York and London 1990, s. 242.
- 56 **B. Wierzchowska**, op.cit., s. 61.
- 57 **W. Jassem**: Podstawy fonetyki akustycznej, Warszawa 1973, str. 156.
- 58 **R. Rodman, D. McAllister, D. Bitzer, L. Cepeda, P. Abbitt**: Forensic speaker identification based on spectral moments, „Forensic Linguistics” 2002, nr 9 (1), s. 22–43.
- 59 Szerzej na temat interpretacji ilorazu wiarygodności patrz: **C. Aitken, F. Taroni**: Statistical and evaluation of evidence for forensic science, Nottingham 2004, s. 157–170.
- 60 **D.A. Reynolds, T.F. Quatieri, R.B. Dunn**: Speaker verification using adapted gaussian mixture models”, Digital Signal Processing 2000, nr 10, s. 19–41.
- 61 **C. Aitken, F. Taroni**, op.cit. s. 330–339; **P. Rose**: Forensic speaker recognition at the beginning of the twenty first century – an overview and a demonstration, „Australian Journal of Forensic Sciences” 2005, nr 37, s. 49–72.
- 62 **J. Dolecki, J. Rzeszotarski**: Zastosowanie metody językowo-pomiarowej do identyfikacji osób w badaniach fonoskopijskich, Z Zagadnień Nauk Sądowych, t. LII, 2002, s. 108–123; **K. Klus, A. Trawińska**: Identyfikacja mówców z wykorzystaniem metody językowo-pomiarowej w KEU i IES, „Problems of Forensic Sciences” 2009, LXXVIII, s. 160–174.
- 63 Przykładowo, w procesie niemieckim za niedozwolone, podstępne działanie, które skutkuje niemożnością wykorzystania dowodu z opinii fonoskopijskiej, uznano „przesłuchanie podejrzanego w celu uzyskania od niego głosu do celów identyfikacyjnych (okazanie głosu) lub badań porównawczych, o ile jest ono dokonane wyłącznie z tej przyczyny” (BGH Urt. z 24.02.1994, 4 StR 317/93 LG Saarbruecken, BGHSt 40, 66–72; por. szerzej: **P. Girdwoyń**: Zarys niemieckiego procesu karnego, Białystok 2006, s. 242).
- 64 W Stanach Zjednoczonych można na podstawie nakazu sądowego zmusić podejrzanego do złożenia przez niego próbek głosu. Już w 1967 r. w znanym precedensie *United States v. Wade* (388 U.S. 218, 1967) SN USA stwierdził, że taka procedura nie łamie V Poprawki do konstytucji zakazującej zmuszania do zeznawania w sprawie karnej na swoją niekorzyść, gdyż, jak stwierdzono *expressis verbis*, przywilej ten nie chroni m.in. przed wymuszaniem dostarczenia własnego głosu dla celów identyfikacyjnych.
- 65 LEX nr 50904; „Prokuratura i Prawo” 2001/1/11.
- 66 **D.A. Reynolds, T.F. Quatieri, R.B. Dunn**, op.cit.
- 67 **H. Hollien**: The acoustic of crime, op.cit., s. 144.
- 68 PN-EN ISO/IEC 17025: 2005. Ogólne wymagania dotyczące kompetencji laboratoriów badawczych i wzorcujących.
- 69 Wprowadzenie problematyki niepewności pomiaru w badaniach w związku z wejściem do stosowania normy ISO/IEC 17025, ILAC-G17:2002.
- 70 **J. Rzeszotarski, W. Witkowska-Pawlak, O. Komosa**: Test biegłości ekspertów fonoskopii spisujących treść rozmów, „Problemy Współczesnej Kryminalistyki” 2008, tom XII, pod red. **E. Gruzy, T. Tomaszewskiego, M. Goca**, s. 297–302; **T. Tomaszewski, J. Rzeszotarski**: Weryfikacja kwalifikacji biegłych wydających opinie kryminalistyczne (na przykładzie opinii fonoskopijskich), [w:] Kryminalistyka i nauki penalne wobec przestępczości, Księga pamiątkowa dedykowana Profesorowi Mirosławowi Owocowi, Poznań 2008, s. 255–265.

### Streszczenie

W 2010 roku mija 50 lat od orzeczenia Sądu Najwyższego, które stało się przyczynkiem do rozwoju w Polsce nowej dziedziny kryminalistyki – fonoskopii. Pierwsza w kraju placówka została utworzona w Zakładzie Kryminalistyki KG MO, którego spadkobiercą jest obecnie CLK KGP.

Polscy przedstawiciele dziedziny – na czele ze S. Blasikiewiczem – chlubnie zapisali się w historii badań fonoskopijskich na świecie. Mieli znaczący wpływ na ukształtowanie metodyki badawczej, kreowali kierunki rozwoju fonoskopii, wskazywali na trudności praktyczne w działalności opiniodawczej.

W artykule przedstawiono twórców polskiej szkoły fonoskopii oraz zaprezentowano podstawowe metody badawcze użyteczne w procesie karnym.

Słowa kluczowe: fonoskopia, spisanie treści, autentyczność nagrań, identyfikacja mówców.

### Summary

In 2010, is 50th anniversary of the Supreme Court's decision that gave start to the development a new field of the forensic science: phonoscopy. The first laboratory in Poland was created in Militia Headquarters Forensic Institute, being the predecessor of the Central Forensic Laboratory of the Polish Police.

Polish representatives of the field lead by Stanisław Blasikiewicz played a significant role in the world history of voiceprint examinations. They considerably influenced methodology, created new directions in the development of voiceprint examinations and identified practical difficulties in forensic casework.

The article presents the founding fathers of Polish school of phonoscopy and the basic methods of examination used in criminal investigation.

Keywords: phonoscopy, writing down script, authenticity of recording, speaker identification.