

Stereofonia jako czynnik rozszerzający pole badań fonoskopii

Postulatem eksperta fonoskopii wobec podmiotu realizującego nagranie – w dążeniu do pełniejszego wykorzystania śladu akustycznego zdarzenia – jest uzyskanie możliwie wysokiej jakości zapisu. Stereotypowe mniemanie, że „wystarczy zarejestrować ekspozycję nagrania” jest obciążone zasadniczym błędem. Raz utracone informacje nie zostaną przywrócone nawet w warunkach laboratoryjnych, z wykorzystaniem profesjonalnego sprzętu.

W dobrej jakości nagraniu mowy częściej współwystępują cechy i parametry językowe oraz pozajęzykowe, dlatego konkluzje z badań kryminalistycznych lepszego zapisu cechuje precyzja stwierdzeń i szerszy zakres wniosków. Jednym z aspektów polepszenia jakości nagrania jest rejestracja stereofoniczna, umożliwiająca m.in. przeprowadzenie analizy akustycznej pod kątem przestrzennej lokalizacji źródła dźwięku, tj. określenia kierunku i odległości od mikrofonów [3]. Słuchacz może dokonać poprawnej lokalizacji, o ile pozwala na to specyfika odbieranego dźwięku. Dokładność oceny słuchowej jest niższa niż w przypadku postrzegania wzrokowego, które jednak nie mieści się w zakresie badawczym fonoskopii [3]. Bodźce percypowane audytywnie są przesyłane w postaci impulsów nerwowych do mózgu, gdzie powstaje model przestrzeni. Rekonstrukcja odbywa się na podstawie trzech wyznaczników [2]:

- stosunku opóźnienia, z jakim sygnał dociera do uszu z punktu znajdującego się w różnej odległości od nich (układ nerwowy wykrywa różnice w pobudzeniu membrany obu uszu tym samym dźwiękiem rzędu 10 mikrosekund),

- różnicy głośności między tymi sygnałami,

- różnicy częstotliwości dźwięków docierających z innych kierunków.

Ta fizjologiczna właściwość aparatu słuchu i centralnego układu nerwowego może mieć pomocnicze, a niekiedy wręcz decydujące znaczenie w procesie kryminalistycznej identyfikacji mówców w obrębie materiału dowodowego [4]. Do badań fonoskopijnych trafiają bowiem nagrania rozmów bezpośrednich, przeprowadzonych przez mówców w tym samym wieku, władających językiem nacechowanym tymi samymi naleciałościami regionalnymi czy grupowymi, o podobnej barwie głosu. Jedni w rozmowie dowodowej mówią dużo, inni niewiele, ale zwykle na podobne tematy. W takim nagraniu, gdzie ekstrakcja indywidualności sposobu mówienia jest bardzo trudna, parametry techniczno-elektroakustyczne – w konsekwencji zakamuflowania mikrofonu – ulegają obniżeniu, dlatego dodatkowo zawęża się zbiór cech odróżniających mówców. Wówczas to właśnie przestrzenna lokalizacja pozwala precyzyjnie ustalić liczbę uczestników zdarzenia.

W Zakładzie Kryminalistyki i Chemii Specjalnej Agencji Bezpieczeństwa Wewnętrznego przeprowadzo-

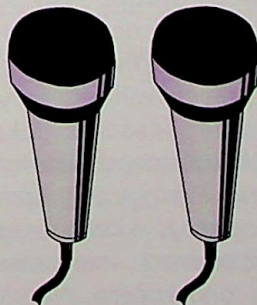
no eksperyment będący próbą wykazania wpływu stereofonii na możliwość rozróżniania lokalizacji przemieszczającego się mówcy, w zestawieniu z rejestracją monofoniczną. Formuła doświadczenia pozwoliła wyeliminować pozostałe czynniki, brane pod uwagę przez słuchacza w procesie oceny usytuowania źródła dźwięku.

Materiały i metody

Warunki nagrania wypowiedzi testowych

Strona techniczna eksperymentu polega m.in. na właściwym ustawieniu mikrofonów użytych do rejestracji wypowiedzi testowych. Literatura opisuje dwa podstawowe sposoby usytuowania mikrofonów stereo [1]:

1) para rozstawiona – dwa odległe od siebie mikrofony (zazwyczaj o charakterystyce dookolnej), ustawione przed lewą i prawą stroną zbieranej przestrzeni. Zwłaszcza ustawienie dwóch oddzielnych mikrofonów w odległości nieprzekraczającej kilkudziesięciu centymetrów nazywane jest ustawieniem A-B i tworzy obraz stereofoniczny, łącząc różnice fazowe i natężeniowe (ryc. 1),



Ryc. 1. Rozmieszczenie mikrofonów w układzie A-B
Fig. 1. Arrangement of microphones in A-B scheme



Ryc. 2. Rozmieszczenie mikrofonów w układzie X-Y
Fig. 2. Arrangement of microphones in X-Y scheme

2) technika koincydencyjna – dwa, zazwyczaj kardioidalne mikrofony ustawione w tym samym miejscu na środku, skrzyżowane pod kątem 90-120 stopni, o niestykających się membranach, przy czym prawy jest wycelowany w lewą stronę przestrzeni, a lewy w prawą stronę (ryc. 2). Ten sposób rejestracji jest nazywany kombinacją X-Y.

Podczas stereofonicznej rejestracji wypowiedzi testowych mikrofony były ustawione w kombinacji koincydencyjnej X-Y.

Do utrwalenia testów użyto następującego zestawu:

- magnetofon cyfrowy PORTA-DAT, firmy HHB, model PDR 1000,
- mikrofony firmy SCHOEPS, o charakterystyce kardioidalnej,
- taśmy cyfrowe firmy HHB, typu DAT,
- okablowanie firmy KLOTZ.

Parametry rejestracji: częstotliwość próbkowania 44,1 kHz/16 bitów, pasmo częstotliwości 50-20 000 Hz. Warunki eksperymentu przybliżono do specyfiki zapisów dowodowych przez zastosowanie urządzenia generującego zakłócenia szumowe w pełnym zakresie częstotliwości słyszalnych. Odstęp sygnał mowy – szum w nagraniach testowych wyniósł średnio 12 dB.

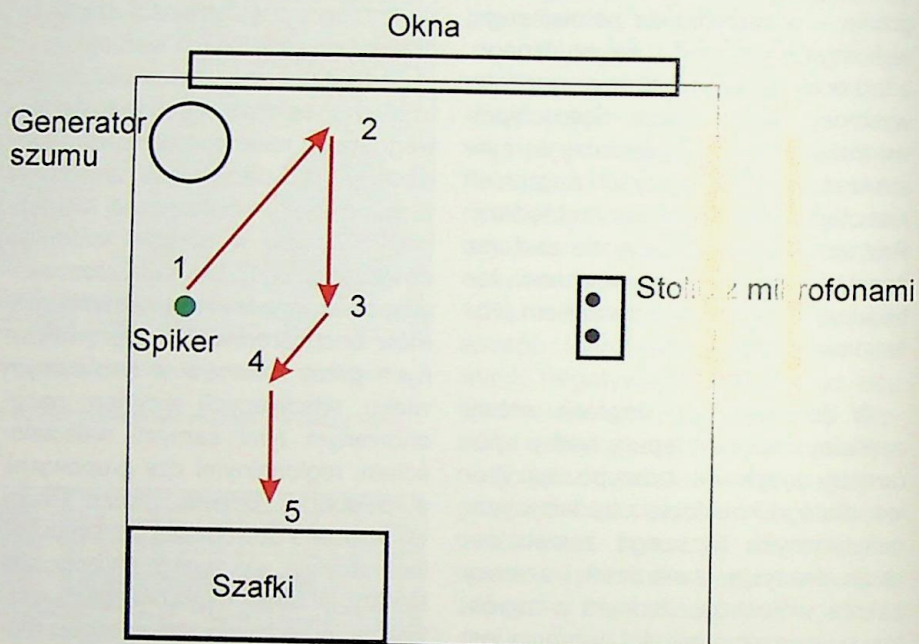
Magnetofon z mikrofonem/mikrofonami umieszczono na drewnianym stoliku, a ten w pomieszczeniu o wymiarach dł. 5 m/szer. 5 m/wys. 3 m (ryc. 3), ścianach otynkowanych i pomalowanych farbą emulsyjną, trzech oknach na jednej ścianie, zastawionym dwoma drewnianymi szafkami biurowymi o zamkniętych drzwiczkach (współczynnik pochłaniania dźwięku 1024 Hz wynosi w tym pomieszczeniu ok. 0,33).

Podczas nagrania w pomieszczeniu znajdował się realizator nagrania i spiker, którego zadaniem było odczytywanie tekstu ze stałym tempem i głośnością wypowiedzi, bez modulacji barwy głosu. Przechodził on po

pokoju w pięć miejsc, zgodnie z wytyczoną trajektorią (ryc. 3). W trakcie zmiany usytuowania spikera następowała przerwa w czytaniu. Po zajęciu miejsca odczytywał on tekst przez dziesięć sekund, przerwa na zmianę pozycji trwała ok. trzech sekund.

znaly założenia spisywania treści nagrań zgodnie z metodą odsłuchu mowy intensywnie zakłóconej i zniekształconej.

Wszystkim uczestnikom testy prezentowano przez słuchawki. Ich zastosowanie sprawia, że pojawia się



Ryc. 3. Trajektoria ruchu spikera w pomieszczeniu, w którym dokonano nagrań testowych
Fig. 3. Speaker movement trajectory in test recording room

Czynność spikera powtórzono i zarejestrowano dwukrotnie, najpierw monofonicznie, później stereofonicznie.

Prezentacja nagrań testowych słuchaczom

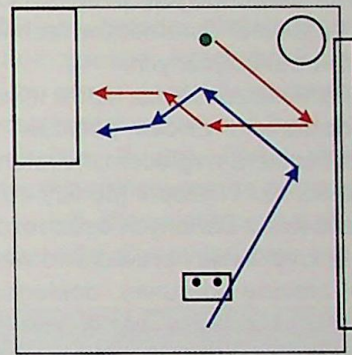
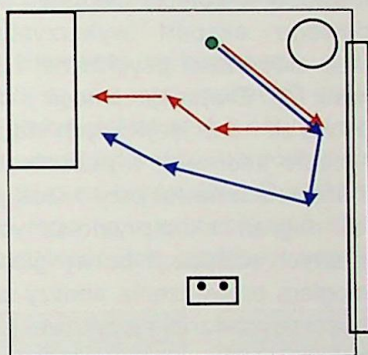
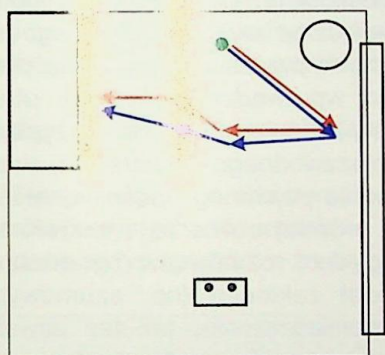
Utrwalone cyfrowo testy posłużyły jako przedmiot oceny audytywnej przez osoby nieznające celu i założeń teoretycznych eksperymentu. Potencjalnie właściwa lokalizacja źródła dźwięku w przestrzeni wymaga jedynie fizjologicznej sprawności narządu, dlatego o doborze uczestników eksperymentu nie przesądzała praktyka fonoskopijna. Nagranie zaprezentowano siedemnastu słuchaczom o percepcji słuchowej w granicach normy, tylko dwie z tych osób

efekt internalizacji. Słuchacz ma wrażenie, że dźwięk pojawia się wewnątrz głowy, względnie przemieszcza się po osi łączącej uszy (lateralizacja) [3]. Słuchacze nie uczestniczyli w rejestracji testowej, nie znali trajektorii spikera. Nie poinformowano ich, że zaprezentowane nagrania różnią się liczbą mikrofonów użytych do rejestracji. Prezentacja nagrań odbywała się w kolejności: jako pierwsze mono-, później stereofoniczne. Zależnie od indywidualnych potrzeb słuchacza, nagranie analizowano kilkakrotnie. Do odsłuchu zastosowano słuchawki firmy Sennheiser, typ HD580, podłączone do zestawu nagrywającego. Zadaniem słuchacza było samodzielne ustalenie, skąd dochodzi głos spikera (kierunek i odległość), a następnie sporządzenie

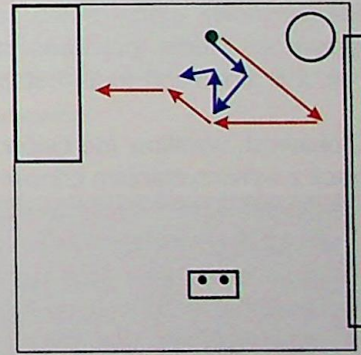
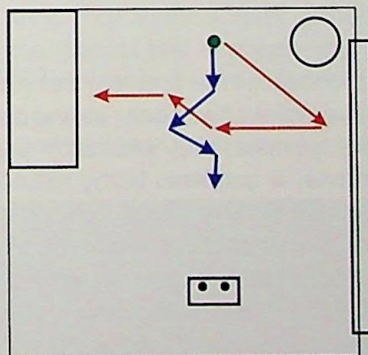
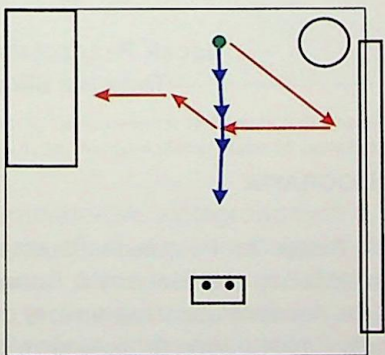
szkiców trajektorii, pierwszego po wysłuchaniu zapisu mono-, drugiego po wysłuchaniu zapisu stereofonicznego. Słuchacz, który choć raz usłyszał testy stereofoniczne, nie mógł już nanosić poprawek na szkicu wykonanym na podstawie nagrania monofonicznego.

Wyniki i dyskusja

Uzyskane wyniki przedstawiono na sześciu schematach reprezentatywnych dla badanej grupy (ryc. 4–9).



Ryc. 4–6. Postrzeganie trajektorii ruchu spikera w odsłuchu nagrania stereo – kolor niebieski, faktyczna trajektoria – kolor czerwony
Fig. 4–6. Perception of speaker movement when hearing the stereo-recording – blue; real trajectory – red



Ryc. 7–9. Postrzeganie trajektorii ruchu spikera w odsłuchu nagrania mono – kolor niebieski, faktyczna trajektoria – kolor czerwony
Fig. 7–9. Perception of speaker movement when hearing the mono-recording – blue; real trajectory – red

Wyniki uzyskane na podstawie nagrania stereofonicznego są następujące:

1. Lokalizacja kierunku usytuowania źródła sprawiała problem jedynie wtedy, gdy mówca znajdował się w osi na wprost mikrofonów (punkty

1 i 3 trajektorii) – 5 błędnych wskazań.

2. W pozostałych pozycjach (punkty 2, 4 i 5) kierunek dobiegania dźwięku był wskazywany zgodnie z rzeczywistym.

3. Bez względu na relacje odległości punktów 1–5 od siebie i od mikrofonów zasadniczo nie zostały oddane poprawnie.

4. Słuchacze wychwycili względne różnice odległości punktów trajektorii, np. że punkt 1 jest najdalej od mikrofonów, że odległość od punktów 2 i 4 jest podobna.

5. Uczestnicy eksperymentu nie postulowali, że szum wytwarzany przez generator – intensywniejszy na prawe ucho – utrudniał im zadanie.

6. Nie stwierdzono, żeby pogłos pomieszczenia wpłynął negatywnie na lokalizację przestrzenną mowy.

Wyniki analizy nagrania monofonicznego:

7. Żaden układ odwzorowany przez słuchaczy nie odzwierciedlał rzeczywistej trajektorii spikera.

8. Nikt nie zlokalizował pięciu pozycji spikera w jednym punkcie przestrzeni.

9. W przeważającej grupie wyników (12 osób) trajektorię ruchu spikera przedstawiano jako linię prostą do stolika ze sprzętem audio.

10. Część osób umiejscowiła kolejne pozycje spikera w skupisku wokół osi głowy.

Wyniki w dużej mierze potwierdziły możliwość przestrzennej lokalizacji źródła mowy – zgodnej z rzeczywistością – pod warunkiem zastosowania zapisu stereofonicznego.

Nagranie monofoniczne pozwala jedynie oceniać różnice natężenia sygnału mowy dobiegającego z kolej-

nych miejsc przestrzeni. Efektem – zaobserwowanym w wynikach doświadczenia – była przypadkowość wskazań.

Zmiana natężenia dźwięku (punkty 1 i 3) nie była wystarczającą cechą różnicującą dla dwóch słuchaczy, także w przypadku odsłuchu nagrania stereofonicznego.

Szum zakłócający percepcję słuchową wypowiedzi spikera nie wpłynął istotnie na lokalizację przestrzenną mówcy, co dobrze rokuje dla możliwości prowadzenia analogicznych analiz w przypadku stereofonicznych nagrań dowodowych o stosunkowo niskiej jakości parametrów technicznoelektroakustycznych.

Najmniejsza różnica kątów między usytuowaniem dwóch punktów trajektorii spikera względem mikrofonów wynosiła ok. 11 stopni (punkty 4 i 5). Szesnaście z badanych osób rozróżniło te lokalizacje i prawidłowo wskazało zmianę kierunku dobiegania dźwięku. Uzyskany wynik jest natomiast precyzji działania narządu słuchu i mózgu człowieka. Piśmiennictwo podaje, że minimalna_percypowana różnica kątów – dla sygnału w paśmie mowy – przy wolno poruszającym się źródle (prędkość kątowa ok. 2,8 stopnia/s) to 2–4 stopnie [3].

Ciekawych wyników (do celów rejestracji z wykorzystaniem odpowiednio zakamuflowanych w pomieszczeniu mikrofonów) mogłoby dostarczyć porównanie lokalizacji źródła sygnału w nagraniach stereo, z zastosowaniem dwóch różnych technik: X–Y i pary mikrofonów rozstawionych na odległość istotnie większą niż szerokość głowy. Zagadnienie to wykraczało poza cel niniejszego eksperymentu.

Podsumowanie

Przeprowadzone doświadczenie wykazało zalety stereofonii jako techniki rejestracji umożliwiającej przestrzenną lokalizację tego samego

mówcy w sytuacji modelowej, istotnie odbiegającej od rzeczywistości utrwalanej w zapisach dowodowych. Badany czynnik został wyizolowany z szeregu współwystępujących okoliczności nagrania, ocenianych przez eksperta jako użyteczne w badaniach śladu akustycznego bądź utrudniające ich przeprowadzenie.

Cechą nagrań dowodowych jest, udział w konkretnym zdarzeniu kilku-kilkunastu rozmówców. Gdy rośnie liczba uczestników dyskusji, narasta problem przypisania wypowiedzi poszczególnym mówcom. W identyfikacji mówców w obrębie materiału dowodowego ekspert wykorzystuje analizę aktywności psychicznej i językową [4]. Ekspresja emocji i odstępstwa od norm językowych stanowią źródło informacji o poszczególnych mówcach nawet przy niskiej jakości nagrań skomprimowanych, o zatartych różnicach barwy głosu. Wymogiem prowadzenia analizy aktywności psychicznej i językowej jest jednak dysponowanie obszernymi wypowiedziami spontanicznymi. Gdy warunek ten jest spełniony, biegły może traktować lokalizację przestrzenną jako czynnik pomocniczy. Ale gdy nagranie jest krótkie, a ekspert fonoskopii ma problem z ekstrakcją kompleksów cech indywidualnych, wzrasta rola lokalizacji przestrzennej w ustaleniu liczby rozmówców. Właściwie przeprowadzona identyfikacja w obrębie materiału dowodowego jest podstawą późniejszej identyfikacji z wykorzystaniem wypowiedzi porównawczych, gdzie wykazanie różnic międzyosobniczych dodatkowo jest możliwe z zastosowaniem analizy korelacji częstotliwości formantowych [4].

Praktyka eksperta fonoskopii pokazuje, że każde nagranie cechuje indywidualność jakości zapisu, tematu rozmowy, relacji interpersonalnych między uczestnikami dyskusji, sposobu mówienia i myślenia rozmówców, ilości i rodzaju ich wypowiedzi itd. Przed wstępnymi badaniami fo-

noskopijnymi, a tym bardziej przed utrwaleniem konkretnego zdarzenia akustycznego, nie sposób przesądzić, czy materiał badawczy spełni wymogi techniczne i merytoryczne do uzyskania odpowiedzi na pytania sformułowane w postanowieniu o zasięgnięciu opinii. Dlatego jeszcze przed zrealizowaniem nagrania nie należy ograniczać pola badawczego ekspertom.

Realizator nagrania utrwalanego do celów sądowych nie ma wpływu na szereg aspektów technicznych, komplikujących późniejszą badanie fonoskopijne. Uwzględniając jednak ewentualne wykorzystanie tego dowodu w procesie karnym nie powinien wprowadzać kolejnych utrudnień. Tak jak ważne jest korzystanie z niezawodnego rejestratora, odpowiednio pojemnego nośnika, właściwe umiejscowienie mikrofonów względem rozmówców, ograniczenie źródeł zakłóceń (np. szumowych) w pomieszczeniu, tak też zawsze, gdy to możliwe, utrwalanie strony dźwiękowej zdarzenia powinno odbywać się z wykorzystaniem rejestracji stereofonicznej.

Jacek Rzeszotarski
Tadeusz Sikora

BIBLIOGRAFIA

1. Butler T.: Połączenia. Podstawy profesjonalnej elektroakustyki i nagłaśniania, Fender Musical Instruments Corporation 1994, wydanie polskie: Mega Music Ltd, Sopot 1996.
2. Lindsay P.H., Norman D.A.: Procesy przetwarzania informacji u człowieka, Państwowe Wydawnictwo Naukowe, Warszawa 1984.
3. Ozimek E.: Dźwięk i jego percepcja, Wydawnictwo Naukowe PWN, Warszawa–Poznań 2002.
4. Rzeszotarski J., Witkowska-Pawlak W.: Identyfikacja mówców w obrębie materiału dowodowego, „Problemy Kryminalistyki” 2004, nr 246, s. 15.