

Identyfikacja mówcy celowo zniekształcającego wypowiedzi

Obecnie do Zakładu Kryminalistyki i Chemii Specjalnej ABW do badań fonoskopijnych najczęściej dostarczane są cyfrowe nagrania rozmów telefonicznych. W przeważającej większości przypadków rozmówcy w trakcie prowadzenia dialogu są spontaniczni, nie odpowiadają się w sposób namyślony, nie uświadamiając sobie, że treść ich wypowiedzi może się stać w przyszłości materiałem dowodowym. W efekcie biegli ZKiChS ABW w wydanej opinii formułują kategorię wniosków z przeprowadzonych badań identyfikacji mówców (fakt ten jest godny podkreślenia, gdyż eksperci z ośrodków zachodnioeuropejskich i krakowskiego Instytutu Ekspertyz Sądowych – bez wglądu w dorobek fonoskopii – odchodzą od sformułowań kategorię wniosków na rzecz stopniowania prawdopodobieństwa).

W niewielkim procencie nagrań dowodowych sprawca, przewidując rejestrację rozmów, usiłuje celowo zniekształcić wypowiedzi (w praktyce ZKiChS ABW w latach 2003–2005 zjawisko to odnotowano w sześciu sprawach), a ma to miejsce w przypadku takich przestępstw z wykorzystaniem transmisji telefonicznej, jak wymuszenia okupu czy kradzież środków pieniężnych z kont bankowych. Na potrzeby niniejszego opracowania intoksykacja substancjami psychoaktywnymi (w tym alkoholem etylowym) nie jest uznawana za celowe zniekształcanie wypowiedzi.

Praktyka fonoskopijna ZKiChS ABW jednoznacznie wskazuje, że celowe zniekształcanie wypowiedzi nie jest okolicznością bezwzględnie dyskwalifikującą stosowaną metodykę identyfikacji osób na podstawie mowy. Decyduje – jak zawsze w przypadku kryminalistycznych badań po-

równawczych – możliwość ekstrakcji zespołu cech indywidualnych, wynikająca z obszerności i jakości materiału dowodowego.

W badaniach wypowiedzi celowo zniekształcanych znajduje zastosowanie rutynowa metodyka identyfikacji osób na podstawie mowy.

Metodyka identyfikacji mówców

Metoda psycholingwoakustyczna i językowo-pomiarowa zostały ostatnio zaprezentowane na łamach polskich czasopism kryminalistycznych¹. Obydwie opierają się na cechach i parametrach konstytuujących odmienność sposobu mówienia poszczególnych osób. I chociaż niektóre wyróżniki indywidualizujące mówców – w związku ze specyfiką rozmówców, konkretnego nagrania i charakteru rozmowy – występują niekiedy jedynie w małym nasileniu (lub wręcz nie występują), to jednak wieloaspektowość badań czyni metodykę wysoce użyteczną. Każdy produkt artykulacji niesie bowiem określone informacje, np. o treściach przekazywanych przez nadawcę, jego stosunku do tematu rozmowy, kompetencji językowej, poziomie napięcia psychoruchowego itd. Możliwa jest więc ekstrakcja cech i parametrów ze strumienia informacji, a następnie odpowiednie klasyfikowanie ich jako właściwości grupowych lub indywidualnych.

W przypadku celowego zniekształcania wypowiedzi mówcy jest skupiony wyłącznie na wycinku zakresu badań identyfikacyjnych – zmianie barwy głosu. Nie uświadamia sobie, że w wypowiedziach przejawiają się jego nawyki językowe i specyfika aktywności psychicznej. Obie te sfery, ujawniające się w trak-

cie mówienia, są kształtowane przez całe życie i niosą wiele cech indywidualnych, które uwidaczniają się tym lepiej, im więcej zostało zgromadzonych wypowiedzi dowodowych.

Praktyka fonoskopijna ZKiChS ABW pokazuje, że w przypadku obszernej wypowiedzi dowodowej fałszerstwo barwy głosu nie stwarza problemu przy identyfikacji mówców. Co jednak, jeśli wypowiedzi dowodowych jest mało, a ocena kompetencji językowej i aktywności psychicznej nie prowadzi do kategorię wniosków ustaleń? Przy nagraniach krótkich, dobrej jakości, niezniekształconych celowo, dominującym elementem identyfikacji mówców jest analiza formantowa. Jak wiadomo, rozkład formantów podlega wielu fluktuacjom, na które wpływa m.in. kontekst fonetyczny, intonacja i ekspresja wypowiedzi. Fluktuacje rozmieszczenia formantów – warunkowane wieloczynnikowo – bywają bardzo znaczne, dlatego dla wykazania indywidualności budowy anatomicznej aparatu mowy konieczne jest porównywanie struktur formantowych w analogicznych ciągach alofonów, wypowiedzianych w możliwie zbliżonym stanie emocjonalnym². Kolejnym warunkiem wykorzystania korelacji pomiarów struktur formantowych jest relatywnie dobra jakość nagrania, gwarantująca odpowiednio szerokie pasmo częstotliwości zarejestrowanego sygnału (ujawniające kolejne formanty) i na tyle duży odstęp sygnał-szum, że formanty nie zacierają się w tle akustycznym.

Jeśli wypowiedzi dowodowe zostały celowo zniekształcone, praktycznie niemożliwe jest uzyskanie wypowiedzi porównawczych w analogicznych realizacjach. Nie można spodziewać się, że sprawca zechce

współpracować w tym względzie z prowadzącym czynność i podejmować próby mówienia w sposób zniekształcony³. Czy oznacza to, że wobec takich przypadków metodyka badań fonoskopijnych jest bezradna? Do uzyskania odpowiedzi na to pytanie posłużył opisany poniżej eksperyment.

Materiał i metody eksperymentu

W eksperymencie wykorzystano transmisję telefoniczną. Każdy z mówców trzykrotnie wypowiadał ciąg słów: „mama”, „lolo”, „lipy”, „zęby”, „syński”, „tromso”, „gaderypoluki”. Wszystkie trzy realizacje słowne były utrwalane z wykorzystaniem taśm analogowych firmy TDK, typu D60 i rejestratora telefonicznego firmy Panasonic, model KX-T2470B. Pierwsza realizacja słowna odbywała się w sposób naturalny, druga z nosem zaciśniętym palcami, trzecia ze zwartymi zębami.

Nagrania – na potrzeby analizy wyników – zostały przetworzone do postaci cyfrowej z wykorzystaniem programu Adobe Audition 2.0.

W eksperymencie wzięło udział trzydziestu mówców. Przed nagraniem zostali oni poinstruowani o konieczności zachowania jednolitego tempa i głośności wypowiedzi. Wypowiadali się wolno, tak aby można by-

ło utrwalić mowę staranną, bez ewentualnych uproszczeń czy przekształceń głosek. W konsekwencji formanty wypowiedzi testowych w obrazie spektrograficznym były możliwie najbardziej wyraziste, a częstotliwości uprzywilejowane energetycznie łatwe do wyznaczenia. Powtarzanie w każdej serii tych samych słów, w krótkim odstępie czasu, miało służyć uzyskaniu analogicznych ciągów alofonów, tak aby został spełniony pierwszy warunek korelacyjnej analizy rozmieszczenia formantów.

Uzyskane trzy próbki wypowiedzi od każdego z mówców były porównywane jedynie w obrębie wypowiedzi tej samej osoby. Porównania dokonano na obrazie spektrograficznym, uzyskanym dzięki zastosowaniu spektrografu Sona-Graph, model 5500, i programu Praat 4.4.24.

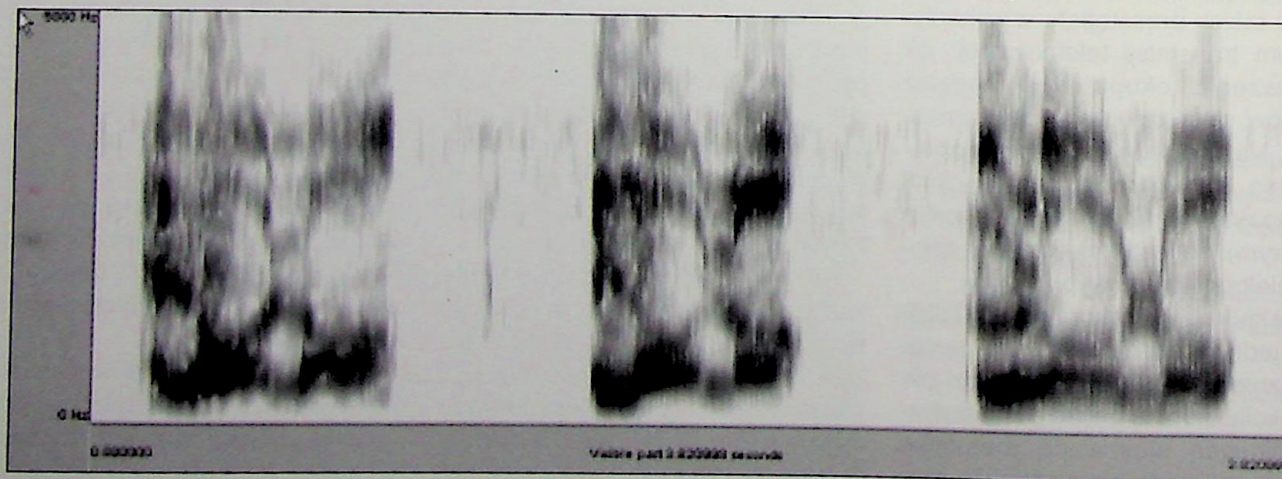
Wyniki i dyskusja

Odpowiednie dobranie zawartości fonetycznej testów miało służyć zminimalizowaniu trudności artykulacyjnych. W efekcie wszyscy badani zrealizowali cały zestaw wypowiedzi testowych. Po nagraniu mówcy zgodnie stwierdzali, że wypowiadanie się ze zwartymi zębami jest mniej wygodne niż mówienie z zaciśniętym nosem. Dlatego tempo mówienia ze

zwartymi zębami – objawiającego się utrudnieniem oddychania – było spowolnione (w przypadku wypowiedzi przedstawionych na rycinie 1 realizacja naturalna trwała 0,78 s, a ze zwartymi zębami 0,89 s). Najszybciej w prezentowanym przykładzie z rycin 1 zrealizowano wypowiedź z zaciśniętym nosem (tu czas wyniósł 0,66 s), jednak nie było to regułą. Niekiedy treść wypowiedzi ze zwartymi zębami była nie do zrozumienia, wówczas badany powtarzał całą serię do uzyskania zadowalających – z punktu widzenia zrozumiałości mowy – wyników. Wyraźność to cecha wypowiedzi dowodzących – sprawca bowiem zwraca uwagę, żeby jego wypowiedzi były właściwie odebrane przez rozmówcę telefonicznego; jest też jednym z podstawowych warunków komunikacji za pośrednictwem transmisji telefonicznej, czyli konieczności przekazywania zrozumiałych informacji.

Następnie zgromadzone wypowiedzi zostały poddane analizie spektrograficznej.

Na wstępie ustalono, że obie metody zniekształcania wypowiedzi nie powodują zawężenia pasma sygnału przesłanego siecią telefoniczną, tj. 300–3400 Hz. Dzięki temu możliwe było wyznaczenie położenia formantu pierwszego, drugiego i trzeciego. Spełniony zatem został także drugi



Ryc. 1. Spektrogram słowa „lolo” wypowiedzianego przez jednego mówcę trzykrotnie. Kolejno: naturalnie, z zaciśniętym nosem i zwartymi zębami
Fig. 1. Spectrogram of word „lolo” uttered 3 times by the same speaker naturally, holding the nose and through set teeth, respectively

z warunków zastosowania analizy formantowej.

Nie stwierdzono istotnych zmian ogólnego rozkładu struktur formantowych, jednak wykazano, że zachodzą przesunięcia częstotliwości formantowych w wypowiedziach celowo zniekształczanych w porównaniu z wypowiedziami naturalnymi.

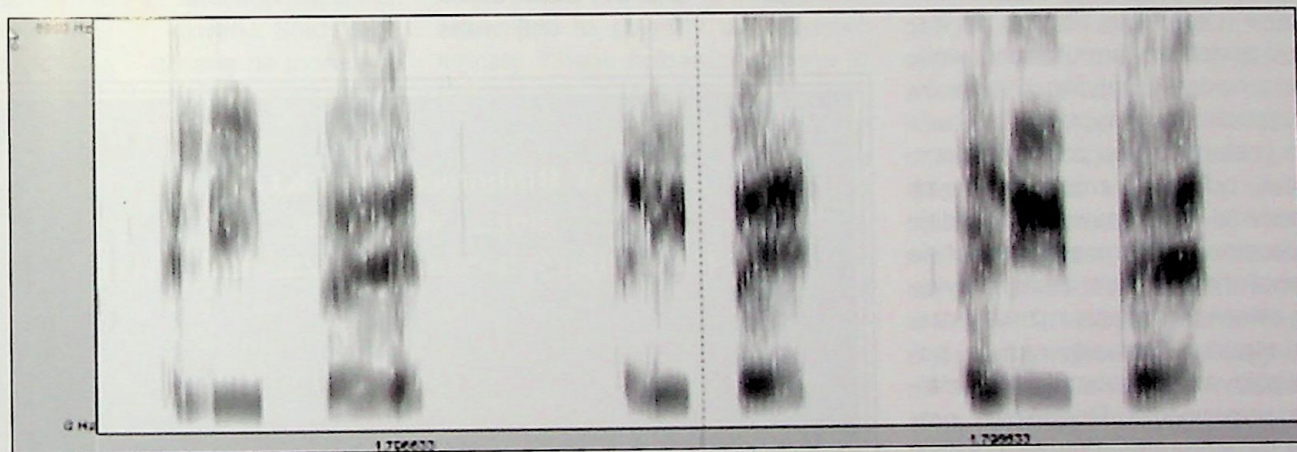
I tak na rycinie 1 widoczne jest podwyższenie drugiego formantu pierwszej głoski „o” w wypowiedzi ze zwartymi zębami (o 66 Hz, tj. 1101 Hz wobec 1035 Hz). W realizacji z zaciśniętym nosem położenie drugiego formantu głoski „o” jest analogiczne do wypowiedzi naturalnej (1035 Hz).

Z kolei na rycinie 2 formant drugi w głosce „y” jest obniżony w wypowiedzi ze zwartymi zębami (o 97 Hz, tj. 1881 Hz wobec 1978 Hz), zaś drugi formant tej głoski w wypowiedzi z zaciśniętym nosem podwyższył się (o 136 Hz, tj. 2114 Hz wobec 1978 Hz).

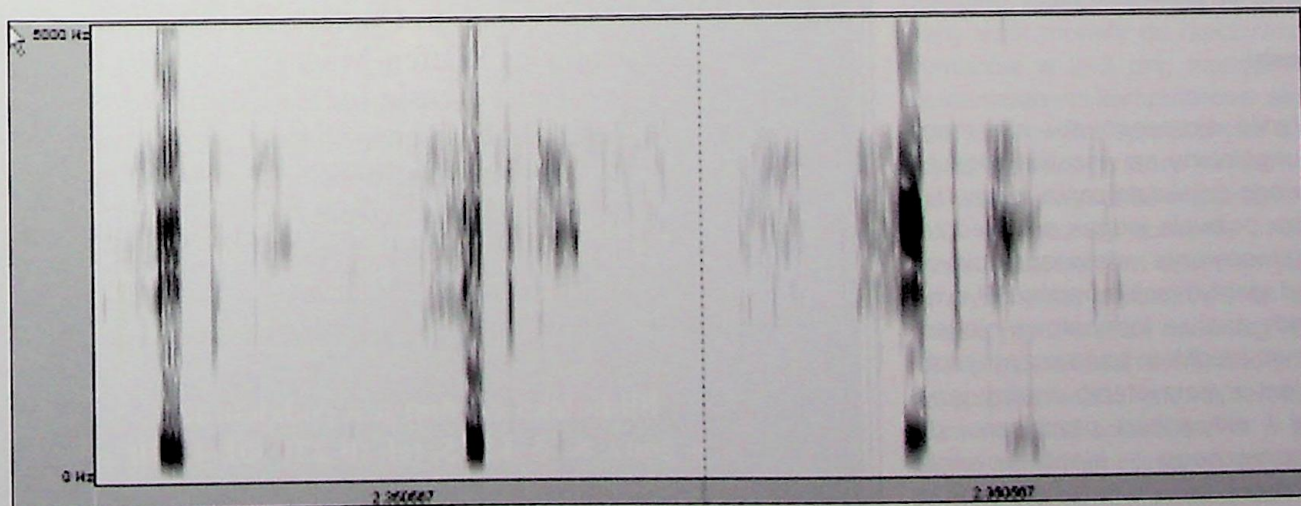
Z ryciny 3 wynika, że słowo „sypki” zostało źle dobrane do testów. Kontekst fonetyczny sprawia, że iloczasy samogłosek są krótkie, przez co obraz spektrograficzny nie spełnia wymogu analizy formantowej.

Przesunięcia położenia formantów nie były jednolite, co być może zostało spowodowane różnicami budowy anatomicznej rezonatorów nadkrtka-

niowych poszczególnych mówców – wówczas ten sam sposób zniekształcania powodowałby różnego stopnia zmianę objętości jam powietrznych. Ponieważ nie wykryto regularności w przesunięciach formantów, niemożliwe jest precyzyjne określenie wpływu sposobów celowego zniekształcania wypowiedzi na poszczególne alofony. To zaś oznacza, że wykorzystywanie w badaniach porównawczych celowo zniekształconych wypowiedzi dowodowych i naturalnych wypowiedzi porównawczych – nawet jeśli będą to analogiczne ciągi alofonów – wiąże się z wprowadzaniem do wyników analizy formantowej błędów o nieznanym poziomie.



Ryc. 2. Spektrogram słowa „lipy” wypowiedzianego przez jednego mówcę trzykrotnie. Kolejno: naturalnie, z zaciśniętym nosem i zwartymi zębami.
Fig. 2. Spectrogram of word „lipy” uttered 3 times by the same speaker naturally, holding the nose and through set teeth, respectively



Ryc. 3. Spektrogram słowa „sypki” wypowiedzianego przez jednego mówcę trzykrotnie. Kolejno: naturalnie, z zaciśniętym nosem i zwartymi zębami.
Fig. 3. Spectrogram of word „sypki” uttered 3 times by the same speaker naturally, holding the nose and through set teeth, respectively

I choć są to zwykle błędy kilkuprocentowe, to jednak interpretacja takich wyników nie upoważnia do formułowania wniosków w zakresie identyfikacji mówców jedynie na podstawie analizy formantowej.

Uzyskano następujący współczynnik korelacji rozmieszczenia formantów dla wszystkich mówców:

- 1) w zestawieniu wypowiedzi naturalnych z wypowiedziami z zaciśniętym nosem współczynnik korelacji wyniósł 0,9826;
- 2) w zestawieniu wypowiedzi naturalnych z wypowiedziami ze zwartymi zębami współczynnik korelacji wyniósł 0,9876.

Są to słabe – jak na warunki fonoskopijne – korelacje, dlatego nie mogą być podstawą formułowania wniosków z badań identyfikacji mówców w przypadku, gdy wypowiedzi dowodowe zostały celowo zniekształcone. Z kolei pobranie zniekształconych wypowiedzi porównawczych wydaje się niecelowe, gdyż w piśmiennictwie kryminalistycznym nie została wykazana zależność między rozmieszczeniem struktur formantowych a poszczególnymi sposobami zniekształcania wypowiedzi. Dodatkowo prowadzący czynność pobrania wypowiedzi porównawczych nie dysponuje wiedzą o sposobie zniekształcenia wypowiedzi dowodowych.

Wnioski

Wyniki eksperymentu nie mogą być uogólnione na wszelkie sposoby celowego zniekształcania mowy. Ich analiza pozwala jednak dostrzec zalety stosowania wieloaspektowych metod identyfikacji mówców. W sytuacji gdy analiza formantowa nie jest pewnym środkiem badawczym, możliwe jest wykorzystanie analizy językowej i aktywności psychicznej do kategorycznego identyfikowania mówców.

Konkludując, z eksperymentu można wyciągnąć następujące wnioski:

1. Porównanie rozkładu formantów dowodowych wypowiedzi celowo zniekształconych z naturalnymi wypowiedziami porównawczymi nie może być podstawą formułowania wniosków w zakresie identyfikacji mówców.
2. Kategoryczna identyfikacja mówców celowo zniekształcających wypowiedzi powinna się opierać na analizie kompetencji językowej i ocenie aktywności psychicznej.

Jacek Rzeszotarski
zdj.: autor

PRZYPISY

- ¹ J. Dolecki, J. Rzeszotarski: Zastosowanie metody językowo-pomiarowej do identyfikacji osób w badaniach fonoskopijnych, „Z Zagadnień Nauk Sądowych” 2002, tom LII, s. 108–123; zob. także J. Rzeszotarski, W. Witkowska-Pawlak: Identyfikacja mówców w obrębie materiału dowodowego, „Problemy Kryminalistyki” 2004, nr 246, s. 15–18;
- ² J. Dolecki, J. Rzeszotarski: Zastosowanie metody fonoskopijnej, op. cit.;
- ³ J. Rzeszotarski, W. Witkowska-Pawlak: Pobranie wypowiedzi porównawczych do badań fonoskopijnych, „Problemy Kryminalistyki” 2004, nr 243, s. 56–66.

**W Bibliotece CLK KGP
można skorzystać
z następujących czasopism zagranicznych**

Archiv für Kryminologie
Computers and Security
Elektor Electronics
Fingerprint Whorld
Forensic Science International
Forensic Science Review
Guide to Copiers
Internationales Waffen-Magazin – Visier
Journal of Forensic Identification
Journal of Forensic Sciences
Kriminalistik
Masterruzhie
Science and Justice
Verkehrsunfall und Fahrzeugtechnik
Wirtschaftsschutz und Sicherheitstechnik

**Biblioteka czynna jest codziennie
w godz. 8.15–12.15, tel. (0-22) 601-45-30**