

PHÂN TÍCH NGỮ ÂM ĐỂ NHẬN DIỆN NGƯỜI NÓI TRONG CÔNG TÁC ĐIỀU TRA

CHIÊM MINH TRƯỜNG¹
NGUYỄN PHƯƠNG TRANG²

Abstract: Based on the theories of phonetics, dialects, forensic science assessment, phonetic apparatus and application of “Voice Workstation VS4.0” software, this article presents an analysis of sound assessment (speaker identification) at the Criminal Science Sub-Institute of the Ministry of Public Security in Ho Chi Minh City. The surveyed voices are those of the cases who used high technology to defraud many people in provinces and cities across the country. Research results show that: (1) Each person has his/her own unique timbre; (2) Pronunciation disorders are individually based; (3) The formants of high-pitched vowels are individually distinctive.

Key words: *voice, pronunciation, inspection, evidence.*

1. Đặt vấn đề

Khoa học hình sự là khoa học của những chứng cứ vật chất đóng vai trò quan trọng trong việc phát hiện, thu lượm đánh giá dấu vết và giám định đưa ra kết luận chính xác, khách quan, đáp ứng yêu cầu của việc điều tra, truy tố, xét xử. Một trong các loại chứng cứ vật chất đó là giọng nói, đối tượng của công tác giám định âm thanh, từ đó có thể truy nguyên người nói trong các vụ án mà tội phạm sử dụng công nghệ cao để thực hiện hành vi phạm tội. Bài viết này trình bày việc ứng dụng lý thuyết ngữ âm học cấu âm, ngữ âm học âm học, kết hợp với phần mềm “Voice Workstation VS4.0” vào việc phân tích tiếng nói trong công tác giám định âm thanh (nhận diện người nói) tại Phân Viện Khoa học Hình sự - Bộ Công an tại Thành phố Hồ Chí Minh.

2. Phương pháp nghiên cứu

Đề tài sử dụng lý thuyết ngữ âm học âm học và ngữ âm học cấu âm, lý thuyết giám định kỹ thuật hình sự, lý thuyết giám định tiếng nói và phương pháp phân tích phổ tiếng nói trên phần mềm.

2.1. Lý thuyết về giám định kỹ thuật hình sự và giám định âm thanh tiếng nói

Các chỉ dẫn về mặt lý thuyết xung quanh cao độ, trường độ, cường độ, âm sắc của giọng nói, cùng cơ chế phát âm, bộ máy cấu âm..., sẽ soi sáng cho việc phân tích, đối chiếu các đặc trưng ngữ âm của giọng nói đang khảo sát.

Nội dung cơ bản của giám định âm thanh (tiếng nói) được xác định là việc giám định tư pháp kỹ thuật hình sự về âm thanh, sử dụng kiến thức, phương tiện, phương pháp khoa học, kỹ thuật, nghiệp vụ của giám định kỹ thuật hình sự nói chung và chuyên ngành giám định âm thanh nói riêng để đưa ra kết luận về chuyên môn về những vấn đề được yêu cầu giám định.

Giám định âm thanh về bản chất là nghiên cứu các đặc điểm của tiếng nói để xác định các thuộc tính của tiếng nói từ đó cho phép truy nguyên người nói.

¹ Viện Khoa học Hình sự - Bộ Công an.

² Trường Đại học Khoa học xã hội và Nhân văn - Đại học Quốc gia Thành phố Hồ Chí Minh.

Giám định âm thanh để truy nguyên đồng nhất người, phải dựa vào lí luận truy nguyên hình sự. Đây là một trong những lí luận căn bản của khoa học hình sự, đó là quá trình xác định và chứng minh sự đồng nhất của các đối tượng vật chất có mối quan hệ với vụ việc hình sự đã xảy ra với các đối tượng vật chất đã được xác định. Nói cách khác, đây là quá trình chứng minh sự đồng nhất.

Truy nguyên hình sự không đòi hỏi sự đồng nhất ở tất cả các thuộc tính của đối tượng. Để chứng minh sự đồng nhất, chỉ cần các thuộc tính cơ bản của đối tượng truy nguyên, tức là các thuộc tính ổn định, bền vững quyết định bản chất của đối tượng.

Trong giám định tư pháp kĩ thuật hình sự về âm thanh, để xác định tiếng nói trong mẫu cần giám định so với tiếng nói trong mẫu so sánh có phải do cùng một người nói ra hay không, trước hết phải phân tích đặc điểm chung, đặc điểm riêng của tiếng nói để xác định các thuộc tính chung, thuộc tính riêng của đối tượng truy nguyên. Từ các thuộc tính của đối tượng truy nguyên (tiếng nói cần giám định, tiếng nói so sánh) giám định so sánh để đưa ra kết luận về sự đồng nhất. Như vậy, lí luận truy nguyên hình sự là cơ sở lí luận cơ bản để chứng minh sự đồng nhất trong giám định âm thanh.

Dựa vào giọng nói cá nhân, chúng ta có thể biết được trạng thái tình cảm của người nói khi nói, hay những đặc tính cá nhân của người nói (giới tính, phương ngữ, cao độ, cường độ, trường độ, đặc điểm ngữ điệu, thói quen phát âm, tật phát âm). Các đặc trưng này tuy không phải là bất biến nhưng nó có tính khá ổn định trong những giai đoạn dài của cuộc đời. Đây cũng là cơ sở quan trọng để người nghiên cứu có thể khoanh vùng, nhận diện giọng nói.

2.2. Phương pháp phân tích phổ tiếng nói trên phần mềm

- Phần mềm Voice Workstation 4.0 phân tích tiếng nói ở dạng sóng hình sin, phân tích phổ, lọc tạp âm nâng cao chất lượng mẫu, xử lí chính xác 1/1000 (một phần nghìn) giây.

- Đánh giá độ giống nhau của các âm tiết đồng âm khi so sánh các vệt formant trên phổ ba chiều của nguyên âm trong âm tiết này dựa trên diễn tiến của các formant, bề rộng và tỉ lệ tương đối giữa chúng. So sánh định lượng formant của nguyên âm trong âm tiết này với nhau, nếu sai khác giữa các formant 1, 2, 3, 4 lần lượt nhỏ hơn 15,4%; 10%; 6,3%; 5,2% thì có thể kết luận hai âm tiết đồng âm này là đồng nhất.

- Nếu số lượng âm tiết đồng nhất vượt quá một ngưỡng nhất định thì có thể *kết luận khẳng định* (hai mẫu tiếng nói này do cùng một người nói), ngược lại là *kết luận khả năng* hoặc *phủ định* nếu số âm tiết đồng nhất quá ít.

- Phân tích phổ âm thanh bằng phần mềm chuyên dụng để phân tích các formant và vân phổ của các âm tiết tương ứng tiếng nói của trong các file âm thanh mẫu cần giám định và tiếng nói của đối tượng trong mẫu so sánh.

3. Tiến hành phân tích tiếng nói

Việc phân tích tiếng nói sẽ dựa trên các tệp âm thanh mà cơ quan điều tra cung cấp. Các tệp âm thanh này thường lẫn tạp âm, nhiễu nền, âm lượng nhỏ do chúng được ghi âm từ hiện trường xảy ra vụ việc. Chúng tôi lần lượt phân tích đặc điểm phát âm, giọng nói, tật phát âm (nếu có) liên quan đến thói quen phát âm hoặc đặc điểm riêng của bộ máy cấu âm của người nói. Những yếu tố “bị đánh

dấu” sẽ được phân tích trên phần mềm qua các thông số kỹ thuật, từ đó xác nhận sự tương đồng hay khác biệt của các giọng cần phân tích. Tất nhiên, dữ liệu so sánh càng nhiều thì nhận định càng có cơ sở.

Năm 2020, Cơ quan Cảnh sát điều tra Bộ Công an điều tra vụ án “Lừa đảo chiếm đoạt tài sản” xảy ra tại Thành phố Hà Nội, Thành phố Hồ Chí Minh và một số địa phương khác. Cơ quan điều tra xác định đối tượng sử dụng điện thoại di động mạo danh Phó Chủ tịch UBND tỉnh N. gọi điện cho ông L.V.N. ngụ tại Hà Nội yêu cầu ông L.V.N. chuyển 550 triệu đồng vào tài khoản mở tại ngân hàng Sacombank. File âm thanh thu được có số hiệu 033643_1918.wav.

Chúng tôi tiến hành nghiên cứu trên mẫu tiếng nói trong file âm thanh 033643_1918.wav, cụ thể là phân tích tiếng nói của người tên Đ.N.H. kí hiệu Nam2 và người nam còn lại kí hiệu Nam1.

Các đặc trưng ngữ âm của giọng nói trong file đã được so sánh phân tích theo bảng sau:

STT	Các đặc trưng nhận dạng	
1	Giới tính	Giọng nam
2	Phương ngữ	Nói giọng Tây Nam bộ
3	Cao độ giọng nói	Trầm
4	Cường độ giọng nói	Trung bình
5	Trường độ âm tiết	Ngắn
6	Đặc điểm ngữ điệu	Tốc độ phát âm chậm
		Phát âm đều giọng
		Nói không lưu loát
		Nói dính hai âm tiết <i>dạ-em; hai-sáu</i>
		Âm lượng bình thường
		Phát ngôn ngắn
7	Đặc điểm ngữ âm	Phát âm không phân biệt thanh hỏi - thanh ngã
		Phát âm sai <i>kiểm tra giùm em → kim cha dùm em; tài khoản → tài phán; trăm triệu → chảm chịu</i>

Nhận xét: Các đặc trưng ngữ âm của giọng nói cho thấy đối tượng này nói phương ngữ Tây Nam Bộ, từng sinh sống ở miền Tây Nam Bộ.

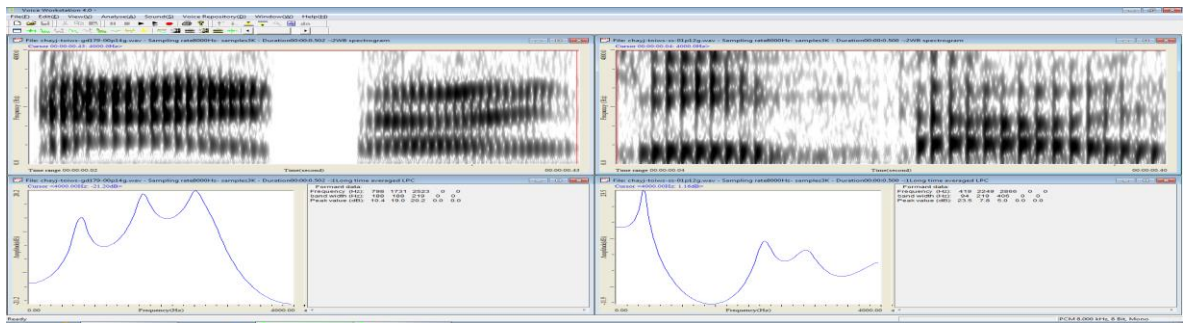
- Ngày 9 tháng 11 năm 2020, Cơ quan điều tra cung cấp cho chúng tôi file âm thanh *Thoại 003.m4a* có tiếng nói của người tên Đ. N. H. (tên người trong file âm thanh 033643_1918.wav).

Tiếng nói của đối tượng kí hiệu Nam2 và anh tên H được miêu tả dưới bảng sau:

Đặc trưng	“033643_1918.wav”	“Thoại 003.m4a”
Giới tính	Giọng nam	Giọng nam
Phương ngữ	Nói giọng Tây Nam Bộ Phát âm: <i>năm sanh; tài phán</i>	Nói giọng Bắc Bộ Phát âm: <i>năm sinh; tài khoản</i>

Nhận xét: Qua phân tích đặc trưng ngữ âm và phân tích phổ tiếng bằng phần mềm chuyên dụng, chúng tôi khẳng định tiếng nói của đối tượng có kí hiệu Nam2 trong file âm thanh 033643_1918.wav và người tên Đ.N.H. trong file âm thanh *Thoại 003.m4a* không đồng nhất.

Phổ tiếng nói của hai âm tiết “tài khoản” ở hai file âm thanh 033643_1918.wav và *Thoại 003.m4a* được thể hiện như sau:



Ảnh phổ âm tiết “tài khoản” của Nam2

Ảnh phổ âm tiết “tài khoản” của H.

Nhận xét: Formant của các nguyên âm trong các âm tiết được phân tích ở trên và ảnh phổ biểu diễn tương ứng cho thấy các formant và vân phổ khác nhau, tiếng nói của người nam có kí hiệu Nam2 trong file âm thanh *033643_1918.wav* và tiếng nói của Đ.N.H. trong file *Thoại 003.m4a* không đồng nhất. Từ đó, chúng tôi khẳng định hai mẫu tiếng nói không phải của cùng một người nói ra.

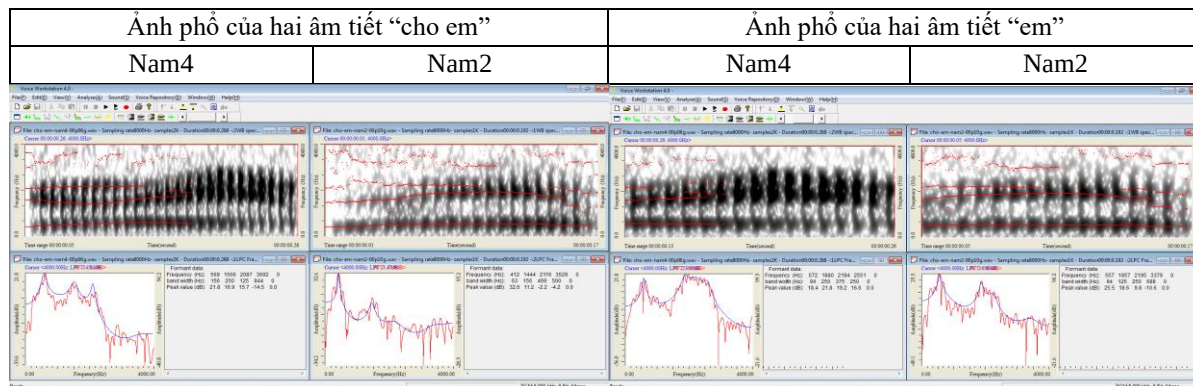
- Ngày 11 tháng 1 năm 2021, Cơ quan điều tra thu được 02 file âm thanh *02.12.2020-14.55-094884.wav* và *24.11.2020-17.07-094884.wav*. Chúng tôi tiếp tục phân tích đặc điểm tiếng nói của người xưng tên P.V.P có kí hiệu Nam4 trong 02 file âm thanh, kí hiệu Nam3 trong file *02.12.2020-14.55-094884.wav*, kí hiệu Nữ trong file *24.11.2020-17.07-094884.wav*.

Tiếng nói của người nam có kí hiệu Nam4 trong hai file âm thanh trên với tiếng nói của đối tượng kí hiệu Nam2 có các đặc điểm như sau:

Các đặc trưng			
STT	Đặc điểm	Nam2	Nam4
1	Giới tính	Giọng nam	Giọng nam
2	Phương ngữ	Nói giọng Tây Nam bộ	Nói giọng Tây Nam bộ
3	Cao độ giọng nói	Trầm	Trầm
4	Cường độ giọng nói	Trung bình	Trung bình
5	Trường độ âm tiết	Ngắn	Ngắn
6	Đặc điểm ngữ điệu	Tốc độ phát âm chậm	Tốc độ phát âm chậm
		Phát âm đều giọng	Phát âm đều giọng
		Nói không lưu loát	Nói không lưu loát
		Nói dính hai âm tiết <i>dạ-em, hai-sáu</i>	Nói dính hai âm tiết <i>hai-sáu</i>
		Âm lượng bình thường	Âm lượng bình thường
		Phát ngôn ngắn	Phát ngôn ngắn
7	Đặc điểm ngữ âm	Phát âm không phân biệt thanh hỏi - thanh ngã	Phát âm không phân biệt thanh hỏi – thanh ngã
		Phát âm <i>kiểm tra giùm em → kim cha dùm em; tài khoản → tài phẩn; trăm triệu → chăm chịu</i>	Phát âm <i>khoản vay → phản dây</i>
Các âm tiết có phát âm tương đồng nhau: <i>em, tiền, khoản, hai-sáu, ba, không, bốn, rồi.</i>			

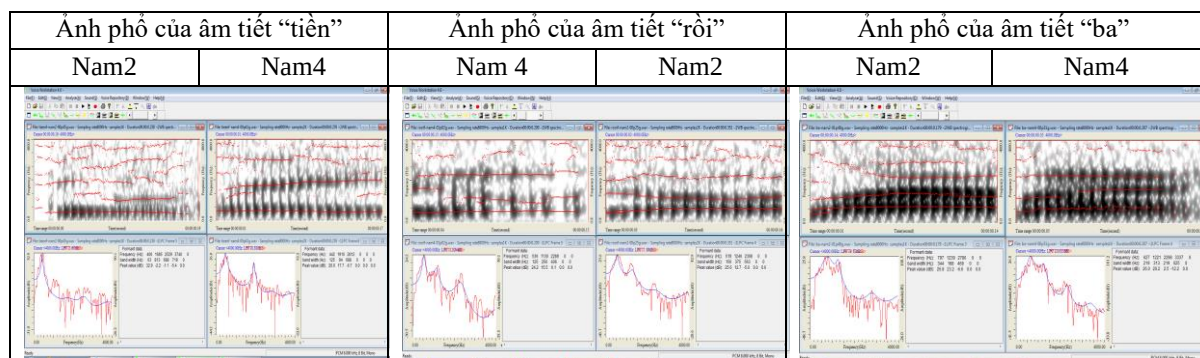
Chúng tôi nhận thấy hai mẫu tiếng nói kí hiệu Nam2 tương đồng với giọng Nam4. Tuy nhiên khả năng lặp lại âm tiết ở hai mẫu còn quá ít nên chưa đủ cơ sở để đưa ra khẳng định mà chỉ nhận định có khả năng đó là tiếng nói của người xưng P.

Phân tích các phổ tiếng của Nam2 và Nam4:



Ảnh phổ của hai âm tiết “cho em” có đặc điểm chung là kết cấu chặt, các vết formant và giá trị tại các điểm tương đồng nhau, tuy nhiên có đặc điểm riêng là hai âm tiết “cho em” của Nam4 có trường độ dài hơn 21/100 giây so với 16/100 giây.

Ảnh phổ âm tiết “tiền” có đặc điểm là trường độ như nhau, vân phổ như nhau, các F2, F3 tương đương nhau, tuy nhiên phổ tiếng nói của Nam2 không thể hiện rõ.



Ảnh phổ âm tiết “rời” có đặc điểm chung là các F1, F2, F3 giống nhau, tuy nhiên đặc điểm riêng là vân phổ có sự khác nhau, trường độ âm tiết có sự khác nhau.

Ảnh phổ âm tiết “ba” có đặc điểm chung là các vân phổ đều nhau, tuy nhiên F3 của Nam2 rất cao lúc mở đầu âm tiết sau đó đi xuống, F3 của Nam4 đi ngang không đổi.

Nhận xét: Formant của các nguyên âm trong các âm tiết được phân tích ở trên và ảnh phổ biểu diễn tương ứng cho thấy các formant và vân phổ tiếng của đối tượng có kí hiệu Nam2 và tiếng nói của người nam có kí hiệu Nam4 có nhiều đặc điểm tương đồng. Bên cạnh đó cũng có nhiều đặc điểm riêng, tiếng nói của người nam có kí hiệu Nam4 trong hai file âm thanh ít, âm tiết lặp lại quá ít, tiếng nói nhỏ nên chưa đủ cơ sở kết luận đồng nhất. Chúng tôi cho rằng, có khả năng tiếng nói của Nam4 và Nam2 là của một người, nếu như số âm tiết được phát âm đồng nhất giữa Nam4 và Nam2 nhiều hơn.

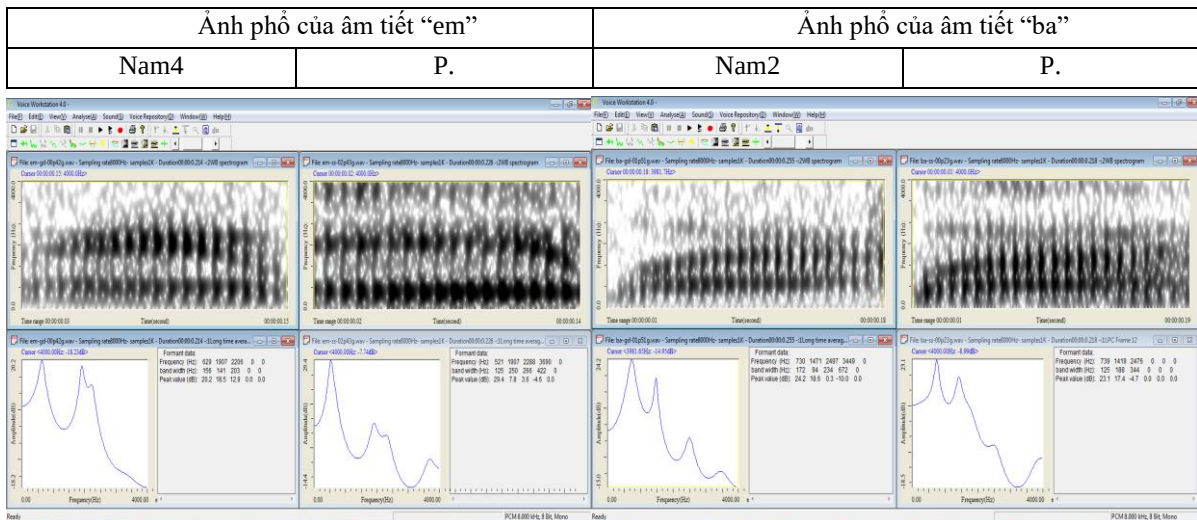
Từ nhận định của chúng tôi, Cơ quan điều tra đã tiến hành các biện pháp nghiệp vụ và đến ngày 05 tháng 4 năm 2021 thì cung cấp cho chúng tôi tiếng nói của P.V.P qua file âm thanh 210319_1539.mp3, đồng thời so sánh với tiếng nói của đối tượng kí hiệu Nam2, Nam4 với các đặc trưng ngữ âm như sau:

Các đặc trưng			
STT	Đặc điểm	Nam2	P. V. P.
1	Giới tính	Giọng nam	Giọng nam
2	Phương ngữ	Nói giọng Tây Nam bộ	Nói giọng Tây Nam bộ
3	Cao độ giọng nói	Trầm	Trầm
4	Cường độ giọng nói	Trung bình	Trung bình
5	Trường độ âm tiết	Ngắn	Ngắn
6	Đặc điểm ngữ điệu	Tốc độ phát âm chậm	Tốc độ phát âm chậm
		Phát âm đều giọng	Phát âm đều giọng
		Nói không lưu loát	Nói không lưu loát
		Nói dính hai âm tiết <i>dạ-em, hai-sáu</i>	Nói dính hai âm tiết <i>dạ-em; hai-sáu</i>
		Âm lượng bình thường	Âm lượng bình thường
		Phát ngôn ngắn	Phát ngôn ngắn
7	Đặc điểm ngữ âm	Phát âm không phân biệt thanh hỏi - thanh ngã	Phát âm không phân biệt thanh hỏi - thanh ngã
		Phát âm <i>kiểm tra giùm em → kim cha giùm em; tài khoản → tài phân; trăm triệu → chăm chịu; hai sáu → hai sáu;...</i>	Phát âm <i>tài khoản → tài phân; trăm triệu → chăm chịu; hai sáu → hai sáu;...</i>
Các âm tiết phát âm tương đồng nhau: <i>em, tiền, khoản, hai-sáu, ba, không, bốn, rồi, dạ-em, đồng ý</i>			

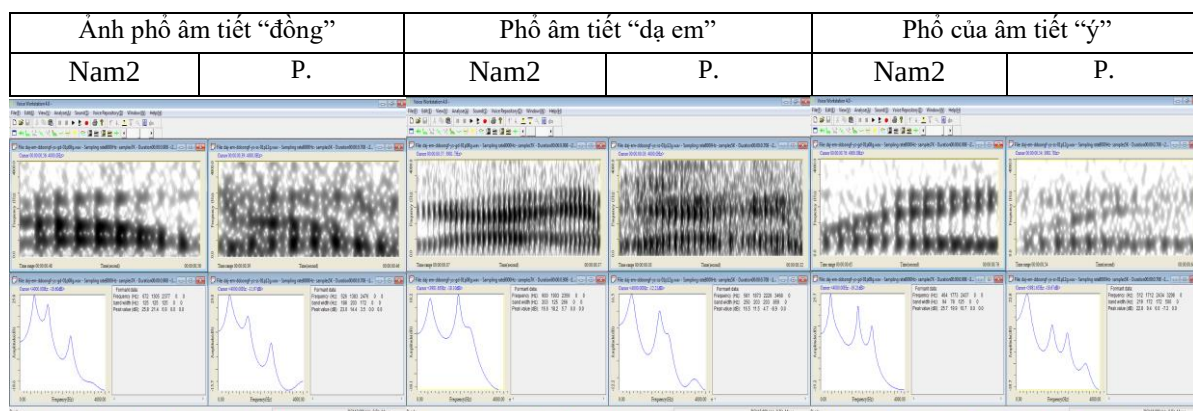
Nhận xét: Qua phân tích ngữ âm, chúng tôi thấy hai mẫu tiếng nói kí hiệu Nam2, Nam4 phát âm có nhiều âm tiết đồng nhất với tiếng nói của P.V.P.

Đồng thời, phân tích phổ tiếng nói của Nam2, Nam4 so sánh với tiếng nói P. thì có thể thấy:

- Ảnh phổ âm tiết “em” có vân phổ như nhau, các vệt formant và giá trị tại các điểm tương đồng nhau.



- Ảnh phổ âm tiết “ba” có vân phổ tương đồng nhau, F3 rất cao khi mở đầu âm tiết sau đó có chiều hướng giảm dần, F1, F2 có giá trị tương đồng nhau.



- Ảnh phổ âm tiết “đồng” có vân phổ tương đồng nhau, các vệt formant và giá trị tại các điểm tương đồng nhau.

- Ảnh phổ hai âm tiết “dạ em” có vân phổ tương đồng nhau, hai âm tiết có kết cấu chặt, các vệt formant và giá trị tại các điểm tương đồng nhau.

- Ảnh phổ âm tiết “ý” có vân phổ tương đồng nhau, các vệt formant và giá trị tại các điểm tương đồng nhau.

Nhận xét: Formant của các nguyên âm trong các âm tiết được phân tích ở trên ảnh phổ biểu diễn tương ứng cho thấy tiếng nói của người nam có kí hiệu Nam2, Nam4 trong 03 file âm thanh và tiếng nói của P.V.P. trong file 210319_1539.mp3 là đồng nhất.

4. Kết quả nghiên cứu

Thông qua phân tích ngữ âm và ứng dụng phần mềm Voice Workstation 4.0 để phân tích formant của tiếng nói, chúng tôi đã thu được kết quả:

Về đặc điểm formant

- Hai âm tiết do cùng một người nói ra có từ hai hoặc ba formant giống nhau.
- Khoảng cách giữa hai formant F2 và F3 có tính ổn định cao.

Về đặc điểm trường độ âm tiết

- Trường độ và vân phổ tiếng nói có tính ổn định cao.
- Chuỗi hai hay ba âm tiết giống nhau khoảng cách giữa các âm tiết của một người nói có tính ổn định cao.
- Kết cấu giữa hai âm tiết có tính ổn định rất cao.

Kết quả nghiên cứu là cơ sở để khẳng định: (1) Mỗi người khi nói có âm sắc riêng mang tính cá biệt; (2) Tập phát âm mang tính cá biệt; (3) Các formant của âm tiết có âm vực cao có tính cá biệt.

Căn cứ vào kết quả nghiên cứu trên, chúng tôi khẳng định tiếng nói người nam có kí hiệu Nam2, Nam4 và tiếng nói của P. V. P. là của cùng một người. Kết luận này đã là một trong những chứng cứ quan trọng để Cơ quan điều tra xác định đúng tội phạm và khép lại vụ án nói trên.

4. Kết luận

Có thể nói, ngữ âm học là lĩnh vực được ứng dụng ngày càng hiệu quả trong công tác giám định âm thanh để truy nguyên người nói chứng minh hành vi phạm tội cũng như hành vi vi phạm pháp luật, góp phần đảm bảo an ninh chính trị và trật tự an toàn xã hội trên địa bàn các tỉnh, thành phố phía nam nước ta. Đây là một lĩnh vực giám định tư pháp kỹ thuật hình sự vô cùng quan trọng và không thể thiếu của khoa học hình sự thế giới nói chung và của Việt Nam nói riêng trong tiến trình cải cách tư pháp và đấu tranh với các loại tội phạm sử dụng công nghệ cao để thực hiện hành vi phạm tội.

Trên thế giới, ngành giám định tiếng nói đã được nghiên cứu từ khá lâu và giành được rất nhiều thành tựu quan trọng, hỗ trợ cho công tác điều tra tội phạm của các quốc gia phát triển như: Hoa Kỳ, Anh, Pháp, Nhật Bản... Ở Việt Nam, giám định tiếng nói bước đầu đạt được kết quả đáng khích lệ. Từ năm 2016, tại Phân viện Khoa học Hình sự - Bộ Công an tại Thành phố Hồ Chí Minh đã triển khai công tác giám định âm thanh, giám định nhanh chóng kịp thời chính xác các vụ án kinh tế, tham nhũng, hối lộ, cưỡng đoạt tài sản, đánh bạc, chống phá nhà nước, tội phạm công nghệ cao, v.v. Vai trò của cơ quan khoa học hình sự ở phía nam bước đầu được khẳng định. Kết luận giám định âm thanh của Phân viện Khoa học Hình sự - Bộ Công an tại Thành phố Hồ Chí Minh đã giúp xác định truy nguyên người nói và xác định nội dung trao đổi giữa các đối tượng, đây là nguồn chứng cứ vật chất rất quan trọng để đấu tranh hiệu quả với loại tội phạm này. Trên cơ sở đó, các cơ quan tố tụng các cấp đã giải quyết nhanh quá điều tra, truy tố, xét xử các vụ án xảy ra trên địa bàn các tỉnh, thành phố phía Nam, các tỉnh miền Trung và Tây Nguyên.

TÀI LIỆU THAM KHẢO

Tiếng Việt

1. Phạm Đăng Diệu, *Giản yếu giải phẫu người*, Nxb Y học, 2019.
2. Cao Xuân Hạo, *Âm vị học và tuyến tính*, Nxb Khoa học xã hội, 2017.
3. Nguyễn Quang Hồng, *Âm tiết và loại hình ngôn ngữ*, Nxb Đại học Quốc gia Hà Nội, 2001.
4. Quốc hội Nước CHXHCN Việt Nam, *Luật Giám định tư pháp*, Nxb Công an Nhân dân, Hà Nội, 2012.
5. Đoàn Thiện Thuật, *Ngữ âm tiếng Việt*, Nxb Đại học Quốc gia Hà Nội,
6. Tư liệu của Phân viện Khoa học Hình sự - Bộ Công an tại Thành phố Hồ Chí Minh.
7. Viện Khoa học Hình sự - Bộ Công an, *Nghiệp vụ giám định Kỹ thuật hình sự*, Nxb Công an nhân dân, Hà Nội, 2018.

Tiếng Anh

8. *Handbook of forensic services*, U. S. Department of Justice Federal Bureau of Investigation Laboratory Division, 2007.
9. Philip Rose, *Forensic speaker identification*, Taylor & Francis, Forencis Science Series, 2002.
10. Amy Neustein, Hemant A. Patil, *Forensic speaker recognition: Law enforcement and counter - terrorism*, Springer, 2012.