

범죄음성의 식별

한림대학교 자연대학 언어청각학부

고 도 홍

서 론

범인의 음성을 녹취하여 본격적으로 컴퓨터를 이용하여 음향분석을 하는 것은 스펙트로그래프가 등장하는 1940년대 말 이전에는 없었다고 해도 과언이 아니다. 그러나 청지각적인 문제는 아마 수 백년 전부터 법정에서 논란이 있었을 것으로 추정한다. 영국에서는 이미 1660년에 법정에서 청지각적인 검사를 시행하였다는 기록이 있으며, 미국에서는 20세기 초엽에 플로리다 주에서 이미 청지각적인 증거를 법정에서 채택하였다는 기록이 있다. 오늘날 범죄음성에 대한 연구는 무고한 어린이의 유괴와 폭발물 설치협박 등과 같은 죄질이 무거운 범죄에 이용되기 때문에 문제가 심각한 실정이다.

현재 전문가들에 의해 판별된 범죄음성의 분석결과가 법정에서 무죄와 유죄를 결정짓는 단서를 제공할 수 있다는 데에 있어서는 아직도 많은 논란의 여지가 있다. 그것은 손가락지문(finger print)과 같이 음성지문(voice print)이 정교하지 못하다는 데 있다. 그럼에도 불구하고 현재 미국 연방수사국(FBI)과 같은 곳에서 거짓말탐지기(lie detector)와 함께 범인의 색출에 기여를 하고 있다.

우리는 화자식별(speaker identification)과 화자확인(speaker verification)을 혼동하는 경우가 많다. 화자식별은 어떤 음성샘플로부터 미지의 화자를 식별해 내는 과정이라고 정의할 수 있다. 이에 대하여 화자확인(tele-banking)과 같이 예금자의 안전(security)을 본인임을 확인하는 과정이라고 정의할 수 있다. 따라서 범인 음성을 가려내는 것은 화자식별의 한 가지 방법이다. 물론 화자식별이나 화자확인은 음성인식(speech recognition)의 두 가지 방법이라고 할 수 있다.

여기서는 범인음성의 식별을 위한 두 가지 방법인 음향분석과 청지각검사에 대하여 알아보기로 한다.

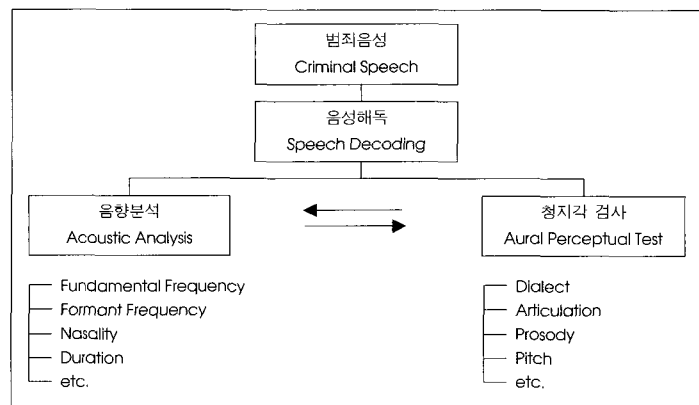


Fig. 1. 범죄음성의 식별체계.

범죄음성 식별 방법

범행에 이용된 녹취 증거자료로서 수집된 음성을 과학적으로 식별하는 것은 범인의 색출에 크게 기여할 수 있기 때문에 서구에서는 오래 전부터 이에 대한 연구가 활발히 진행되고 있다. 범행에 사용된 음성은 특히 공중전화를 이용하기 때문에 최상의 음질을 확보하기가 쉽지 않다. 그럼에도 한정된 양의 음성샘플을 이용하여 동일인의 여부, 특정방언, 교육정도, 초범여부, 등을 밝혀야 하기 때문에 범죄음성의 식별은 매우 까다로울 수밖에 없다.

범행에 이용된 음성이라고 일반인의 음성식별과 크게 다르지는 않지만 분석의 결과가 재판에 영향을 미칠 수 있기 때문에 매우 신중한 접근이 요구된다. 범죄음성의 분석이 음향실험에 의한 결과이든 청지각검사에 의한 결과이든 음향적인 요인과 청지각적인 요인을 동시에 고려하여야 한다. 또한 이러한 연구를 주도하는 책임자는 실험음성학, 방언학 및 언어심리학 분야에 정통한 고도로 숙련된 전문가이어야 한다.

범죄음성의 식별을 위해서는 고성능 녹음기와 마이크가 필수적이다. 녹음에는 현장녹음(live recording)와 전화녹음(telephone recording) 두 가지 경우가 있다. 현장녹음의 경우에는 가급적 마이크나 녹음기를 혐의자(suspect) 바로 앞에 놓고 날짜, 시간, 장소를 기록해야 한다. 녹음테이프를 정해진 녹취시간에 맞추어야 하고, 녹음이 다 끝나면 재생상태를 다시 확인한다. 만약을 위해 같은 음성샘플을 복사해 둘 필요가 있다. 전화녹음에는 주로 도청장치(wiretapping device)에 의존해야 하기 때문에 고성능(high fidelity) 녹음기가 필수적이다. 이때 녹음하는 동안 혐의자의 음성은 당사자의 수화기와 수사관(investigator) 헤드폰을 통해 동시에 수용된다. 음향분석이나 청지각검사를 하기 전에 일단 녹취된 내용을 그대로 옮긴 전사자료(transcribed text)를 준비하여야 한다. 끝으로 범인들은 자신의 목소리를 위장(disguise)하기 위해서 흔히 아주 부드러운 목소리(soft voice)나 속삭임(whispers)을 이용할 수 있으나 음성을 왜곡(distortion)시키는 음성변조기를 사용하기 전에는 위장발화가 성공하기는 어렵다.

범죄음성의 식별 방법으로는 녹취된 자료에 대하여 음향분석(acoustic analysis)과 청지각검사(aural-perceptual test) 두 가지 방법이 있다. 그러나 이 두 가지 방법은 각각 독립적으로 기능하지 않고 상호보완적으로 작용한다. 즉, 음향분석을 할 때도 지각검사가 선행될 필요가 있고 지각검사의 결과도 음향분석의 결과와 일치되는가를 확인할 필요가 있다.

1. 음향분석

범죄음성 식별을 위한 음향분석을 위해서는 몇 가지 고려해야 할 요소들이 있다. 첫째, 분석 대상으로 이용될 수 있는 음성샘플(speech sample)이 충분하여야 한다. 음성샘플의 수가 많으면 많을 수록 신뢰도가 높아질 수 있다. 그러나 음향분석에 있어서는 음성샘플의 양(quantity)보다는 질(quality)이 더 중요하다. 우선 녹음상태에 잡음이 없고 아주 깨끗해야 한다. 뿐만 아니라, 범인의 음성을 녹취한 자료와 수사선상에 오른 혐의자(suspect)의 음성녹취에 가급적이면 많은 공통된 단어(words), 구절(phrase, clause), 문장(sentence)들을 추출할 수 있어야 한다. 특히 공통된 단어(paired words)의 수가 적어도 10개 이상이 되어야 신뢰도를 유지할 수 있다. 음향분석을 하기에 앞서서 반드시 고려해야 할 사항은 기초지각검사가 선행되어야 한다. 범인과 혐의자의 음성을 비교할 때, 음성샘플 간에 감정의 변화가 지나치게 크거나, 사투리가 현격하게 다르거나, 발화의 속도가 크게 차이가 나면 음향분석으로 신뢰할 수 있는 결과를 추출해 내기가 어렵게 된다.

일반적으로 음향분석에 이용되는 주요 음향파라미터를 열거하면 다음과 같다:

1) 발화속도(speech rate)

분당 음절 발화수(syllables per minutes : SPM)를 계산한다. 이때, 일반적으로 2초 이상 되는 휴지(pause)는 발화속도의 계산에서 제외시킨다.

2) 단어 또는 문장의 길이(duration)

3) 기본주파수(fundamental frequency : Fo)

4) 발성개시시간 비교(voice onset time : VOT)

5) 마찰주파수 및 마찰구간(frequency & duration of frication)

6) 비성(nasality)

7) 특정 모음의 포먼트 주파수

8) 모음 포먼트 공간(vowel space)

9) 주파수 변인(jitter)

10) 진폭 변인(shimmer)

11) 문강세(intonation)의 비교

2. 청지각검사

범죄음성 식별을 위한 청지각검사를 위해서 가장 중요한 선행조건으로 분석 대상으로 이용될 수 있는 음성샘플(speech sample)이 충분하여야 한다. 말할 것도 없이, 음성샘플의 수가 많으면 많을 수록 비례해서 신뢰도가 높아진다. 즉, 음성샘플의 양은 음향검사의 경우보다 더 중요하다고 할 수 있다. Bricker & Pruzansky(1966)의 연구에

따르면, “자극(stimulus)이 충분한 경우에는 청지각검사가 98%까지 향상될 수 있는 반면에 자극이 부족하면 신뢰할 수 없는 수준인 56%까지 떨어지게 된다.”고 한다.

청지각검사에서 집단으로 지각검사에 참여하는 청자들에게 잘 알려진 화자의 음성식별이 더 용이하지만 범인음성을 다루는 경우에는 적용되지 않는다. 간혹 증인에게 자극음성을 들려주고 진위를 판단하게 하는 경우는 종종 있다. 또한 청지각검사에서 전체되어야 하는 것은 검사에 참여하는 청자들은 범인의 음성을 식별해야 하는 화자식별에 능력차가 있다는 사실을 알아야 한다.

음향분석의 경우에는 결과를 객관적으로 판단하는데 비하여, 청지각검사의 경우에는 결과를 검사에 참여하는 청자들의 주관적인 판단에 의존하게 된다. 일반적으로 판단척도(scale)는 각 항목별로 0~10까지로 한다. Hollien (2002)에 따르면, 0~10 척도에서 0~3은 완전히 다른 사람용, 7~10의 경우에는 동일인으로, 그리고 4~6은 둘의 음성이 완전히 다르지 않을 수도 있다는 단서를 보여준다고 한다. 또한 검사에 참여하는 청자의 수는 결과에 대한 신뢰도를 평가하는데 중요하다고 할 수 있다. 다만, 청자들은 음성을 식별하는 방법에 대하여 사전에 훈련이 필요하다. 훈련이 잘된 집단과 훈련이 덜 된 집단 간에는 신뢰도 면에서 커다란 차이를 보일 수도 있다. 특히 검사를 주관하는 책임자는 고도로 훈련된 전문가이어야 한다.

일반적으로 청지각검사에 이용되는 주요 변인을 열거하면 다음과 같다:

1) 피치(pitch)와 강도(intensity)

2) 방언(dialect)

(가) 방언어휘 선별, (나) 억양 테스트

방언어휘와 억양을 분석하여 ① 특정지역 출신인가 ② 출생지와 성장지가 다른가를 추정한다. 이때 가급적 청지각 검사에 참여하는 검사자

3) 조음(articulation)

(가) 모음(vowels), (나) 자음(consonants)

(다) 오조음(misarticulation), (라) 비유창성(nonfluencies)

4) 운율

(가) 속도(speech rate), (나) 장단(length)

5) 음질(voice quality)

(가) 기식성(aspiration), (나) 비성(nasality)

6) 교육정도(education)

(가) 어휘선택 능력, (나) 교양정도

7) 정서상태(emotion)와 태도(attitude)

(가) 초범여부, ① 안정성, ② 대담성

8) 발화습관(speaking habits)

9) 기타

(가) 음성장애(speech disorders)

(나) 연령추정(age), 예) 20대 초반, 20대 중반, 20대 후반

(다) 지인(지인)여부, ① 가족사항, ② 주변 지리(위치추적)

청지각검사를 위해서는 각 항목별로 점수를 매겨야 하기 때문에 동일한 음성샘플을 여러 차례 들려주어야 한다. 그리고 무엇보다도 청지각검사에 참여하는 청자들은 선입견을 가져서는 안 된다. 또한 앞의 문항이 뒤의 문항에 영향을 주지 않도록 한다.

결 어

요즘 세대가 각박해지고 힘악해짐에 따라 범죄는 갈수록 대담해지고, 충동적이고, 점점 지능화되어 가고 있다. 어린이를 유괴(kidnapping)하여 부모를 위협하여 금품을 요구하거나, 폭발물을 설치해놓았다고 협박을 하는 사례가

늘어나고 있다. 지금까지 우리나라에서 범죄음성을 식별하는 체계적인 연구는 거의 없었지만, 실용적인 측면에서 음성과학자들이 관심을 가져야 하는 분야임에 틀림없다고 생각한다.

REFERENCES

- 고도흥. 개별화자의 음성파라미터 추출에 관한 연구. 음성과학 2003;10(2):129-43.
Abberton E. Aspects of Voice Qualities in Women. Proceedings (The 1st International Conference on Speech Sciences (ICSS), Korean Association of Speech Sciences;2002. p.128-31.
Bricker P. Pruzansky. Effects of Stimulus Content and Duration on Talker Identification. J Acousti Soc Amer 1966;40:1441-50.
Ko DH. Voice Similarities between Sisters. Speech Sciences 2001;8(3):43-50.
Ko DH. Voice Similarities between Brothers. Speech Sciences 2002;9(2):1-11.
Hollien H. The Acoustics of Crime. Plenum Press: New York and London;1990.
Hollien H. Forensic Voice Identification. Academic Press;2002.