
Assinatura Vocal

1. Introdução

A análise de vestígios de materiais relacionados à anatomia e à biologia humana, a qual permite indicar a autoria de um delito ou a identificação de uma vítima, é muito comum nos processos forenses, tanto na área cível como na criminal. Impressões digitais, fios de cabelo ou resíduos de sangue e de epitélio e mesmo registros em vídeo de um rosto ou fisionomia são exemplos de objetos de exame na área da identificação humana para fins forenses. Todos fornecem características anátomo-biológicas que são constantes em cada ser humano e diferentes entre os seus pares, permitindo que se definam particularidades do indivíduo sob investigação.

Utilizar indícios de elementos anatômicos é estratégia natural e foi a base da identificação na Criminalística, quando Alphonse Bertillon, no final do século XIX, desenvolveu o Sistema Antropométrico para a identificação humana. Esta e demais estratégias de reconhecimento através de características físicas ou comportamentais compõem a área da Identificação Biométrica, sendo uma dessas estratégias a análise da voz, da fala e da linguagem dos seres humanos.

É bom saber !!!

Quer se aprofundar sobre a história da Antropometria na Identificação Humana?

Fonte: García Ferrari e Galeano (2016).
<https://www.scielo.br/j/hcsm/a/hLZmDLNdqDMMtLsSQPx6gVq/?lang=pt&format=pdf>

Polícia, antropometria e dactiloscopia: história transnacional dos sistemas de identificação, do rio da Prata ao Brasil

A manifestação verbal oral diferencia o ser humano das demais espécies e pode servir como elemento de particularização de um indivíduo, a ponto de torná-lo único. O reconhecimento através de elementos da comunicação oral fundamenta-se na ideia de que, ao se comunicar deste modo, o indivíduo transmite informações que permitem caracterizá-lo com relação a

suas formações cultural, educacional e social, constituição orgânica, região de origem geolinguística e condições de saúde ou emocionais.

É bom saber !!!

Por falta de valor jurídico da prova, a Sexta Turma do Superior Tribunal de Justiça (STJ) anulou uma condenação baseada apenas na identificação da voz do réu em gravação exibida na delegacia de polícia, sem observância das formalidades do artigo 226 do Código de Processo Penal (CPP) e sem nenhum tipo de perícia técnica. (HC 461709).¹

Fonte: <https://www.youtube.com/watch?v=K53ibS6Kn4>



Imaginemos a situação em que haja a convicção de se reconhecer a voz ouvida, porém, em não se saiba a identidade do dono da voz. Nesse caso, podemos dizer que houve o reconhecimento, pois se conhece a voz, mas não a identificação propriamente dita, uma vez que não se definiu a autoria.

Essa situação, ocorrida talvez com a maioria das pessoas, indica uma capacidade natural do ser humano para o reconhecimento, chamado de *reconhecimento leigo do falante* (em inglês, *naive speaker recognition*) (NOLAN, 1983). Em outras palavras, é a capacidade que as pessoas têm para realizar o reconhecimento apenas pelas características do sinal acústico transmitido, naturalmente, sem o apoio de recursos tecnológicos ou de conhecimento técnico.

Isso pode levar o público sem uma formação específica para avaliar a comunicação humana para fins forenses a imaginar que um indivíduo possui uma voz única, como se fosse um elemento próprio característico, de maneira análoga à impressão digital, o que seria uma simplificação. Primeiro, porque a voz de uma pessoa não é constante como o é a impressão digital. Além disso, não se sabe ao certo em que medida a variação na fala de um indivíduo causa uma sobreposição de aspectos de sua voz em relação aos outros indivíduos de sua comunidade de fala (NOLAN, 2002).

Este importante fato relacionado à comunidade de fala é chamado pelos avaliadores de voz de variações intrapessoais e acontecem a tal ponto que o falante, dependendo do lapso temporal e de fatores como o estado físico ou emocional, pode-se distanciar inclusive das características que lhe eram próprias anteriormente.

1 Disponível em: <https://www.stj.jus.br/sites/porta1p/Paginas/Comunicacao/Noticias/28042021-Sexta-Turma-anula-condenacao-baseada-em-reconhecimento-de-voz-sem-contraditorio-judicial.aspx>

Importante se faz esclarecer alguns termos utilizados pelos estudiosos envolvidos no campo do conhecimento em que se insere esta obra.

Reconhecimento é o termo mais genérico e que engloba os demais. Pode ser dividido em Reconhecimento **Leigo**, referido acima, ou **Técnico** (NOLAN, 1983), o qual evidentemente é o foco desta obra.

Identificação é o procedimento específico de definir um indivíduo juridicamente como único, por isso cada cidadão tem um número de Registro Geral em sua cédula de identidade. Nesse caso, percebemos a relevância da impressão digital para essa finalidade.

Curiosidade

“[...] as identificações pessoais equivocadas são a principal causa de erros judiciais, presentes em 69% dos casos” (INNOCENCE PROJECT, 2019, apud MACHADO, BARILLI, 2019).

ACADEMIA DE POLÍCIA

O reconhecimento de pessoas como fonte de injustiças criminais

Fonte: Machado e Barilli (2019). <https://www.conjur.com.br/2019-jul-16/academia-policia-reconhecimento-pessoas-fonte-injusticias-criminais>

A **Identificação do Falante**, por suas características de voz, fala e linguagem, pode ser efetivada por um sistema automático e, neste caso, é denominado mais comumente de **verificação**, o qual vemos em alguns sistemas bancários ou em empresas. Porém, o foco desta obra recairá sobre o processo de **comparação**, que é a tarefa de se tomar um registro de fala cuja autoria é desconhecida e comparar com um ou mais registros de origem conhecida para determinar se há compatibilidade com algum dos conhecidos (NOLAN, 1997; ROSE, 2002; HOLLIEN, 2002). Repare, leitor, que o processo de comparação sempre será binário, ou seja, se houver uma fala questionada e dois suspeitos, duas comparações binárias serão realizadas para a identificação. Isso poderá ser observado a seguir.

2. Identificação de Falantes

A Identificação de Falantes visa determinar:

- os parâmetros de voz e de fala indicadores de características anatomo-biológicas e referentes ao idioleto, que dizem respeito ao modo particular como cada falante usa a língua, segundo os seus hábitos discursivos e aos aspectos próprios do indivíduo, particularizando-o;

- as características socioculturais e referentes ao dialeto, que dizem respeito a aspectos contextuais e que, mesmo que não definam as características particulares do indivíduo, determinam seu perfil de inserção em uma comunidade de fala, auxiliando em procedimentos de triagem ou corroborando dados individualizantes.

Apesar de haver alguns elementos análogos a outras áreas da identificação humana, a fala está longe de ser um fenômeno constante em um indivíduo, como o são o DNA e as impressões digitais.

É possível afirmar que uma impressão digital é única, ou seja, varia consideravelmente quando tomado o âmbito interindividual, pois dois indivíduos não apresentarão a mesma digital, mas não varia quando considerado o âmbito intraindividual, já que a conformação das papilas de um dedo do indivíduo será sempre a mesma. Tal fato, somado à relativa facilidade de obtenção de amostras para este processo, quando se compara a outras técnicas, tornou a identificação através de impressões digitais a forma mais difundida e tradicional na área da identificação civil e criminal.

Já o processo de identificação pela fala consiste em comparar as amostras de fala produzidas por duas pessoas, atendo-se à variabilidade intrafalante, ou seja, aquela existente em um mesmo parâmetro registrado repetidas vezes de um mesmo falante. A partir das análises realizadas pelo especialista, as amostras gravadas de voz, tornam-se então em prova técnico-científica (SILVA, 2011).

A analogia entre os processos de identificação através da fala e das impressões digitais foi apresentada por Kersta (1962). Defendeu-se o princípio de que a imagem espectrográfica obtida pelo sonógrafo seria equivalente à impressão digital no que diz respeito à sua condição individualizadora. O achado foi batizado de *voiceprint*, ou impressão vocal, numa tentativa de remeter à expressão *fingerprint*, ou impressão digital.

Isso ocorreu após investimentos com objetivos milita-

New Crime Detector Takes 'Pictures' of Human Voice

'Voiceprints' Said to Be More Accurate Than Handwriting Comparisons, More Infallible Than Fingerprints

By THEON WRIGHT
United Press Staff Writer

NEW YORK, Feb. 1.—Fingerprint experts, most deadly trackers down in acentric crime detection, may as well shove over and make room for the latest gadget in exposing criminal skullduggery—the "voiceprint" or voice picture.

In fact, if the other scientific fellows had showed over a bit earlier, "voiceprints" could have solved the Lindbergh kidnaping case, as far as the ransom negotiations were concerned.

All a sleuth, or a ransom bargainer, needs to get a "voiceprint" is a button-microphone, which he cuff wear in the lapel of his coat, a radio broadcasting station, which he can carry in his pocket, and a sound recording machine, which he can leave in his auto.

The idea is that every time you say, "Father, get the work bench out," you are writing a signature which nobody can break. Of course, every kidnaper might not say "Father, get the work bench out"—which happens to be a very revealing speech for "voiceprints"—but he probably would say something nearly as good.



Fonte: <https://languagegel.idc.upenn.edu/nll/?p=36412>

res, pois se pretendia reconhecer as vozes nas mensagens interceptadas por rádio. Assim, em meados da década de 1940, surgiu o sonógrafo, esse aparelho que permitia “enxergar” os sons (HOLLIEN, 2002).

A técnica de comparação através da impressão vocal proposta por Kersta sofreu severas críticas, pois essa analogia ao termo *fingerprint* estabelecia um paralelo entre os dois vestígios como se a imagem espectrográfica apresentasse uma consistência como vestígio tão evidente e intuitiva quanto a impressão digital, o que oferece uma noção equivocada a respeito dos princípios da Identificação de Falantes (NOLAN, 1997; HOLLIEN, 1990; KÜNZEL, 1994; BONASTRE *et al.*, 2003).

Tal equívoco de perspectiva consiste no fato de que a variabilidade da fala respeita a dicotomia “orgânico/adquirido”, ou seja, as diferenças entre as amostras de fala podem ser separadas em dois grupos fundamentais: as que podem ser determinadas pelo aparato orgânico, as diferenças orgânicas, e aquelas determinadas pela aprendizagem, ou adquiridas. A impressão digital, por outro lado, não sofre, ou sofre muito pouco, a ação do adquirido (NOLAN, 1993; GLENN; KLEINER, 1968; WOLF, 1972) e, dessa maneira, varia pouco sua morfologia. Assim, a variabilidade individual inerente ao vestígio linguístico deveria ser considerada de maneira diversa em relação ao vestígio das papilas digitais.

Divertido e fácil de compreender



Há diversas maneiras de falarmos sobre um mesmo assunto, por meio da variação estilística. Saiba mais no vídeo Variação estilística no dia a dia, do canal Sabrina Vieira Teixeira, do Youtube.

Fonte: Moreira, D. de L., Costa, M. da S.; Teixeira, S. V.; Batistela, V. P. (2017).
<https://www.youtube.com/watch?v=NbULhQAeX0I>



As críticas ao termo *voiceprint* (NOLAN, 1997; HOLLIEN, 1990; KÜNZEL, 1994; BONASTRE *et al.*, 2003) também serviram como um alerta à tendência de que os operadores do Direito se deixassem levar pelo óbvio apelo da evolução tecnológica, confiando mais no resultado visual do espectrograma do que na formação e no conhecimento do profissional que interpretava os dados oferecidos pelo recurso tecnológico.

Críticos mais enfáticos à técnica de *voiceprint identification*, Hollien *et al.* (2014) e Hollien (2013) chegam a afirmar que a notoriedade rapidamente obtida pela técnica, em função do alegado “sucesso” da mesma, representou um entrave ao desenvolvimento da pesquisa em Identificação de Falantes, pois causou uma retração das agências oficiais de fomento.

Porém, um ponto de acordo entre os estudiosos é a ênfase na Linguística como formação de base para a atuação na área Identificação de Falantes, apesar de haver o entendimento da multidisciplinaridade requerida (ROSE, 2002; HOLLIEN, 1990).

3. Objetivo

O conteúdo defendido nesta obra é que a analogia da manifestação através da fala e outras manifestações da anátomo-biologia humana não é uma ideia absurda, motivo pelo qual tem sido citada em outras oportunidades.

A proposta desta obra, expressa em seu título, é que a comparação com a ação manuscrita (assinatura) seria mais apropriada, visto que a determinação de autoria através da Grafoscopia aparenta oferecer mais elementos em comum com a identificação pela fala e linguagem, se observados os elementos que guardam semelhança com a determinação de autoria na análise dos lançamentos manuscritos de um indivíduo.

A mesma palavra escrita diversas vezes nunca será exatamente igual, da mesma forma que a palavra proferida oralmente diversas vezes não o é. Os condicionadores emocionais que alteram a velocidade da escrita, a pressão do traçado sobre a superfície e eventuais distorções ortográficas são os mesmos que atuam sobre a velocidade de fala, a intensidade da voz e que determinam trocas e omissões dos sons que compõem a fala. Também podemos imaginar que elementos morfológicos, sintáticos, semânticos e pragmáticos que permeiam o texto possam estar presentes no enunciado/discurso. Além disso, temos a variabilidade estilística e contextual a depender do nosso interlocutor: falamos diferente a depender da pessoa com quem dialogamos e do assunto que abordamos, mesmo que com o mesmo interlocutor. Ou seja, nossa comunicação oral realmente se modifica o tempo todo.

É bom saber !!!



Sabia que a estatística nos traz uma ótima solução para as medidas de variabilidade intrapessoal? Com ela, conseguimos entender o quão homogênea ou heterogênea uma amostra é. Veja mais no vídeo *Medidas de VARIABILIDADE (amplitude, variância, desvio-padrão e intervalo interquartil)*, do Canal Pesquisa, do Youtube.

Fonte: Honório (2020).
<https://www.youtube.com/watch?v=slSPcOlwXs>



A complexidade que envolve os mecanismos anátomo-fisiológicos, bem como todo o processamento cortical, guarda extensas semelhanças entre a

produção da fala e da escrita manuscrita. Assim, pode-se melhor entender a atribuição do título deste livro e transitar pelos capítulos que compõem a obra.

Isso posto, convidamos o leitor a percorrer as páginas desta obra, que procurou trazer cada ponto passível de análise, com detalhes descritivos e exemplificados para uma leitura mais proveitosa e compreensível.

No capítulo 2 abordamos todo o cuidado que se deve ter ao tratar e preservar as mídias digitais e, portanto, evidências digitais. A apresentação dos casos midiáticos elencados para a presente obra está no capítulo 3, junto ao guia de consulta dos áudios analisados. Do capítulo 4 ao 6 discorremos sobre todas as possíveis análises, sendo em ordem: aspectos perceptivos-auditivos da voz e da fala; aspectos fonéticos-fonológicos, com destaque aos fenômenos linguísticos, bem como aos aspectos sintáticos, semânticos e prosódicos; análise da linguagem e análise acústica com os principais parâmetros para extração de dados mais utilizados na tarefa de estabelecer elementos indicativos da particularidade do indivíduo.

Devemos alertar o leitor que apresentamos análises da manifestação oral de indivíduos notórios que, porém, não estão sendo submetidos ao processo de comparação binária para fins de identificação forense, uma vez que não há uma fala questionada e uma fala padrão, ou, como referimos antes, um material desconhecido e outro conhecido sobre os quais se deseja atestar a semelhança ou a diferença. A despeito dessa peculiaridade, todos os conceitos e procedimentos podem e, efetivamente, são utilizados na prática diária dos autores na tarefa de comparação binária e, assim, representam grande valia para aquele leitor interessado no assunto.

4. Referências bibliográficas

- BONASTRE, J. F.; BIMBOT, F.; BOË, L. J.; CAMPBELL, J. P.; REYNOLDS, D. A.; MAGRIN-CHAGNOLLEAU, I. Person Authentication by Voice: A Need for Caution. *Eurospeech*, 2003.
- BRASIL. Superior Tribunal de Justiça. Comunicação. Notícias. *Sexta Turma anula condenação baseada em reconhecimento de voz sem contraditório judicial*. Brasília, DF: Portal do Superior Tribunal de Justiça, 28 abr. 2021. Disponível em: <<https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/28042021-Sexta-Turma-anula-condenacao-baseada-em-reconhecimento-de-voz-sem-contraditorio-judicial.aspx>>. Acesso em: set. 2021.
- GARCÍA FERRARI, Mercedes; GALEANO, Diego. Polícia, antropometria e datiloscopia: história transnacional dos sistemas de identificação, do rio da Prata ao Brasil. *História, Ciências, Saúde – Manguinhos*, Rio de Janeiro, v. 23, supl., dez. 2016, p. 171-194. Disponível em: <<https://www.scielo.br/j/hcsm/a/hLZmDLNdqDMMtLsSQPx6gVq/?lang=pt&format=pdf>>. Acesso em: set. 2021.

- GLENN, J. W.; KLEINER, N. Speaker identification based on nasal phonation. *Journal of the Acoustical Society of America*, v. 43 n. 2, p. 368-372, 1968.
- HOLLIEN, H. An Approach to Speaker Identification. *Journal of Forensic Sciences*, v. 61, n. 2, p. 334-344, 2016.
- HOLLIEN, H. *Forensic Voice Identification*. London: Academic Press, 2002. 240p.
- HOLLIEN, H. *The Acoustics of Crime*. New York: Springer Science+Business Media, 1990. 370 p.
- HOLLIEN, H.; HUNTLEY BAHR, R.; HARNSBERGER, J. D. Issues in forensic voice. *Journal of Voice*, v. 28, n. 2, p. 170-184, 2014.
- HOLLIEN, H.; HUNTLEY BAHR, R.; KUNZEL, H.; HOLLIEN, P. A. Criteria for earwitness lineups. *International Journal of Speech Language and the Law*, v. 2, n. 2, p. 143-153, 2013. Disponível em: <https://doi.org/10.1558/ijssl.v2i2.143>.
- HONÓRIO, H. M. *Medidas de VARIABILIDADE (amplitude, variância, desvio-padrão e intervalo interquartil)*. Youtube. Publicado no Canal Pesquise, 1 vídeo (47 min), 2020. Disponível em: <<https://www.youtube.com/watch?v=sISpOlcwXs>>. Acesso em: set. 2021.
- INNOCENCE PROJECT. *Dna exonerations in the United States*. Disponível em: <<https://www.innocenceproject.org/dna-exonerations-in-the-united-states>>. Acesso em: 13/7/2019, *apud* MACHADO, BARILLI. *O reconhecimento de pessoas como fonte de injustiças criminais*, 2019. Disponível em: <<https://www.conjur.com.br/2019-jul-16/academia-policia-reconhecimento-pessoas-fonte-injusticas-criminais#sdfootnote5sym>>. Acesso em: set. 2021.
- KERSTA, L. G. Voiceprint identification. *Nature*, v. 196, p. 1253-1257, 1962.
- KÜNZEL, H. Current approaches to forensic speaker recognition. *Automatic Speaker Recognition, Identification and Verification*, p. 135-141, abril 1994.
- MOREIRA, D. de L.; COSTA, M. da S.; TEIXEIRA, S. V.; BATISTELA, V. P. *Variação estilística no dia a dia*. Youtube. Publicado no canal Sabrina Vieira Teixeira, 1 vídeo (4m34s), 2017. Disponível em: <<https://www.youtube.com/watch?v=NbULhQAeX0I>>. Acesso em: set. 2021.
- NOLAN, F. Intonation in speaker identification: an experiment on pitch alignment features. *International Journal of Speech Language and the Law*, v. 9, n. 1, p. 1-21, 2002. Disponível em: <https://doi.org/10.1558/sll.2002.9.1.1>.
- NOLAN, F. Speaker recognition and forensic phonetics. In: HARDCASTLE, W.; LAVER, J. (eds.). *A Handbook of Phonetic Science*. Oxford: Blackwell, 1997.
- NOLAN, F. *The phonetic bases of speaker recognition*. Cambridge: Cambridge University Press, 1983. 221 p.
- ROSE, P. *Forensic Speaker Identification*. London: Taylor & Francis, 2002. 364p.
- WOLF, J. J. Efficient acoustic parameters for speaker recognition. *Journal of the Acoustical Society of America*, v. 51 n. 6, p. 2044-2056, 1972.
- WRIGHT, T. New crime detector takes “pictures” of human voice. *The Pittsburgh Press*, 2 jan. 1935, *apud* LIBERMAN, M. “Voiceprint” springs eternal. *Language and the law, Language and the media*, 29 jan. 2018. Disponível em: <<https://languageolog ldc.upenn.edu/nll/?p=36412>>. Acesso em: set. 2021.