

# Proposta de Construção de um Banco de Dados de Amostras de Fala para Uso Forense em um Arcabouço Bayesiano

G.A. Silva \*

*Instituto de Criminalística, Superintendência da Polícia Técnico-Científica do Estado de São Paulo, São Paulo (SP), Brasil.*

\*Endereço de e-mail para correspondência: [albuquerque2012@hotmail.com](mailto:albuquerque2012@hotmail.com). Tel.: +55-98739-2379.

Recebido em 22/02/2016; Revisado em 07/04/2016; Aceito em 08/04/2016

---

## Resumo

Há pouco mais de dez anos, um novo paradigma para algumas disciplinas forenses foi pela primeira vez descrito. Nesse novo modelo, o arcabouço Bayesiano para avaliação da evidência foi proposto como um modelo teórico comum para a interpretação da evidência. Nos exames de Comparação Forense de Locutor, os esforços para a adoção dessa abordagem já haviam sido iniciados bem antes do novo paradigma proposto, seja na metodologia automática, seja na metodologia combinada clássica. No Brasil, onde se tem optado pela adoção de um modelo binário de decisão e pela adoção de uma escala clássica de probabilidades para expressão dos resultados, os peritos têm, no arcabouço bayesiano, uma boa alternativa. A adoção desse arcabouço requer, porém, a obtenção de uma base de dados suficientemente representativa. O objetivo do presente artigo é fazer uma revisão do arcabouço teórico bayesiano nas metodologias utilizadas em Comparação Forense de Locutor, bem como propor um esboço de uma base de dados de fala forense a ser utilizada no Estado de São Paulo.

*Palavras-Chave:* Fonética Forense; Estatística Forense, Metrologia Forense.

---

## Abstract

Just over ten years, a new paradigm for some forensic disciplines was first described. In this new model, the Bayesian framework was proposed as a common theoretical framework for the interpretation and assessment of evidence. In the Forensic Speaker Comparison examination, efforts to adopt this approach had already been started long before the new paradigm had been proposed, whether it be in the automatic methodology or within a phonetic - based approach. In Brazil, where the option is to use either a binary decision model or a classical probability scale for the expression of the outcomes, the experts could count on a good alternative with the Bayesian framework. In order to adopt this framework, however, it would be necessary to obtain a speech database sufficiently representative. The purpose of this article is to review the theoretical framework of the Bayesian methods used in Forensic Speaker Comparison and propose a forensic speech database to be used in the State of São Paulo.

*Keywords:* Forensic Phonetics; Forensic Statistics, Forensic Metrology.

---

## 1. INTRODUÇÃO

Compelidas por forças transformadoras [1], muitas das disciplinas da Criminalística vêm passando por uma mudança de paradigma. Essa mudança de paradigma não deve ser entendida no sentido tal qual foi descrito por Thomas Kuhn [2], mas sim como uma metáfora que serve para descrever a inevitável transformação que tais disciplinas vêm experimentando em direção a uma metodologia verdadeiramente científica [3]. Para que essa mudança se efetive, a avaliação da evidência deve consistir em três elementos, quais sejam:

- uso obrigatório do LR (razão de verossimilhança – do inglês likelihood ratio);
- emprego de abordagens fundamentadas em medidas quantitativas, em base de dados representativos de uma população relevante e em modelos estatísticos.
- testabilidade, tanto em termos de acurácia quanto de precisão.

É cada vez maior o uso do LR para expressão dos resultados de suas análises. Prova disso é que, em 2011, mais de trinta peritos foram signatários de um documento que incluía uma afirmação segundo a qual eles consideram ser o framework bayesiano (com o uso do

LR) o mais apropriado arcabouço teórico para avaliação da evidência [4]. Esse posicionamento foi endossado pela ENFSI (European Network of Forensic Science Institutes), a qual representa cinquenta e oito laboratórios em mais de trinta países.

As abordagens baseadas em medidas quantitativas e em bases de dados que reflitam uma população relevante são bem mais preferíveis que metodologias baseadas na experiência humana não apenas pelo fato de as primeiras serem mais transparentes, mas também por serem mais facilmente replicáveis e testáveis em termos de precisão e acurácia [3].

As metodologias usadas em fonética forense dividem-se em combinada clássica (ou tradicional) e automática. Na primeira, privilegiam-se as informações relacionadas à estrutura da linguagem, seja no plano puramente físico do som - fonético - seja em níveis mais abstratos da linguagem: de cunho fonológico, dialetal, etc. Na abordagem automática, o processo de decisão se vale de algoritmos baseados em coeficientes derivados do sinal da fala, os quais são submetidos a um tratamento estatístico. É bastante consensual o fato de que os sistemas automáticos, quando utilizados, devem ser de forma complementar à análise combinada clássica.

Qualquer exame de Comparação de Locutor compreenderá a comparação de duas ou mais gravações: o material questionado (referido como traço, amostra criminal ou disputada), obtido, por exemplo, de uma interceptação telefônica; e o material suspeito (também chamado de material padrão ou de referência), obtido, geralmente, por meio de coleta de material de fala do indivíduo suspeito.

Pode-se constatar que são de longa data os esforços visando à incorporação do arcabouço bayesiano nos exames de Comparação Forense de Locutor, independentemente da abordagem adotada. Deste modo, o objetivo do presente artigo é apresentar, em linhas gerais, esse arcabouço bayesiano no qual o resultado é expresso por meio da razão de verossimilhança, doravante LR, fazendo uma revisão acerca de como tal arcabouço vem sendo usado nas abordagens fonético-forense. O uso rigorosamente científico do LR e as exigências do novo paradigma levam necessariamente à construção de uma base de dados de fala. Tendo isto em vista, o presente artigo apresenta, em linhas gerais, uma proposta de criação de uma base de dados de fala para a variedade do Estado de São Paulo. A criação de uma base de dados de fala é o primeiro e mais importante passo para se iniciar a mudança de paradigma, permitindo o uso do arcabouço bayesiano, a transparência, a testabilidade e a quantificação rigorosa da evidência nos exames de Comparação de Locutor.

## 2. O LIKELIHOOD RATIO

A fórmula de Bayes é uma maneira de atualizar a opinião do observador/estatístico/perito a respeito da probabilidade de ocorrência de um evento. Em termos estatísticos [5-7], tem-se:

$$\Pr(H|\Delta) = \frac{\Pr(\Delta|H)}{\Pr(\Delta)} * \pi \quad (1)$$

sendo que  $\pi$  é chamada a priori, a opinião do estatístico antes de conhecer os dados. Geralmente, a determinação de uma a priori adequada é difícil, opta-se por consultar um especialista no processo em questão, ou utilizar uma a priori que não favoreça de antemão alguns dos eventos considerados, ou apresentar como resultado a razão de verossimilhança. A dificuldade em tal determinação se dá pelas razões mostradas a seguir:

- $\Delta$  representa a informação obtida através de um ou mais experimentos;
- H representa o evento de interesse – para o qual se deseja inferir a probabilidade;
- $\Pr(H|\Delta)$ , às vezes denominado  $L(H|\Delta)$ , é a chamada verossimilhança proporcionada a H por  $\Delta$ , é a probabilidade de se observar os dados  $\Delta$ , supondo H conhecido;
- $\Pr(\Delta|H)$  é a probabilidade a posteriori;
- $\Pr(\Delta)$  é o termo que pondera que os dados foram observados. Operacionalmente, é um normalizador que faz a probabilidade a posteriori estar no intervalo [0,1].

A razão de verossimilhança, por construção, informa o quão mais provável é um evento se comparado com sua alternativa. É um valor de fácil interpretação e não propaga viés inserido pela a priori. Por estas razões, pode ser preferido em relação à probabilidade a posteriori.

Suponha que um evento H assuma um dos valores {Verdadeiro, Falso}. Por ora, assume-se que os dados são representados por  $\Delta$ , de forma que:

$$\Pr(H|\Delta) = \frac{L(H|\Delta)}{\Pr(\Delta)} * \pi \quad (2)$$

Calculando a razão entre as probabilidades a posteriori de cada evento {Verdadeiro, Falso}, obtém-se:

$$\frac{\Pr(H=Verdadeiro|\Delta)}{\Pr(H=Falso|\Delta)} = \frac{L(H=Verdadeiro|\Delta)}{L(H=Falso|\Delta)} * \frac{\pi}{1-\pi} \quad (3)$$

Define-se a razão de verossimilhança como:

$$LR = \frac{L(H=Verdadeiro|\Delta)}{L(H=Falso|\Delta)} \quad (4)$$

e obtém-se:

$$\frac{\Pr(H=Verdadeiro|\Delta)}{\Pr(H=Falso|\Delta)} = LR * \frac{\pi}{1-\pi} \quad (5)$$

Assim, a razão de verossimilhança (LR) tem duas características importantes:

- é desnecessário calcular o denominador da fórmula de Bayes em (1), na medida em que ele se cancela no cálculo;
- independe da probabilidade a priori, podendo ser considerada uma representação da informação que os dados oferecem, sem outras influências.

Outro aspecto muito importante é a possibilidade de se combinar vários experimentos nesse arcabouço.

Deste modo, considere os resultados de dois experimentos estatisticamente independentes  $\Delta_1$  e  $\Delta_2$  e a variável aleatória cuja probabilidade se quer atualizar -  $\Pr(H)$ . Se  $\Delta$  representar a informação resultante das duas realizações,  $\Delta = (\Delta_1, \Delta_2)$ , segundo a fórmula de Bayes, tem-se que:

$$\begin{aligned} \Pr(H|\Delta) &= \frac{\Pr(\Delta_1, \Delta_2|H)}{\Pr(\Delta_1, \Delta_2)} * \pi \\ &= \frac{\Pr(\Delta_1|H) * \Pr(\Delta_2|H)}{\Pr(\Delta_1) * \Pr(\Delta_2)} * \pi \\ &= \frac{\Pr(\Delta_1|H)}{\Pr(\Delta_1)} * \frac{\Pr(\Delta_2|H)}{\Pr(\Delta_2)} * \pi \end{aligned} \quad (6)$$

A fórmula (6) significa que se pode aplicar a fórmula (2) sequencialmente para cada uma das realizações e, assim, obter a probabilidade *a posteriori*. Esse fato é muito usado nos casos de testes de paternidade por meio de DNA, nos quais vários marcadores são analisados e, para cada um, resulta um valor de probabilidade de o indivíduo ser pai [5]. Nesses exames específicos de testes de paternidade levam-se em conta duas hipóteses: **H = sim**, correspondente à hipótese do alegado pai ser o pai biológico da criança, e **H = não**, significando que existe outro homem na população que é o pai biológico da criança. Os dois eventos têm probabilidade *a priori* 0.5 (50%), ou seja, consideram-se equiprováveis os eventos (**H = sim** e **H = não**):  $\Pr(H= sim) = \Pr(H= não) = 0.5$ . Com essa probabilidade *a priori* previamente determinada, chega-se, pois, nesses exames, à determinação da probabilidade *a posteriori*.

Do mesmo modo, a razão de verossimilhança será o produto de cada experimento individual. Para a Criminalística, tal propriedade é importante na medida em que os resultados dos vários exames periciais podem ser unificados segundo um arcabouço comum [8,9]. No caso de Comparação de Locutor, vários parâmetros podem ser combinados e o LR resultante será o produto de cada razão de verossimilhança obtida.

Os conceitos de variabilidade da fala e das proposições concorrentes são fulcrais nesse modelo e se relacionam com a necessidade de bases de dados de fala que reflitam esses conceitos, os quais serão discutidos, em linhas gerais, a seguir.

### 3. A VARIABILIDADE DA FALA

Sob condições ideais, as pessoas podem ser facilmente reconhecidas por meio de suas vozes. Vozes familiares são facilmente reconhecidas por qualquer indivíduo [10-12] e essa habilidade faz com que se assuma que diferentes falantes têm, de fato, diferentes vozes, ou, em outras palavras, assume-se que existe uma intervariabilidade na fala. Tal intervariabilidade se dá tanto em termos das diferenças anatômicas e fisiológicas existentes entre um indivíduo e outro quanto em termos do ambiente linguístico de aquisição da fala e de hábitos articulatórios aprendidos [13]. Por outro lado, embora se assumam que diferentes falantes têm diferentes vozes, outro desafio enfrentado por aquele que lida com Comparação de Locutor é o fato de que a voz de um mesmo falante também varia em relação a ele próprio. Essas variações podem estar ligadas tanto a fatores de ordem psicológica e emocional como também à alternância de códigos de expressão formal (variação diafásica) [14] e à contemporaneidade, dentre outros. Ninguém diz uma mesma palavra exatamente da mesma maneira duas vezes e, ao menos que se esteja lidando com gravações de uma mesma emissão, é certo que se deparará com o fato de que a variabilidade da fala se manifesta em um mesmo indivíduo, ou, em outras palavras, a par da variação entre diferentes locutores, há uma intravariabilidade da fala. Assim, o sinal acústico de duas palavras, sintagmas ou sentenças nunca é o mesmo, seja produzido pelo mesmo falante, seja produzido por diferentes falantes.

A existência da variabilidade da fala orienta na escolha de um dado parâmetro acústico em termos de seu poder discriminativo: quanto maior a razão entre a intervariabilidade e a intravariabilidade de um dado parâmetro acústico, tanto maior será seu poder discriminativo [15]. Além disso, no arcabouço bayesiano, exige-se a construção de base de dados contendo amostras de fala que deem conta dessa variabilidade.

### 4. A NECESSIDADE DE SE CONSIDERAR UMA PROPOSIÇÃO ALTERNATIVA

O LR é um fator que expressa numericamente o peso da evidência, convertendo uma probabilidade *a priori* em probabilidade *a posteriori*. Esse peso é sempre relacionado a duas hipóteses concorrentes. É uma relação que quantifica a hipótese da evidência com respeito a duas proposições alternativas. Dentro do processo legal, a

probabilidade *a posteriori* é, por outro lado, função do juiz. Não cabe ao perito fornecer probabilidade *a posteriori*. Pela fórmula (1), isso só faria sentido se o perito tivesse acesso a todas as informações referentes ao caso que antecedem ou independem de sua análise. Mas ele não tem acesso a essa probabilidade *a priori*. Além disso, no plano do processo, tais probabilidades *a priori* são subjetivas e, portanto, sujeitas a enviesamento. Ao fornecer probabilidade *a posteriori* o perito estaria usurpando o papel do juiz cuja função, dentro do processo legal, é determinar a culpabilidade ou inculpabilidade do réu. Este costuma formular suas questões no sentido de se determinar a probabilidade de uma das hipóteses (de acusação ou defesa) dada a evidência. Do ponto de vista do perito, a quem interessa tão somente a avaliação da evidência com respeito a duas hipóteses alternativas, respostas a essas questões, nos exatos moldes formulados pelo juiz, serão sempre lacunares, na medida em que se ressentem de informações relevantes. O exemplo a seguir, extraído de [16], serve para ilustrar como tais questões conduzem às chamadas falácias (da acusação ou da defesa). Falácias são argumentos enganadores que se utilizam de proposições estatísticas enviesadas transvestidas em argumentos legais cujos desdobramentos podem levar à condenação de inocentes ou livrar criminosos da punição.

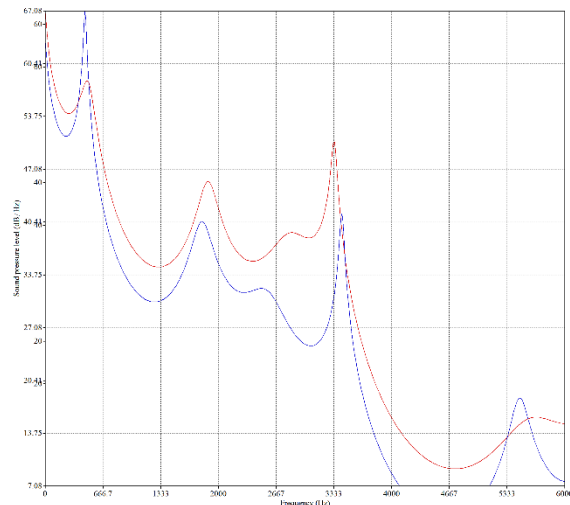
Suponha-se que é sabido que 80% das crianças que sofreram abuso sexual adquirem a mania roerem suas unhas. Uma criança da qual se suspeita ter sido vítima de abuso é levada ao juiz, perante o qual ela, de fato, manifesta a mania de roer unhas. O juiz então, pergunta qual a probabilidade de essa criança ter sido vítima de abuso sexual.

Uma afirmação do perito acerca da probabilidade questionada pelo juiz se constitui numa falácia da acusação, ou, em termos estatísticos, transposição da condicional: a hipótese da probabilidade dada a evidência é colocada no lugar da probabilidade da evidência dadas as hipóteses alternativas. É lógico, portanto, que a fim de avaliar esta evidência (a mania de roer as unhas), é necessário obter mais uma informação: qual a probabilidade de crianças manifestarem a mania de roer as unhas sem terem sido vítimas de abuso sexual? Se esta porcentagem for menor do que a probabilidade relativa às crianças que sofreram abuso sexual, então, a evidência oferece algum apoio à hipótese de que houve abuso sexual. Por outro lado, se a probabilidade de as crianças roerem as unhas e não terem sofrido abuso sexual for maior do que a probabilidade daquelas que roem e foram vítimas de abuso, então, nesse caso, roer as unhas favorece a hipótese de que não houve abuso sexual.

Outro exemplo análogo-este, de cunho fonético-acústico [15] - é dado a seguir. Suponha-se que se tenham dados suficientemente robustos e confiáveis que garantam que, em 80% dos casos, o valor médio do segundo

formante da vogal média baixa seja bastante próximo quando a amostra de fala provém de um mesmo falante, como ilustrado na Figura 1.

Suponha-se também que o perito foneticista está lidando com duas amostras muito similares em relação ao segundo formante da média baixa. O juiz, então, quer saber qual a probabilidade de essas amostras provirem de um mesmo falante. Novamente, qualquer afirmação em termos de probabilidade *a posteriori* nos coloca, também neste caso, diante de uma falácia da acusação, a qual se expressa na probabilidade da hipótese da acusação ser verdadeira dada a evidência. Essa inversão de probabilidades implica na atribuição de um peso indevido à hipótese da acusação [8].



**Figura 1.** Compatibilidade nos valores dos formantes relativos à vogal média baixa para um mesmo falante.

Da mesma forma, para que se dê uma resposta satisfatória a essa questão, outra informação é imprescindível, a saber: qual seria a probabilidade de os valores médios do segundo formante da vogal baixa ser bastante próximos quando se originam de falantes diferentes? A falácia da acusação consistiria em nos induzir a pensar que esta probabilidade seria de 20%.

Duas informações são, portanto, necessárias, quais sejam: a probabilidade da evidência (mania de roer as unhas; grau de similaridade em relação ao segundo formante) assumindo que a hipótese (abuso sexual; mesmo falante) seja verdadeira; e a probabilidade da mesma evidência assumindo que a hipótese não é verdadeira.

É, destarte, fundamental o princípio segundo o qual não é possível para o cientista forense fazer considerações acerca de uma proposição sem considerar ao menos uma proposição concorrente. No âmbito forense, simplifica-se todo esse processo ao se reduzir todas as proposições possíveis a apenas duas: uma representando a hipótese da acusação e outra, o ponto de vista da defesa. Essas proposições são, habitualmente, designadas por  $H_p$  e  $H_d$ .

A avaliação da evidência científica (E) considera a probabilidade da evidência, dada cada uma das proposições. Em Comparação de Locutor, é o grau de similaridade oriundo das medidas acústicas das amostras comparadas. Assim, a evidência não é a voz em si, mas a relação entre o material padrão e o material questionado. Além disso, há um conjunto de circunstâncias, designadas genericamente por I. Essas circunstâncias envolvem as investigações, os testemunhos, o enviesamento e crenças, dentre outros fatores, em geral, inacessíveis ao cientista forense.

Colocando todas essas informações relacionadas ao contexto forense na fórmula (5), obtém-se a fórmula de Bayes adaptada para as disciplinas criminalísticas:

$$\frac{f(H_p|E,I)}{f(H_d|E,I)} = \frac{f(E|H_p,I)}{f(E|H_d,I)} \times \frac{f(H_p|I)}{f(H_d|I)} \quad (7)$$

A razão de verossimilhança é dada por:

$$LR = \frac{f(E|H_p,I)}{f(E|H_d,I)} \quad (8)$$

sendo:

$H_p$  é a hipótese da acusação (o suspeito é a fonte do áudio atribuído a ele no material questionado);

$H_d$  é a hipótese da defesa (outro indivíduo que não o suspeito é a fonte do áudio atribuído ao suspeito no material questionado);

I representa as informações do contexto, como por exemplo, os testemunhos e as investigações;

E são as medidas extraídas do confronto entre o material oriundo do suspeito e o material questionado atribuído a ele;

f() são funções densidade de probabilidade.

O numerador da fórmula (8) fornece a medida da similaridade entre o material questionado e o suspeito (uma estimativa da intravariabilidade); e o denominador fornece uma medida da tipicidade da evidência em relação a uma população relevante (uma estimativa da intervariabilidade). No caso de um exame de Comparação de Locutor, a população relevante é formada é aquela formada por áudios os mais similares possíveis da voz do suspeito. A fórmula (8) assevera que o perito deve, necessariamente, considerar os respectivos pontos de vista da acusação e da defesa, o que é bastante diferente de ser chamado ao tribunal para expressar a sua opinião em relação ao ponto de vista da acusação e, posteriormente, responder às proposições da defesa. Tal fórmula também afirma que, se por um lado, ao juiz interessa respostas a questões do tipo:

- qual a probabilidade de que a proposição da acusação/defesa seja verdadeira, dada a evidência?;

Ao perito, por seu turno, interessa responder a questões como:

- qual a probabilidade da evidência, considerando dadas as respectivas proposições da defesa/acusação?.

Assim, dentre as vantagens do uso do LR, pode-se destacar:

- permite ao perito obter e apresentar ao tribunal um valor significativo do peso da evidência, na forma de uma razão de verossimilhança. Tais valores podem ser apresentados de forma numérica ou em uma escala verbal [17];
- torna evidente qual é o verdadeiro papel- no processo penal - do perito, da prossecução penal, da defesa e do juiz, deixando a cargo deste último a prerrogativa, no sistema legal adotado no Brasil, do uso das probabilidades *a priori* na tarefa de decisão;

O que se pretende, ao longo deste artigo, é descrever o arcabouço bayesiano, colocando-o como uma alternativa à disposição dos peritos brasileiros, não importando os métodos de abordagem de que se valem e mostrando que a adoção deste framework requer a construção de uma base de dados suficientemente representativa. Não se pretende aqui listar as vantagens e desvantagens desse modelo em relação a uma abordagem estatística clássica (ou frequencista) – discussões nesse sentido podem ser encontradas em [8], por exemplo. É importante, entretanto, destacar que uma das principais objeções à abordagem bayesiana repousa em sua interpretação subjetivista de probabilidade. Uma vez que as probabilidades *a priori* são determinadas por meio das crenças e preconceitos, suas escolhas envolvem procedimentos intrincados, dependentes do contexto e com forte carga psicológica. Como o juiz deve impor essa informação *a priori* antes ou de forma independente da análise da evidência, colocamo-nos diante de uma interpretação subjetivista da probabilidade, o que se constitui num dos maiores empecilhos para a implantação do modelo. Não há dúvidas, entretanto, de que a abordagem bayesiana tem muitas vantagens que a colocam como um arcabouço bastante lógico e apropriado de interpretação da evidência científica dentro do processo legal.

Nos itens que se seguem, far-se-á uma breve revisão bibliográfica acerca do uso do arcabouço bayesiano em ambas as metodologias da Fonética Forense.

## 5. COMPUTAÇÃO DO LR EM UM SISTEMA AUTOMÁTICO

No Brasil, a forma de se apresentar conclusões se dá de modo binária ou por meio de uma escala probabilística clássica [18,19]. Algumas instituições contam com sistemas automáticos para Comparação de Locutor em seus respectivos parques tecnológicos.

Embora se possam encontrar importantes estudos datados do final do século passado [20,21], foi principalmente nos últimos quinze anos que os métodos automáticos de identificação de falante para fins forenses experimentaram uma vertiginosa evolução, tornando-se um dos mais convincentes meios de identificação do indivíduo [22-28]. Parte dessa evolução pode ser creditada aos esforços realizados pelo NIST, o qual promove avaliações de desempenho dos sistemas automáticos - conhecidas como SRE (Speaker Recognition Evaluation) - e que ocorrem anualmente desde 1996. Tais esforços têm contribuído para que se tenham tecnologias cada vez mais robustas e desempenhos cada vez melhores por parte dos sistemas automáticos, mesmo em ambientes desafiadores em termos da relação sinal/ruído e quantidade [29-34].

A comparação da voz, no contexto forense, pode ser implementada por meio da teoria de processamento digital de sinal [35-37]. O desenvolvimento de um Sistema Automático de Reconhecimento de Locutor (ASR, do inglês Automatic Speaker Recognition), também chamado de Biometria de Voz, envolve três áreas: ciências da fala, reconhecimento de padrões e engenharia. As ciências da fala contribuem com os conhecimentos acerca de como as pessoas produzem e percebem o sinal de fala, além de estudar as técnicas de processamento digital de sinal a serem usadas para analisar os correlatos acústicos do sinal de fala. A área de reconhecimento de padrões fornece algoritmos para modelar o falante por meio das características mais salientes encontradas no sinal de fala. A área da engenharia permite a construção desses sistemas e é usada para lidar com as variações que ocorrem nas aplicações em contextos reais [34].

Ao invés de um resultado qualitativo, a abordagem automática se refere a um procedimento essencialmente quantitativo. No contexto forense, fala-se em Sistema Automático de Reconhecimento Forense de Locutor (FASR em inglês - Forensic Automatic Speaker Recognition). Também chamado de Biometria Forense de Voz, consiste na utilização de algoritmos computacionais para reconhecer pessoas pela voz [37]. Nesses sistemas, uma dada amostra de voz é comparada com modelos estatísticos do locutor, resultando, assim, em uma medida de similaridade. Em um contexto forense, no qual se compara um material de origem conhecida (material padrão) com outro de origem desconhecida (material

questionado), as medidas acústicas das similaridades/diferenças resultam na evidência [38].

O objetivo de qualquer sistema automático é indicar se a voz de uma gravação de origem desconhecida pertence a determinado indivíduo suspeito. A evidência aqui será medida pelo grau de similaridade entre as características distintivas extraídas do material questionado e aquelas extraídas do material suspeito, sempre representado por meio de um modelo matemático. A metodologia para o cálculo da evidência em um sistema automático pode ser encontrada em [38] e [39] e está ilustrada no esquema que pode ser visto na Figura 02 [36].

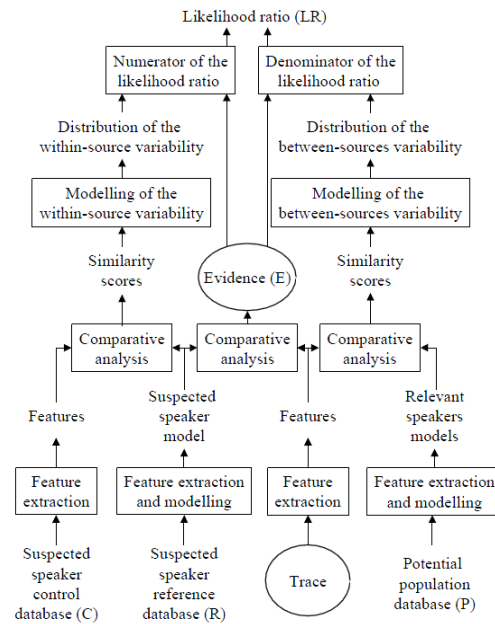


Figura 2. Esquema geral de um sistema automático [36].

No esquema acima, trace é o áudio questionado—obtido por interceptação ou gravação — que será comparado com um material de origem conhecida—chamado aqui de suspeito—em geral, obtido em uma sala especialmente desenhada para tal. Além do áudio questionado, os sistemas automáticos necessitam de três bases de dados:

- a base de dados de uma população relevante (P), contendo vozes similares às do material questionado. É essa base de dados que é usada para calcular a intervariabilidade;
- a base de dados de referência do suspeito é indicada por (R); essa base contém gravações do suspeito tão próximas quanto possível de (P), tanto em condições de gravação quanto em termos linguísticos;
- a base de dados de controle do suspeito é indicada por (C) e consiste em gravações similares ao material questionado. Essa base é usada a fim de calcular a intravariabilidade.

Nos sistemas automáticos, a razão de verossimilhança é calculada da seguinte maneira:

- as características acústicas do material questionado são comparadas com o modelo estatístico do suspeito –criado usando-se a base de dados de referência do falante (R) – e a pontuação obtida (score) informa o valor estatístico da evidência E. Note-se que a evidência não é a voz, mas a medida acústica da similaridade/ diferença entre as amostras;
- as características do material questionado são comparadas com modelos estatísticos da população relevante (P). Daqui se obtém a função de distribuição de probabilidade relacionada à intervariabilidade;
- a base de dados de controle da gravação do suspeito (C) é comparada com a base de dados de referência (R) resultando na função relacionada à intravariabilidade.
- a razão de verossimilhança (LR) – ou seja, a proporção em que a evidência contribui com cada uma das hipóteses descritas no item anterior – é dada pela razão das alturas das funções densidade de probabilidade no ponto E. Isto está ilustrado na Figura 2, extraída de [39], na qual um LR de 9.165 é obtido para  $E = 9.94$ .

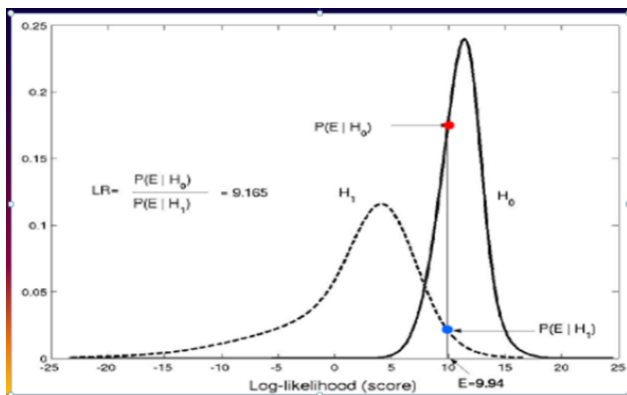


Figura 3. Ilustração da estimativa do LR [39] em um sistema automático.

## 6. O USO DO LR EM ABORDAGENS CLÁSSICAS DA FONÉTICA FORENSE

Os primeiros estudos para incorporar o arcabouço bayesiano na abordagem clássica datam de 1999 [40,41]. Desde então, numerosos trabalhos foram publicados, sendo a maioria deles com enfoque exclusivamente em vogais. As metodologias variam desde medidas no ponto médio dos formantes [42-44] até estudos dinâmicos de trajetórias em ditongos [45-47], passando por análise de formante de longo termo por meio de misturas gaussianas [48]. Parâmetros não vocálicos também têm sido estudados segundo o arcabouço bayesiano. Assim, frequência fundamental, VOT (*Voice Onset Time*), nasais,

laterais, fricativas, taxa de articulação, espectro de longo termo, *clicks*, entre outros parâmetros, têm recebido atenção de vários estudiosos [40,49-52].

Além desses, vários outros parâmetros acústicos – tanto em nível segmental quanto suprasegmental [53,54] – podem ser usados no arcabouço bayesiano, como qualidade de voz, parâmetros dialetais, ritmo, escolhas gramaticais e lexicais, processos fonológicos, entoação, expressividade da fala, *baseline F0*, ênfase espectral, etc.

Paralelamente a essa literatura do uso do LR para avaliar a capacidade discriminativa de parâmetros fonéticos, muita literatura tem sido dedicada ao aprimoramento metodológico nos cálculos da razão de verossimilhança. Assim, é possível encontrar artigos sobre técnicas de modelagem estatística de dados para cálculo de LR [55,56], identificação da população relevante [57,58], combinação de parâmetros [59] e sistemas de calibração [60].

Tudo o que foi apresentado nos itens 4 e 5 mostra que já é bem consolidado o uso do LR em exames de Fonética Forense, não importando qual a metodologia utilizada. Para que os peritos do Brasil possam incorporar esse arcabouço em suas análises, é necessária a construção de uma base de dados de fala.

A construção de uma base de dados de fala é uma tarefa extremamente complexa. Envolve uma base de *background* (que cria um modelo estatístico de distribuição das propriedades acústicas em uma população relevante), uma base de desenvolvimento (usada para otimização e calibração) e uma base para testes de precisão e acurácia.

A proposta que se segue, por seu turno, é bastante incipiente e foi elaborada para uso no IC de São Paulo, servindo, portanto, como uma primeira aproximação para constituição de uma base a ser utilizada no Estado, não havendo nenhuma dúvida, por parte do autor, de que o projeto pode ser melhorado por meio de sugestões e mesmo expandido ao incluir variedades de outros Estados do Brasil.

## 7. A PROPOSTA DE UMA BASE FORENSE DE DADOS DE FALA PARA O ESTADO DE SÃO PAULO

Do que foi exposto até aqui, fica evidente que o LR é a razão entre similaridade e tipicidade das características extraídas das amostras de fala, independentemente da abordagem adotada. O numerador consiste na probabilidade de se obter a evidência dada a hipótese que as amostras comparadas são oriundas de um mesmo falante. O denominador é a probabilidade de se obter a evidência dada a hipótese de as amostras serem oriundas de outro locutor que não o suspeito. A hipótese do numerador é determinada pela comparação entre o material questionado e o material padrão, estabelecendo

daí o grau de similaridade. A hipótese do denominador é determinada pela comparação da amostra questionada com uma população relevante, estabelecendo dali uma medida de tipicidade. A estimativa da tipicidade envolve a construção de base de dados de amostras de fala que sejam representativas de uma população relevante. A variabilidade da fala exige uma população suficientemente grande que permita a avaliação de desempenho do sistema. Assim, o primeiro passo em direção a uma abordagem bayesiana consiste na construção de uma base de dados de amostras de fala suficientemente ampla sob situações realísticas de fala em contexto forense. Essa população poderá ser usada em quaisquer umas das abordagens e os resultados de suas respectivas análises pode ser combinado com os obtidos em outras análises forenses relacionadas com o crime em questão, por meio da multiplicação das razões de verossimilhança de cada exame, conforme fórmula (6). Tendo em vista a inexistência de bases de dados de fala para uso forense no Brasil, propõe-se aqui uma base para o Estado de São Paulo, num esforço inicial para algo mais amplo.

Grande parte das instituições de pesquisa e das instituições policiais que lidam com exames de Comparação de Locutor possui bases de dados de fala. Uma das mais conhecidas, a *D<sub>y</sub>ViS (Dynamic Variability in Speech Database)* foi gravada na Universidade de Cambridge [60][61] e é composta de 100 amostras de fala de locutores de sexo masculino simulando condições normalmente encontradas em contexto forense em canal telefônico e em entrevista em estúdio. Mas há várias bases de dados em diversas línguas [62-67].

Do mesmo modo, há uma vasta literatura acerca de procedimentos para a construção de bases de dados de fala. A base de dados de voz proposta seguirá os quesitos estabelecidos em [67] e terá as seguintes características:

- consistirá em três amostras não contemporâneas - curto, médio e longo prazo;
- a base de dados contempla as variedades dialetais do Estado;
- será composta por áudios obtidos por meio do sistema de interceptação da Polícia Civil do Estado de São Paulo.

A seguir, seguem alguns dos componentes da proposta. Essa proposta poderá ser melhorada e adaptada pelas instituições que pesquisam ou que trabalham com Comparação de Locutores.

- **Número, sexo e idade dos locutores**

A população proposta será constituída de 300 falantes, sendo 200 do sexo masculino e 100 do sexo feminino. A população feminina ser constituída de um menor número de amostras justifica-se na medida em que é bem menor a participação feminina em muitos crimes, como por

exemplo, o de tráfico de drogas [68]. Além disso, no Brasil, a maioria penal é de 18 anos [69], idade que se pode adotar como limite mínimo. Levando-se em consideração as respectivas idades para estabilização da voz, pode-se situar os falantes entre 20 e 60 anos (locutores masculinos) e 18 e 55 anos (locutores femininos) [70-72].

- **Dialeto**

Dos duzentos indivíduos do sexo masculino, 100 devem representar os falares da capital e da região metropolitana [73, 74], 50 devem representar a variedade caipira [75,76] e 50 devem ser provenientes do litoral [77]. No caso dos locutores de sexo feminino, a proporção seria 50 falantes da capital e região metropolitana, 25 do interior e 25 do litoral.

- **Canais**

Basicamente, depara-se no dia a dia da atividade pericial com gravações ambientais e interceptações telefônicas. O banco de dados de fala proposto deve ser, portanto, construídos nesses canais, e também por microfone em ambiente tratado acusticamente. A fim de refletir de modo bem mais realístico as condições requeridas pelo novo paradigma, as interceptações telefônicas devem, preferencialmente, passar pelo sistema de interceptação da Polícia Civil. Similarmente, as gravações ambientais podem ser obtidas por meio de *smartphones*, os quais são muito usados atualmente para se fazer esse tipo de gravação

- **Número de Sessões e de Gravações**

As gravações devem ser feitas em três sessões: uma pela manhã, uma à tarde e outra à noite. Uma vez que a intravariabilidade é um aspecto inerente ao sinal de fala, o aspecto da contemporaneidade das amostras deve ser considerado. Alguns estudos têm demonstrado que, quanto maior o *gap* temporal entre duas amostras de um mesmo sujeito, maior a diferença entre eles, tanto em nível acústico quanto perceptivo [78 - 80]. Gravações de interesse policial não são necessariamente contemporâneas. Em casos de tráfico de drogas, há procedimentos de interceptação que duram meses. Assim, define-se uma população de curto, médio e longo termo, sendo que se considera a primeira com, no máximo, uma semana de intervalo; a segunda, dois meses; e a terceira, seis meses.

- **Eliciação**

Para cada locutor, é desejável dez minutos de fala líquida em eliciação espontânea. A fim de se obter amostras espontâneas e, ao mesmo tempo, controladas, técnicas como *map task* poderão ser empregadas [81]. Outras técnicas para obtenção de amostras espontâneas



podem ser encontradas em [67, 82]. Os segmentos que devem ser extraídos incluem as vogais, os ditongos e as sonorantes.

#### • Vantagens do Projeto

Dentre as vantagens do projeto, encontra-se maior confiabilidade de uso dos sistemas automáticos e de parâmetros tradicionais. O projeto promoverá o robustecimento da prova em crimes nos quais são utilizadas conversações telefônicas – tais como o tráfico de drogas, extorsões, ameaças – e em evidências coletadas por meio de gravações ambientais. O projeto também pode estimular outras criações de bases de dados de fala, permitindo o compartilhamento entre as instituições, tanto em nível nacional como internacional.

#### • Parcerias e Cronograma

A construção da base de dados proposta demandaria um esforço multi-institucional e multidisciplinar. Estima-se três anos como um prazo razoável para a construção de uma base de dados que atenda às demandas do novo paradigma.

Para que o projeto logre sucesso, faz-se necessária a colaboração de entidades diversas, tais como universidades, Secretaria de Segurança Pública, Secretaria Nacional de Segurança Pública (SENASP) e do Judiciário. A verba oriunda para tal projeto poderia vir das agências de fomento (FAPESP, CAPES), da SENASP e do Fundo Estadual de Segurança Pública.

#### AGRADECIMENTOS

Agradeço aos membros do corpo editorial da Revista Brasileira de Criminalística, uma ferramenta importantíssima à disposição dos Peritos Oficiais do Brasil.

#### REFERÊNCIAS BIBLIOGRÁFICAS

[1] M.J. Saks, J.J. Koehler. The coming paradigm shift in forensic identification science. *Sci.* **309**, 892-895, 2005.  
 [2] T.S. Kuhn. *The structure of scientific revolutions*. University of Chicago Press, 2012.  
 [3] G.S. Morrison. Distinguishing between forensic science and forensic pseudoscience - Testing of validity and reliability, and approaches to forensic voice comparison. *Sci. Justice* **54(3)**, 245-256, 2014.  
 [4] C. Aitken Berger, C.E., Buckleton, J.S. Champod, C. Curran, J. Dawid, A.P. Kloosterman. Expressing evaluative opinions: a position statement. *Sci. Justice* **51(1)**, 1-2, 2011.  
 [5] C.A. Pereira, F. Nakano, J.M. Stern, M.R. Whittle. Genuine Bayesian multiallelic significance test for the Hardy-Weinberg equilibrium law. *Genet. Mol. Res.* **5(4)**, 619-631, 2006.

[6] A.N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933. English translation, Chelsea, New York, 1950.  
 [7] C.A.B. Pereira, J.M. Stern. Evidence and credibility: full Bayesian significance test for precise hypotheses. *Entropy* **1(4)**, 99-110, 1999.  
 [8] C.G.G. Aitken, F. Taroni. *Statistics and the evaluation of evidence for forensic scientists*. Chichester: Wiley, 2004.  
 [9] D. Lucy. *Introduction to statistics for forensic scientists*. John Wiley & Sons, 2013.  
 [10] D. Van Lancker, J. Kreiman, K. Emmorey. Familiar voice recognition: patterns and parameters. Part I: Recognition of backward voices. *J. Phon.* **13**, 19-38, 1985.  
 [11] A.D. Yarmey Yarmey, A. D., Yarmey, A. L., Yarmey, M. J., & Parliament, L. Commonsense beliefs and the identification of familiar voices. *Applied Cognitive Psychology* **15(3)**, 283-299, 2001.  
 [12] R.M. Warren. *Auditory perception: A new synthesis*. Elsevier, 2013.  
 [13] K.N. Stevens. Sources of Inter - and Intra- speaker Variability in the Acoustics. Properties of Speech Sounds. In: A. Rigault (Ed.). *Actes Du Septième Congrès Int Des Sci.s Phonétiques, Tenu À L'Université de Montréal Et À L'Université McGill, 22-28 Août 1971*. Mouton, 1972.  
 [14] C.D. Romero. La identificación de locutores en el ámbito forense. *Tese de Doutorado*. Universidad Complutense de Madrid, 2001.  
 [15] P. Rose. *Forensic speaker identification*. CRC Press, 2003.  
 [16] B. Robertson, G.A. Vignaux. *Interpreting evidence: evaluating forensic science in the courtroom*. 1995.  
 [17] C. Champod, I.W. Evett. Commentary on APA Broeders (1999) 'Some observations on the use of probability scales in forensic identification', *Forensic Linguistics* 6 (2): 228-41. *Int. J. Speech Lang. Law* **7(2)**, 239-243, 2007.  
 [18] Gold, Erica Ashley, French Peter. Int practices in forensic speaker comparison. *Int. J. Speech Lang. Law* **18**, 2, 2011.  
 [19] A. Eriksson. Aural/Acoustic vs. Automatic Methods in Forensic Phon. Case Work. In: *Forensic Speaker Recognition*. Springer New York, 41-69, 2012.  
 [20] L. Rabiner, B.H. Juang. *Fundamentals of speech recognition*. 1993.  
 [21] J.P. Campbell Jr. Speaker recognition: A tutorial. *Proc. IEEE* **85(9)**, 1437-1462, 1997.  
 [22] D. Meuwly. Voice analysis. *Encyclopedia of Forensic Sci.* **3**, 1413-1420, 2000.  
 [23] D. Meuwly. Forensic individualization from biometric data. *Sci. Justice* **46(4)**, 205-213, 2006.  
 [24] F. Bimbot, J.F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meignier, T. Merlin, J. Ortega-Gacia, D. Petrovska-Delacrétaz, D.A. Reynolds. A tutorial on text-independent speaker verification. *EURASIP J. Appl. Signal Process.* **4**, 430-451, 2004.  
 [25] C. Champod, D. Meuwly. The inference of identity in forensic speaker recognition. *Speech Commun.* **31(2)**, 193-203, 2000.  
 [26] D.A. Reynolds, T.F. Quatieri, R.B. Dunn. Speaker verification using adapted Gaussian mixture models. *Digital Signal Processing* **10(1)**, 19-41, 2000.

- [27] Jain, Anil; Flynn, Patrick; Ross, Arun A. (Ed.). Handbook of biometrics. Springer Sci. & Business Media, 2007.
- [28] D. Reynolds. An overview of automatic speaker recognition technology. In: *Proc. Int Conf. Acoust., Speech and Signal Processing* **4**, 4072-4075, 2002.
- [29] J. Gonzalez-Rodriguez, A. Drygajlo, D. Ramos-Castro, M. Garcia-Gomar, J. Ortega-Garcia. Robust estimation, interpretation and assessment of likelihood ratios in forensic speaker recognition. *Comp. Speech Lang.* **20(2)**, 331-355, 2006.
- [30] N. Brümmer, L. Burget, J.H. Černocký, O. Glembek, F. Grezl, M. Karafiat, A. Strasheim. Fusion of heterogeneous speaker recognition systems in the STBU submission for the NIST speaker recognition evaluation 2006. *IEEE Transactions on Audio, Speech, and Language Processing* **15(7)**, 2072-2084, 2007.
- [31] D.A. van Leeuwen, N. Brümmer. An introduction to application-independent evaluation of speaker recognition systems. In: *Speaker classification I*. Springer Berlin Heidelberg, 330-353, 2007.
- [32] G.R. Doddington. The NIST speaker recognition evaluation—overview, methodology, systems, results, perspective. *Speech Commun.* **31(2)**, 225-254, 2000.
- [33] Pearce, N. J., Perkins, W. T., Westgate, J. A., Gorton, M. P., Jackson, S. E., Neal, C. R., & Chenery, S. P.. A compilation of new and published major and trace element data for NIST SRM 610 and NIST SRM 612 glass reference materials. *Geostandards Newsletter*, **21(1)**, 115-144, 1997.
- [34] D.E. Sturim, W.M. Campbell, D.A. Reynolds. Classification methods for speaker recognition. In: *Speaker Classification I*. Springer Berlin Heidelberg, 278-297, 2007.
- [35] D. Meuwly, A. Drygajlo. Forensic speaker recognition based on a Bayesian framework and Gaussian mixture modeling (GMM). In: *2001: A Speaker Odyssey—The Speaker Recognition Workshop*. 2001.
- [36] A. Drygajlo, D. Meuwly, A. Alexander. Statistical methods and Bayesian interpretation of evidence in forensic automatic speaker recognition. In: *INTERSPEECH*, 2003.
- [37] J.P. Campbell, W. Shen, W.M. Campbell, R. Schwartz, J.F. Bonastre, D. Matrouf. *Forensic speaker recognition*. Institute of Electrical and Electronics Engineers, 2009.
- [38] F. Botti, A. Alexander, A. Drygajlo. An interpretation framework for the evaluation of evidence in forensic automatic speaker recognition with limited suspect data. In: *ODYSSEY04—The Speaker and Lang Recognition Workshop*. 2004.
- [39] A. Alexander. Forensic automatic speaker recognition using Bayesian interpretation and statistical compensation for mismatched conditions. *Tese de Doutorado*. Institut de traitement des signaux. Section de Génie Électrique et Électronique, Indian Institute of Technology., Madras, 2005.
- [40] Gold, Erica Ashley. Calculating likelihood ratios for forensic speaker comparisons using Phon. and linguistic parameters. *Tese de Doutorado*. The University of York, 2014.
- [41] P. Rose. Long-and short-term within-speaker differences in the formants of Australian hello. *J. Int. Phon. Ass.* **29(1)**, 1-31, 1999.
- [42] T. Alderman. The Bernard data set as a reference distribution for Bayesian likelihood ratio-based forensic speaker identification using formants. In: *Proc. of the 10th Australian Int. Conf. Speech Sci. and Tech.* 510-515, 2004.
- [43] P. Rose. Forensic speaker discrimination with Australian English Vowel Acoustics. *ICPhS XVI Saarbrücken* **6**, 10, 2007.
- [44] C. Zhang, G.S. Morrison, P. Rose. Forensic Speaker Recognition in Chinese: a multivariate likelihood ratio discrimination on/i/and/y/. In: *Interspeech*, 1937-1940, 2008.
- [45] E. Enzinger. Characterizing formant tracks in Viennese diphthongs for forensic speaker comparison. In: *Audio Engineering Soc. Conf.: 39th Int. Conf. Audio Forensics: Practices and Challenges*. Audio Engineering Soc, 2010.
- [46] Y. Kinoshita, T. Osanai. Within speaker variation in diphthongal dynamics: What can we compare. In: *Proc. of the 11th Australasian Int Conf. on Speech Sci. & Tech., Auckland, New Zealand. Australia: Australasian Speech Sci. & Tech. Assoc., Canberra*, 112-117, 2006.
- [47] G.S. Morrison. Likelihood-ratio forensic voice comparison using parametric representations of the formant trajectories of diphthongs). *J. Acoust. Soc. Am.* **125(4)**, 2387-2397, 2009.
- [48] T. Becker, M. Jessen, C. Grigoras. Forensic speaker verification using formant features and Gaussian mixture models. In: *Interspeech*, 1505-1508, 2008.
- [49] C.M. Kavanagh. New consonantal acoustic parameters for forensic speaker comparison. *Tese de Doutorado*. University of York, 2012.
- [50] Y. Kinoshita. Does Lindley's LR estimation formula work for speech data? Investigation using long-term F0. *Int. J. Speech Lang. Law* **12(2)**, 235-254, 2007.
- [51] Y. Kinoshita, S. Ishihara, P. Rose. Exploring the discriminatory potential of F0 distribution parameters in traditional forensic speaker recognition. *Int. J. Speech Lang. Law* **16(1)**, 91-111, 2009.
- [52] Y. Kinoshita. Testing realistic forensic speaker identification in Japanese: A likelihood ratio based approach using formants. *Tese de Doutorado*. Australian National University, 2001.
- [53] A. Machado. Uso de técnicas acústicas para verificação de locutor em simulação experimental. *Língua, Literatura e Ensino*. ISSN 1981-6871, **6**, 2012.
- [54] A.C. Constantini, P.A. Barbosa. Prosodic characteristics of different varieties of Brazilian Portuguese. *Rev. Bras. Crimin.* **4(3)**, 44-53, 2015.
- [55] G.S. Morrison. A comparison of procedures for the calculation of forensic likelihood ratios from acoustic-Phon. data: Multivariate kernel density (MVKD) versus Gaussian mixture model—universal background model (GMM—UBM). *Speech Communication* **53(2)**, 242-256, 2011.
- [56] G.S. Morrison. Tutorial on logistic-regression calibration and fusion: converting a score to a likelihood ratio. *Aust. J. Forensic Sci.* **45(2)**, 173-197, 2013.
- [57] V. Hughes, P. Foulkes. Effects of variation on the computation of numerical likelihood ratios for forensic

- voice comparison. In: *Int Assoc. of Forensic Phon. and Acoust. Conf., Universidad Int. Menedez Pelayo, Santander, 5th-8th August*. 2012.
- [58] P. Rose. The effect of correlation on strength of evidence estimates in Forensic Voice Comparison: uni- and multivariate Likelihood Ratio-based discrimination with Australian English Vowel Acoustics. *Int. J. Biometrics* **2(4)**, 316-329, 2010.
- [59] P. Rose. More is better: likelihood ratio-based forensic voice comparison with vocalic segmental cepstra frontends *Int. J. Speech Lang. Law* **20(1)**, 77-116, 2013.
- [60] F. Nolan, McDougall, K., de Jong, G., & Hudson, T. A forensic Phon. study of dynamic sources of variability in speech: the DyViS project. In: *Proc. of the 11th Australasian Int. Conf. Speech Sci. and Tech.*, 13-18, 2006.
- [61] G. de Jong, K. McDougal, F. Nolan. Sound change and speaker identity: an acoustic study. In: *Speaker Classification II*. Springer Berlin Heidelberg, 130-141, 2007.
- [62] C. Zhang, G.S. Morrison, T. Thiruvaran. Forensic voice comparison using Chinese/iau. In: *Proc. 17th Int. Cong. Phon. Sci.*, 2280-2283, 2011.
- [63] P. Rose. Catching criminals by their voice—combining automatic and traditional methods for optimum performance. In: *Forensic Speaker Identification*, 2006.
- [64] G.S. Morrison Morrison, G. S., Rose, P., Zhang, C., Epps, J., Ambikairajah, E., González-Rodríguez, J., & Roux, C.I. Making Demonstrably Valid and Reliable Forensic Voice Comparison a Practical Everyday Reality in Australia-Database Collection Protocol. In: *Invited presentation at 13th Australasian Int. Conf. Speech Sci. Tech.* 14-16, 2010.
- [65] H. Melin. Databases for speaker recognition: Activities in COST250 working group 2. *COST 250-Speaker Recognition in Telephony, Final Report 1999*, 1999.
- [66] C.Y. Wang, P. Rose. Likelihood Ratio-Based Forensic Voice Comparison with Cantonese/i/F-Pattern and Tonal F0. In: *Proc. of the 14th Australasian Int Conf. on Speech Sci. and Tech. (SST 2012)*, 209-212, 2012.
- [67] G.S. Morrison, P. Rose, C. Zhang. Protocol for the collection of databases of recordings for forensic-voice-comparison research and practice. *Aust. J. Forensic Sci.* **44(2)**, 155-167, 2012.
- [68] Mapa do Encarceramento. Os jovens do Brasil. Secretaria Geral da Presidência da República, Brasília, 2014.
- [69] R. Soares. *A maioria penal no Brasil e em outros países*. Consultoria Legislativa. Fevereiro, 2006.
- [70] D.R.S. Lorenzi, E.C. Baracat, B. Saciloto, Padilha Jr. Fatores associados à qualidade de vida após menopausa. *Rev. Assoc. Med. Bras.* **52(5)**, 312-7, 2006.
- [71] A.O. Pedro, A.M. Pinto Neto, L.H. Paiva, M.J. Osis, E. Hardy. Age at natural menopause among Brazilian women: results from a population-based survey. *Cad. Saude Publica* **19(1)**, 17-25, 2003.
- [72] A.C. Bonaccorsi. Andropausa: insuficiência androgênica parcial do homem idoso. Uma revisão. *Arq. Bras. Endocrinol Metab.* **45(2)**, 123-133, 2001.
- [73] C.B.G. Fouquet. A influência no dialeto nordestino frente ao dialeto paulista. *Tese de Doutorado*. Universidade de São Paulo, 2013.
- [74] L. Oushiro, R.B. Mendes. A pronúncia do (-r) em coda silábica no português paulistano. *Rev. GEL.* **8(2)**, 66-95, 2013.
- [75] C.M.B. Leite. O/R/EM posição de coda silábica na capital do interior paulista: uma abordagem sociolinguística. *Sínteses - ISSN 1981-1314*, 15, 2012.
- [76] A.A.M. Silveira, L.E. Tenani. Elevação vocálica no diletto do interior paulista: contribuições para os estudos de variação fonológica do português do Brasil. *Estudos Linguísticos* **36(1)**, 18-26, 2007.
- [77] L.M.C. Imaguire. Estudo geolinguístico de alguns municípios do litoral sul paulista: abordagem de aspectos semântico-lexicais. *Tese de Doutorado*. Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2004.
- [78] P. Rose. Long-and short-term within-speaker differences in the formants of Australian “hello”. *J. Int. Phon. Assoc.* **29(1)**, 1-31, 1999.
- [79] F. McGhee. An experimental study of voice recognition. *J. Gen. Psychol.* **31(1)**, 53-65, 1944.
- [80] F. McGhee. The reliability of the identification of the human voice. *J. Gen. Psychol.* **17(2)**, 249-271, 1937.
- [81] A.H. Anderson, M. Bader, E.G. Bard, E.H. Boyle, G.M. Doherty, S.C. Garrod, S.D. Isard, J.C. Kowtko, J.M. McAllister, J. Miller, C.F. Sotillo, H.S. Thompson, R. Weinert. The HCRC map task corpus. *Lang. Speech* **34(4)**, 351-366, 1991.
- [82] G.S. Morrison, F. Ochoa, T. Thiruvaran. Database selection for forensic voice comparison. In: *Proc. Odyssey*, 62-77, 2012.