

Uma Proposta de Protocolo Técnico para Leitura Labial Forense

Renata Christina Vieira¹

¹Universidade Federal da Paraíba, João Pessoa, Paraíba, Brasil,
renata_c_vieira@yahoo.com.br

Resumo

A leitura labial tem sido empregada como recurso técnico em contextos forenses, especialmente em vídeos sem áudio ou com gravação comprometida. Apesar da aplicação crescente, faltam protocolos sistematizados que orientem a análise da fala silenciosa sob critérios técnicos, linguísticos e periciais. Este artigo apresenta um protocolo técnico-descritivo para a leitura labial forense, organizado em cinco etapas e fundamentado em práticas periciais documentadas. O roteiro considera parâmetros como resolução da imagem, taxa de quadros por segundo, ângulo da câmera, visibilidade dos visemas, variação geolinguística e contexto pragmático do enunciado. Discutem-se também os limites da inferência visual e os critérios de admissibilidade da mídia, com foco na integridade e rastreabilidade das informações analisadas. Ancorada na vertente da Linguística Forense voltada à linguagem como prova, a proposta busca consolidar a leitura labial como prática legítima e auditável, oferecendo fundamentos para formação profissional, atuação técnica e futura validação empírica em estudos aplicados.

Palavras-chave: leitura labial; fala; fonética forense; fonética articulatória; visemas; linguagem como prova; prova técnica; linguística forense.

Abstract

Lipreading is used as a technical tool in forensic contexts, especially in cases involving videos without audio or with compromised recordings. Despite its growing application, systematized protocols are still lacking to guide the analysis of silent speech using technical, linguistic, and forensic criteria. This article presents a technical-descriptive protocol for forensic lipreading, structured in five stages and grounded in documented practices. The analytical framework considers parameters such as image resolution, frame rate, camera angle, viseme visibility, geolinguistic variation, and pragmatic context. The discussion also addresses the limits of visual inference and the criteria for media admissibility, emphasizing the integrity and traceability of the analyzed material. Anchored in the forensic linguistics perspective that treats language as evidence, the proposal aims to establish lipreading as a legitimate and auditable practice. It provides a methodological foundation

for professional training, practical use, and future empirical validation in applied forensic communication studies.

Keywords: *lip reading; speech; forensic phonetics; articulatory phonetics; visemes; language as evidence; technical evidence; forensic linguistics.*

1. Introdução

A leitura labial é uma prática comumente associada à reabilitação de pessoas com perda auditiva (Bernstein, Auer, & Eberhardt, 2022; Tye-Murray et al., 2022), mas que vem ganhando relevância em contextos forenses, especialmente em casos que envolvem análise de vídeos sem áudio ou com gravação comprometida (Kaur et al., 2024; Theobald et al., 2006). Apesar de seu potencial como recurso técnico-pericial, a área carece de protocolos estruturados voltados à análise da movimentação articulatória em registros audiovisuais, com fins probatórios (Brendel, 2023).

A leitura labial forense se insere no campo da Linguística Forense na perspectiva da “linguagem como prova”, uma das três grandes vertentes da área, ao lado da análise da linguagem do processo judicial e das interações em contexto jurídico. No sentido restrito, “a definição de Linguística Forense limita a disciplina à linguagem como prova” (Sousa-Silva & Coulthard, 2016, p. 2), compreendida aqui como toda forma de manifestação linguística (falada, escrita ou visual), passível de análise técnica em contextos legais.

No caso da leitura labial, o objeto de análise é o movimento articulatório visível da fala, registrado em vídeo, cujo conteúdo pode ser interpretado por peritos com formação em fonética, fonologia, linguagem e comunicação audiovisual. Trata-se de uma evidência visual cuja análise envolve múltiplos níveis linguísticos, como o fonético-articulatório, o sintático, o discursivo e o pragmático, dependendo da extensão e do contexto da fala silenciosa. Como destacam Caldas-Coulthard (2014, p. 4), “o exame de evidência ou prova exemplifica as diversas ferramentas usadas por peritos/as na análise linguística que será usada por advogados/as em todos os tipos de conflitos jurídicos ou em tribunais de júri”.

A literatura sobre percepção da fala demonstra que o processamento linguístico não é exclusivamente auditivo, mas multimodal, envolvendo a integração entre informação visual e auditiva. Evidências clássicas dessa interação são descritas no chamado *efeito McGurk*, no qual estímulos visuais articulatórios influenciam a percepção do sinal sonoro (Green & Gerdeman, 1995; McGurk & MacDonald, 1976). Embora o efeito, em sentido estrito, dependa da presença simultânea de estímulos auditivos e visuais, esses achados corroboram a relevância do canal visual na ativação de representações fonético-articulatórias, aspecto diretamente relacionado ao fundamento perceptivo da leitura labial.

Esse tipo de perícia tem sido cada vez mais demandado em processos judiciais em que o áudio está ausente, ininteligível ou comprometido por fatores externos, sendo considerado, nesses casos, um recurso complementar para a elucidação de eventos captados por câmeras. No entanto, os critérios técnicos utilizados na leitura labial pericial

ainda são pouco sistematizados na literatura especializada. São raros os trabalhos que descrevem detalhadamente as etapas do processo analítico, os fundamentos fonético-articulatórios mobilizados, os parâmetros técnicos considerados ou os limites da inferência visual. Isso dificulta a replicação metodológica, a formação de novos profissionais e a avaliação crítica dos laudos periciais apresentados em juízo. A ausência de um protocolo claro compromete a possibilidade de reproduzir os procedimentos analíticos e de aplicar o mesmo raciocínio técnico a novos dados — aspecto que, para Grant (2010), é central em qualquer método forense que se pretenda confiável.

Além da transparência metodológica, é necessário que a validade de determinados marcadores ou escolhas analíticas não dependa exclusivamente da eficácia demonstrada em um caso específico. Como adverte Turell (2010), a robustez de um método pericial exige que ele seja testável fora do julgamento em que foi aplicado.

Embora existam avanços significativos na leitura labial automatizada com base em redes neurais profundas (Assael et al., 2016; Oghbaie et al., 2021), esses sistemas ainda enfrentam dificuldades em vídeos espontâneos, com ruído visual, baixa resolução ou ângulos desfavoráveis, justamente os tipos de material mais frequentes em contexto forense (Feng et al., 2020; Gimeno-Gómez & Martínez-Hinarejos, 2023). Isso reforça a importância da análise visual realizada por profissionais qualificados, baseada em critérios linguísticos e técnicos observáveis.

Este artigo apresenta uma proposta metodológica para a realização de análises técnicas em leitura labial forense, baseada na experiência profissional da autora em perícias e assistências técnicas realizadas sob demanda judicial. A proposta busca contribuir com um modelo sistematizado de análise visual da fala, voltado à atuação pericial, à formação técnica e à futura validação por estudos controlados (Brendel, 2023; Dörnyei, 2007).

2. Aplicações forenses documentadas

A leitura labial tem sido empregada como técnica auxiliar em diferentes sistemas judiciais, especialmente em situações em que o áudio está ausente, comprometido ou inteligível. Casos registrados em tribunais de diversos países demonstram seu potencial como elemento probatório, tanto em contextos de acusação quanto de defesa.

No Brasil, há registros de perícias em leitura labial que ganharam repercussão midiática e jurídica. Um dos exemplos mais citados é o caso envolvendo Marcelo Arruda, ocorrido em Foz do Iguaçu (Paraná), no qual imagens de videovigilância captaram o momento em que um agente penitenciário invadiu uma celebração e efetuou disparos. A defesa recorreu à leitura labial para analisar os segundos que antecederam os disparos, a fim de verificar a existência de possíveis provocações verbais por parte da vítima. A análise baseou-se na extração de quadros do vídeo, aplicação de zoom, ajustes de contraste e lentificação. Embora o conteúdo do parecer permaneça sob sigilo, o caso ilustra a importância da técnica na interpretação de comportamentos verbais em vídeos sem som.

Outro episódio ocorreu no âmbito do processo nº 0000072-17.2016.8.24.0005, julgado pelo Tribunal de Justiça de Santa Catarina. As imagens analisadas mostravam um detento a conversar com outros indivíduos durante o transporte prisional. A promotoria alegou tratar-se de tentativa de articulação criminosa, e a defesa solicitou perícia para

comprovar que o conteúdo das falas não correspondia à alegação. A análise técnica, conduzida segundo protocolo descritivo, contribuiu para a manutenção da sentença e validação do material como prova (Migalhas, 2024).

Em outros países, há também decisões judiciais que reconheceram a leitura labial como técnica válida, desde que aplicada com critérios rigorosos. É o caso de *Luttrell v. United Kingdom*, julgado pela Corte Europeia de Direitos Humanos em 2004, que discutiu a admissibilidade da técnica no âmbito da acusação. A Corte confirmou sua validade, desde que observados o contraditório e a transparência pericial (CaseMine, 2004). No Reino Unido, o caso de Arlene Fraser ganhou destaque por incluir uma perícia em leitura labial que identificou, em gravações de vigilância, expressões ameaçadoras inaudíveis na faixa sonora original. Os resultados dessa análise foram citados como evidência no processo (Northern Scot, 2014).

A diversidade de contextos em que a leitura labial tem sido aplicada reforça a necessidade de normatização técnica e fundamentação metodológica clara, capazes de garantir a rastreabilidade dos procedimentos, a transparência dos laudos e sua adequada valoração judicial. Diante desse cenário, torna-se fundamental a apresentação de protocolos sistematizados que orientem a prática pericial da leitura labial em contextos forenses, tema desenvolvido na seção seguinte.

3. Protocolo técnico de leitura labial forense

Apesar da prática crescente em contextos judiciais, os critérios técnicos utilizados na leitura labial pericial ainda são pouco sistematizados na literatura especializada. São raros os trabalhos que descrevem detalhadamente as etapas do processo analítico, os fundamentos fonéticos mobilizados, os parâmetros técnicos considerados ou os limites da inferência visual. Isso dificulta a replicação metodológica, a formação de novos profissionais e a avaliação crítica dos pareceres técnicos apresentados em juízo.

3.1. Finalidade e escopo do protocolo

O protocolo técnico-descritivo apresentado neste artigo tem como objetivo sistematizar os procedimentos de análise utilizados em perícias de leitura labial, especialmente em vídeos sem áudio ou com qualidade sonora comprometida (Brendel, 2023; Kaur et al., 2024). A proposta parte da experiência da autora em demandas judiciais, com a finalidade de auxiliar outros profissionais na execução das análises, na documentação dos procedimentos e na justificativa técnica das conclusões.

O escopo do protocolo abrange exclusivamente a análise técnica perceptiva-visual dos movimentos articulatórios registrados em vídeo. Não inclui o uso de inteligência artificial, nem apresenta técnicas de reconhecimento facial ou outras análises externas ao movimento da fala (Assael et al., 2016; Oghbaie et al., 2021). Sua finalidade é orientar a leitura labial com base em critérios fonéticos, linguísticos e técnicos da imagem, assegurando rastreabilidade, coerência metodológica e fundamentação técnico-pericial (Brendel, 2023; Peymanfard, Chatzarrin, & Reda, 2023).

Trata-se de uma proposta aberta à revisão, complementação e futura validação empírica, cuja contribuição principal reside na organização das etapas analíticas e na explicitação dos critérios decisórios utilizados pelo perito durante o exame técnico em leitura labial, elementos centrais em propostas metodológicas de base aplicada (Dörnyei, 2007).

3.2. Aspectos técnicos da imagem

A análise técnica em leitura labial forense depende fortemente da qualidade do material audiovisual. Os parâmetros técnicos da imagem impactam diretamente a observação dos visemas e, conseqüentemente, a precisão da inferência visual (Home Office UK, 2009; International Organization for Standardization, 2012). Nesse sentido, a seguir são apresentados os principais parâmetros técnicos da imagem relevantes para a análise em leitura labial forense.

Resolução: refere-se à quantidade de *pixels* que compõem a imagem. Quanto maior a resolução, maiores as chances de se observar com nitidez os contornos dos lábios e outros movimentos relevantes. Embora não exista um valor mínimo fixo, recomenda-se preferência por vídeos em 1080p ou superior, especialmente para análise quadro a quadro.

Taxa de quadros (*fps*): indica quantas imagens por segundo compõem o vídeo. Taxas mais altas proporcionam maior fluidez e permitem captar transições rápidas entre visemas. Abaixo de 15 *fps*, a análise tende a ser prejudicada, especialmente em fala espontânea (Pessoa et al., 2024).

Iluminação e contraste: referem-se à clareza da imagem e à distinção entre o rosto e o fundo. O rosto deve estar iluminado de forma uniforme, sem sombras sobre a região da boca e com contraste suficiente para destacar os contornos articulatórios (Pessoa et al., 2024).

Ângulo de visão e enquadramento: dizem respeito à posição da câmera em relação ao rosto. O ideal é que o vídeo apresente visão frontal ou levemente lateral, com a região da boca visível e sem obstruções (Home Office UK, 2009).

Compressão e *codec*: compressão é o processo de redução do tamanho do arquivo de vídeo. *Codecs* interframe (como H.264 em perfil *baseline*) descartam parte dos quadros intermediários e podem comprometer a observação de transições sutis entre gestos articulatórios. *Codecs* intraframe (como ProRes 422 ou DNxHD) mantêm todos os quadros completos, favorecendo a análise detalhada e o beneficiamento da imagem (Gloe, Fischer, & Kirchner, 2014; International Organization for Standardization, 2012).

Integridade da mídia: diz respeito à garantia de que o arquivo analisado é o mesmo que foi recebido, sem alterações ao longo da cadeia de custódia. Para esse fim, o perito deve proceder ao cálculo do *hash* do arquivo original, bem como das versões eventualmente derivadas, assegurando a rastreabilidade do material e a documentação de todas as etapas de processamento.

De forma complementar, a extração e o registro das características técnicas da mídia são essenciais para a descrição objetiva da qualidade do material audiovisual e para a avaliação de sua adequação à análise de leitura labial forense. Ferramentas como o MediaInfo destinam-se especificamente a essa finalidade descritiva, não se confundindo com os procedimentos de verificação de integridade do arquivo (Associação Brasileira de Normas Técnicas, 2022; Pessoa et al., 2024).

Em contexto forense, eventuais procedimentos de beneficiamento da imagem devem ser estritamente limitados a ajustes que não introduzam novos elementos visuais, como contraste, brilho, nitidez ou correção de cor. Não é admissível o emprego de ferramentas baseadas em inteligência artificial (IA) que gerem ou preencham *pixels*

inexistentes na gravação original, uma vez que tais intervenções violam o princípio da autenticidade da evidência e comprometem a admissibilidade probatória. O beneficiamento deve, portanto, preservar integralmente a estrutura da imagem, atuando apenas sobre informações efetivamente captadas no registro original.

Em conjunto, os parâmetros técnicos descritos nesta seção constituem critérios fundamentais para a avaliação da viabilidade e da confiabilidade da leitura labial em contexto forense. A inadequação de um ou mais desses aspectos pode restringir a observação dos visemas, introduzir ambiguidades interpretativas ou limitar o alcance inferencial da análise. Assim, a consideração sistemática desses parâmetros é indispensável tanto para a triagem técnica do material quanto para a interpretação prudente dos resultados obtidos em análises de leitura labial forense.

3.3. Etapas do roteiro de leitura labial forense

A análise técnica em leitura labial forense, conforme sistematizada neste protocolo, é organizada em cinco etapas interdependentes. Esse roteiro permite que cada decisão interpretativa seja fundamentada em critérios observáveis, técnicos e documentáveis, o que contribui para a rastreabilidade da análise e para a redução de subjetividade (Brendel, 2023; Gloe et al., 2014).

A adoção de um roteiro estruturado em etapas sucessivas também atua como estratégia de mitigação de vieses cognitivos, em particular do chamado *efeito priming*, amplamente descrito na literatura sobre transcrição e interpretação de emissões discutidas. Esse efeito refere-se à influência, muitas vezes inconsciente, de expectativas prévias do analista — derivadas do contexto do caso, da sua repercussão midiática ou de informações externas — sobre a interpretação de sinais ambíguos, sejam eles auditivos ou visuais. Em contextos forenses sensíveis, como casos de grande exposição pública, tal influência pode levar à sobrevalorização de hipóteses interpretativas compatíveis com narrativas pré-existentes. Por essa razão, diversos autores recomendam que análises desse tipo sejam conduzidas por profissionais com formação específica em fonética e percepção da fala, bem como que se explicitem hipóteses alternativas e o grau de certeza associado a cada interpretação, conforme proposto para emissões discutidas (*disputed utterances*) (Fraser, 2003).

3.3.1. Etapa 1: Preparação e avaliação inicial da mídia

Antes de qualquer tentativa de leitura labial, o perito deve verificar se a mídia audiovisual atende aos critérios técnicos mínimos para uma análise visual confiável. Isso inclui observar se o vídeo possui resolução adequada (preferencialmente 720p ou superior), taxa de quadros constante (mínimo de 24 *fps*), iluminação uniforme e enquadramento frontal da face. Também devem ser considerados fatores como compressão excessiva, artefatos visuais e visibilidade da região bucal, incluindo lábios, mandíbula e parte inferior do rosto (Home Office UK, 2009; IBM Corporation & Microsoft Corporation, 1991; International Organization for Standardization, 2012; Pessoa et al., 2024).

A ausência desses parâmetros pode inviabilizar a análise ou exigir intervenções técnicas na imagem (beneficiamento), que devem estar plenamente documentadas. O objetivo desta etapa é garantir a admissibilidade técnica da mídia, evitando que a perícia se baseie em evidências visuais inaptas para observação articulatória.

3.3.2. Etapa 2: Análise fonética e articulatória inicial

Cumpridos os critérios de qualidade da mídia, o perito inicia a análise da movimentação articulatória visível. Isso envolve a observação dos gestos dos lábios, mandíbula, bochechas e, quando possível, da língua, com o intuito de identificar os visemas produzidos. O foco recai sobre a identificação de vogais e consoantes visíveis, considerando as transições entre segmentos e avaliando se a sequência gestual observada forma uma cadeia fonético-articulatória coerente com os padrões da fala natural (Barbosa & Albano, 2004; Bear & Harvey, 2019; Peymanfard et al., 2023).

Essa etapa exige domínio das categorias de visemas e de suas ambiguidades, principalmente em casos em que diferentes fonemas compartilham gestos visuais semelhantes, os chamados *homofemas* (Brendel, 2023; Thangthai, Bear, & Harvey, 2018). O perito deve registrar hipóteses interpretativas concorrentes, incluindo alternativas possíveis, bem como explicitar o grau de confiança associado a cada hipótese, de modo a refletir as incertezas inerentes à análise de material visual ambíguo.

Essa prática encontra respaldo nas recomendações de Fraser (2003), segundo as quais, em materiais discutidos, a interpretação de segmentos específicos não pode ser dissociada do contexto discursivo mais amplo, sendo fundamental apresentar interpretações alternativas e indicar o nível de certeza do perito em relação a cada uma delas. Tal procedimento contribui para a transparência analítica, a prudência inferencial e a adequada valoração probatória da leitura labial em contexto forense.

3.3.3. Etapa 3: Análise contextual, sintática e geolinguística

Na sequência, o perito deve avaliar se as palavras ou frases inferidas são compatíveis com o contexto pragmático, a situação comunicativa e a estrutura sintática da fala. Isso envolve considerar aspectos como cenário, interações não verbais, expressões faciais e elementos situacionais visíveis (Cascone, Battaglini, & Galasso, 2023; Kim et al., 2023).

Quando necessário, o perito pode recorrer à comparação com outros vídeos do mesmo indivíduo ou à coleta de material padrão (gravações controladas da fala do sujeito). Assim como na comparação de locutores, essa coleta pode ajudar a verificar a compatibilidade entre os padrões articulatórios. Para evitar distorções causadas pelo “paradoxo do observador” (Labov, 1966, 1972), é recomendável que a coleta ocorra em ambiente neutro, com elicitación de frases espontâneas e contextualizadas, minimizando a influência da monitorização sobre os gestos naturais.

Em perícias vocais, esse princípio já foi adaptado para obtenção de amostras semispontâneas, guiadas pela escuta prévia do material questionado (Vieira, 2017). A proposta deste artigo é aplicar a mesma lógica metodológica à leitura labial, permitindo que se capte a variabilidade gestual do indivíduo, com respeito aos limites éticos da atuação pericial.

3.3.4. Etapa 4: Inspeção temporal detalhada (análise quadro a quadro)

Com as hipóteses fonético-contextuais em mente, o perito realiza uma inspeção precisa do vídeo com o auxílio de *softwares* que permitam lentificação, congelamento e extração de quadros individuais. Essa análise quadro a quadro é essencial para observar microtransições entre visemas, pausas, hesitações, prolongamentos articulatórios e outros

marcadores discursivos que sustentam a interpretação (Bernstein et al., 2022; Macedo & Silva, 1996).

Nesta etapa, é altamente recomendável a inserção de um cronômetro digital sobreposto ao vídeo, de modo que cada gesto articulatório possa ser associado a um tempo exato. Essa marcação temporal permite correlacionar segmentos interpretados com quadros específicos do vídeo e oferece suporte técnico à organização das evidências, além de facilitar a replicação do parecer por outros especialistas. A documentação visual precisa, com imagens extraídas, tempo marcado e justificativas fonéticas para cada inferência, compõe o material probatório que acompanha o laudo técnico (Gloe et al., 2014).

3.3.5. Etapa 5: Validação cruzada e colaboração entre especialistas

Por fim, recomenda-se que as hipóteses interpretativas sejam, sempre que possível, revistas por outro especialista ou por um leitor labial experiente. Essa prática de validação cruzada atua como mecanismo de controle metodológico e de mitigação de vieses interpretativos. Em casos ambíguos ou sensíveis, essa colaboração pode ocorrer por meio de revisão independente, trabalho em dupla ou discussão consensual das hipóteses, podendo incluir, quando pertinente, surdos oralizados com uso funcional da leitura labial no português, desde que sua participação se restrinja ao plano perceptivo e que todas as etapas sejam devidamente registradas no parecer (Dörnyei, 2007; Tye-Murray et al., 2022).

A sistematização dessas cinco etapas contribui para que a leitura labial forense deixe de ser uma prática intuitiva e passe a ser compreendida como procedimento técnico fundamentado, auditável e compatível com os parâmetros exigidos em perícias judiciais.

4. Limitações técnicas e epistêmicas da leitura labial

Apesar de sua utilidade prática, a leitura labial apresenta limitações importantes, tanto do ponto de vista técnico quanto epistemológico. Como técnica baseada na observação visual da fala, ela está sujeita a zonas de ambiguidade visual decorrente da própria articulação da fala. Muitos fonemas compartilham gestos articulatórios semelhantes, como os pares homorgânicos /p/ e /b/, /t/ e /d/, dificultando a diferenciação apenas pelo movimento dos lábios. Além disso, sons como vogais médias e vogais orais não arredondadas tendem a gerar visemas pouco contrastivos, o que aumenta o risco de sobreposição interpretativa.

Do ponto de vista técnico, a qualidade da imagem interfere diretamente na confiabilidade da análise. Baixa resolução, taxa de quadros inferior a 24 *fps*, iluminação deficiente e ângulo desfavorável da câmera comprometem a identificação precisa dos movimentos labiais e das transições articulatórias. A compressão excessiva por *codecs* interframe pode apagar microgestos essenciais à leitura, e a perda de nitidez ou presença de artefatos digitais pode introduzir contornos artificiais ou atrasos visuais que afetam a percepção de sincronia entre gestos.

Além das limitações perceptivas, a leitura labial também enfrenta desafios epistemológicos. Considerando que o conteúdo analisado não possui faixa sonora audível, o perito recorre exclusivamente a sinais visuais para inferir unidades fonéticas e construções sintáticas, o que exige cautela redobrada. A ausência de elementos acústicos

impede a recuperação de pistas como entonação, intensidade, timbre e prosódia fina, embora algumas dessas pistas possam, em determinados casos, ser parcialmente inferidas a partir de sinais visuais associados ao ritmo articulatório ou à expressividade facial.

Nesse contexto, é necessário que o laudo técnico explicita com clareza os limites da análise, sinalizando as condições que afetaram a confiabilidade da interpretação e evitando afirmações categóricas. Conforme argumenta Gascón Abellán (2010), a prova técnica deve ser compreendida no paradigma da verossimilhança, ou seja, como um contributo plausível, passível de questionamento e de reconstrução argumentativa, e não como portadora de verdade absoluta.

5. Discussão

A proposta metodológica apresentada neste artigo representa um esforço de sistematização técnica da leitura labial aplicada ao contexto forense, campo ainda desprovido de protocolos estruturados, acessíveis e passíveis de replicação (Brendel, 2023). Embora o uso dessa técnica em perícia tenha crescido nos últimos anos, especialmente em vídeos sem áudio ou de difícil escuta (Kaur et al., 2024; Theobald et al., 2006), as abordagens permanecem, em grande medida, baseadas em práticas individuais, pouco documentadas ou sem respaldo publicado.

Uma das contribuições centrais deste protocolo está em tornar explícitas etapas analíticas recorrentes na prática técnica, muitas vezes realizadas de forma intuitiva. Ao nomear essas etapas, definir sua função e estabelecer critérios, a proposta permite que outros profissionais compreendam a lógica da análise, repliquem o procedimento e reflitam sobre suas próprias decisões periciais. A padronização favorece também a rastreabilidade e a justificativa técnica de cada conclusão, exigência cada vez mais presente no sistema de justiça brasileiro (Associação Brasileira de Normas Técnicas, 2022).

Outro ponto relevante da proposta é a articulação entre aspectos fonéticos, parâmetros técnicos de gravação e análise linguística e contextual da fala. Essa abordagem integrada valoriza a experiência do perito como observador qualificado, sem depender de algoritmos automatizados, cujos resultados ainda apresentam limitações importantes em contextos reais (Assael et al., 2016; Feng et al., 2020).

A literatura atual aponta que, mesmo em ambientes controlados e com material de alta qualidade, modelos de leitura labial automática enfrentam dificuldades na generalização para fala espontânea, variação linguística e gravações fora de estúdio (Oghbaie et al., 2021). Isso reforça o valor da análise perceptiva criteriosa como prática técnica, especialmente em vídeos forenses com múltiplas interferências.

A prova técnica, nesse cenário, deve ser compreendida como elemento interpretativo e não como mecanismo de certeza. A leitura labial pericial assume uma natureza perceptiva e linguística, exigindo clareza metodológica, explicitação dos limites e transparência na apresentação das hipóteses analíticas. Essa perspectiva está alinhada ao paradigma da verossimilhança descrito por Gascón Abellán (2010), segundo o qual a prova científica contribui com a plausibilidade de uma hipótese, sem assumir função de decisão judicial.

Como reforça Gascón Abellán (2010, p. 100), o papel do perito é fornecer ao processo uma interpretação tecnicamente fundamentada dos dados, mas a valoração das hipóteses processuais cabe ao juiz, e não ao laudo técnico:

El informe pericial —no importa repetirlo una vez más— no valora directamente estos enunciados. No habla de la probabilidad de las hipótesis judiciales en juego a la luz de los datos analíticos resultantes de la prueba, sino de la probabilidad de esos resultados a la luz de esas hipótesis. Quien tiene la tarea de valorar esos enunciados es el juez, que debe hacerlo además tomando en consideración no sólo los datos del informe sino también los demás datos que obren en la causa.

Essa compreensão reforça a importância de que a leitura labial forense seja apresentada com responsabilidade epistêmica, argumentação clara e base observacional sólida, respeitando os limites da inferência visual e o papel do perito no processo.

Essa mesma perspectiva permite compreender a leitura labial forense como uma modalidade de reconhecimento técnico visual. O reconhecimento de falantes, no campo pericial, é tradicionalmente dividido em duas categorias: o reconhecimento simples, realizado por qualquer ouvinte com base em memória auditiva, e o reconhecimento técnico, realizado por peritos com uso de procedimentos sistematizados (Rose, 2002). Enquanto o primeiro tem valor testemunhal, o segundo visa produzir prova técnico-científica por meio de análises perceptivo-auditiva e acústica (Broeders, 2001; Nolan, 1983). No âmbito pericial, esse reconhecimento técnico pode ocorrer por verificação, ou seja, quando se compara uma amostra questionada com uma amostra padrão do mesmo indivíduo, em processo de escolha binária (Figueiredo, 1994).

Essa lógica metodológica pode ser estendida à leitura labial, desde que conduzida por profissional qualificado, com base em parâmetros observáveis e replicáveis. Trata-se de um reconhecimento visual especializado que, ao comparar registros questionados e padrões visuais, busca sustentar hipóteses com base na consistência articulatória. Essa abordagem visa reduzir a ambiguidade perceptiva e conferir maior confiabilidade à análise. Quando apresentada com clareza metodológica, fundamentação linguística e responsabilidade epistêmica, a leitura labial forense se insere no campo das perícias técnicas passíveis de controle, revisão e validação.

A colaboração entre especialistas em leitura labial, embora nem sempre prevista de forma obrigatória, representa uma estratégia valiosa para aumentar a confiabilidade das análises periciais. Em contextos forenses, especialmente naqueles marcados por baixa qualidade de imagem, ângulos desfavoráveis ou presença de múltiplos interlocutores, a interpretação visual da fala pode apresentar ambiguidade, sobretudo quando influenciada por inferências contextuais e pela sensibilidade linguística do perito. Nessas situações, a consulta a outro especialista pode funcionar como forma de validação cruzada, oferecendo uma segunda interpretação técnica com base em critérios semelhantes (Dörnyei, 2007).

Essa colaboração pode ocorrer de diferentes maneiras: revisão técnica independente do parecer; trabalho em dupla desde o início da análise; ou participação de um terceiro perito para desempate interpretativo. Em casos que envolvam falas com regionalismos marcantes ou traços de identidade social visíveis nos gestos articulatórios, a contribuição de especialistas com experiência sociolinguística ou com vivência em determinada comunidade de fala pode agregar precisão à análise. Além disso, a leitura labial pode

ser complementada por análise não verbal, ampliando o escopo de compreensão da cena analisada.

Mais do que uma sobreposição de competências, a colaboração entre especialistas deve ser entendida como uma salvaguarda metodológica: um mecanismo que protege o processo de vieses individuais, ao mesmo tempo em que contribui para a transparência do parecer técnico. Quando documentada com clareza — ou seja, explicitando quem participou, quais interpretações foram propostas e quais consensos ou divergências surgiram — a colaboração entre especialistas fortalece a legitimidade epistêmica da perícia e demonstra compromisso ético com a robustez das conclusões.

6. Conclusão

Como destaca Caldas-Coulthard (2014, p. 6), “a Linguística Forense é uma área de conhecimento que precisa constantemente legitimar-se perante os operadores do direito, demonstrando que seus métodos de análise são sistemáticos, científicos e reproduzíveis.” Nesse cenário, a leitura labial forense vem se consolidando como uma ferramenta relevante em contextos nos quais o áudio é ausente, ininteligível ou insuficiente para análise acústico-linguística (Brendel, 2023; Kaur et al., 2024; Theobald et al., 2006; Vieira & Troian de Souza, 2025). No entanto, a área ainda carece de protocolos metodológicos amplamente descritos e validados, o que compromete a uniformidade dos procedimentos e a possibilidade de revisão crítica entre especialistas.

Este artigo apresentou um protocolo técnico-descritivo construído a partir da prática pericial, com o objetivo de organizar as etapas da análise e explicitar os critérios técnicos, linguísticos e contextuais que orientam a atuação em leitura labial. A descrição das etapas, da avaliação inicial do vídeo à possibilidade de colaboração entre especialistas, busca contribuir para a rastreabilidade, a transparência e a justificabilidade das conclusões periciais (Associação Brasileira de Normas Técnicas, 2022; Dörnyei, 2007).

Embora ainda não validado empiricamente, o protocolo oferece uma base estruturada útil tanto para a formação profissional quanto para a documentação técnica de análises em pareceres ou laudos. Estudos futuros poderão incorporar testes de confiabilidade interavaliadores, construção de bancos de dados visuais em português e comparações sistemáticas com registros visuais controlados do mesmo falante (Bernstein et al., 2022).

Para além da análise perceptiva dos movimentos articulatórios captados em vídeo, o protocolo proposto contempla estratégias complementares que ampliam a robustez da interpretação, como o uso de material padrão, a formulação de hipóteses linguísticas situadas e a consideração de traços geolinguísticos e sociolinguísticos. A colaboração entre especialistas também é destacada como recurso valioso para validação cruzada.

Como proposta aberta, a sistematização aqui apresentada constitui um ponto de partida metodológico para que possa ser debatido, refinado e validado por meio de estudos empíricos e da colaboração da comunidade científica (Dörnyei, 2007).

Referências

Assael, Y. M., Shillingford, B., Whiteson, S., & de Freitas, N. (2016). *LipNet: End-to-End Sentence-level Lipreading*.

- Associação Brasileira de Normas Técnicas. (2022). *NBR ISO/IEC 27002:2022 – Tecnologia da informação – Técnicas de segurança – Código de prática para a gestão da segurança da informação* (Tech. Rep.). Rio de Janeiro: ABNT.
- Barbosa, P. A., & Albano, E. C. (2004). Brazilian Portuguese. *Journal of the International Phonetic Association*, 34(2), 227–232.
- Bear, H. L., & Harvey, R. (2019). Alternative visual units for visual speech recognition. *Applied Sciences*, 9(18), 3870.
- Bernstein, L. E., Auer, E. T., & Eberhardt, S. P. (2022). Lipreading training with sentence stimuli. *American Journal of Audiology*, 31(1), 57–77.
- Brendel, O. (2023). Some aspects of visual-auditory perception of oral speech while video and sound recordings examination. *Theory and Practice of Forensic Science and Criminalistics*, 28, 13–22.
- Broeders, A. P. A. (2001). *Forensic speech and audio analysis forensic linguistics 1998 to 2001: A review*. (Published: Paper apresentado no 13th INTERPOL Forensic Science Symposium)
- Caldas-Coulthard, C. R. (2014). ReVEL na escola: O que é a linguística forense? *ReVEL*, 12(23).
- Cascone, L., Battaglino, C., & Galasso, F. (2023). *Language identification for lip-based biometric systems*.
- CaseMine. (2004). *R v Luttrell & Ors*. Retrieved from <https://www.casemine.com/judgement/uk/5a8ff7b460d03e7f57eb1590>
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Feng, D., Yang, S., Shan, S., & Chen, X. (2020). *Learn an effective lip reading model without pains*.
- Figueiredo, R. M. (1994). *Identificação de falantes: Aspectos teóricos e metodológicos* (Tese de Doutorado). Universidade Estadual de Campinas.
- Fraser, H. (2003). Issues in transcription: Factors affecting the reliability of transcripts as evidence in legal cases. *Forensic Linguistics*, 10(2), 203–226.
- Gascón Abellán, M. (2010). Prueba científica: Mitos y paradigmas. *Anales de la Cátedra Francisco Suárez*, 44. Retrieved from <https://revistaseug.ugr.es/index.php/acfs/article/view/500/590>
- Gimeno-Gómez, D., & Martínez-Hinarejos, C.-D. (2023). *Continuous Lip Reading in Spanish*.
- Gloe, T., Fischer, A., & Kirchner, M. (2014). Forensic analysis of video file formats. *Digital Investigation*, 11, S68–S76.
- Grant, T. (2010). Text messaging forensics: Txt 4n6 – Idiolect free authorship analysis? In M. Coulthard, A. Johnson, & D. Wright (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 517–533). London; New York: Routledge.
- Green, K. P., & Gerdeman, A. (1995). Cross-modal discrepancies in coarticulation and the integration of speech information: The McGurk effect with mismatched vowels. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6), 1409–1426.
- Home Office UK. (2009). *CCTV Operational Requirements Manual* (Tech. Rep.). Home Office UK.
- IBM Corporation, & Microsoft Corporation. (1991). *Multimedia Programming Interface*

- and Data Specifications 1.0 (Tech. Rep.). IBM Corporation; Microsoft Corporation. International Organization for Standardization. (2012). *ISO/IEC 14496-12:2012 – Coding of audiovisual objects – Part 12: ISO base media file format* (Tech. Rep.). Geneva: ISO.
- Kaur, M., Rastogi, D., Sharma, A., Mahajan, P., Dahiya, A., & Nagrath, P. (2024, November). Crime Investigation Using Lip Reading. In *SCCTT-2024*. Delhi, India.
- Kim, M., Yeo, J. H., Choi, J., & Ro, Y. M. (2023). *Lip reading for low-resource languages by learning and combining general speech knowledge and language-specific knowledge*. arXiv. Retrieved from <https://arxiv.org/abs/2308.09311> doi: 10.48550/ARXIV.2308.09311
- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Macedo, A. T., & Silva, G. M. d. O. e. (1996). Análise sociolinguística de alguns marcadores conversacionais. In A. T. Macedo, C. N. Roncarati, & M. C. Mollica (Eds.), *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–748. doi: 10.1038/264746a0
- Migalhas. (2024). *TJ-SC confirma júri após prova recuperada por leitura labial*. Retrieved from <https://www.migalhas.com.br/quentes/420959/tj-sc-confirma-juri-apos-prova-recuperada-por-leitura-labial>
- Nolan, F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge: Cambridge University Press.
- Northern Scot. (2014). *Feature: A decade since final chapter of Arlene Fraser murder*. Retrieved from <https://www.northern-scot.co.uk/news/feature-decade-since-final-chapter-of-arlene-fraser-murder-329389/>
- Oghbaie, M., Sabaghi, A., Hashemifard, K., & Akbari, M. (2021). *Advances and Challenges in Deep Lip Reading*. (arXiv:2110.07879)
- Pessoa, A. F., Vieira, R. C., Sanches, A. B., & Gonzalez, R. C. S. (2024). Admissibilidade de amostras forenses. In L. Lopes et al. (Eds.), *Tratado de Fonoaudiologia* (pp. 598–605). São Paulo: Manole.
- Peymanfard, J., Chatrzzarrin, H., & Reda, H. T. (2023). *Leveraging Visemes for Lip Reading*. (arXiv:2307.10157)
- Rose, P. (2002). *Forensic Speaker Identification*. London: Taylor & Francis.
- Sousa-Silva, R., & Coulthard, M. (2016). Linguística forense. In R. J. Dinis-Oliveira & T. Magalhães (Eds.), *O que são as ciências forenses? – Conceitos, abrangência e perspectivas futuras*. Lisboa: Pactor.
- Thangthai, K., Bear, H. L., & Harvey, R. (2018). *Comparing phonemes and visemes for lipreading*.
- Theobald, B. J., Harvey, R., Cox, S. J., Lewis, C., & Owen, G. P. (2006, September). Lip-reading enhancement for law enforcement. In C. Lewis & G. P. Owen (Eds.), . Stockholm, Sweden. Retrieved from <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.689960> doi: 10.1117/12.689960
- Turell, M. T. (2010). The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *The International Journal of Speech, Language and the*

Law, 17(2), 211–250.

- Tye-Murray, N., Spehar, B., Sommers, M., Mauzé, E., Barcroft, J., & Grantham, H. (2022, January). Teaching children with hearing loss to recognize speech: Gains made with computer-based auditory and/or speechreading training. *Ear and Hearing*, 43(1), 181–191. Retrieved from <https://journals.lww.com/10.1097/AUD.0000000000001091> doi: 10.1097/AUD.0000000000001091
- Vieira, R. C. (2017). Contribuições dos estudos sociofonéticos para a identificação de falantes. *Revista Intercâmbio – Especial Expressividade*, XXXVI, 86–102.
- Vieira, R. C., & Troian de Souza, L. (2025). Lipreading in speech-language-hearing forensic analysis: Perspectives and challenges. *CoDAS*, 37(6). Retrieved from <https://www.scielo.br/j/codas/a/KGTkmrFzh9R7TZ9rMdMfpgc/>