Renata Christina Vieira[1]
Letícia Troian de Souza[2]

Lipreading in speech-language-hearing forensic analysis: perspectives and challenges

A técnica de leitura labial na perícia fonoaudiológica: perspectivas e desafios

Dear Editors-in-Chief,

The speech-language-hearing (SLH) sciences have been expanding their reach in the legal field through SLH forensic analysis. Lipreading is a promising technique whose applications, potential, and limitations merit discussion. It is used to identify words uttered by a speaker based on the movements of the articulatory organs, consistent with the context of analysis[1,2]. Although lipreading identifies only 50% of speech, it may be the only viable source of evidence in certain cases, which justifies its use. This technique seeks to understand speech through the visual analysis of lip movements and facial expressions, an area of SLH expertise that integrates knowledge of articulation, phonetics, language, and orofacial dynamics[3,4]. In SLH, this skill helps to validate testimonies and clarify situations that require investigation and depend on understanding the speech of someone who has been filmed and whose audio cannot be verified, only the image. The reliability of the analysis depends on several factors, such as the quality of the images and the expert's mastery of phonetic and linguistic knowledge. Lipreading expertise lies at the crossroads between forensic phonetics and forensic linguistics, integrating concepts from both fields to provide a more accurate interpretation of speech based on visual-perceptual evaluation. There are several studies on lipreading, but they typically focus on assisting deaf individuals, focusing on improving speech comprehension through visual training[5-8], and on the possibility of automatic lipreading. The literature on lipreading highlights the importance of its use in investigations and in supporting law enforcement and argues for its relevance to the legal field[9].

Although lipreading has been studied to assist deaf people, its forensic use has not been adequately addressed. Lipreading expertise can play an important role in investigations and legal support, but it is also important to analyze its limitations and potential.

LIMITATIONS

The following limitations stand out in forensic lipreading: admissibility, it lacks robust methodology, leading to potential uncertainty in the analysis results. SLH sciences contribute to the initial stage of the forensic analysis, verifying whether the samples meet the minimum criteria for producing evidence[10]. Integrity, quantity, and quality are the elements of this stage that most effectively affect the use of lipreading. Integrity refers

Correspondence address:
Renata Christina Vieira
Hospital Central Aristarcho Pessoa – CBMERJ
Av. Paulo de Frontin, 876, Rio Comprido, Rio de Janeiro (RJ), Brasil, CEP: 20261-243.
E-mail: renatachristinavieira@gmail.com

to the analysis of the material to determine whether there has been any alteration, damage, or modification to the video. The quantity aspect of lipreading relates to the number of frames in the image, as poor image quality affects the visualization of articulatory movements. Quality in the context of lipreading is related to resolution, brightness, focus, camera positioning, and the quality of the equipment. Inadequate quality affects image clarity, making the task impossible. Furthermore, it's important to ensure the face is visible; wearing hats and masks can impede visualization of articulatory movements. Another aspect that limits the technique's use is the lack of a robust methodology, established protocols, and studies in the field, reducing the reliability of expert analyses. The absence of a method can lead the expert to biased interpretations, subjecting them to error. This brings us to the third limiting factor in the technique's use: the potential uncertainty in the analysis results. As mentioned previously, lipreading allows us to visualize 50% of speech, as it can only capture phonemes produced in the anterior part of the vocal tract. Furthermore, some phonemes are homorganic (i.e., produced similarly), which prevents visual distinction between them. It's also important to consider that lipreading doesn't capture important information for speech comprehension, such as prosody and situational context, which are crucial for fully understanding what is being said.

POTENTIALS

The strengths of lipreading include the knowledge of hypotheses and context, the expert's comprehensive knowledge of phonetics and linguistics, the use of videos of the suspect as standard material, and partnership with lipreaders. Understanding hypotheses allows one to validate what witnesses claim to have heard and verify whether this is consistent with the observed lip movements[11]. This knowledge helps reduce ambiguities by comparing what was said with the allegations, which makes the analysis more accurate. Knowledge of context facilitates speech interpretation and allows one to predict more expected words or phrases, reducing the margin of error and increasing the reliability of lipreading. Phonetic knowledge is important for identifying visible and invisible sounds, facilitating the prediction of what was said by observing lip movements[12-15]. Linguistic knowledge involves understanding grammatical structure and vocabulary, helping to correctly associate lip movements with words[12,16,17]. Proficiency in auditory-perceptual evaluation contributes to the identification of a speaker's vocal characteristics, especially vocal tract adjustments[18], improving lipreading accuracy. Furthermore, geolinguistic knowledge facilitates the identification of regional variations, considering differences in pronunciation and local expressions, which enhances the analysis. It is important to investigate the individual's background, including their birthplace and place of residence, as these factors can influence how speech is articulated and perceived. Studying videos of the suspect speaking allows us to understand their vocal and linguistic profile and identify individual articulation patterns[19]. If access to videos of the suspect is not possible, the court may request the collection of standard material. Regarding partnerships with lipreaders,

qualitative studies largely depend on the investigator's analytical (subjective) ability. Reliability checks can be performed[20], meaning another researcher can analyze the data and verify whether the results agree or disagree. Partnering with individuals skilled in lipreading can validate interpretations and strengthen the analysis, as they can make pertinent observations about lip movements and help confirm hypotheses.

FINAL CONSIDERATIONS

Lipreading is an important SLH analytical tool, essential for validating testimonies and analyzing situations where audio from video recordings is inaudible or nonexistent. As a technique, it requires the SLH pathologist to have in-depth knowledge of speech articulation, orofacial dynamics, and contextual language analysis, integrating phonetic and linguistic aspects to ensure greater accuracy.

REFERENCES

1. Sanches AP, Cazumbá LAF, Telles IFC. Introdução à fonoaudiologia forense. In: Rehder MI, Cazumbá LAF, Cazumbá M, eds. Identificação de falantes: uma introdução à fonoaudiologia forense. Rio de Janeiro: Revinter; 2015.

2. Oghbaie M, Sabaghi A, Hashemifard K, Akbari M. Advances and challenges in deep lip reading. arXiv. 2021:arXiv:2110.07879. http://doi.org/10.48550/arXiv.2110.07879.

3. Nunes EL, Menzen L, Cardoso MCAF. Assessment protocols in orofacial motricity: a systematic review. RSD. 2022;11(14):e25111435896. http://doi.org/10.33448/rsd-v11i14.35896.

4. Feng D, Yang S, Shan S, Chen X. Learn an effective lip reading model without pains. arXiv. 2020:arXiv:2011.07557. http://doi.org/10.48550/arXiv.2011.07557.

5. Bernstein LE, Auer ET, Eberhardt SP. During lipreading training with sentence stimuli, feedback controls learning and generalization to audiovisual speech in noise. Am J Audiol. 2022;31(1):57-77. http://doi.org/10.1044/2021_AJA-21-00034. PMid:34965362.

6. Bernstein LE, Jordan N, Auer ET, Eberhardt SP. Lipreading: a review of its continuing importance for speech recognition with an acquired hearing loss and possibilities for effective training. Am J Audiol. 2022;31(2):453-69. http://doi.org/10.1044/2021_AJA-21-00112. PMid:35316072.

7. Tye-Murray N, Spehar B, Sommers M, Mauzé E, Barcroft J, Grantham H. Teaching children with hearing loss to recognize speech: gains made with computer-based auditory and/or speechreading training. Ear Hear. 2022;43(1):181-91. http://doi.org/10.1097/AUD.0000000000001091. PMid:34225318.

8. Chaves MS, da Rocha EMS S, Castro HC. A surdez e sua diversidade: comparação de materiais para o atendimento visando surdos sinalizantes e oralizados. Revista Arqueiro. 2024;19(2):1-15.

9. Theobald BJ, Harvey R, Cox SJ, Lewis C, Owen GP. Lip-reading enhancement for law enforcement. In: SPIE 6402, Optics and Photonics for Counterterrorism and Crime Fighting II; 2006; Stockholm. Proceedings. Bellingham: Society of Photo-Optical Instrumentation Engineers; 2006. p. 640205. http://doi.org/10.1117/12.689960.

10. Pessoa AF, Vieira RC, Sanches AB, Gonzalez RCS. Admissibilidade de amostras forenses. In: Lopes L, Machado APL, Azoni CAS, Benatti JF, Santos RS, Ribeiro VV, et al, editors. Tratado de fonoaudiologia. São Paulo: Manole; 2024. p. 598–605.

11. Kaur M, Rastogi D, Sharma A, Dahiya A, Nagrath P. Crime investigation using lip reading. In: Third International Symposium on Smart Cities Challenges Technologies (SCCTT); 2024 Nov 29; Delhi. Proceedings. Aachen: CEUR; 2024. p. 103-115.

12. Cascone L, Nappi M, Narducci F. Language identification as improvement for lip-based biometric visual systems. arXiv. 2023:arXiv:2302.13902. http://doi.org/10.1109/ICIP49359.2023.10222415.

13. Peymanfard J, Saeedi V, Mohammadi MR, Zeinali H, Mozayani N. Leveraging visemes for better visual speech representation and lip reading. arXiv. 2023:arXiv:2307.10157. http://doi.org/10.48550/arXiv.2307.10157.

14. Bear HL, Harvey R. Alternative visual units for an optimized phoneme-based lipreading system. Appl Sci. 2019;9(18):3870. http://doi.org/10.3390/app9183870.

15. Thangthai K, Bear HL, Harvey RW. Comparing phonemes and visemes with DNN-based lipreading. arXiv. 2018:arXiv:1805.02924. http://doi.org/10.48550/arXiv.1805.02924.

16. Kim M, Yeo JH, Choi J, Ro YM. Lip reading for low-resource languages by learning and combining general speech knowledge and language-specific knowledge. IEEE/CVF International Conference on Computer Vision (ICCV); 2023 Oct 1-6; Paris. Proceedings. New York: Institute of Electrical and Electronics Engineers; 2023. p. 15313-25. http://doi.org/10.1109/ICCV51070.2023.01409.

17. Gimeno-Gómez D, Martínez-Hinarejos C-D. Analysis of visual features for continuous lipreading in spanish. arXiv. 2023:arXiv:2311.12468. http://doi.org/10.48550/arXiv.2311.12468.

18. Vieira RC, Pereira TR. Analysis of VPAS and simplified VPAS in speaker comparison forensic. Revista de Ciências Jurídicas e Sociais. 2020;5(1):6-23. http://doi.org/10.47595/cjsiurj.v5i1.149.

19. Lalitha SD, Thyagharajan KK. A study on lip localization techniques used for lip reading from a video. arXiv. 2020:arXiv:2009.13420. http://doi.org/10.48550/arXiv.2009.13420.

20. Dörnyei Z. Research methods in applied linguistics. Oxford: Oxford University Press; 2007.

Author contributions

RCV: conceptualization of the study; RCV and LTS: analysis and interpretation of data; RCV and LTS; writing of the article; RCV and LTS: approval of the final version of the article.