

Celso Marcos Rocha de Melo Junior

A concordância da frequência fundamental média na síntese de vogais entre um microfone e um transdutor piezoelétrico:  
uma experiência em laboratório com vogais sintetizadas

Escola Superior de Saúde Fernando Pessoa

Porto, 2025



Celso Marcos Rocha de Melo Junior

A concordância da frequência fundamental média na síntese de vogais entre um microfone e um transdutor piezoelétrico: uma experiência em laboratório com vogais sintetizadas

---

(Assinatura)

Trabalho apresentado à Escola Superior de Saúde Fernando Pessoa, sob orientação do Professor Doutor Pedro Pestana e coorientação da Professora Doutora Rita Alegria, como parte dos requisitos para obtenção do grau de Licenciatura em Terapia da Fala.

## **Resumo**

**Objetivo:** Avaliar a concordância entre um transdutor piezoelétrico e um microfone na análise da frequência fundamental de vogais sintetizadas.

**Métodos:** Foram sintetizadas seis vogais usando três frequências distintas e duas intensidades diferentes. Durante o processo, dois sinais foram gravados simultaneamente: um captado por um microfone posicionado próximo de um altifalante, de forma a simular a posição da cabeça humana, e outro obtido por um transdutor piezoelétrico acoplado ao cone da mesma fonte sonora. Para calibração do sistema foi utilizado um medidor de pressão sonora. A frequência fundamental média foi analisada para cada vogal e um gráfico de Bland-Altman foi gerado para comparar o par microfone-transdutor piezoelétrico.

**Resultados:** A abordagem Bland-Altman permitiu a avaliação da concordância da frequência fundamental média dos dois métodos de captação, o que nos indicou que a ausência de viés sistemático e inexistência de valores discrepantes.

**Conclusão:** Os resultados demonstram que a captação com um microfone e com um transdutor podem ser tecnicamente substituíveis um pelo outro para a análise da frequência fundamental em contexto não-humano, sem necessidades de correções ou ajustes adicionais nos dados recolhidos. Porém, a interpretação clínica destes resultados necessita investigação adicional, uma vez que este estudo foi realizado num contexto de laboratório e com vogais sintetizadas.

**Palavras chave:** Frequência fundamental, transdutor piezoelétrico, microfone, acústica, concordância, análise Bland-Altman.

## **Abstract**

**Aim:** To evaluate the agreement between a piezoelectric transducer and a microphone in the analysis of the fundamental frequency of synthesized vowels.

**Methods:** Six vowels were synthesized using three different frequencies and two different intensities. During the process, two signals were recorded simultaneously: one captured by a microphone positioned close to a loudspeaker in order to simulate the position of the human head, and the other obtained by a piezoelectric transducer coupled to the cone of the same sound source. A sound pressure meter was used to calibrate the system. The average fundamental frequency was analyzed for each vowel and a Bland-Altman graph was generated to compare the piezoelectric transducer-microphone pair.

**Results:** The Bland-Altman approach allowed us to assess the agreement of the mean fundamental frequency of the two pickup methods, which indicated that there was no systematic bias and no discrepant values.

**Conclusion:** The results show that microphone and transducer recording can be technically substituted for each other for fundamental frequency analysis in a non-human context, without the need for additional corrections or adjustments to the data collected. However, the clinical interpretation of these results requires further investigation, since this study was carried out in a laboratory setting and with synthesized vowels.

**Key words:** Fundamental frequency, piezoelectric transducer, microphone, acoustics, concordance.

## **Agradecimentos**

A conclusão deste projeto de graduação representa o fim de uma etapa muito importante da minha vida, e não poderia deixar de expressar a minha gratidão a todos que, de alguma forma, contribuíram para esta conquista.

Em primeiro lugar, agradeço aos meus pais, pelo amor incondicional, pelo apoio constante e por acreditarem em mim, mesmo em momentos que eu duvidei. Sem eles, nada disto teria sido possível.

Aos meus avós, que, mesmo que alguns já não estejam fisicamente entre nós, permanecem vivos no meu coração e na minha memória. Foram e continuarão a ser uma fonte de inspiração, sabedoria e força.

Agradeço, também, a todos os Professores que nos acompanharam ao longo desta caminhada. Cada ensinamento, orientação e incentivo fizeram a diferença na minha formação pessoal e profissional.

Aos amigos e colegas que estiveram ao meu lado durante estes anos, o meu sincero agradecimento. Pelas conversas, pelas brincadeiras, pelo companheirismo e até mesmo pelos desabafos nos momentos difíceis, que tornaram esta caminhada mais leve, mais rica e, acima de tudo, memorável.

E não posso deixar de agradecer ao meu fiel companheiro de quatro patas, o Enzo, o meu cão, que esteve ao meu lado em todos os momentos, dos bons aos mais difíceis, com alegria e uma presença silenciosa, mas profundamente significativa.

A todos, muito obrigado.

## **Índice**

1. Introdução.....	1
2. Metodologia.....	2
3. Resultados.....	3
4. Discussão.....	5
5. Conclusão.....	7
6. Bibliografia.....	8

## **Índice de Gráficos**

**Gráfico1**– Gráfico de *Bland-Altman*: concordância entre dois métodos de análise vocal: o microfone e o transdutor, utilizados para a medição da frequência fundamental média.

## 1. Introdução

O processo de avaliação da voz, tanto ao nível da estrutura quanto ao nível da função, é essencial nas mais variadas áreas, como por exemplo na medicina, na terapia da fala e até no desenvolvimento de tecnologias de comunicação. A avaliação da voz deve ser multimodal incluindo a visualização da estrutura e biomecânica da laringe, a avaliação audioperceptiva da voz, a avaliação aerodinâmica e a análise acústica. Por norma para a análise acústica, o sinal vocal é captado através de um microfone. A análise do sinal é feita através de um programa de *software* no qual várias medidas são calculadas, sendo as mais comuns: a frequência fundamental ( $f_0$ ), o *jitter*, o *shimmer* e proporção harmónico-ruído (PHR) (Felippe et al., 2006). Porém, a fiabilidade da análise captada por microfone é, muitas vezes, comprometida por fatores externos pelos mais variados motivos, como por exemplo a incorreta calibração do sinal ou a impossibilidade de controlar o ambiente acústico durante o processo de gravação, o que pode influenciar as medidas vocais que são utilizadas na pesquisa e avaliação clínica, principalmente se forem microfones de baixa qualidade, ou projetados para o uso não profissional (Parsa et al., 2001). Estas situações levam a resultados inconsistentes e pouco precisos. Além disso, por vezes os clínicos não possuem formação diferenciada na utilização de diferentes microfones e na compreensão das suas características. Para avaliações vocais, o ideal é que o microfone tenha uma resposta de frequências plana, ou seja, que apresente a mesma sensibilidade para todas as frequências, para preservar as características originais da voz do participante (Švec & Granqvist, 2010). Todo o espectro da voz humana deve ser coberto pela resposta de frequência do microfone, desde as frequências mais baixas (10Hz - vozes roucas e crepitantes) até às mais altas (entre 16 000 e 20 000Hz – limites da audição humana) (Švec & Granqvist, 2010).

De forma a evitar resultados comprometidos por fatores externos, uma possível solução poderá ser a utilização de equipamentos de captação de som que não dependam exclusivamente de ondas sonoras conduzidas pelo ar. Os transdutores piezoelétricos, por exemplo, convertem as vibrações mecânicas, como as refletidas pelas pregas vocais na superfície da pele, em sinais elétricos, tendo ainda como vantagens a elevada sensibilidade na deteção de vibrações e uma grande amplitude de frequências, o que é essencial para captar as pequenas variações das vibrações vocais (Javel et al., 2003). Este processo é realizado através de elétrodos ligados ao elemento piezoelétrico, que

A concordância da frequência fundamental média na síntese de vogais entre um microfone e um transdutor piezoelétrico: uma experiência em laboratório com vogais sintetizadas.

transfere a carga elétrica convertida das vibrações mecânicas sob forma de sinal elétrico (Mathew et al., 2021).

Neste âmbito, o presente projeto de graduação utilizou a frequência fundamental média como parâmetro de comparação. A  $f_0$  da voz é definida como o primeiro harmônico e é frequentemente usada para caracterizar a voz humana – esta medida é apresentada em *hertz* (Hz) (Oliveira et al., 2021). Do ponto de vista fisiológico é determinada pelo número de ciclos glóticos que as pregas vocais realizam num segundo, sendo uma consequência natural das características destas estruturas (Oliveira et al., 2021).

O objetivo deste projeto é analisar a concordância entre um microfone e um transdutor piezoelétrico de vogais sintetizadas relativa ao parâmetro  $f_0$ . Poderá contribuir, assim, para a compreensão das potencialidades e limitações dos transdutores piezoelétricos na análise acústica da voz, suscitando futuras aplicações em contexto clínico e de investigação, uma vez que, a utilização de transdutores piezoelétricos pode representar uma opção promissora para a captação de sinais vocais em situações onde o controlo acústico é insuficiente, o que poderá levar à recolha de dados mais fiáveis e consistentes.

## **2. Metodologia**

O estudo consistiu na avaliação do grau de concordância entre o parâmetro acústico  $f_0$  média, obtido através de um microfone convencional e um transdutor piezoelétrico na análise de vogais sintetizadas.

No laboratório de voz da Escola Superior de Saúde Fernando Pessoa foram sintetizadas seis vogais através do *software* Madde, utilizando três frequências distintas e duas intensidades variáveis. Dois sinais foram gravados simultaneamente: um captado por um microfone Shure MX153 T posicionado próximo a um altifalante M-Audio BX5, de forma a simular a posição da cabeça humana, e outro captado por um transdutor piezoelétrico Shadow SH712 colocado diretamente no cone do mesmo altifalante. A gravação foi realizada com o auxílio do *software* LogicPro X e da interface de áudio Behringer FCA1616. O sistema foi calibrado usando o decibelímetro AZ 8922.

A concordância da frequência fundamental média na síntese de vogais entre um microfone e um transdutor piezoelétrico: uma experiência em laboratório com vogais sintetizadas.

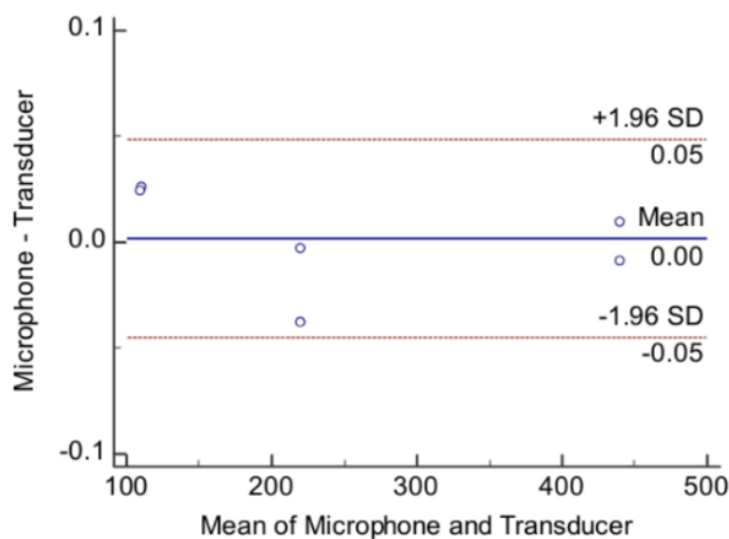
Por fim, a medida acústica de cada amostra de som foi analisada com o recurso ao *software* VOXplot e, posteriormente, foi gerado um gráfico de *Bland-Altman* através do *software* MedCalc para comparar os resultados obtidos pelos dois dispositivos.

A utilização da abordagem *Bland-Altman* deve-se ao facto de este ser um dos métodos de eleição quando a questão de interesse do estudo é a comparação de métodos, o que acontece neste estudo, onde se procura comparar a utilização de um microfone tradicional e um transdutor piezoelétrico (Mansournia et al., 2021).

### 3. Resultados

Com base na análise de um gráfico de *Bland-Altman* (Gráfico 1), foi possível constatar que os resultados deste estudo demonstram uma ausência de viés sistemático entre os dois métodos de captação vocal, visto que revelou uma diferença média (viés) de 0.

**Gráfico1**– Gráfico de *Bland-Altman* sobre a concordância entre dois métodos de análise vocal: o microfone e o transdutor, utilizados para a medição da  $f_0$  média (Hz).



O eixo X do gráfico representa a média aritmética entre os valores obtidos pelo microfone e pelo transdutor para cada gravação individual.

No eixo Y, é expressada a diferença entre os dois métodos (microfone-transdutor), ou seja, a diferença direta de cada par de vogais captado, com uma amplitude de 0.1.

A concordância da frequência fundamental média na síntese de vogais entre um microfone e um transdutor piezoelétrico: uma experiência em laboratório com vogais sintetizadas.

A diferença média observada foi de 0, o que nos indica a ausência de viés. Isto significa que, em média, os valores obtidos pelo microfone não tendem a ser superiores nem inferiores aos obtidos pelo transdutor.

Através do cálculo da média das diferenças  $\pm 1.96$  vezes o desvio padrão foi possível situar os limites de concordância entre -0.05 e +0.05, o que nos permite ver que existe uma excelente concordância entre os métodos de captação, isto porque todos os pontos observados no gráfico encontram-se dentro destes limites, não havendo *outliers*, o que reforça a robustez da concordância observada.

É possível apurar, também, que o pressuposto de homocedasticidade (variação constante) para o cálculo dos parâmetros do gráfico de Bland-Altman está cumprido, uma vez que os pontos de dados estão próximos da linha de polarização (Haghighyegh et al., 2020).

A utilização do gráfico de Bland-Altman permitiu a avaliação de forma abrangente da concordância da frequência fundamental média dos dois métodos de captação com amplitude de 0.1, o que nos indicou que a ausência de viés sistemático e inexistência de valores discrepantes, demonstram que a captação com um microfone e com um transdutor podem ser tecnicamente substituíveis um pelo o outro para a análise de  $f_0$  em contexto não-humano, sem necessidades de correções ou ajustes adicionais nos dados recolhidos.

#### 4. Discussão

Com o gráfico de *Bland-Altman* foi possível avaliar de forma abrangente a concordância da  $f_0$  média entre os dois métodos de medição – o microfone e o transdutor - com uma amplitude de 0.1. A ausência de viés revela que os dois métodos produzem resultados equivalentes, não havendo predomínio de um sobre o outro.

Devido à pequena variação dos dados, a amplitude dos limites de concordância ( $\pm 0.05$ ) é estreita. Associado a isto, a ausência de dados fora dos limites (*outliers*) traduz-se numa variação reduzida entre métodos e com parâmetros estatísticos esperados.

É possível constatar, também, uma variância constante das diferenças ao longo do intervalo dos valores medidos e, com isto, podemos afirmar que a utilização do gráfico de *Bland-Altman* foi adequada, visto que a homogeneidade da variância ao longo da faixa de medição (homocedasticidade) é um princípio fundamental na análise de *Bland-Altman*, o que assegura a reprodução e validade dos resultados em diferentes contextos clínicos (Gerke, 2020).

É de salientar que a utilização de transdutores piezoelétricos na região cervical dá-nos acesso a um conjunto único de dados, que os microfones tradicionais são incapazes de nos fornecer, como é o caso da deteção de atividade traqueal, onde estão incluídas a mastigação, deglutição e tosse (Olubanjo & Ghovanloo, 2014).

Existem, todavia, relatos de uma melhor conversão do sinal mecânico em elétrico pelo transdutor piezoelétrico nos homens, facto que pode ser explicado pela proeminência laríngea, uma vez que esta estrutura atua como um condutor mais eficiente de altas frequências aquando da vibração das pregas vocais (Mathew et al., 2021).

Sobre a fiabilidade dos transdutores piezoelétricos existem poucos estudos e, por isso, foram feitas pesquisas sobre tecnologias idênticas, como é o caso do acelerómetro, que também é um dispositivo de contato aplicado diretamente sobre a pele, e que só capta vibração mecânica. Servindo assim como uma referência teórica útil para sustentar a viabilidade dos transdutores piezoelétricos como alternativa aos microfones tradicionais. Segundo Lien & Stepp (2014), foi demonstrado que acelerómetros são capazes de estimar a frequência fundamental relativa com resultados fiáveis em

comparação com microfones tradicionais, com o benefício adicional de menor suscetibilidade ao ruído ambiente, como acontece no transdutor piezoelétrico.

Outra tecnologia similar são os microfones de contato, como é o caso do sensor FBG-PDMS, estudado por Surapaneni et al. (2024). Com este estudo foi apresentada uma análise de similaridade do espectro do sinal captado pelo sensor e aquilo que é esperado para a fala humana, onde se destacou a preservação da estrutura harmónica, que é fundamental na identificação da  $f_0$ . Neste estudo foram analisadas amostras de fala captadas em condições variadas de ruído e volume e os resultados sugerem que os microfones de contato mantêm melhor a consistência da  $f_0$ , o que nos indica que a  $f_0$  média pode ser mais fiável em transdutores piezoelétricos do que em microfones tradicionais. Importa referir, também, o estudo de Castellana et al. (2015) que demonstrou que microfones de contato, onde se incluem os transdutores piezoelétricos e acelerómetros, preservam melhor a estrutura harmónica da voz, uma vez que evitam a filtragem do trato vocal. Com isto podemos afirmar que os dispositivos, sejam eles transdutores piezoelétricos, acelerómetros ou microfones de contato, que estão em contato com a pele e que captam apenas a vibração são capazes de captar a  $f_0$  da voz com fidelidade de um microfone tradicional, com a vantagem de fazê-lo de forma consistente mesmo em ambientes não controlados e ruidosos.

## 5. Conclusão

Com este estudo podemos concluir que os dois métodos de análise acústica da voz, o microfone tradicional e o transdutor piezoelétrico, podem ser considerados tecnicamente intercambiáveis neste tipo de análise em contexto não-humano para a análise do parâmetro  $f_0$ . Isto é um fato que pode ser particularmente relevante em contextos onde a substituição de um método por outro seja necessária por razões logísticas, económicas ou tecnológicas, sem comprometer a fiabilidade das avaliações acústicas.

Porém, como este estudo foi realizado num contexto controlado e com vogais sintetizadas, a interpretação clínica destes resultados ainda necessita de investigação adicional, nomeadamente em situações reais, com vozes humanas e em ambientes menos controlados.

## 6. Bibliografia

- Castellana, A., Carullo, A., Casassa, F., Astolfi, A., Pavese, L., & Puglisi, G. E. (2015, julho). *Performance comparison of different contact microphones used for voice monitoring*. Proceedings of the 22nd International Congress on Sound and Vibration (ICSV22), Florence, Italy. <https://doi.org/10.13140/RG.2.1.1364.3280>
- Felippe, A. C. N. de, Grillo, M. H. M. M., & Grechi, T. H. (2006). Normatização de medidas acústicas para vozes normais. *Revista Brasileira de Otorrinolaringologia*, 72.
- Gerke, O. (2020). Reporting Standards for a Bland–Altman Agreement Analysis: A Review of Methodological Reviews. *Diagnostics*, 10(5). <https://doi.org/10.3390/diagnostics10050334>
- Haghighyegh, S., Kang, H.-A., Khoshnevis, S., Smolensky, M. H., & Diller, K. R. (2020). A comprehensive guideline for Bland–Altman and intra class correlation calculations to properly compare two methods of measurement and interpret findings. *Physiological Measurement*, 41(5), 055012. <https://doi.org/10.1088/1361-6579/ab86d6>
- Javel, E., Grant, I. L., & Kroll, K. (2003). In Vivo Characterization of Piezoelectric Transducers for Implantable Hearing Aids. *Otology & Neurotology*, 24(5). [https://journals.lww.com/otology-neurotology/fulltext/2003/09000/in\\_vivo\\_characterization\\_of\\_piezoelectric.16.aspx](https://journals.lww.com/otology-neurotology/fulltext/2003/09000/in_vivo_characterization_of_piezoelectric.16.aspx)
- Lien, Y.-A., & Stepp, C. (2014). Comparison of voice relative fundamental frequency estimates derived from an accelerometer signal and low-pass filtered and unprocessed microphone signals. *The Journal of the Acoustical Society of America*, 135, 2977. <https://doi.org/10.1121/1.4870488>

- Mansournia, M. A., Waters, R., Nazemipour, M., Bland, M., & Altman, D. G. (2021). Bland-Altman methods for comparing methods of measurement and response to criticisms. *Global Epidemiology*, 3, 100045. <https://doi.org/10.1016/j.gloepi.2020.100045>
- Mathew, L. R., Priya, G., & Gopakumar, K. (2021). Piezoelectric Throat Microphone Based Voice Analysis. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1, 1603–1608. <https://doi.org/10.1109/ICACCS51430.2021.9441880>
- Oliveira, R. C., Gama, A. C. C., & Magalhães, M. D. C. (2021). Fundamental Voice Frequency: Acoustic, Electroglottographic, and Accelerometer Measurement in Individuals With and Without Vocal Alteration. *Journal of Voice*, 35(2), 174–180. <https://doi.org/10.1016/j.jvoice.2019.08.004>
- Olubanjo, T. & Ghovanloo, M. (2014). Tracheal activity recognition based on acoustic signals. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1436–1439. <https://doi.org/10.1109/EMBC.2014.6943870>
- Parsa, V., Jamieson, D. G., & Pretty, B. R. (2001). Effects of Microphone Type on Acoustic Measures of Voice. *Journal of Voice*, 15(3), 331–343. [https://doi.org/10.1016/S0892-1997\(01\)00035-2](https://doi.org/10.1016/S0892-1997(01)00035-2)
- Surapaneni, H., Chalackal, A. T., & Kanakambaran, S. (2024). Spectral Similarity Estimation and ML-Based Classification of Speech Using PDMS-Embedded FBG Sensor-Based Contact Microphone. *IEEE Sensors Journal*, 24(21), 34487–34497. <https://doi.org/10.1109/JSEN.2024.3454462>

Švec, J. G. & Granqvist, S. (2010). Guidelines for Selecting Microphones for Human Voice Production Research. *American Journal of Speech-Language Pathology*, 19(4), 356–368. [https://doi.org/10.1044/1058-0360\(2010/09-0091\)](https://doi.org/10.1044/1058-0360(2010/09-0091))