

Prosódia da paródia

Um piloto sobre imitação humorística

Gabriel Catani¹, Leticia Kaori Hanada²
Instituto de Estudos da Linguagem
Universidade Estadual de Campinas
Campinas, SP, Brasil
sci@catani.me, leticiahanada@hotmail.com

Resumo: Observamos aqui a manipulação de parâmetros acústicos por parte de um imitador profissional ao parodiar a fala do atual Presidente da República do Brasil. Realizamos uma comparação entre 3 amostras de fala (em português brasileiro) retiradas da internet, com cerca de 100 segundos cada: uma do comediante Marcelo Adnet em uma entrevista televisiva, uma com a fala do político Jair Bolsonaro em um *streaming* na internet e, por fim, uma *sketch* de Adnet imitando Bolsonaro. Os resultados apontam para um alto grau de sucesso da imitação especialmente em medidas associadas à frequência fundamental: o imitador distancia-se da sua fala habitual e aproxima-se significativamente da fala alvo.

Prosódia; paródia; imitação; Marcelo Adnet; Jair Bolsonaro.

I. INTRODUÇÃO

O termo "imitação" é polissêmico: de maneira ampla, podemos distinguir a mera imitação vocal, ou gestual, da imitação personificadora ou "impersonamento" (do inglês, *impersonate*), na qual o imitador não só emula aspectos verbais, mas também características extralinguísticas individualizantes do sujeito imitado, como vestuário, expressões faciais e afins. A imitação pode ser compreendida como um processo mais ou menos consciente: ainda que haja a observação e tentativa de replicação da voz e/ou da imagem do alvo, o imitador não necessariamente tem plena consciência ou controle sobre os recursos mobilizados para tal. O processo de imitação, portanto, depende de uma espécie de intuição por parte do imitador, de modo que certas tarefas podem até mesmo evocar associações entre certos traços linguísticos e estereótipos sociais subjacentes: em pesquisa prévia, por exemplo, já se observou uma imitadora, falante de sotaque britânico, adotar um sotaque americano quando o experimento pedia simplesmente para que ela imitasse uma mulher com fala anasalada [1].

Quanto aos objetivos da imitação, visa-se retratar um sujeito-alvo de maneira parcial, total, ou caricata. Nesse processo, pode-se buscar representar o sujeito ou suas características com fins meramente ilustrativos, cômicos, dramáticos, ou até mesmo tentando enganar o ouvinte quanto à sua identidade. O estudo da imitação, portanto, mostra-se relevante para diversas áreas, desde as artes até as ciências forenses.

Na fonética, há alguns estudos sobre o tema, principalmente em sueco [11, 23, 29-33, *inter alia*], mas também em inglês [1, 12], francês [9], japonês [15] e alemão [10]. Em geral, utilizam-se amostras de fala breves em sentenças idênticas proferidas pelo imitado e pelo imitador. Há certa divergência quanto às conclusões, mas é certo que na imitação há uma notável alteração nos parâmetros acústicos correspondentes aos aspectos suprasegmentais da fala na voz do imitador.

¹ Gabriel Catani é Bolsista CAPES/BRASIL (CAPES/PROCAD-SPCF, 88887.608060/2021-00).

² Processo nº 2020/13522-9, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Neste trabalho, portanto, analisamos acusticamente e auditivamente a fala do premiado imitador brasileiro Marcelo Adnet [24] ao realizar uma imitação jocosa do atual presidente do Brasil, Jair Bolsonaro. O objetivo foi analisar com qual precisão os parâmetros acústicos foram manipulados pelo comediante em sua imitação, utilizando como comparação a fala habitual gravada tanto de Adnet quanto de Bolsonaro. Dessa forma, buscamos explorar a plasticidade da fala e compreender como o falante pode utilizá-la para representar outrem, através de similaridades acústicas e associações semânticas. Apresentamos aqui resultados relacionados a medidas de frequência fundamental (F0) e de qualidade vocal dessas imitações no português brasileiro, língua na qual o fenômeno foi pouco explorado – a saber, apenas em [19].

II. CORPUS

As amostras de fala foram retiradas de vídeos publicados no *Youtube* e convertidas para *wav* através do *software* *ffmpeg* [25]. Julgando a realização dessa pesquisa não danosa aos proprietários dos material analisado, sem interesses financeiros, tampouco de apropriação ou deturpação do material, nos fundamos na premissa de uso justo, agindo em conformidade com a Lei Geral de Proteção de Dados Pessoais, que no art. 7, IV, outorga o uso de dados pessoais para fins de pesquisa [34] e na Lei de Direitos Autorais, art. 46, III, que permite a citação de trechos de qualquer obra para fins de estudo [35]. Diferentemente de pesquisas prévias, optamos por trabalhar apenas com dados previamente gravados e de fala pública. Apesar disso sabidamente implicar em uma menor comparabilidade entre as amostras, considerando a diferença no conteúdo das gravações, consideramos que há ganhos na validade ecológica dos dados: todas as amostras de fala foram retiradas de seu contexto habitual, ao invés de serem eliciadas pelos pesquisadores. Além disso, o desenvolvimento de comparações entre amostras com conteúdo discrepante pode ser interessante para propósitos forenses, em que a autoria de áudios interceptados poderia ser questionada, contrapondo fala usual e possíveis contextos de imitações, por exemplo. Uma melhor compreensão de parâmetros acústicos associados a esse tipo de imitação também pode contribuir para o aumento da segurança relacionada a tecnologias de reconhecimento de fala, especialmente em sujeitos com alta exposição, cujo conteúdo publicado pode servir como base de treino para imitações fraudulentas, potencialmente aliadas a vídeos manipulados usando técnicas de *deepfake* [27].

O material utilizado foi:

- (1) Fala habitual de Adnet em entrevista ao programa de TV Roda Viva [22]
- (2) Fala habitual de Bolsonaro em uma transmissão ao vivo pela internet [20];
- (3) Adnet imitando Bolsonaro em uma *sketch* humorística [26].

A amostra (3) consiste na íntegra de uma *sketch* de imitação publicada pelo próprio Adnet, com 105 segundos. As amostras (1) e (2) consistem em trechos de 108 segundos, provenientes de vídeos com mais de uma hora de duração. Buscamos selecionar trechos de fala ininterruptos com uma duração próxima da amostra da imitação. Evitamos trechos com ruídos e interrupções e momentos marcados por fortes emoções, com o intuito de mitigar diferenças entre as amostras. Apesar de potenciais diferenças estilísticas na voz dos falantes, decorrentes de diferentes contextos de enunciação, também evitamos selecionar trechos muito marcados pelo gênero, como comentários sobre a situação de entrevista, sobre estar falando publicamente etc. Consideramos que, apesar da situação de fala pública, como os trechos selecionados foram retirados do meio de transmissões longas, há menos atenção à fala que na amostra de imitação [17]. Apesar disso, não consideramos que as amostras correspondam à fala vernácula dos sujeitos. O emprego dos termos "habitual", "natural" e "normal" são utilizados aqui meramente em oposição a "imitação".

A. Perfis analisados

Conforme Abercrombie (1967), existem algumas restrições anatomofisiológicas que impossibilitam o que poderia ser considerada uma imitação "perfeita" (uma replicação total) [3]. Alguns dos recursos necessários para tal feito estariam fora do controle do falante e até mesmo um bom imitador tem limites em sua habilidade de alterar sua configuração articulatória [30]. No que diz respeito à idade, no ano da análise, Adnet tem 39 e Bolsonaro 66 anos (cerca de 27 anos de diferença), essa diferença se faz relevante, portanto, uma vez que o trato vocal tende a se modificar com a idade: a saber, o processo de envelhecimento da voz, ou presbifonia, se caracteriza por um aumento nos valores de F0, *jitter* e *shimmer* para o sexo masculino [6]. Além disso, especula-se que o político seja acometido por algum nível de anquiloglossia [14], condição popularmente conhecida como "língua presa", que associa-se a problemas de articulação, restringindo o movimento da língua e afetando a realização de sons sibilantes [13].

Além de questões de ordem fisiológica, há também diferenças de âmbito social entre eles. Enquanto Adnet é nativo e residente do Rio de Janeiro, RJ, exibindo um comportamento linguístico associado à fala na região, Bolsonaro nasceu na cidade de Glicério, no noroeste do estado de São Paulo, tendo residido por muitos anos no estado do Rio de Janeiro e, recentemente, passou a ocupar a residência presidencial, em Brasília. Seu sotaque poderia ser descrito como um misto de traços associados à fala paulista e carioca. Quanto à formação, Adnet é formado em jornalismo, enquanto Bolsonaro seguiu carreira militar.

B. O processo de imitação, segundo Adnet

Uma imitação bem-sucedida está associada a um bom domínio da entoação, do ritmo e da qualidade de voz [30], mas também a uma boa observação por parte do imitador. Ao comentar sobre o seu processo, em um trecho da entrevista (que não faz parte da amostra), Adnet atribui certa importância para o reconhecimento de traços físicos e linguísticos característicos do imitado, e aponta que sua imitação ocorre de maneira "instintiva", a seguir:

“Esse processo [da imitação] é muito instintivo cara, é uma baixada de espírito mesmo que dá, eu dou as credenciais para o meu ouvido, nem tanto para a corda vocal, porque o ouvido ele tá lá e ele que trabalha para sacar as nuances de cada voz, se a voz é grave, se a voz é mais aguda, a boca da pessoa influencia bastante também [...] as pessoas tem vício de linguagem muitas vezes, é uma coisa muito abrangente e você vai sacando, eu saco primeiro isso, a coisa do som, como é que a pessoa fala, o formato do rosto e da boca influencia também e depois os vícios de linguagem da pessoa, qual é a palavra que ela escolhe, qual é a linha de raciocínio dela [...]”.

III. METODOLOGIA

Com as amostras selecionadas, foram segmentados *chunks* com enunciados entoacionais, usando o *software* Praat [7] e seguindo o exemplo disponibilizado na documentação do *script* utilizado para as aferições, o *Prosody Descriptor Extractor* [5]. Independentemente, também foram realizadas transcrições ortográficas do material utilizando o *software* ELAN [28], seguindo o padrão estabelecido no projeto SP2010 [18]. Utilizamos tais transcrições como base para o alinhamento fonético forçado usando o Alinha-PB [16], ferramenta desenvolvida para facilitar tal alinhamento,

oferecendo modelos já treinados em Português Brasileiro. O alinhamento foi revisado e corrigido manualmente pelos pesquisadores.

Posteriormente, foi aplicado o *script* sobre os áudios e TextGrids, extraíndo as medidas associadas a F0 e qualidade vocal. Dos arquivos alinhados automaticamente, foram também coletadas medidas associadas a sílabas VV e grupos acentuais. Os dados resultantes do alinhamento automático, no entanto, ainda não foram plenamente analisados e, portanto, não serão aqui abordados.

A fase seguinte consistiu no processamento dos dados extraídos e na análise estatística. Para tanto, foi utilizada a linguagem de programação R [21], e realizadas ANOVA (*Analysis of Variance*) de um fator, quando atendidos os devidos pré-requisitos (normalidade, homocedasticidade e independência) e testes de Kruskal-Wallis nos casos de não normalidade dos dados [2].

IV. RESULTADOS

A. ANÁLISE QUANTITATIVA

Por conta da extensão, apresentamos na Tabela 1 apenas dados onde se observaram diferenças significativas. Mais detalhes sobre os parâmetros de extração estão disponíveis no manual do *script* [4].

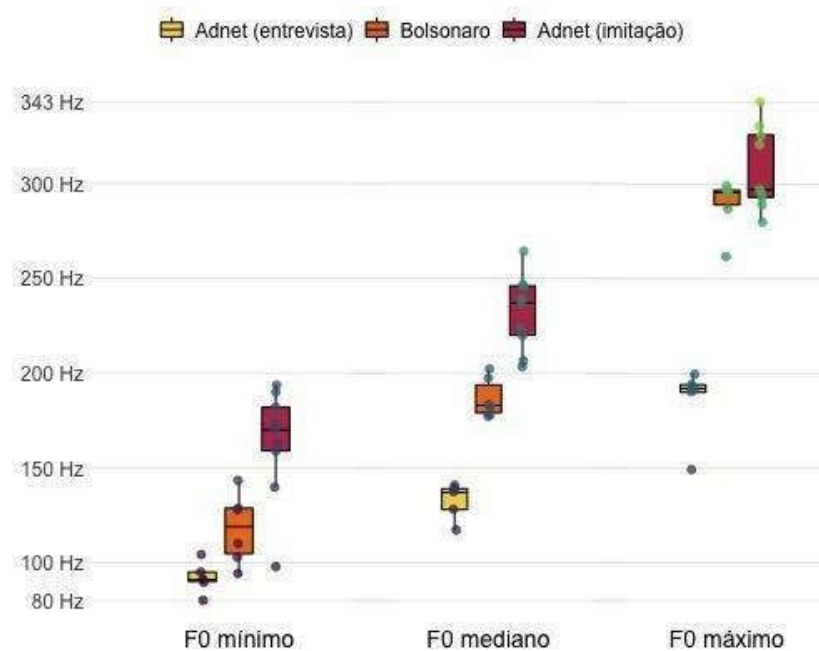


Figura 1. Medidas globais de F0 na fala de Adnet (entrevista e imitação) e Bolsonaro

O valor mínimo de F0 ($F0_{min}$) da fala natural de Adnet não teve diferença significativa em relação ao mínimo de Bolsonaro. No entanto, durante as imitações, o mínimo de Adnet é significativamente mais alto que o de Bolsonaro e cerca de 70 Hz mais alto do que em sua fala natural. Já o valor máximo de F0 ($F0_{max}$) da fala natural de Adnet é significativamente menor que o de Bolsonaro. Na imitação também se aproxima do máximo do presidente, sem diferença significativa entre imitação e imitado.

| Parâmetros ($\alpha = 0.05$) | Adnet entrevista (AE) | Bolsonaro (BO) | Adnet imitação (AI) | Diferenças ($\alpha = 0.05$) |
|--|--------------------------|-------------------|------------------------|---|
| F0med (Hz) (F = 65.4, p < 0.001) | 132 | 187 | 231 | AE-AI (-99.3, p < 0.001) BO-AE (54.2, p < 0.001) BO-AI (-66.3, p < 0.001) |
| F0sd (Hz) (H = 11.2, p < 0.003) | 22 | 39 | 33 | AE-AI (-11, p = 0.021) BO-AE (17, p = 0.013) |
| F0min (Hz) (F = 17, p < 0.001) | 92 | 118 | 163 | AE-AI (-71.3, p < 0.001) BO-AI (-45.5, p = 0.004) |
| F0max (Hz) (F = 67.1, p < 0.001) | 185 | 289 | 308 | AE-AI (-123, p < 0.001) BO-AE (104.3, p < 0.001) |
| F0peakwidth (Hz) (H = 6, p < 0.03) | 7.5 | 6.8 | 10.3 | BO-AI (-3.5, p = 0.024) |
| F0peakrate (pico/seg) (F = 11.3, p < 0.001) | 1.2 | 1.3 | 1.6 | AE-AI (-0.37, p = 0.003) BO-AI (-0.35, p = 0.002) |
| Df0posmean (Hz/frame) (F = 29.3, p < 0.001) | 2.9 | 5.9 | 5.8 | AE-AI (-2.94, p < 0.001) BO-AE (3.01, p < 0.001) |
| Df0negmean (Hz/frame) (F = 28, p < 0.001) | -2.6 | -6.3 | -6.4 | AE-AI (3.80, p < 0.001) BO-AE (-3.67, p < 0.001) |
| Df0sdpos (Hz/frame) (F = 8.2, p < 0.001) | 2.8 | 4.9 | 5.0 | AE-AI (-2.27, p = 0.003) BO-AE (2.13, p = 0.011) |
| Df0sdneg (Hz/frame) (F = 15.9, p < 0.001) | 2.6 | 5.5 | 5.6 | AE-AI (-2.96, p < 0.001) BO-AE (2.92, p < 0.001) |
| Emph (dB) (F = 55.5, p < 0.001) | 5.32 | 3.9 | 8.8 | AE-AI (-3.50, p < 0.001) BO-AI (-4.93, p < 0.001) |
| Cvint (F = 17.3, p < 0.001) | 12.4 | 14.3 | 15.9 | AE-AI (-3.48, p < 0.001) BO-AE (1.93, p = 0.021) BO-AI (-1.55, p = 0.034) |
| sLTASmed (H = 11.7, p = 0.002) | -14.2 | -19.3 | -18.9 | AE-AI (4.7, p = 0.009) BO-AE (-5.1, p = 0.022) |
| sLTAShigh (F = 56, p < 0.001) | -25.6 | -34.3 | -41.0 | AE-AI (15.39, p < 0.001) BO-AE (-8.68, p < 0.001) BO-AI (6.71, p < 0.001) |
| Hnr (F = 8.8, p = 0.002) | 8.6 | 7.7 | 9.3 | BO-AI (-1.64, p < 0.001) |
| Shimmer (%) (F = 19, p < 0.001) | 11.4 | 15.0 | 13.7 | AE-AI (-2.34, p = 0.001) BO-AE (3.65, p < 0.001) |
| Jitter (%) (F = 31, p < 0.001) | 2.0 | 3.6 | 2.8 | AE-AI (-0.79, p < 0.001) BO-AE (1.52, p < 0.001) BO-AI (0.73, p = 0.001) |

Tabela 1: Resultados da comparação de parâmetros acústicos entre Adnet (entrevista e imitação) e Bolsonaro

A taxa de picos de F0 por segundo (*F0peakrate*) na fala natural de Adnet é similar à de Bolsonaro. Durante a imitação, porém, a taxa de Adnet é significativamente maior que a do falante-alvo. O mesmo padrão se observa na

largura dos picos de F0 ($F0_{peakwidth}$): a fala natural de Adnet tem um valor similar ao de Bolsonaro, e na imitação há um aumento significativo da largura em comparação a Bolsonaro.

Nos parâmetros associados às derivadas de F0, que representam a taxa de mudança da curva entoacional, Adnet, ao imitar, deixa de se diferenciar significativamente da fala de Bolsonaro tanto nas médias das derivadas positivas ($df0_{posmean}$) quanto nas negativas ($df0_{negmean}$), tendo subidas e descidas de F0 mais bruscas que seu habitual, se aproximando da fala de Bolsonaro. O mesmo acontece com os valores do desvio padrão das derivadas positivas ($df0_{sdpos}$) e negativas de F0 ($df0_{sdneg}$), ou seja, as suas subidas e descidas de F0 tornam-se mais variáveis que na sua fala normal, também se aproximando da fala de Bolsonaro.

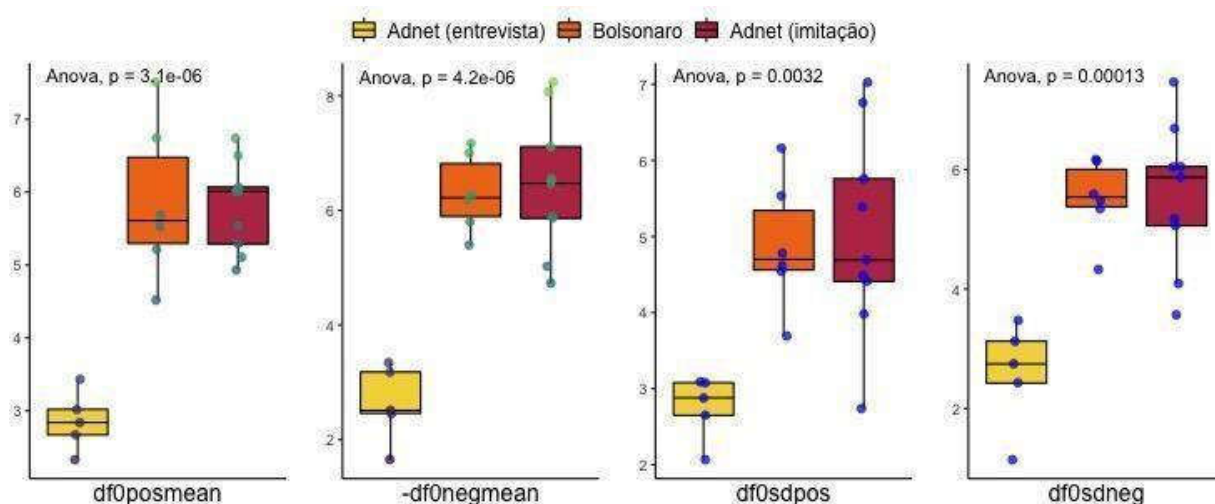


Figura 2. Medidas de derivadas de F0 na fala de Adnet (entrevista e imitação) e Bolsonaro

Ao imitar o Bolsonaro, a fala de Adnet tem ênfase espectral ($Emph$) significativamente maior que sua fala normal e que a fala de Bolsonaro. Esse resultado, no entanto, pode estar relacionado à duração das gravações originais: podemos considerar que a ênfase tenderia a ser menor em um trecho mais longo de fala, por conta do cansaço vocal e como estratégia de autopreservação. O coeficiente de variação de intensidade (C_{vint}) de Adnet na entrevista é menor que o de Bolsonaro, mas também ultrapassa o alvo durante a imitação. Quanto ao esforço vocal nas frequências médias, entre 0-1000 Hz e 1000-4000 Hz, ($sLTASS_{med}$), a fala habitual de Adnet apresenta menor esforço que a fala de Bolsonaro, e em contexto de imitação seu esforço é semelhante ao do político. Nas frequências altas, entre 0-1000 Hz e 4000-8000 Hz ($sLTASS_{high}$), o esforço de Adnet, normalmente menor que o de Bolsonaro, ultrapassa o alvo na imitação.

A razão de harmônico/ruído (hnr) na fala natural de Adnet é similar à da fala de Bolsonaro. Na imitação, porém, a fala de Adnet torna-se menos soprosa, se distanciando significativamente de Bolsonaro. Em sua fala habitual, Adnet tem 2% de $jitter$, subindo para 2.8% nas imitações, mas ainda significativamente menor que o de Bolsonaro (3.6%). Há também aumento do $shimmer$ de Adnet, indo de 11.4% para 13.7% na imitação, deixando se diferir da voz alvo (15%).

B. OBSERVAÇÕES QUALITATIVAS

Durante o processo de transcrição dos dados notamos algumas modificações com impacto potencial na tarefa de evocar fala e comportamento do imitado. A principal delas percebida foi o uso de certos itens lexicais atípicos na sua fala habitual, associados à fala de Bolsonaro, como "aí" ("isso daí", "olha aí ó"), "pô" ("faz o três pô"), o uso de perguntas-marcadores de modo mais frequente ("tá certo?", "okay?") etc. Em um breve levantamento desses usos, observamos que a frequência dessas perguntas (n = 6) na imitação se assemelha à fala do imitado, sendo ausentes na entrevista. Pela brevidade das ocorrências e o tamanho da amostra, julgamos que sua presença não impactou significativamente os resultados discutidos. Também observamos que a escolha das temáticas abordadas na imitação visa relacionar a personagem a um cenário associado, referindo-se a políticos e assuntos associados.

V. CONCLUSÃO

Consideramos que a imitação feita por Adnet teve êxito na replicação da maioria dos parâmetros observados. Dos parâmetros acústicos mobilizados de maneira significativa, a maior parte pode ser descrita como certa, ou seja, sem diferença significativa entre a imitação e a fala de Bolsonaro. Alguns outros parâmetros são modificados de maneira mais caricata, ultrapassando os valores correspondentes a Bolsonaro. Tais exageros, no entanto, não necessariamente consistem em uma "falha" na performance do imitador. Ao contrário: podem ser vistos como um aspecto constituinte da imitação humorística, em que o objetivo não é replicar fielmente a fala do alvo, mas produzir efeito cômico e/ou crítico em relação ao imitado. Assim, ao se distanciar de sua fala habitual em direção à fala de Bolsonaro de maneira tão drástica, a ponto de "ultrapassar" o alvo, o imitador demonstra grande flexibilidade linguística e um potencial para realizar imitações ainda mais distantes de sua fala habitual, satirizando aspectos salientes da voz imitada.

Uma dificuldade associada ao estudo do tipo de imitação aqui avaliado é o papel do conteúdo no reconhecimento da imitação como fala autêntica do alvo ou como imitação da fala do alvo. Além dos aspectos fonéticos mobilizados pelo imitador, observamos o emprego de estratégias semânticas sabidamente eficazes para veicular a imagem do imitado, reforçando elementos simbólicos que remetem ao sujeito-alvo [33]. No entanto, o conteúdo desse tipo de imitação talvez seja uma das características mais marcantes do gênero, de modo que a deslexicalização pode distorcer o valor das amostras. Nessa direção, julgamos que a comparação entre imitações humorísticas e imitações com fins de cópia plena pode revelar importantes especificidades de cada gênero a serem consideradas no desenvolvimento de estudos na área. Além disso, consideramos que testes de percepção associados ao estudo de imitações de qualquer espécie são fundamentais para garantir a robustez dos resultados, já que, as imitações são pautadas e até mesmo definidas por critérios perceptivos definidos socialmente e potencialmente variáveis.

REFERÊNCIAS

- [1] Amin T.B, Marziliano P, German JS. Nine voices, one artist: Linguistic and acoustic analysis. In: *IEEE International Conference on Multimedia and Expo*. 2012, pp. 450-454.
- [2] Arantes, P, Barbosa PA. *ANOVA. Análise de Variância*. Não publicado..
- [3] Abercrombie, D. Elements of general phonetics. Edinburgh: Edinburgh University Press. 1967.
- [4] Barbosa, P.A. Manual of Praat Script Prosody Descriptor Manual. 2020.
- [5] Barbosa, P.A. Prosody Descriptor Extractor. Script em R. 2014. Disponível em: <<https://github.com/pabarbosa/prosody-scripts/tree/master/ProsodyDescriptorExtractor>>
- [6] Benito P., A. B. A voz do idoso: características clínicas, endoscópicas, vocais e morfológicas. Tese de Doutorado. Faculdade de Medicina de Botucatu, UNESP. São Paulo. 2015.
- [7] Boersma, P., Weenink, D. Praat: doing phonetics by computer [Computer program]. 2021. Version 6.1.50, retrieved 20 June 2021 from <http://www.praat.org/>

- [8] Cole, J., Shattuck-Hufnagel, S. The phonology and phonetics of perceived prosody: What do listeners imitate?. In: *Twelfth Annual Conference of the International Speech Communication Association*. 2011.
- [9] Delvaux V et al. Voice Disguise vs. Impersonation: Acoustic and Perceptual Measurements of Vocal Flexibility in Non Experts. In: *INTERSPEECH*. 2017, pp. 3777-3781.
- [10] Endres W, Bambach W, Flösser G. Voice spectrograms as a function of age, voice disguise, and voice imitation. In: *The Journal of the Acoustical Society of America*. 1971, 49(6B):1842-8.
- [11] Eriksson A, Wretling P. How flexible is the human voice?-a case study of mimicry. In: *Fifth European Conference on Speech Communication and Technology*. 1997.
- [12] Evans BE. An acoustic and perceptual analysis of imitation. *Handbook of perceptual dialectology*. 2002, 2:95-112.
- [13] Flint PW et al. Cummings otolaryngology-head and neck surgery. In: *7th Edition. Elsevier Health Sciences*. 2020.
- [14] Galdea JG. República do trava-línguas: entenda os problemas de fala de Bolsonaro e Lula. *Correio 24 horas*. 2018. Disponível em: <<https://www.correio24horas.com.br/noticia/nid/republica-do-trava-linguas-entenda-os-problemas-de-fala-de-bolsonaro-e-lula/>>
- [15] Kitamura T. Acoustic analysis of imitated voice produced by a professional impersonator. In: *Ninth Annual Conference of the International Speech Communication Association*. 2008.
- [16] Kruse JS, Barbosa PA. Alinha-PB. In: *Journal of Communication and Information Systems*. 2021, 36(1):192-9.
- [17] Labov W. Sociolinguistic patterns. University of Pennsylvania press. 1972.
- [18] Mendes, R. B., Oushiro, L. Documentação do Projeto SP2010: Construção de uma amostra da fala paulistana. 2013.
- [19] Passetti RR, Barbosa PA. Análise fonético-acústica de disfarces vocais com aplicações para a fonética forense. In: *Cadernos de Estudos Linguísticos*. 2018, 60(3):567-83.
- [20] Pingos nos Is. Integra da live de Jair Bolsonaro de 10/06/2021. Disponível em <<https://www.youtube.com/watch?v=BHotKAFglYE>>.
- [21] R Core Team. R: A Language and Environment for Statistical Computing. *Sign Language & Linguistics*, Viena. 2013. Disponível em: <<https://benjamins.com/catalog/sll>>. Acesso em: março de 2021.
- [22] Roda Viva. 2020. Marcelo Adnet 17/08/2020. Disponível em: <<https://www.youtube.com/watch?v=hcONxsW75Ys>>
- [23] Sullivan KP, Schlichting F. A perceptual and acoustic study of the imitated voice. *Journal of the Acoustical Society of America*. 1998, 103:2894.
- [24] Telaviva. APCA entrega Troféu aos melhores de 2018 nas artes. 2019. Disponível em: <<https://telaviva.com.br/14/06/2019/apca-entrega-trofeu-aos-melhores-de-2018-nas-artes/>>
- [25] The FFmpeg developers. ffmpeg 4.4 (software). 2021. Disponível em: <<https://ffmpeg.org/>>
- [26] Vozes da democracia - Senador Humberto. Marcelo Adnet ironiza a maquiagem dos números de Bolsonaro. 2020. Disponível em: <https://www.youtube.com/watch?v=AQ8dPa_8vdc>
- [27] Westerlund M. The emergence of deepfake technology: A review. *Technology Innovation Management Review*. 2019, 9(11).
- [28] Wittenburg P et al. ELAN: A professional framework for multimodality research. In: *5th International Conference on Language Resources and Evaluation (LREC 2006)*. 2006, pp. 1556-1559..
- [29] Zetterholm E. A phonetic case study of impersonation. Dept. of Phonetics, Umeå Univ.. In: *Phonum* 1997, pp. 33-36.
- [30] Zetterholm, E. Impersonation: a phonetic case study of the imitation of a voice. Working Papers, Lund University, Dept. of Linguistics; Vol. 46. 1997.
- [31] Zetterholm, E. Impersonation - reproduction of speech. Working Papers, Lund University, Dept. of Linguistics; Vol. 49. 2001
- [32] Zetterholm E. Same speaker–different voices. A study of one impersonator and some of his different imitations. In: *Proceedings of the 11th Australian International Conference on Speech Science & Technology*. 2006, pp. 70-75.
- [33] Zetterholm E, Sullivan KP, van Doorn J. The impact of semantic expectation on the acceptance of a voice imitation. *Proceedings of SST2002*. Melbourne, Australia. pp. 2002: 379-84.
- [34] Brasil, Lei nº 13.709/2018, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm>.
- [35] Brasil, Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. 1998. Disponível em: <http://www.planalto.gov.br/ccivil_03/Leis/L9610.htm>