

De quem é a voz? Narração de audiolivros em tempos de inteligência artificial¹

Eliane Hatherly Paz²
Universidade Federal do Rio de Janeiro

RESUMO

O mercado de audiolivros ainda engatinha no Brasil e uma questão já se impõe: a adoção da inteligência artificial (IA) na narração do texto impresso. Ao substituir a voz humana pela voz sintética, as obras narradas por IA podem acabar se tornando um padrão? Neste trabalho, refletimos sobre a aplicação dessa ferramenta no mercado editorial de audiolivros e como ela afeta a fruição das obras literárias pelos leitores-ouvintes. Partindo da análise do fenômeno à luz do conceito de *oralização* de Bajard (2014), concluímos que a IA cria outras, e diferentes, formas de fruição, ampliando o acesso a obras literárias. Porém, ainda sem a conectividade leitor-obra causada pela narração humana.

PALAVRAS-CHAVE

Inteligência Artificial (IA); Oralização; Audiolivros digitais; Narração de audiobooks; Mercado editorial.

Introdução

O mercado de audiolivros engatinha no Brasil e uma questão já se impõe para os profissionais do setor editorial: a adoção da inteligência artificial (IA) na narração do texto impresso. Alguns *players* começam a adotar a ferramenta, com a perspectiva de que seu baixo custo amplie a publicação de títulos em domínio público e de fundo de catálogo, aumentando a circulação de obras literárias e democratizando o acesso a elas. Já editoras com insuficiente capital para custear a produção de audiolivros veem o recurso como meio de entrada a um mercado com perspectivas otimistas de crescimento.

Por outro lado, a integração da IA em processos editoriais de audiolivros já começa a impactar a prática de narração nesse formato, e circulam no mercado milhares de audiolivros gerados a partir de bancos de dados vocais. Diante dessa realidade,

¹ Trabalho apresentado no GP Produção Editorial, XIX Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 47º Congresso Brasileiro de Ciências da Comunicação.

² Profa. Adjunta A do curso de Produção Editorial da ECO-UFRJ. E-mail: eliane.paz@eco.ufrj.br

devemos indagar: ao substituir a voz humana por vozes artificiais, os títulos narrados por IA podem acabar se tornando um padrão no mercado editorial? Neste trabalho, nosso objetivo é refletir brevemente sobre como essa ferramenta afeta a fruição, pelos leitores-ouvintes, de obras literárias no formato sonoro digital³.

Não cabe aqui, diante das limitações do formato deste artigo, historiografar as diferentes mídias que já serviram de suporte aos audiolivros ao longo de mais de um século. Neste primeiro momento é importante, pois, mapear alguns dos usos da IA pelos maiores *players* do setor. Em números reais, uma recente matéria⁴ divulgou que a Audible – plataforma de audiolivros da norte-americana Amazon, há oito meses ao país –, já possui mais de 40 mil títulos narrados em inglês por inteligência artificial disponíveis para seus assinantes brasileiros. Retrospectivamente, em dezembro de 2022, a também norte-americana Google lançou por aqui sua ferramenta de narração automática de audiolivros. Baseada em inteligência artificial, na fase beta ela só convertia títulos de *backlist* de não ficção. No ano seguinte, em janeiro, sua conterrânea Apple lançou um catálogo (ainda indisponível no Brasil) de audiolivros de romance e ficção, em inglês, narrados sinteticamente a partir de arquivos de e-books.

Já em setembro de 2023, o Projeto Gutenberg⁵ – uma plataforma digital colaborativa de livros eletrônicos gratuitos baseada igualmente nos EUA – disponibilizou, em parceria com a Microsoft e o MIT, um catálogo com mais de 5 mil títulos em domínio público narrados por vozes artificiais. Por fim, na mais recente Feira de Londres⁶, em painel que discutiu o mercado do formato em áudio, participantes declararam que a utilização da IA – tema presente em quatro dos nove palcos principais do evento – na narração de livros de grandes editoras inglesas já é total.

Do lado de cá, a *startup* brasileira de livros digitais Skeelo lançou seu primeiro audiolivro narrado por IA em maio de 2023: o best-seller *A lei da atração: ‘O Segredo’*, de Rhonda Byrne, colocado em prática, de Michael J. Losier, impresso pela editora LeYa. Segundo André Palme, CMCO da Skeelo, “com a ajuda da IA, damos início a um

³ O conceito de audiolivro com o qual trabalhamos aqui é o de narração, por uma ou mais vozes, gravada de obra geral já impressa em livro.

⁴ Disponível em: <https://www.terra.com.br/byte/40-mil-audiolivros-narrados-por-ia-dominam-o-ranking-do-audible,32ba8e14caf91ef50bbdef7856d592deaow6igr.html>. Acesso em 24 jun.2024.

⁵ Disponível em: <https://medium.com/@cybernodes/5-000-audiolivros-do-project-gutenberg-est%C3%A3o-na-blockchain-da-likecoin-2127196c2a89>. Acesso em 24 jun.2024.

⁶ Disponível em: <https://www.publishnews.com.br/materias/2024/03/19/topicos-que-movimentaram-as-discussoes-na-feira-do-livro-de-londres-2024>. Acesso em 24 jun.2024.

processo muito mais rápido e funcional, o que nos dá a chance de disponibilizar, de maneira ágil, uma ampla variedade de títulos em formato de áudio, (...), democratizando cada vez mais o acesso ao conhecimento”⁷. Assim como a Google, a empresa utilizará a ferramenta inicialmente apenas em livros de não-ficção.

Todos esses dados nos fazem indagar qual será o impacto da IA no campo editorial de audiolivros. No curto prazo, há os benefícios elencados pelos entusiastas da nova ferramenta: diminuição nos custos e na complexidade da produção, diversidade de vozes, fim das barreiras de idiomas, aumento do catálogo e da circulação de títulos, acesso à auto publicação por autores independentes. Vamos nos estender um pouco mais sobre eles.

De fato, os custos de criação de audiolivros com IA são significativamente mais reduzidos, e sua produção menos demorada, em comparação com os processos que utilizam narradores humanos. Em média, um audiolivro de 10 a 15 horas (equivalentes a 300 páginas) leva até quase dois meses para cumprir todas as etapas que envolvem a sua confecção – escolha do título, da voz certa, adaptação do texto impresso ao narrado, gravação, edição, revisão, mixagem e masterização –, ao custo médio de R\$ 15 a 20 mil nas produções menos complexas.

Em paralelo, a narração por IA permite que os ouvintes escolham, entre um amplo banco de vozes, a que lhes proporciona uma fruição auditiva mais rica e diversificada, com uma variedade de gêneros, idiomas, sotaques, timbres e estilos vocais. Essa ‘customização’ dos audiolivros promove, ainda, a acessibilidade de pessoas com deficiências de visão ou com barreira de idiomas ao consumo de conteúdo literário.

Já para editoras comerciais, a utilização da inteligência artificial aumenta a disponibilidade de títulos em formato auditivo, possibilitando que convertam seus catálogos de obras impressas em audiolivro, alguns subaproveitados como geradores de receita, como os títulos de *backlist*. No caso dos autores independentes, a ferramenta viabiliza rapidez na publicação e a comercialização de suas obras nas diversas plataformas de streaming presentes no país, promovendo a bibliodiversidade.

Mas, e a médio e longo prazos?

O uso da ferramenta de inteligência artificial nos processos produtivos do setor de audiolivros, entretanto, pode causar um impacto mais disruptivo do que em outras áreas

⁷ Disponível em: https://www.linkedin.com/posts/opalme_audiobook-narrado-por-ia-activity-7065353324060901377-BR8v/ Acesso em 20 jul. 2024.

editoriais: em termos econômicos, ela atinge diretamente a fonte de renda de diversos profissionais que trabalham com o formato (narradores, locutores, atores, dubladores). Não à toa, artistas de voz têm se manifestado pela regulamentação do uso da IA na indústria criativa. Já no campo jurídico, a disponibilidade ampla e irrestrita da ferramenta põe em risco os direitos autorais, já que a IA pode converter em áudio qualquer texto impresso disponível digitalmente. Já há diversas opções de serviços de conversão na internet: em uma rápida pesquisa no Google, foram listadas quase seis milhões de entradas, desde grandes empresas – como Google, Microsoft, Amazon, Bookwire e IBM, que cobram por sua utilização –, até opções gratuitas de leitores de textos.

E no que tange à experiência propriamente dita de escuta do texto narrado, qual a qualidade das narrações sintéticas? A presença cada vez mais constante da inteligência artificial no mercado editorial de audiolivros vem promovendo o debate sobre os limites da interação homem-máquina assim como sobre os limites éticos de seu uso.

Você nunca saberia que é digital?

Paul Zumthor, em resposta a um questionário realizado em 1986 sobre o “impacto dos meios sobre a vocalidade”, previu que, no futuro, a “abstração vocal será tanto maior que já não se tratará de gravação, mas de voz fabricada” (2018, p. 16). Passadas quase quatro décadas, a predição do linguista, historiador e crítico literário suíço se tornou realidade na indústria editorial, foco de nosso artigo. E nos coloca a seguinte pergunta: estarão os ouvintes de audiolivros dispostos a aceitar uma voz digital no lugar de uma voz humana em sua experiência de escuta?

Antes de entrarmos nessa discussão, precisamos definir se o audiolivro narrado por IA de fato pode ser classificado como *leitura*, *leitura em voz alta*, ou se é algo completamente diverso. Vejamos: para promover uma maior imersão na história, narradores conseguem, ao explorar o enredo dos livros, dar a entonação correta ao conteúdo e criar uma ambientação para o ouvinte. Logo, a figura do narrador é determinante para a experiência de quem ouve o audiolivro: ele é *a voz* do texto. No caso da inteligência artificial, há uma série de questionamentos sobre o resultado final, e ainda não há certeza de que ela consiga um efeito parecido: uma narração com som natural, similar à de um narrador humano.

Nesse sentido, apesar de a atual tecnologia de IA buscar humanizar ao máximo a experiência do ouvinte no que diz respeito à fruição sonora de audiolivros, a “leitura” que

ela realiza é, de fato, uma “oralização”, como explica Bajard (2014): nada mais do que uma decodificação sonora das palavras, uma transmissão sem compreensão e sem emoção, e, portanto, sem um fundo notavelmente humano – não deixemos de lembrar, também, que a ferramenta “lê” o audiolivro sempre da mesma maneira. Em contraste a essa constatação, como argumenta Roger Chartier (1988, p. 77), “a leitura é sempre apropriação, invenção, produção de significados”. Algo que a IA – mesmo com seus bancos de vozes emprestadas de matrizes humanas – (ainda) não é capaz de realizar.

Dito isso, como a IA afeta a fruição, pelos leitores-ouvintes, de obras literárias no formato sonoro digital? É consenso que atributos de linguagem como timbre de voz, entonação, volume, clareza, velocidade, ritmo, e mesmo representatividade, pesam na avaliação da *voz do livro* – tanto por quem o produz como por quem o escuta. Na relação do leitor-ouvinte com o narrador, a *voz certa* é fundamental para conduzir o primeiro para dentro da narrativa, para tornar o texto *vivo*. Em outras palavras, a voz humana, com sua corporeidade – “o peso, o calor, o volume real do corpo, do qual a voz é apenas expansão” (Zumthor, 2018, p. 17) – produz um efeito sobre o próprio sentido do texto, e, conseqüentemente, sobre a recepção desse mesmo texto.

Recorro a Roland Barthes para propor que a fruição de audiolivros gravados com vozes humanas se fundamente no que ele identifica, em artigo sobre a linguagem da crítica musical, como o “grão da voz”, ou seja, o “espaço – isto é, ‘tudo aquilo que, na execução, está a serviço da comunicação, da representação, da expressão’ – muito preciso onde *uma língua encontra uma voz*” (grifos do autor), esse “lado aparentemente abstrato, o impossível relato de um prazer individual que experimento continuamente ao ouvir cantar” (1972, p. 238); “o grão seria: a materialidade do corpo falando sua língua materna” (idem, p. 239), ou seja, “o ‘grão’ é o corpo na voz que canta, na mão que escreve, no membro que executa” (ibidem, p. 244); em nosso contexto, a voz que narra um audiolivro: aquilo que lhe dá textura, volume, materialidade, naturalidade, taticidade.

Assim sendo, a decodificação (ou oralização) de um audiolivro por IA parece não conseguir romper a barreira da superficialidade e da automação, e já há dados que indicam que leitores-ouvintes estão insatisfeitos com a quantidade de vozes artificiais nos catálogos das plataformas de streaming do formato. Habilidades como modulação de voz, interpretação, dramatização – todos recursos acionados pela narração humana, por uma voz ‘encorpada’ –, e mesmo respiração, suspiro, risadas, pausas narrativas, são relevantes e valorizadas por esses leitores-ouvintes, que buscam estabelecer uma relação emocional

com o texto vocalizado, com o corpo que narra, desde o início da audição, sendo esta, inclusive, critério para muitos deles continuarem ou não a escuta da obra sonora.

Conclusão

É fato que o uso da inteligência artificial na narração de audiolivros cria outras – e diferentes – formas de acesso às obras literárias, ampliando sua fruição; porém, sem se igualar à complexidade multidimensional da narração humana.

Sendo o processo para transformar um texto em áudio caro e demorado, pois é preciso contratar narradores profissionais, realizar a gravação e a correção, além de outras etapas de pós-produção sonora, a substituição de narradores humanos por IA pode interferir positivamente na sustentabilidade do negócio editorial. Audiolivros narrados por IA também podem beneficiar indivíduos com deficiências de visão ou barreiras de idioma, permitindo-lhes consumir conteúdo literário de uma forma mais acessível.

Por outro lado, os críticos ao uso de vozes sintéticas acreditam que, ao reduzir a expressividade da narração humana à simples oralização, estas podem causar o empobrecimento da economia do livro como um todo. Logo, títulos narrados por IA estão longe de se tornarem um padrão no mercado editorial, pois o ato de contar histórias é o mais essencial dos contatos sociais, e estes só são possíveis aos seres humanos.

É ponto pacífico, entretanto, que precisamos pensar nas consequências do uso da inteligência artificial na narração de obras literárias, para que o mercado editorial continue a promover a bibliodiversidade assim como a variedade de vozes, sotaques, timbres, garantindo que o uso da ferramenta seja feito de forma ética em todas as etapas da produção de audiolivros. Caso contrário, “narrado por humano” pode se tornar, em breve, um grande argumento de marketing editorial.

Referências

Bajard, Élie. **Ler e dizer**: compreensão e comunicação do texto escrito. São Paulo, SP: Cortez, 2014.

Barthes, Roland. O grão da voz. In: Barthes, Roland. **O óbvio e o obtuso**: ensaios críticos III. Rio de Janeiro: Nova Fronteira, 1990. p. 237-245.

Chartier, Roger. **A aventura do livro**: do leitor ao navegador. São Paulo: Editora da UNESP, 1988.

Zumthor, Paul. **Performance, recepção, leitura**. São Paulo: Ubu Editora, 2018 [1986]. 122p.