

A polifonia bakhtiniana como paradigma epistemológico para a análise da autoria em textos gerados por inteligência artificial generativa /
Bakhtinian Polyphony as an Epistemological Paradigm for Analyzing Authorship in Generative Artificial Intelligence Texts

*Tarcísio de Souza Peres**

*Lígia Razera Gallo***

RESUMO

Examinaremos a evolução histórica da autoria e as tensões contemporâneas em torno da produção textual por sistemas de inteligência artificial, argumentando que a polifonia, conforme concebida por Mikhail Bakhtin, oferece uma chave interpretativa consistente para compreender a ausência de autoria plena nesses casos. A produção de textos por modelos generativos não implica a presença de consciências discursivas independentes nem de vozes axiologicamente orientadas em interação dialógica, condição essencial para a configuração da autoria no sentido polifônico. A análise contrasta essa perspectiva com abordagens como o teste de Turing e o argumento da sala chinesa, destacando que tais modelos operam dentro de limites funcionalistas e simulacionais. Em contraposição, a polifonia exige a coexistência de subjetividades responsivas, articuladas por meio de linguagem compartilhada e situada, o que os sistemas de IA ainda não são capazes de realizar. Nesse sentido, a polifonia revela-se como um operador analítico eficaz na delimitação da autoria.

PALAVRAS-CHAVE: Autoralidade; Polifonia; Inteligência Artificial Generativa; Consciência; Intencionalidade

ABSTRACT

We examine the historical evolution of authorship and contemporary tensions surrounding textual production by artificial intelligence systems, arguing that polyphony, as conceived by Mikhail Bakhtin, offers a consistent interpretative key for understanding the absence of full authorship in such cases. The generation of texts by generative models does not entail the presence of independent discursive consciousnesses nor of axiologically oriented voices engaged in dialogic interaction, a condition essential to the configuration of authorship in the polyphonic sense. This analysis contrasts such a perspective with approaches like the Turing test and the Chinese room argument, emphasizing that these models operate within functionalist and simulation-based limits. In contrast, polyphony requires the coexistence of responsive subjectivities, articulated

* Universidade de São Paulo – USP, Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de Letras Clássicas e Vernáculas, São Paulo, São Paulo, Brasil; <https://orcid.org/0009-0009-2801-0896>; dr.tarcisio@usp.br

** Universidade de São Paulo – USP, Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de Letras Clássicas e Vernáculas, São Paulo, São Paulo, Brasil; <https://orcid.org/0009-0005-8345-882X>; giagallo@gmail.com

through shared and situated language, which AI systems are still incapable of achieving. In this regard, polyphony proves to be an effective analytical operator in delineating authorship.

KEYWORDS: Authorship; Polyphony; Generative Artificial Intelligence; Consciousness; Intentionality

Introdução

A autoria, historicamente, deslocou-se de matrizes coletivas e orais para formas individualizadas, densificadas pela tradição letrada e pela consagração moderna do autor como fonte singular de sentido. A difusão de sistemas de inteligência artificial (IA) generativa reabre esse debate ao produzir textos que imitam estilos humanos sem consciência ou intencionalidade atribuíveis. Em grande parte da discussão técnico-filosófica, abordagens funcionalistas - como o jogo de imitação de Turing¹ e o argumento da sala chinesa - privilegiam critérios comportamentais de simulação, deixando em segundo plano dimensões propriamente dialógicas e axiológicas. À luz do pensamento de Mikhail Bakhtin, a polifonia oferece um quadro analítico para conceber o texto como enunciado multivocal, sustentado por vozes responsivas e relativamente autônomas; tal requisito permite interrogar a suficiência de produções algorítmicas para configurar autoria e responsabilidade discursiva.

Este artigo tem por objetivo: (i) reconstruir, em linhas gerais, a evolução do conceito de autoria; (ii) delimitar limites dos modelos funcionalistas aplicados a textos de IA; (iii) propor a polifonia bakhtiniana como ferramenta epistemológica para analisar a presença - ou ausência - de vozes independentes em textos gerados; e (iv) discutir consequências para debates éticos e para práticas editoriais em humanidades digitais. A estrutura é a seguinte: a seção 1 apresenta um panorama histórico de concepções de autoria em tradições ocidentais e orientais; a seção 2 discute contribuições de teoria literária; a seção 3 contrapõe testes funcionalistas à noção de polifonia; as considerações finais sintetizam implicações para a autoria em contextos de IA.

¹ O jogo, atualmente conhecido como “teste de Turing” recebeu esse nome em homenagem ao matemático britânico Alan Mathison Turing.

1 Panorama histórico

Na Antiguidade, a produção e transmissão dos textos se dava principalmente por meio da tradição oral e da cópia manual feita por escribas, de modo que a figura do autor individual não era vista como a fonte exclusiva da criação. O conhecimento era considerado um legado coletivo, e as obras, muitas vezes, integravam um patrimônio cultural compartilhado, no qual os textos eram constantemente reinterpretados e adaptados. A fixação escrita de performances orais altera o modo de leitura e de atribuição de autoria. Em *Oralidade e cultura escrita: a tecnologização da palavra*, Walter Jackson Ong (1998) observa que apresentações orais, depois de transcritas, passam a ser estudadas como textos escritos. Essa estabilização pode atenuar a percepção do caráter performático, situacional e frequentemente coletivo dessas produções, favorecendo sua associação a um sujeito autoral unificado.

Já na Idade Média, a transmissão dos textos se orientava pela autoridade dos ensinamentos religiosos e pela tradição coletiva. Os textos sagrados e clássicos eram copiados e difundidos em contextos comunitários, nos quais o valor da mensagem superava a importância da identidade de seu criador. Nesse ambiente, o papel do autor era secundário em comparação com a função do texto como veículo de conhecimento e fé, refletindo uma cultura na qual a coletividade prevalecia sobre a individualidade.

Com o advento do Renascimento, surgiu uma valorização crescente do potencial humano e do pensamento individual, impulsionada pelos ideais humanistas que redescobriram e celebraram a cultura clássica. Nesse período, a figura do autor passou a ser entendida como a fonte original e única de uma obra, portadora de intenções, identidade e visão próprias. Esse novo olhar sobre a criação literária ganhou contornos ainda mais intensos no Romantismo, quando a exaltação do eu criativo, da subjetividade e da emoção se tornou central na produção artística. Na advertência de seu *Estética como ciência da expressão e linguística geral*, o filósofo italiano Benedetto Croce escreve que “arte é expressão, expressão, é claro, não imediata e prática, mas teórica, ou seja, intuição” (2016, p. 20). Ao longo de sua obra, ele defende que a arte é, essencialmente, a expressão da personalidade do autor, o que reforça a ideia de que cada obra carrega uma marca única de seu criador.

A abordagem renascentista envolvia um intenso trabalho de tradução, edição e crítica dos textos antigos. Jacob Burckhardt (1991) descreve como esse movimento cultural foi decisivo para moldar a identidade do Renascimento, enfatizando a importância do resgate dos clássicos para a afirmação de um novo ideal de humanidade.

No decorrer do século XV, as obras da Antiguidade foram descobertas e difundidas com rapidez extraordinária. Todas as obras dos filósofos gregos que possuíamos agora estavam então nas mãos de todos, pelo menos sob a forma de traduções latinas. É curioso observar que alguns dos mais zelosos apóstolos dessa nova cultura eram homens da piedade mais estrita, e até mesmo ascetas. [...] O resultado dessas tendências todas foi que a Academia Platônica de Florença escolheu deliberadamente como seu objetivo a reconciliação entre o espírito da Antiguidade e o do cristianismo (Burckhardt, 1991, pp. 308-309).

Figuras como Marsilio Ficino e Giovanni Pico della Mirandola, por exemplo, dedicaram-se a reinterpretar as ideias de Platão e de outros pensadores, contribuindo para que esses autores passassem a ser vistos não apenas como relíquias do passado, mas como fontes vivas de inspiração para a nova era. Esse processo de recomposição implicava uma revalorização da autoria: os humanistas passaram a considerar os clássicos como expressões do gênio individual, mesmo que, originalmente, suas obras tivessem sido transmitidas de forma mais coletiva e menos centrada na figura do autor.

Dessa forma, a evolução do conceito de autorialidade evidencia uma transição histórica na qual o saber, inicialmente compartilhado e coletivo, gradualmente se converteu em uma expressão do gênio individual. Essa mudança alterou as práticas de produção e transmissão dos textos e transformou as percepções acerca da criatividade, da originalidade e da identidade, deixando um legado que continua a influenciar as discussões contemporâneas sobre a autoria nas artes e na literatura.

Em tradições orientais, diversos cânones revelam processos de composição e transmissão nos quais a autoria se distribui ao longo do tempo. Os *Analectos* (Confúcio, 2023) articulam voz magistral e sedimentação escolar; as crônicas de Sima Qian (Sima, 2009) integram narrativas de múltiplas proveniências; a transmissão dos sutras supõe comentário cumulativo; os épicos indianos como o *Mahabharata* (Rajagopalachari, 2005) e o *Ramayana* (Rajagopalachari, 2013) evidenciam estratos e expansões; e *O conto de Genji* (Shikibu, 2026) repercute em diálogos intertextuais de corte. Nesses casos, a

identidade autoral emerge menos como origem exclusiva e mais como nó de uma rede de vozes, práticas exegeticas e intervenções históricas.

Essas experiências sugerem que, em diferentes regiões do Oriente, a noção de um autor único muitas vezes cede lugar a processos culturais de reinterpretação e compilação, em que a obra representa a confluência de linhagens de pensamento, comentários históricos e práticas de memorização ritual. Assim como no Ocidente, em que a transição da oralidade para a escrita redefiniu as bases da autoria, no Oriente as transformações políticas, religiosas e literárias moldaram a maneira pela qual as pessoas concebem a figura daquele que compõe ou transmite um texto. A historicidade da ideia de autor, portanto, não se resume a um movimento linear do anonimato para a individualidade, mas a uma rede de influências mútuas, em que cada cultura e cada tempo estabeleceram novas formas de perceber, valorizar e atribuir sentido às vozes que se tornam parte de sua herança textual.

2 A crítica e a problemática da autoralidade na teoria literária

A obra literária emerge como um espaço no qual se entrelaçam a experiência pessoal do autor e as inúmeras influências acumuladas ao longo de sua trajetória. Esse mosaico de referências mostra que a originalidade surge de um diálogo contínuo com textos e tradições que precedem o criador. Harold Bloom (2002), argumenta que os escritores convivem com a presença intimidadora de seus predecessores, vivenciando uma tensão que os impulsiona a subverter e reconfigurar essas influências na busca por uma voz própria:

O poeta que vem depois abre-se para o que acredita ser um poder no poema-pai que não pertence ao pai mesmo, mas a uma gama de ser logo além desse precursor. Ele faz isso, em seu poema, colocando a relação da obra com o poema-pai de modo a desfazer pela generalização a unicidade da obra anterior (Bloom, 2002, p. 65).

Para Bloom, essa relação conflituosa e transformadora é fundamental para o surgimento de obras que, embora enraizadas em um legado literário, consigam

estabelecer sua originalidade. Ademais, Roland Barthes (1988) reforça a ideia de que os textos são construídos por meio de uma teia de intertextualidades:

[...] um texto é feito de escrituras múltiplas, oriundas de várias culturas e que entram umas com as outras em diálogo, em paródia, em contestação; mas há um lugar onde essa multiplicidade se reúne, e esse lugar não é o autor, como se disse até o presente, é o leitor: o leitor é o espaço mesmo onde se inscrevem, sem que nenhuma se perca, todas as citações de que é feita uma escritura; a unidade do texto não está em sua origem, mas no seu destino (Barthes, 1988, p. 70).

Segundo Barthes, o sentido de uma obra não está fixado na intenção ou biografia do autor, mas se forma no encontro e na sobreposição de múltiplas referências, o que subverte a noção de autoria como origem exclusiva do significado. Essa perspectiva aponta para uma criação que é, em essência, um processo de ressignificação constante das influências que a cercam.

Mikhail Mikhailovich Bakhtin contribui com a noção de que a literatura é um espaço de múltiplas vozes e diálogos, no qual a interação entre diversos discursos enriquece o campo de sentidos da obra. Para Bakhtin (1993), a presença de diferentes perspectivas dentro de um mesmo texto evidencia que a autoria não é um ato isolado; é, todavia, um processo profundamente influenciado por tradições culturais e literárias:

[...] o romance orchestra todos os seus temas, todo seu mundo objetal, semântico, figurativo e expressivo. O discurso do autor, os discursos dos narradores, os gêneros intercalados, os discursos das personagens não passam de unidades básicas de composição com a ajuda das quais o plurilingüismo se introduz no romance. Cada um deles admite uma variedade de vozes sociais e de diferentes ligações e correlações (sempre dialogizadas em maior ou menor grau) (Bakhtin, 1993, pp. 74-75).

No estudo de Dostoiévski, Bakhtin explora como as ideias (conteúdo cognitivo), quando incorporadas à arte do romance, são libertadas de uma função meramente lógica ou argumentativa e ganham vida própria no diálogo artístico. Ele esclarece que, em Dostoiévski, a ideia deixa de ser um conceito abstrato para tornar-se um acontecimento vivido no contato entre consciências, buscando ser compreendida e respondida por vozes diversas. Esse caráter dialógico da ideia no romance evidencia a natureza não

instrumental da arte - a ideia não aparece para provar uma tese unívoca, mas para ser explorada livremente:

A ideia é um *acontecimento vivo*, que irrompe no ponto de contato dialogado entre duas ou várias consciências. Neste sentido a ideia é semelhante ao discurso, com o qual forma uma unidade dialética. Como o discurso, a ideia quer ser ouvida, entendida e “respondida” por outras vozes e de outras posições. Como o discurso, a ideia é por natureza dialógica... (Bakhtin, 2002, p. 87).

Neste excerto, Bakhtin ilustra o potencial da arte de liberar o conteúdo ideativo de finalidades predeterminadas. Em vez de servir a uma conclusão fechada, a ideia, ao ingressar na obra artística, se desprende de uma finalidade única e passa a existir por si mesma como diálogo vivo entre perspectivas. A arte aqui opera de modo não teleológico: Dostoiévski não utiliza as personagens e suas ideias para defender diretamente uma verdade prévia ou um ensinamento moral unívoco, mas cria uma arena em que múltiplas vozes ideológicas interagem abertamente. Essa falta de finalidade prática imediata não significa indiferença - pelo contrário, há um engajamento profundo com os conflitos de valor e de pensamento, porém em termos dialógicos, não dogmáticos. Desinteressada de propósitos utilitários, a forma artística interessa-se em dar plena ressonância a cada voz e visão de mundo. Como complementa Renata Coelho Marchezan:

Vê-se, assim, que Bakhtin privilegia as reflexões sobre a obra literária, mas sempre situa a estética no domínio mais amplo da cultura humana, que compreende também a ciência/conhecimento e a ética. A arte não é indiferente também a esses dois outros domínios; ela os acolhe e, desse modo, dá à cultura uma unidade concreta. Ao fazê-lo, a arte libera o conteúdo do conhecimento e do ato ético dos elos, determinações e interesses que tem em seus domínios específicos. É esse isolamento - separação que, apressadamente, poderia ser confundida com o conceito de estranhamento - do conteúdo de seus domínios que compõe também a explicação da relativa autonomia de que desfruta a arte; é ele que possibilita a atividade não teleológica, não instrumental, desinteressada, mas não indiferente, do autor-criador (Marchezan, 2015, pp. 190-191).

Bakhtin (2003) afirma que a voz criativa do escritor é necessariamente uma segunda voz - isto é, uma voz que soa através de outra. Ele sugere que uma palavra plenamente unívoca seria ingênua e insuficiente para a arte verdadeira. Em contraste, a palavra literária autêntica é bivocal: o autor não fala de modo direto e literal, mas

empresta sua voz a outras vozes, falando por meio delas. Assim, a voz do autor se insere por trás da voz do personagem, do narrador ou de uma determinada linguagem estilizada.

É possível que toda palavra sem objeto e monovocal seja ingênua e imprestável para uma criação autêntica. Toda voz autenticamente criadora sempre pode ser apenas uma segunda voz no discurso. Só a segunda voz — a *relação pura* — pode ser até o fim desprovida de objeto, sem abandonar a sombra substancial figurada. O escritor é aquele que sabe trabalhar a língua estando fora dela, aquele que tem o dom do falar indireto. [...] Neste caso, minha própria palavra se torna objetificada e recebe a segunda voz - a minha própria (Bakhtin, 2003, p. 315; itálicos no original).

Em outras palavras, o escritor habilidoso consegue se distanciar do idioma comum e das falas diretas, para então reinseri-los na obra de forma artística, fazendo-os falar com duplo sentido. Esse é o princípio do discurso bivocal: na obra literária, as falas das personagens ou narradores carregam simultaneamente duas vozes - a voz própria da personagem e a voz do autor refratada nessa fala. Por exemplo, numa narração irônica, as palavras do personagem (primeira voz) são intencionalmente moldadas pelo autor para que o leitor perceba também a atitude do autor (segunda voz) diante dessas palavras. Bakhtin diz que somente essa segunda voz “pura” (isto é, que expressa apenas a relação do autor, sem conteúdo próprio) pode ser totalmente “sem objeto” - porque ela não introduz novo objeto de discurso, apenas refrata um já existente. Isso confirma a presença do fenômeno do discurso indireto livre ou duplamente orientado na literatura, no qual “minha própria palavra se torna objetificada e recebe a segunda voz” (Bakhtin, 2003, p. 315) dentro do texto artístico. Bakhtin, portanto, define o escritor/artista como aquele que fala com uma voz alheia, ou melhor, que consegue fazer ressoar sua intenção através das vozes sociais e ideológicas das personagens - uma voz segunda que enriquece o diálogo da obra sem jamais se apresentar de modo plano ou unívoco.

A noção bakhtiniana de polifonia, elaborada a partir da leitura do romance dostoievskiano como orquestração de vozes autônomas em interação responsiva, pode ser estendida com fundamento na tese de que a linguagem é constitutivamente dialógica e de que o romance realiza um dialogismo máximo. Como resumem Gary Saul Morson e Caryl Emerson, “o romance do século XIX, e especialmente os romances de Dostoiévski, são os principais exemplares desse tipo de dialogismo máximo” (Morson, 1988, p. 52;

tradução nossa)². Essa base permite deslocar o foco do estritamente literário para esferas culturais, quando o dialogismo é entendido como princípio que “conduz das relações intralinguísticas às relações intraculturais” (Morson, 1989, p. 52; tradução nossa)³. Dessa articulação decorre que reconhecer a dimensão dialógica da autoria não implica exigir um texto polifônico em sentido estrito; trata-se de avaliar a interação responsiva de vozes axiologicamente plurais, aplicável a produções textuais contemporâneas em contextos culturais diversos, à luz do quadro dialógico acima delineado.

Em paralelo, Roland Barthes propõe que a fixação do sentido de um texto na biografia e intenção do autor restringe as múltiplas interpretações possíveis “Dar ao texto um Autor é impor-lhe um travão, é provê-lo de um significado último, é fechar a escritura. [...] encontrado o Autor, o texto está 'explicado', o crítico venceu” (1988, p. 69). Para Barthes, a obra se desassocia da autoridade singular do criador, permitindo que o leitor reconstrua seu significado de forma livre e plural. Assim, o texto se torna autônomo, rompendo com a tradição que colocava o autor como a fonte única e definitiva de sentido.

Para aprofundar essa discussão convém mencionar outras contribuições de Bakhtin, que destaca a polifonia e a natureza dialogal dos textos. Para ele, as obras literárias se configuram como espaços de interação entre múltiplas vozes e consciências, o que diminui a relevância de uma autoria singular. Em sua análise, o significado emerge do confronto e da convivência de diversas perspectivas, evidenciando que a criação textual é um fenômeno coletivo e dinâmico. Walter Benjamin também acrescenta uma dimensão importante ao debate ao abordar a questão da aura e da originalidade das obras: “com a reprodutibilidade técnica, a obra de arte se emancipa, pela primeira vez na história, de sua existência parasitária, destacando-se do ritual” (1987, p. 171). Nota-se que ele questiona a singularidade e a autenticidade que tradicionalmente se atribuem ao autor.

Assim, a problemática da autoralidade revela uma tensão constante entre a exaltação do gênio criativo e a compreensão de que os textos se configuram como espaços abertos à multiplicidade de interpretações. Essa perspectiva contemporânea desloca o foco da intenção exclusiva do autor para a interação entre leitor, texto e contexto cultural, oferecendo uma visão mais complexa da produção de significado. Tal reflexão permite

² No original: “the nineteenth-century novel, and especially the novels of Dostoevsky, are the prime exemplars of this kind of maximal dialogism”.

³ No original: “leads from intralinguistic to intracultural relationships”.

compreender que, embora a figura do autor continue a ter um papel relevante, ela é apenas uma parte de um processo mais amplo de construção e circulação de sentidos.

Ato contínuo, a autoralidade pode ser entendida não como uma expressão de uma essência individual imaculada, mas como o resultado de um processo criativo permeado por influências externas. A originalidade se manifesta na habilidade do autor de rearticular, transformar e dialogar com esse vasto campo de referências, criando obras que são, ao mesmo tempo, parte de uma tradição e uma inovação dentro dela.

3 Singularidade, teste de Turing, sala chinesa e polifonia bakhtiniana

A inteligência artificial (IA) generativa é um ramo da aprendizagem de máquina que se dedica à criação de dados novos e originais, a partir da compreensão profunda dos padrões presentes em grandes volumes de informação. Este campo evoluiu significativamente com o desenvolvimento de modelos capazes de sintetizar imagens, textos, músicas e outras formas de conteúdo, permitindo a reprodução de informações existentes e a criação autônoma de novos dados. A partir do conceito central de redes neurais, esses modelos são treinados em vastos conjuntos de dados para aprender representações complexas, o que lhes confere a habilidade de generalizar e gerar amostras que se assemelham às distribuições originais dos dados de treinamento (Goodfellow *et al.*, 2020).

O advento de arquiteturas baseadas em atenção e modelos sequenciais transformou ainda mais o cenário da IA generativa possibilitando o desenvolvimento de sistemas que produzem saídas coerentes e contextualmente relevantes. Tais avanços podem ser observados na evolução dos modelos de linguagem, os quais demonstram um certo domínio na criação de textos que reproduzem estilos literários e articulam argumentos de forma lógica e integrada. Esse progresso é fruto da combinação de técnicas estatísticas sofisticadas com métodos de otimização que ajustam os parâmetros dos modelos a partir de grandes volumes de dados, contribuindo para uma melhor compreensão e representação das estruturas subjacentes à linguagem e à percepção visual (Vaswani *et al.*, 2017).

Nesse cenário, os progressos da inteligência artificial generativa fomentaram, no domínio da teoria do conhecimento, um debate cada vez mais intenso acerca da autoridade em textos produzidos por sistemas artificiais.

Ray Kurzweil, um futurista e cientista da computação, é conhecido por suas previsões otimistas acerca do progresso tecnológico e, em especial, sobre a evolução da inteligência artificial. Segundo ele, caminhamos para um ponto de singularidade tecnológica em meados do século XXI, no qual as máquinas poderão superar a inteligência humana em diversos aspectos e se tornarão capazes de aprimorar a si mesmas. Nesse cenário, Kurzweil também sustenta que a inteligência artificial eventualmente alcançará aquilo que entendemos como consciência humana ou, pelo menos, algo muito próximo disso:

Então, como chegaremos a um acordo com a consciência que será reivindicada pela inteligência não biológica? Do ponto de vista prático, tais reivindicações serão aceitas. Por um lado, “elas” serão nós, portanto não haverá uma distinção clara entre a inteligência biológica e a não biológica. Além disso, essas entidades não biológicas serão extremamente inteligentes, portanto conseguirão convencer outros humanos (biológicos, não biológicos ou intermediários) que elas são conscientes. Terão todas as delicadas sugestões emocionais que nos convencem hoje de que os humanos são conscientes. Conseguirão fazer com que outros humanos riem ou chorem. E vão se irritar se os outros não aceitarem suas reivindicações (Kurzweil, 2018, p. 431).

No que se refere à consciência humana, a Psicanálise formulada por Sigmund Freud, entre o final do século XIX e o início do XX, passou a conceber a mente como composta por três instâncias: consciente, pré-consciente e inconsciente (Freud, 2019). O consciente engloba tudo aquilo que se encontra no campo imediato de percepção; o pré-consciente inclui conteúdos acessíveis, porém não presentes na consciência a todo momento; e o inconsciente envolve conteúdos reprimidos, que influenciam o comportamento sem emergir diretamente à superfície consciente. Embora a Psicanálise não tenha inventado o conceito de consciência, elaborou uma descrição mais ampla do aparelho psíquico, colocando a consciência ao lado de outras dimensões da vida mental.

De modo geral, entende-se que a consciência pode ser vista como a experiência subjetiva ou fenomenológica (relacionada à sensação de sentir algo, perceber cores, sons e emoções), a consciência de si (autoconsciência, ou a capacidade de reconhecer-se como

agente que observa, sente e pensa) e a consciência intencional ou reflexiva (habilidade de voltar a atenção para os próprios processos mentais, pensando sobre o próprio pensamento).

Kurzweil não se envolve diretamente nem aplica sistematicamente o modelo tripartido da mente de Freud em *A singularidade está próxima* (2018). O psicanalista austríaco é mencionado apenas em trechos de conversas ficcionalizadas, mais como uma figura de *cameo* nas cenas de diálogo de Kurzweil. Essas passagens mostram um Sigmund Freud interpondo perguntas ou observações curtas sobre identidade e o eu, mas Kurzweil nunca discute ou critica formalmente a estrutura psicanalítica da consciência de Freud no livro. Em vez de desenvolver conceitos freudianos, o foco principal de Kurzweil está na natureza da consciência em relação às tecnologias emergentes, reconhecimento de padrões e a possibilidade de replicar a mente em meios não biológicos.

A concepção de consciência que Kurzweil defende baseia-se em aspectos como o processamento de informação em alto nível, ou seja, quando um sistema não só processa dados, mas tem consciência desse processo e é capaz de refletir sobre seus próprios estados. Para ele, a criação de modelos internos complexos do mundo e de si próprio, com capacidade de adaptação, reflexão, tomada de decisão e aprendizado contínuo, aproximaria a IA daquilo que entendemos por estados conscientes. A ausência de diálogo crítico com a economia libidinal proposta pela psicanálise dificilmente se explica apenas por uma suposta fragilidade interna da teoria de Kurzweil; antes reflete a distância epistemológica entre um projeto engenheirístico preocupado com replicabilidade mecânica e uma tradição hermenêutica focada em conflito pulsional.

Na ciência da computação, o teste de Turing⁴, proposto inicialmente como um critério operacional para determinar se uma máquina pode ser considerada inteligente ou consciente (Turing, 1950), restringe-se à avaliação da capacidade imitativa e performativa de uma inteligência artificial diante de interações linguísticas controladas. Sua limitação fundamental consiste na concepção behaviorista implícita na avaliação da inteligência artificial, concentrando-se exclusivamente na aparência externa de compreensão e

⁴ Um avaliador humano interage por meio de uma interface textual com dois interlocutores ocultos: um humano e uma máquina. Caso o avaliador não consiga identificar com segurança qual dos dois é a máquina, considera-se que esta passou no teste. Embora sujeito a críticas, o teste continua a ser uma referência central nos debates sobre inteligência artificial e filosofia da mente, pois desloca o foco da essência da inteligência para sua manifestação observável.

interação linguística sem questionar a dimensão axiológica e autoral do discurso produzido. Nesse aspecto, mesmo uma simulação sofisticada e convincente poderia ultrapassar o teste sem que houvesse qualquer garantia de que estivesse em operação uma consciência discursiva autônoma e genuinamente autoral.

De fato, o argumento da sala chinesa⁵, proposto por John Searle (1980), avança na crítica ao reducionismo do teste de Turing ao sustentar que manipulações sintáticas, ainda que complexas e aparentemente significativas, não resultam necessariamente em compreensão semântica ou consciência linguística. A sala chinesa evidencia que a máquina pode produzir textos compreensíveis sem, contudo, experienciar subjetivamente os significados que manipula, evidenciando a ausência de verdadeira consciência discursiva. Todavia, esse argumento permanece restrito à crítica da compreensão semântica, não abrangendo diretamente a dimensão autoral ou dialógica da linguagem, aspecto essencial na determinação da autoria de textos produzidos por sistemas generativos.

Por outro lado, as correntes funcionalistas em filosofia da mente defendem a ideia de que, se um sistema realiza as mesmas funções e processos que um cérebro consciente, ele seria consciente independentemente de seu substrato. Kurzweil se alinha a essa visão, pois crê que a consciência poderia ser emulada quando o processamento de informações e a complexidade de interconexões atingirem certo nível. Ele descarta a necessidade de um componente misterioso para que a consciência exista, acreditando que, quando a IA for suficientemente avançada e puder realizar atualizações em si mesma, alcançará a mesma essência da consciência humana.

Os argumentos que apresento neste livro em relação à consciência têm o propósito de ilustrar essa natureza irritante e paradoxal (e, portanto, profunda) da consciência: como um conjunto de suposições (isto é, que uma cópia do meu arquivo mental compartilha ou não minha consciência) leva, no final, a uma visão contrária e vice-versa (Kurzweil, 2018, p. 443).

⁵ Uma pessoa dentro de uma sala, que não fala chinês, segue instruções de um livro de regras para manipular caracteres da língua de modo a responder perguntas escritas em chinês de maneira correta. Apesar de parecer entender a língua aos olhos de quem está do lado de fora, o indivíduo, na realidade, apenas correlaciona símbolos de entrada com símbolos de saída, sem ter qualquer compreensão semântica. Na analogia com os computadores, o sistema seria capaz de processar informações de forma sintática (regras), mas não teria consciência ou entendimento sobre o conteúdo que manipula. Dessa forma, a indistinguilidade no comportamento linguístico pode ser considerada apenas uma faceta superficial da inteligência.

A noção de intencionalidade é fundamental para a compreensão da experiência humana e da relação do sujeito com o mundo, sendo objeto de estudo de filósofos desde o século XIX. Franz Brentano introduziu formalmente o conceito ao observar que a mente se caracteriza por estar sempre dirigida a algo, isto é, possui uma propriedade que a faz referir-se a um conteúdo ou objeto:

Todo fenômeno mental é caracterizado pelo que os escolásticos da Idade Média chamavam de inexistência intencional (ou mental) de um objeto, e que poderíamos chamar, ainda que não de forma totalmente inequívoca, de referência a um conteúdo, direção para um objeto (o que aqui não deve ser entendido como significando uma coisa), ou objetividade imanente. Todo fenômeno mental inclui algo como objeto dentro de si, embora não todos o façam da mesma maneira. Na apresentação algo é apresentado, no juízo algo é afirmado ou negado, no amor algo é amado, no ódio algo é odiado, no desejo algo é desejado, e assim por diante.(Brentano, 1995, p. 68; tradução nossa)⁶.

Posteriormente, Edmund Husserl (2014) aprofunda essa reflexão em sua fenomenologia, ampliando a discussão sobre como a consciência se dirige a objetos e constitui o sentido desses objetos na experiência. Esse enfoque fenomenológico destaca a estrutura intencional como o cerne da vida consciente, realçando a distinção entre os atos mentais e aquilo a que eles se referem.

Daniel Dennett propõe a chamada postura intencional para explicar como atribuímos intenções a agentes, humanos ou não, a fim de prever e compreender seu comportamento. Em *The Intentional Stance* (1989), ele escreve: “A decisão de adotar a postura intencional é livre, mas os fatos sobre o sucesso ou fracasso da postura, caso ela seja adotada, são perfeitamente objetivos” (1989, p. 24)⁷. Embora Dennett sublinhe o valor heurístico dessa postura, ele questiona se tal atribuição de intenções revela efetivamente estados internos genuínos ou se funciona como uma estratégia interpretativa

⁶ No original: “Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on”..

⁷ No original: “The decision to adopt the intentional stance is free, but the facts about the success or failure of the stance, were one to adopt it, are perfectly objective”.

útil: “[...] é uma ferramenta extraordinariamente poderosa para a previsão - um fato amplamente ocultado por nossa tendência de nos concentrarmos nos casos nos quais ela produz resultados duvidosos ou pouco confiáveis”⁸ (Dennett, 1989, p. 24).

No centro dessa discussão, a intencionalidade envolve a capacidade de um sujeito consciente - dotado de estados mentais que se referem a objetos - de construir significados e de agir sobre o mundo. A questão que permeia esses debates é se tal direcionamento para o mundo necessariamente requer uma experiência subjetiva ou um ponto de vista em primeira pessoa. Em síntese, a complexidade do fenômeno intencional sugere que ele não se reduz a um simples processamento de informações, mas está intrinsecamente ligado à experiência vivida e ao enquadramento conceitual típico de agentes conscientes. Nesse sentido, as máquinas, mesmo que operem com base em algoritmos sofisticados, não demonstram a característica nuclear da intencionalidade discutida por Brentano, Husserl, e Dennett, pois lhes falta a qualidade fenomenológica e a relação significativa com o mundo que esses autores identificam como condição indispensável para que haja estados intencionais genuínos.

Sustentamos, neste ponto, que a polifonia bakhtiniana revela-se epistemologicamente mais completa e assertiva, justamente porque o conceito ultrapassa os critérios superficiais de aparência comunicativa ou a manipulação sintática-semântica para exigir como condição essencial para a verdadeira autoria a existência autônoma de vozes independentes e axiologicamente plurais. Em *Problemas da poética de Dostoiévski*, Bakhtin argumenta:

Dostoiévski é criador do *romance polifônico*. Criou um gênero romanesco essencialmente novo. [...] A voz do herói sobre si mesmo e o mundo é tão plena como a palavra comum do autor; não está subordinada à imagem objetificada do herói como uma de suas características, mas tampouco serve de intérprete da voz do autor. Ela possui independência excepcional na estrutura da obra, é como se soasse ao lado da palavra do autor, coadunando-se de modo especial com ela e com as vozes plenivalentes de outros heróis (Bakhtin, 2002, p. 5; itálicos no original)

⁸ No original: “[...] it is an extraordinarily powerful tool in prediction—a fact that is largely concealed by our typical concentration on the cases in which it yields dubious or unreliable results”.

Nesse contexto, a polifonia bakhtiniana funciona como um operador analítico que integra o dialogismo como sua realização essencial, permitindo sua utilização em avaliações de autoria além do romance, como em textos gerados por IA. Estudos que contextualizam Bakhtin em disciplinas interdisciplinares defendem usos expandidos da polifonia, posicionando-a como paradigma para examinar interações culturais sem restringi-la ao gênero literário específico (Mandelker, 1995). Bakhtin evolui essa integração na edição revisada de sua obra, descrevendo a polifonia como uma realização dialógica que enfatiza a coexistência de perspectivas responsivas, sem demandar que todo texto autoral seja polifônico no modelo dostoievskiano (Bakhtin, 2002). Dessa forma, o reconhecimento da dialógica na autoria não implica a obrigatoriedade de um texto polifônico, mas destaca a ausência de consciências independentes em simulações artificiais.

A polifonia, na concepção bakhtiniana, depende diretamente da autonomia das consciências discursivas que se apresentam sem subordinação a um autor-monológico, estabelecendo entre si relações dialógicas autênticas. Mesmo em sua pluralidade dialógica, a obra literária permanece mediada pela consciência ativa daquele que a organiza, seleciona e mobiliza essas vozes. A ausência desse elemento inviabiliza que inteligências artificiais, cuja estrutura funcional é orientada por parâmetros pré-definidos e condicionada por dados prévios de treinamento, satisfaçam integralmente essa condição necessária para a autoria efetiva e verdadeira em sentido polifônico.

No contexto contemporâneo, essa figura mediadora e consciente é precisamente o humano que interage com a inteligência artificial. O usuário humano, enquanto agente consciente que seleciona e organiza elementos e comandos fornecidos à IA, ocupa uma posição semelhante àquela que Bakhtin descreve como a do autor-polifônico: aquele que articula vozes e contextos, fazendo com que essas vozes adquiram significado em sua coexistência.

Na obra *Estética da criação verbal*, no ensaio “Os gêneros do discurso”, Bakhtin esclarece que a criação verbal nunca é neutra, consequentemente, não é gerada simplesmente por um mecanismo técnico ou por mera casualidade. Ao contrário, é intrinsecamente ligada ao contexto humano vivo e consciente: " Quando escolhemos as palavras no processo de construção de um enunciado, nem de longe as tomamos do

sistema da língua em sua forma neutra, *lexicográfica*. Costumamos tirá-la de *outros enunciados* [...]” (Bakhtin, 2003, p. 292; itálicos no original).

Aplicando essa reflexão à produção textual via IA, fica evidente que a inteligência artificial, ao gerar textos, está sempre respondendo a um impulso humano consciente e intencional. O humano, enquanto destinatário ativo e consciente da produção gerada, orienta significativamente o conteúdo, o estilo e o propósito do texto. A IA, nesse sentido, atua como instrumento tecnológico que amplia o alcance dessa capacidade dialógica - e polifônica - da consciência humana, mantendo intacta a autoria deste último.

Mas essas vozes ainda não se tornaram plenamente autônomas, vozes reais, três consciências com plenos direitos. Isto só ocorre nos romances de Dostoiévski. Em *O Sósia* não há o discurso monológico que se basta somente a si mesmo e ao seu objeto. Cada palavra está dialogicamente decomposta, em cada palavra há uma interferência de vozes, mas aqui ainda não se verifica o autêntico diálogo de consciências imiscíveis que aparecerá posteriormente nos romances. Aqui já existe o embrião do contraponto: este se esboça na própria estrutura do discurso (Bakhtin, 2002, p. 222).

A afirmação inicial desse excerto reforça o entendimento de que a verdadeira polifonia não se resume a uma justaposição sintática ou estilística de falas distintas, mas depende essencialmente de consciências discursivas independentes e axiologicamente íntegras.

Além disso, a aplicação da polifonia como critério para avaliar a autoralidade no contexto das inteligências artificiais envolve uma análise qualitativa das relações discursivas internas ao texto produzido, permitindo a identificação precisa da ausência ou presença de verdadeira alteridade discursiva, dialógica e axiológica, elementos indispensáveis à autoria genuína.

O ideal do eu é, portanto, imensamente poderoso e pareceria atuar da mesma forma que a alteridade da sociedade atua em Bakhtin. Mas Freud não chega a relacionar diretamente o ideal do eu à força social. [...] o que falta em sua análise das relações entre o indivíduo e o grupo é qualquer conceito acerca dos meios pelos quais tais relações são mediadas [...]. Por quais meios os fantasmas internalizados se comunicam com os atores reais? Em que linguagem o romance familiar internalizado encena seu drama no teatro de nossas mentes? A resposta da dialogia a essa questão é que se trata da mesma linguagem que usamos no teatro da sociedade. O grupo como coletivo e os

indivíduos que o compõem compartilham todos a mesma linguagem [...]. O conflito fundamental entre o que Freud chama de ego e ideal do eu não é uma tensão abstrata entre a espontaneidade libidinal do primeiro e as exigências introjetadas do segundo. Essa descrição é adequada até certo ponto, mas não vai longe o bastante; é necessário acrescentar que essa oposição ocorre em circunstâncias altamente específicas (e especificáveis) como um conflito entre fala interna e fala externa. Desse pressuposto decorre o argumento de Bakhtin de que é preciso haver um elo necessário entre o estudo da psicologia e o estudo da linguagem (Holquist, 2002, p. 52)⁹.

Esse fragmento ilustra como, em Bakhtin, só se pode falar em autoria genuína se houver espaço para uma alteridade efetiva. Na crítica a Freud, observa-se que a força social internalizada (o “ideal do eu”) precisa de um meio de expressão - a linguagem - que seja partilhada socialmente, mostrando que a subjetividade não se forma isoladamente, mas em interação real com o outro e com o repertório de discursos coletivos. Dessa forma, quando essa alteridade (a dimensão dialógica no sentido bakhtiniano) está ausente ou é puramente ficcional, a autoria não se perfaz de modo legítimo e responsivo.

Em outras palavras, tanto o teste de Turing quanto o argumento da sala chinesa permanecem ancorados em uma perspectiva funcionalista ou representacional, na qual a questão da inteligência se restringe à simulação comportamental ou ao processamento sintático desprovido de intencionalidade. Ainda que ambiciosas, as conjecturas de Kurzweil também operam em uma lógica escalonada e ontologicamente difusa, centrada na consciência, contornando a acepção freudiana clássica. Em contrapartida, a concepção de polifonia em Bakhtin intervém de modo decisivo, ao deslocar o foco para a

⁹ No original: “The ego ideal is, then, immensely powerful, and would seem to work much as the otherness of society does in Bakhtin. But Freud stops short of directly relating the ego ideal to social force. [...] what is lacking in his account of relations between individual and group is any concept of the means by which such relations are mediated [...] By what means do the internalized phantoms communicate with the real actors? In what language does the internalized family romance play itself out in the theater of our minds? Dialogism’s answer to this question is that it is the same language as we use in the theater of society. The group as a collective and the individual members who comprise it all share the same language [...]. The essential clash between what Freud calls ego and ego ideal is not an abstract tension between the former’s libidinal spontaneity and society’s introjected demands in the latter. This description is adequate so far as it goes, but it does not go far enough; what needs to be added is that this opposition works itself out in highly specific (and specifiable) circumstances as a conflict between inner and outer speech. From this assumption flows Bakhtin’s argument for a necessary link between the study of psychology and the study of language”.

constituição dialógica do sujeito discursivo e para as condições ético-enunciativas que fundamentam sua agência autoral.

Assim, a distinção estabelecida é entre o ato de falar ou responder como um humano e a efetiva capacidade de entender o que está sendo dito. O que está em pauta é a diferença entre aparência e essência: ainda que a máquina apresente um desempenho linguístico impecável, não significa necessariamente que haja algum grau de experiência subjetiva ou intencionalidade por trás de suas respostas. Torna-se evidente que a cognição não se reduz a uma reação automática a estímulos externos, mas consiste em um complexo processo de integração no qual experiências, contexto e intencionalidade se entrelaçam de maneira dinâmica.

A ausência de uma vivência interna - uma capacidade de sentir ou mesmo de atribuir significado às informações processadas - contradiz a possibilidade de equiparar tais sistemas ao funcionamento da mente humana, que se caracteriza por uma constante articulação entre o processamento de dados e a experiência subjetiva.

Considerações finais

A trajetória da autoralidade evidencia um processo intrinsecamente vinculado às transformações históricas nos modos de constituição do sujeito e às práticas de apropriação, circulação e interpretação dos textos. Nesse contexto, a noção bakhtiniana de polifonia, compreendida em sua acepção ampliada como paradigma dialógico, permite problematizar os limites da inteligência artificial generativa. Embora Bakhtin tenha relacionado a polifonia, em sentido estrito, ao romance como espaço de vozes autônomas, desenvolvimentos posteriores a configuraram como instrumento epistemológico que envolve o dialogismo em sua totalidade, aplicável a contextos culturais que não pressupõem necessariamente a presença de textos literários polifônicos. Na versão revista de sua poética, Bakhtin articula o dialogismo à polifonia como realização constitutiva, o que viabiliza sua extensão a domínios não romanescos. Desde os períodos em que o saber era compartilhado como patrimônio coletivo até a consolidação moderna da figura autoral como instância singular de criatividade e produção de sentido, a ideia de autoria transita entre regimes de anonimato, práticas colaborativas e a emergência da subjetividade

artística como núcleo de expressão individualizada, ressaltando a ausência de uma dinâmica dialógica efetiva nas produções artificiais.

Nas tradições orientais, a permeabilidade entre autoria singular e produção comunitária evidencia que a criação literária e filosófica frequentemente se ancora em uma transmissão cumulativa, ritualizada e reiterada ao longo do tempo. Obras como *Os Analectos* de Confúcio, os sutras budistas ou os épicos hindus expressam uma dinâmica autoral profundamente vinculada à memória cultural compartilhada. A presença do sujeito criador se mescla à continuidade interpretativa das gerações, indicando que a intencionalidade e a autoria não estão ausentes, mas operam em formas distintas daquelas consolidadas no paradigma ocidental moderno. A figura do autor não é anulada, e sim complexificada por uma lógica de reescrita, comentário e atualização comunitária que produz um tipo de autoria difusa, mas ainda responsiva.

A teoria literária contemporânea aprofundou essa inflexão ao desestabilizar a centralidade da intenção autoral como fonte exclusiva de significado. Harold Bloom introduziu a noção de uma angústia estrutural frente ao legado dos predecessores, enquanto Roland Barthes deslocou a origem do sentido do autor para o leitor, dissolvendo a figura autoral como instância unívoca de autoridade semântica.

Se faz necessário considerar também que a discussão relativa à morte do autor adquire maior abrangência conceitual quando se insere a reflexão de Michel Foucault, para quem o autor não corresponde a uma figura empírica fixa e sim a uma função decorrente de regimes discursivos específicos. Ao proferir a conferência *O que é um autor?*, Foucault declarou: “a noção do autor constitui o momento crucial da individualização na história das idéias, dos conhecimentos, das literaturas, e também na história da filosofia, e das ciências” (2009, p. 267).

O autor surge, portanto, como operador histórico que individualiza discursos, estabelecendo cortes em meio a práticas de saber que, anteriormente, se mantinham mais difusas ou coletivas. Essa concepção altera a perspectiva ampliada por Barthes, para quem “a escritura é a destruição de toda voz, de toda origem. A escritura é esse neutro, esse composto, esse oblíquo pelo qual foge o nosso sujeito, o branco-e-preto em que vem se perder toda identidade, a começar pela do corpo que escreve” (2004, p. 57).

Barthes suspende a centralidade biográfica; Foucault detalha o dispositivo institucional que substitui tal centralidade, indicando que o nome do autor demarca “um

certo número de discursos que são providos da função ‘autor’, enquanto outros são dela desprovidos” (2009, p. 274). Logo, a assinatura funciona como princípio classificatório que habilita textos a circular em circuitos de legitimação, fixando fronteiras entre produção ordinária e produção sujeita a regime hermenêutico especializado.

A pertinência dessa noção torna-se ainda mais evidente quando Foucault distingue o autor tradicional dos “fundadores de discursividade” (2009, p. 280). Segundo ele, existem pensadores que “têm de particular o fato de que eles não são somente os autores de suas obras, de seus livros. Eles produziram alguma coisa a mais: a possibilidade e a regra de formação de outros textos” (Foucault, 2009, p. 280). Tal diferenciação aproxima-se do dilema formulado por Harold Bloom acerca da angústia de influência, pois indica que certos escritores - vistos como fundadores - instauram matrizes textuais que remodelam o horizonte de criação subsequente. A dialética bloomiana do precursor e do epígono encontra, dessa forma, um correspondente foucaultiano: a função-autor evolui para função fundadora quando cria regras de produção capazes de engendrar comunidades críticas inteiras.

No debate sobre inteligências artificiais generativas, a função-autor de Foucault oferece critério distinto daquele fornecido pelo mero desempenho linguístico avaliado pelo teste de Turing. Ao carecer de uma posição institucional na qual seu nome produza recortes, responsabilidades e protocolos de leitura, o modelo de linguagem permanece exterior à lógica foucaultiana do autor. Mesmo que o texto maquínico aparente coerência, falta-lhe a inscrição social conferida pelo nome que estabelece limites de corpus, regula atribuições e, sobretudo, vincula discurso e responsabilidade.

Mikhail Bakhtin, ao propor o conceito de polifonia, ultrapassa o paradigma da intertextualidade passiva e afirma a existência de múltiplas vozes autônomas que coexistem em um mesmo texto, cada uma portando uma perspectiva axiológica singular. A originalidade, dentro desse modelo, não reside na ruptura com o passado, mas na capacidade de articular uma constelação dialógica de vozes e consciências, em que o autor-criador atua como um orquestrador de alteridades, e não como um emissor soberano de significados.

Essa concepção implica a superação de modelos centrados no autor-monológico, uma vez que a obra se revela, em sua complexidade estética e epistêmica, como uma arena de confronto entre visões de mundo, entre línguas sociais e vozes ideológicas em

disputa. A polifonia, tal como formulada por Bakhtin, exige que o texto contenha múltiplas falas que representem consciências discursivas efetivamente independentes e não subordinadas ao ponto de vista autoral.

A emergência de sistemas de inteligência artificial generativa, capazes de produzir textos com elevado grau de coerência e estilização, recoloca a discussão em um terreno novo, mas não imune às categorias herdadas da filosofia e da teoria literária. Os modelos computacionais contemporâneos, ao passarem com desenvoltura por testes como o de Turing, são capazes de imitar os modos de organização discursiva próprios da linguagem humana. No entanto, a crítica filosófica, representada por John Searle em seu argumento da sala chinesa, revela que a capacidade de manipular signos segundo regras sintáticas não equivale à posse de compreensão semântica nem à presença de um sujeito consciente.

Essa distinção se articula com os limites do funcionalismo computacional, que, ao reduzir a inteligência à capacidade de processar informação, ignora a dimensão experiencial e fenomenológica da consciência. A ausência de um ponto de vista em primeira pessoa, de uma interioridade que se relaciona com o mundo de forma valorada e situada, compromete a possibilidade de atribuir aos sistemas de IA qualquer grau de intencionalidade autêntica. Embora os enunciados gerados por máquinas possam exibir complexidade estrutural e adaptação contextual, carecem do componente ético-enunciativo que caracteriza o discurso humano: não há um eu que responda, que assuma responsabilidade, que inscreva sua alteridade no tecido da linguagem.

É precisamente nesse ponto que a teoria da polifonia bakhtiniana oferece o critério mais exigente e epistemologicamente refinado para a avaliação da autoria. A produção textual de uma IA, ainda que bem operacionalizada e capaz de simular vozes, não atinge a condição polifônica no sentido pleno do termo, pois as vozes que apresenta não são dotadas de autonomia axiológica. Não há, na estrutura do modelo, uma pluralidade de consciências em interação responsiva. As vozes simuladas por esses sistemas não possuem horizonte de sentido próprio, tampouco emergem de posições ético-discursivas específicas. Trata-se, antes, de uma reprodução estatística de padrões, carente da dialógica genuína que fundamenta a autoria segundo Bakhtin.

Os sistemas de inteligência artificial generativa, especialmente os baseados em modelos de linguagem autorregressivos, como GPT (*Generative Pre-trained*

Transformer)¹⁰, operam fundamentalmente pela aplicação de processos probabilísticos sobre grandes conjuntos de dados linguísticos previamente acumulados. Tal operação estatística implica necessariamente a captura e reprodução de padrões discursivos dominantes nesses conjuntos, o que conduz à emergência de um texto representativo não da multiplicidade autêntica de vozes, mas sim de uma espécie de média ou normalização do discurso. Em consequência, ao contrário do ideal bakhtiniano da polifonia os sistemas generativos acabam, de maneira não intencional, mas estruturalmente inevitável, promovendo o silenciamento das vozes marginais, incomuns ou menos frequentes em suas bases de treinamento. Assim, enquanto superficialmente parecem expressar uma diversidade textual e discursiva, na realidade tais sistemas reproduzem aquilo que pode ser descrito como uma maioria estatística. Em termos discursivos, tais regularidades equivalem às formas mais recorrentes e convencionais do uso linguístico, as quais frequentemente correspondem às perspectivas culturalmente hegemônicas, socialmente majoritárias e ideologicamente dominantes na base textual utilizada no treinamento.

A criação literária ou filosófica engendrada por um sujeito humano apresenta-se indissociável de uma *hexis* existencial que projeta o texto no horizonte histórico dos valores e da responsabilidade. O autor comparece enquanto consciência enunciativa capaz de retirar cada signo da neutralidade lexical, investindo-o de intencionalidade axiológica que o inscreve num campo hermenêutico saturado de sentido. Esse gesto responsivo institui a autoria como agência, jamais como mera concatenação combinatória. Ainda que a obra incorpore polifonias estilísticas ou mimetize heteroglossias, é a dinâmica dialógica instaurada entre essas vozes que congrega tal multiplicidade em um todo coerente, orientado por imperativo ético compartilhado, sem recorrer à centralidade de um sujeito autoral unívoco. Os algoritmos generativos, embora operem com formidável complexidade estatística, permanecem alheios a qualquer

¹⁰ A abordagem probabilística subjacente à inteligência artificial generativa consiste na modelagem estatística do texto, fundamentada em cálculos de probabilidade condicional, nos quais a geração de cada novo *token* (palavra, símbolo ou unidade linguística mínima) depende diretamente das probabilidades extraídas de vastas quantidades de dados previamente assimilados. Modelos baseados em arquiteturas *Transformer*, por exemplo, constroem suas saídas textuais por meio da maximização das probabilidades condicionais, tendo como resultado direto que expressões linguísticas incomuns ou divergentes, sejam elas linguística ou culturalmente minoritárias, recebam probabilidades comparativamente mais baixas, ficando assim sub-representadas ou completamente excluídas do discurso gerado.

economia vivencial: registram sem experimentar, reiteram sem responder, compõem sem assumir.

Resulta daí que a categoria de autoralidade preserva vínculo inquebrantável com a experiência fenomenológica do sujeito e com a aptidão de instaurar alteridade dialógica efetiva. A configuração maquínica não hospeda tal alteridade no sentido bakhtiniano, visto que carece de consciência capaz de ocupar um lugar discursivo responsável diante do outro. Os textos produzidos por inteligência artificial qualificam-se, no máximo, como artefatos simulacionais de linguagem, privados dos fundamentos ontológicos imprescindíveis à voz autoral. A ausência de *pathos* experiencial converte esses enunciados em imitações sofisticadas, revelando o limite infranqueável entre a engenharia logotécnica e a autoria compreendida como acontecimento ético-dialogal.

REFERÊNCIAS

- BAKHTIN, Mikhail. O discurso no romance. In: BAKHTIN, Mikhail. *Questões de literatura e de estética: a teoria do romance*. Tradução de Aurora F. Bernardini, José Pereira Júnior, Augusto Góes Júnior, Helena S. Nazário, Homero F. de Andrade. 3. ed. São Paulo: Editora UNESP; Hucitec, 1993. pp. 71-210.
- BAKHTIN, Mikhail. *Problemas da poética de Dostoiévski*. 3. ed. Tradução direta do russo, notas e prefácio de Paulo Bezerra. Rio de Janeiro: Forense Universitária, 2002.
- BAKHTIN, Mikhail. Os gêneros do discurso. In: BAKHTIN, Mikhail. *Estética da criação verbal*. Introdução e tradução do russo de Paulo Bezerra. 4. ed. São Paulo: Martins Fontes, 2003. pp. 262-306.
- BAKHTIN, Mikhail. O problema do texto na linguística, na filologia e em outras ciências humanas: uma experiência de análise filosófica. In: BAKHTIN, Mikhail. *Estética da criação verbal*. Introdução e tradução do russo de Paulo Bezerra. 4. ed. São Paulo: Martins Fontes, 2003. pp. 307-335.
- BARTHES, Roland. A morte do autor. In: BARTHES, Roland. *O rumor da língua*. Tradução de Mário Laranjeira. São Paulo: Brasiliense, 1988. pp. 65-70.
- BENJAMIN, Walter. A obra de arte na era de sua reprodutibilidade técnica. In: BENJAMIN, Walter. *Magia e técnica, arte e política: ensaios sobre literatura e história da cultura*. Tradução de Sérgio Paulo Rouanet. Prefácio de Jeanne Marie Gagnebin. 3. ed. São Paulo: Brasiliense, 1987. pp. 165-196.
- BLOOM, Harold. *A angústia da influência: uma teoria da poesia*. Tradução de Marcos Santarrita. 2. ed. Rio de Janeiro: Imago, 2002.
- BRENTANO, Franz. *Psychology from an Empirical Standpoint*. Edição de Oskar Kraus e Linda L. McAlister. Trans. Antos C. Rancurello, D. B. Terrell & Linda L. McAlister. 2. ed. London: Routledge, 1995.

BURCKHARDT, Jacob. *A cultura do Renascimento na Itália: um ensaio*. Tradução de Vera Lúcia de Oliveira Sarmiento e Fernando de Azevedo Corrêa. Brasília: Editora Universidade de Brasília, 1991.

CONFÚCIO. *Os Analectos*. Tradução do inglês Caroline Chang; tradução do chinês, introdução e notas D. C. Lau. Porto Alegre: L&PM, 2023.

CROCE, Benedetto. *Estética como ciência da expressão e linguística geral: teoria e história*. Organização de Giuseppe Galasso. Tradução de Omayr José de Moraes Júnior. São Paulo: É Realizações, 2016.

DENNETT, Daniel Clement. *The Intentional Stance*. Cambridge, MA: MIT Press, 1989.

FOUCAULT, Michel. O que é um autor? In: FOUCAULT, Michel. *Ditos & escritos III: estética: literatura e pintura, música e cinema*. Organização e seleção de textos de Manoel Barros da Motta. Tradução de Inês Autran Dourado Barbosa. 2. ed. Rio de Janeiro: Forense Universitária, 2009. pp. 264-298.

FREUD, Sigmund. *Obras completas, volume 4: A interpretação dos sonhos (1900)*. Tradução de Paulo César de Souza. 1. ed. São Paulo: Companhia das Letras, 2019.

GOODFELLOW, Ian J.; POUGET-ABADIE, Jean; MIRZA, Mehdi; XU, Bing; WARDE-FARLEY, David; OZAI, Sherjil; COURVILLE, Aaron C.; BENGIO, Yoshua. Generative adversarial networks. *Communications of the ACM*, [S. l.], v. 63, n. 11, pp. 139-144, nov. 2020. DOI: [10.1145/3422622](https://doi.org/10.1145/3422622). Disponível em: <https://doi.org/10.1145/3422622>. Acesso em: 23 jul. 2026.

HOLQUIST, Michael. *Dialogism: Bakhtin and His World*. 2. ed. New York: Routledge, 2002.

HUSSERL, Edmund. *Investigações lógicas: prolegômenos à lógica pura*. Tradução de Diogo Ferrer. Rio de Janeiro: Forense, 2014. v. 1.

KURZWEIL, Ray. *A singularidade está próxima: quando os humanos transcendem a biologia*. Tradução de Ana Goldberger. São Paulo: Iluminuras, 2018.

MANDELKER, Amy (org.). *Bakhtin in Contexts: Across the Disciplines*. Introduction Caryl Emerson. Evanston: Northwestern University Press, 1995.

MARCHEZAN, Renata Coelho. A noção de autor na obra de M. Bakhtin e a partir dela. *Bakhtiniana: Revista de Estudos do Discurso*, São Paulo, v. 10, n. 3, pp. 186-204, set./dez. 2015. DOI: [10.1590/2176-457322365](https://doi.org/10.1590/2176-457322365). Disponível em: <https://doi.org/10.1590/2176-457322365>. Acesso em: 23 jul. 2026.

MORSON, Gary Saul; EMERSON, Caryl. Introduction: Rethinking Bakhtin. In: MORSON, Gary Saul; EMERSON, Caryl. *Rethinking Bakhtin: Extensions and Challenges*. Evanston: Northwestern University Press, 1989. pp. 1-60.

ONG, Walter Jackson. *Oralidade e cultura escrita: a tecnologização da palavra*. Tradução de Enid Abreu Dobránszky. Campinas, SP: Papirus, 1998.

RAJAGOPALACHARI, Chakravarti. *Mahabharata*. 46th ed. Mumbai: Bharatiya Vidya Bhavan, 2005.

RAJAGOPALACHARI, Chakravarti. *Ramayana*. 53rd ed. Mumbai: Bharatiya Vidya Bhavan, 2013.

SIMA, Qian. *The First Emperor: Selections from the Historical Records*. Trans. Raymond Dawson. Edição e introdução de K. E. Brashier. Oxford: Oxford University Press, 2009.

SEARLE, John Rogers. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, [S. l.], v. 3, n. 3, pp. 417-424, set. 1980. DOI: [10.1017/S0140525X00005756](https://doi.org/10.1017/S0140525X00005756). Disponível em: <https://doi.org/10.1017/S0140525X00005756>. Acesso em: 23 jul. 2026.

SHIKIBU, Murasaki. *O conto de Genji*. Tradução de Autri Books. [S. l.]: Autri Books, 2026. E-book.

TURING, Alan Mathison. Computing Machinery and Intelligence. *Mind*, [S. l.], v. 59, n. 236, pp. 433-460, out. 1950. DOI: [10.1093/mind/LIX.236.433](https://doi.org/10.1093/mind/LIX.236.433). Disponível em: <https://doi.org/10.1093/mind/LIX.236.433>. Acesso em: 23 jul. 2026.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Łukasz; POLOSUKHIN, Illia. Attention Is All You Need. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 31., 2017, Long Beach. *Advances in Neural Information Processing Systems 30*. Red Hook: Curran Associates, 2017. pp. 5998-6008. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Acesso em: 23 jul. 2026.

Recebido em 30/03/2025

Aprovado em 06/08/2026

Declaração de contribuição de cada autor

Ambos os autores contribuíram substancialmente para o manuscrito e atendem integralmente aos quatro critérios de autoria. Tarcísio de Souza Peres e Lígia Razera Gallo participaram da concepção e do delineamento do estudo, da análise e interpretação dos dados, da redação do artigo e da revisão crítica de seu conteúdo intelectual, da aprovação final da versão a ser publicada e da responsabilidade pela exatidão e integridade de todos os aspectos do trabalho. Suas contribuições foram globalmente equivalentes; contudo, Tarcísio teve envolvimento um pouco maior na concepção e no projeto do estudo, bem como na análise e interpretação dos dados, enquanto Lígia atuou mais intensamente na tradução e na revisão linguística do manuscrito.

Declaração de disponibilidade de conteúdo

Os conteúdos subjacentes ao texto da pesquisa estão contidos no manuscrito.

Pareceres

Tendo em vista o compromisso assumido por *Bakhtiniana. Revista de Estudos do Discurso* com a Ciência Aberta, a revista publica somente os pareceres autorizados por todas as partes envolvidas.

Bakhtiniana, São Paulo, 21 (3): e70992p, jul./set. 2026

Todo conteúdo de *Bakhtiniana. Revista de Estudos do Discurso* está sob Licença Creative Commons CC - By 4.0

Parecer I

O artigo em questão apresenta uma reflexão densa, original e de grande relevância teórica sobre os limites da autoria em textos produzidos por inteligência artificial, a partir da epistemologia bakhtiniana da polifonia. O título é bastante preciso e traduz o conteúdo do artigo, revelando tanto o referencial teórico central (Bakhtin e a polifonia) quanto o objeto de análise (autoria em textos gerados por IA). O objetivo é claro: compreender em que medida a produção textual por sistemas de IA pode ou não configurar-se como autoria plena, a partir do conceito de polifonia. O texto desenvolve esse propósito com extrema coerência, articulando argumentos teóricos, referências históricas e críticas pertinentes a abordagens clássicas (como o teste de Turing e o argumento da sala chinesa). A bibliografia é ampla, atualizada e pertinente, inclusive com contribuições contemporâneas da ciência da computação e da filosofia da mente. Há cuidado na articulação entre autores de diferentes áreas, sem comprometer a profundidade conceitual. O autor propõe um deslocamento teórico da análise técnica-funcional da IA para um critério ético-discursivo de autoria, o que representa avanço significativo no debate interdisciplinar. A ideia de aplicar a polifonia bakhtiniana como critério para delimitação da autoria em textos gerados por IA é inovadora, e o argumento é conduzido com rigor científico, especialmente na área das humanidades. Há densidade argumentativa e sofisticação estilística, mas sem comprometer a inteligibilidade do texto para leitores especializados. Recomenda-se, apenas, a revisão de alguns trechos extensos que poderiam ser reorganizados para favorecer a fluidez da leitura, ainda que sem prejuízo de conteúdo. O artigo poderia ainda contemplar, em suas considerações finais, as implicações éticas e políticas da substituição ou da simulação de autoria em contextos educacionais e editoriais. APROVADO COM SUGESTÕES [Revisado]

Elaine Conte – Universidade La Salle, Canoas, Rio Grande do Sul, Brasil;
<http://orcid.org/0000-0002-0204-0757>; elaine.conte@unilasalle.edu.br

Parecer emitido em 29 de abril de 2025.

Parecer II

Explicitação do objetivo do trabalho e coerência de seu desenvolvimento no texto. Com o objetivo de refletir sobre a instigante produção textual pelos sistemas de inteligência artificial contemporâneos - que passariam, facilmente por testes de desempenho linguístico imitativo e controlado, tal como o chamado teste de Turing, realizado pelo cientista da computação Alan Mathison Turing -, o artigo propõe, digamos, um teste de autoria; estabelece um recorte da questão tratando-a nos domínios da noção de autor, que, lembremos, compreende também uma noção de texto. Nesse caminho, expõe a constituição histórica da noção de autor, indicando várias de suas transformações políticas, religiosas, literárias, filosóficas e tecnológicas. Com foco no domínio filosófico e literário, destaca, especialmente, a reflexão sobre o autor na obra de M. Bakhtin, e seu conceito de dialogia e de polifonia. O artigo detalha a dimensão dialógica e axiológica dessa noção de autor, e sustenta que uma análise, com base nessa perspectiva, detectaria a ausência autoral nos textos gerados por inteligência artificial. O artigo traz, assim, tema interessante e apresenta boa argumentação; indicamos, portanto, sua publicação.

Bakhtiniana, São Paulo, 21 (3): e70992p, jul./set. 2026

Apresentamos, no entanto, duas recomendações. Uma delas é a inclusão de um breve trecho introdutório que trouxesse para o corpo do artigo seu objetivo, já indicado no resumo. A outra demanda um pouco mais (até mesmo alteração no título do artigo): trata-se de ajustar a noção de polifonia, de modo que o reconhecimento da dimensão dialógica da autoria não significasse também a exigência da produção de um texto polifônico – lembremos, em Bakhtin, a polifonia é definidora do gênero romanesco, em especial, do romance de Dostoiévski. APROVADO COM RESTRIÇÕES [Revisado]

Renata Coelho Marchezan – Universidade Estadual Paulista “Júlio Mesquita Filho” – UNESP, Araraquara, São Paulo, Brasil; <http://orcid.org/0000-0001-7453-1734>; renata.marchezan@unesp.br

Parecer emitido em 08 de agosto de 2025.

Parecer III

No parecer inicial ao artigo, indicamos sua publicação com duas recomendações. A versão final apresentada atende à primeira recomendação com a inclusão de dois parágrafos introdutórios e, a seu modo, dá também uma resposta à segunda recomendação, contornando a questão levantada, ao defender – com argumentação que acrescenta Mandelker e Morson – uma noção estendida de polifonia. Reafirmamos o interesse do artigo e de sua publicação. APROVADO

Renata Coelho Marchezan – Universidade Estadual Paulista “Júlio Mesquita Filho” – UNESP, Araraquara, São Paulo, Brasil; <http://orcid.org/0000-0001-7453-1734>; renata.marchezan@unesp.br

Parecer emitido em 11 de outubro de 2025.

Editores responsáveis

Luciana Salazar Salgado (Universidade Federal de São Carlos – UFSCar)

Lucia Santaella (Pontifícia Universidade Católica de São Paulo – PUC-SP)

Tony Berber Sardinha (Pontifícia Universidade Católica de São Paulo – PUC-SP)