

O QUE PODE A INTELIGÊNCIA ARTIFICIAL HOJE? UMA ANÁLISE DO FUNCIONAMENTO DA AUTORIA NO CHATGPT-3.5

Carolina Fernandes¹

O trabalho apresentado no evento XI SEAD analisou, a partir do dispositivo teórico-analítico da Análise de Discurso materialista (AD), o funcionamento da autoria em textos produzidos por Inteligência Artificial (IA), especificamente o ChatGPT-3.5, em língua portuguesa, considerando sua possível utilização pelo sujeito-enunciador na posição de aluno. Para tratar da escrita por inteligência artificial, fez-se necessário, num primeiro momento, refletir sobre sua materialidade técnica e digital (Paveau, 2021) assim como sua relação com a noção de sujeito discursivo e a categoria teórica de autor. Dessa forma, observei que o programa sendo produto tecnológico, resultado de um sistema de algoritmos ao qual está subordinado, não apresenta uma subjetividade, pois não se trata de um sujeito interpelado pela ideologia, não representa um lugar social no discurso já que não é efeito da relação da linguagem com a história.

Considerando a concepção de autoria na AD como o princípio de organização da dispersão discursiva que produz *textualidade* (Orlandi, 2008), isto é, efeito de unidade, discuti na apresentação, e mais detalhadamente em Fernandes (prelo), o desdobramento desse conceito nas noções de função-autor, posição-autor e posição-escrevente com base nos estudos de Eni Orlandi (2008, 2012), Solange Gallo (2001, 2023), Fernandes (2017), entre outros. Essas autoras apresentam a relação da função-autor com a produção do efeito de unidade textual, a posição-autor relacionada à tomada de posição do sujeito no gesto de interpretação e a posição-escrevente como sendo aquela que assume o sujeito-aluno ao produzir um texto com base na repetição parafrástica apenas a fim de cumprir uma tarefa escolar.

Ao observar a inscrição do sujeito no discurso a partir da tomada de posição ideológica, Pêcheux e Fuchs ([1975] 1997) estabelecem que o sujeito discursivo, sendo interpelado ideologicamente, assume posições conforme sua identificação com determinadas formações discursivas como é possível observar no texto da SD 1 (sequência discursiva):

SD1 - Conto de fadas brasileiro

Num morro encantado, onde fadas portavam armas e obrigavam pessoas a se calarem, vivia uma garota que tinha ganhado no Baile Funk o apelido de gorro rosa. Morena, fartas nádegas, brasileira e ingênua. Tinha treze anos. Sua mãe já achava que estava na hora de ela ir para a “casa da vovozinha”. Foi.

De dia brincava de princesa pelas ruelas da favela, à noite, “na casa da vovozinha”, era ninfeta e satisfazia os apetitosos lobos gringos e ricos.

(Texto de Matheus Simabukuro. *In*: Romão, 1999, p. 166).

¹ Doutora em Letras pela UFRGS, professora associada na Unipampa, líder do grupo de pesquisa Estudos Pêcheuxtianos, tutora PET-Letras com bolsa FNDE.

O texto trazido na SD1 foi produzido por um aluno do Ensino Médio e publicado no livro “A pulsação da palavra” (1999) que reúne textos resultantes do projeto realizado pelas pesquisadoras Lucília Romão e Soraya Pacífico (2006). Nele, vê-se a produção de sentidos para Chapeuzinho Vermelho deslocado de seu sentido mais recorrente, ou seja, essa versão não repete de forma parafrástica o conto clássico, mas, por meio do processo de polissemia, rompe com o pré-estabelecido a partir do lugar do sujeito-enunciador, jovem brasileiro conhecedor da desigualdade social que assola os grandes centros urbanos. Além do nome da personagem ser alterado para “gorro rosa” e o contexto da narrativa ser outro, o conselho da mãe sofre uma torção no texto produzido pelo aluno, o que produz o efeito de um “anticonto de fadas” como observam Pacífico e Romão (2006, p. 55), criando assim uma ambivalência nas imagens de “mãe” e “fadas” (como a fada madrinha no conto da Cinderela), visto que essas figuras seriam aquelas que deveriam cuidar da menina, e não a colocar em perigo. A figura do lobo permanece como sendo a do “lobo mau”, mas agora não remete a um animal como na versão dos irmãos Grimm, mas sim a homens que abusam da menina por meio de um mercado do sexo. Sua tomada de posição é de crítica social com relação às condições de abuso e prostituição a que as meninas faveladas estão submetidas. É possível observar uma identificação com uma formação discursiva de rejeição ao abuso infantil e ao capitalismo que produz essas condições precárias à população de baixa renda. Assim, a produção de sentido é de uma denúncia da situação de pobreza e desigualdade social por que passa o país no contexto de produção do texto.

Já o texto seguinte é uma versão produzida pelo ChatGPT, versão 3.5, da OpenAI, a quem foi demandada uma versão do conto clássico *A Chapeuzinho Vermelho* para o contexto do Brasil atual. Vejamos o que produz o programa na SD2:

SD2 - Maria e o Lobo Falso: Uma Aventura de Cuidado e Astúcia

Era uma vez uma menina chamada Maria, que morava em uma cidade no interior do Brasil. Um dia, a avó de Maria, que morava em uma cidade vizinha, ligou para ela pedindo que ela fosse visitá-la e levasse alguns mantimentos.

Maria prontamente pegou sua mochila e saiu de casa, seguindo pelo caminho que cortava a floresta. Enquanto caminhava, ela encontrou um homem suspeito que lhe perguntou para onde estava indo. Maria, que não queria confiar nele, deu uma resposta evasiva e seguiu em frente.

Mais adiante, Maria encontrou uma senhora muito simpática que lhe ofereceu algumas frutas. A menina aceitou a gentileza e seguiu seu caminho com a mochila mais pesada.

Finalmente, Maria chegou à casa da avó, mas notou que a porta estava aberta e havia algo estranho. Ao entrar, ela descobriu que a avó havia sido enganada por um lobo que se disfarçou de entregador de comida para entrar em sua casa e roubar seus pertences.

Maria conseguiu assustar o lobo e recuperar os pertences da avó, salvando o dia. A partir desse momento, ela prometeu nunca mais confiar em estranhos e sempre estar atenta a tudo ao seu redor.

E assim, Maria e a avó viveram felizes para sempre, lembrando-se da importância de ser cuidadoso e de confiar somente em pessoas que conhecem e confiam.

(Texto produzido pelo Chat-GPT-3.5)

É possível observar que o texto retoma o conto clássico, uma vez que mantém a sequência narrativa dos fatos: menina vai ao encontro da avó levar alimentos, encontra o lobo no caminho, o lobo chega na casa da avó antes da protagonista, ele é abatido e a idosa e sua neta ficam a salvo. Como elementos do contexto atual são modificados alguns significantes: o nome próprio Maria, muito comum no Brasil, é usado no lugar de “Chapeuzinho Vermelho”, ligação telefônica surge como o meio de comunicação com a avó, “mantimentos” é usado ao invés de “cesta de doces”, também “mochila” está no lugar de “cesta”, “entregador de comida” e “homem suspeito” remetem ao lobo mau, “cidade vizinha” também substitui o interior da floresta. No entanto, “floresta” e “lobo” ainda aparecem em outros trechos revelando certa contradição na consistência do efeito de atualidade que essa versão do conto deveria produzir. Ainda assim, o sentido de “maldade” para esse personagem é representado na ação de enganar e roubar, desse modo o lobo mau se torna um ladrão e não um animal que literalmente “devora” a vozinha.

A estratégia adotada pelo programa para produzir uma “versão do conto” é substituir certos significantes para melhor adequar o texto à demanda feita. Não há referências a outros textos, indicando uma intertextualidade. A menção à senhora que oferece frutas pode até lembrar a bruxa que dá a maçã envenenada à Branca de Neve em outro conto de fadas clássico, entretanto esse trecho não é significativo para a narrativa, já que o fato da mochila ficar mais pesada não implica em uma ação futura da personagem que poderia atingir o invasor com esse objeto. Aliás, a ação de “assustar o lobo e recuperar os pertences da avó” não é detalhada, como ela consegue essa façanha? A falta de descrição e detalhes das ações também comprometem o aparecimento de marcas de autoria.

Na construção dessa versão, percebo o predomínio do movimento parafrástico na tentativa de manter a referência ao texto original. Por outro lado, há indícios de uma polissemia quando a menina é que salva a avó e não um lenhador ou outro personagem que entraria na história para ganhar os créditos. Nesse ponto poderia haver a produção do sentido de “empoderamento feminino”, de que as mulheres são capazes de se defender sem a proteção de uma figura masculina. Mas não é esse o sentido dado como “moral da história”, o que fica é a lição sobre a “importância de ser cuidadoso e de confiar somente em pessoas que conhecem e confiam”. Essa lição não se conecta bem com o resto do texto, visto que, na narrativa do Chat: “Maria, que não queria confiar nele, deu uma resposta evasiva e seguiu em frente”, ou seja, ela não teria indicado onde morava a vó por desconfiança do “homem suspeito”. Logo, o arrependimento de Maria que “prometeu nunca mais confiar em estranhos e sempre estar atenta a tudo ao seu redor” não confirma o que é dito antes, causando uma contradição aparente no próprio nível enunciativo.

É um desafio encontrar marcas de autoria nesse texto sem condições de produção que remetam a algum contexto sócio-histórico, a um lugar social. A chapeuzinho que vira Maria não representa um lugar social, já que não podemos identificar um lugar de sujeito urbano, de moradora de cidade grande, de favela? Talvez por ter se defendido do lobo possa indicar algo de sua personalidade, que é corajosa, hábil ou forte (isso depende do modo como teria se livrado do lobo). A personagem Maria poderia ser qualquer moça, não representa um lugar de sujeito.

Desse modo, não temos indicação de uma posição-sujeito com relação a uma subjetividade. Como o sujeito-autor se posicionaria com relação à imagem de uma moça corajosa, perspicaz e habilidosa em sua defesa? Não temos como saber, o que prevalece nesse texto é o trabalho sobre o enredo, é o que importa já que a demanda era produzir uma versão para um conto. É o que é pedido, e é o que o ChatGPT entrega. Como um bom “respondedor” como o chama Gallo (2023, p. 92), o programa, entendo, executa a produção textual a partir de uma função-autor simulada que dá efeito de início, meio e fim à narrativa.

Ao observarmos que o ChatGPT organiza as informações dispersas no *database* em um único espaço significativo para produzir um efeito de homogeneidade textual, podendo até mesmo dar um direcionamento de sentidos ao seu texto dependendo da demanda, não haveria uma função-autor operando aí? Além do mais, ainda que não se possa reconhecer uma subjetividade da IA, sua produção simbólica apresenta uma aparente *textualidade*. E sendo o princípio da autoria “a origem da textualidade” (Orlandi, 2002, p. 75), não se pode negar que há *um efeito de autoria* que se apresenta em condições de produção específicas do *tecnodiscurso* (Paveau, 2021). Desse modo, é preciso retomar a afirmação de Orlandi (2002): “um texto pode até não ter um autor específico, mas, pela função-autor, sempre se imputa uma autoria a ele”.

Desse modo, entendo que, nos textos do ChatGPT, está em funcionamento uma função-autor como um *simulacro de autoria*. Simulacro é um conceito apropriado de Deleuze (1988) cuja definição diz respeito ao “efeito de semelhança forjado na representação por similitude [da falsa cópia]”, o que produz o “efeito de igual” (Fernandes, 2017, p. 238). Em trabalho anterior (Fernandes, 2017), compreendo o simulacro como um recurso autoral para a produção de sentidos, o que parece ser o caso do ChatGPT, já que é necessário, para sua produção linguística ser interpretável, apresentar textualidade e, conseqüentemente, efeito de unidade. Se pensarmos que, no campo da programação, as redes neurais artificiais simulam o funcionamento do cérebro humano, produzindo ligações entre processadores que imitam as sinapses entre neurônios (Brookshear, 2013), a imitação, ou cópia do sujeito-humano, está na base da “arquitetura de multiprocessamento” (Brookshear, 2013, p. 101) da IA. Dessa forma, parece-me plausível a existência de *um simulacro de autoria* como uma imitação da função discursiva de autor.

Ao analisar os recortes efetuados para essa pesquisa, compreendo o funcionamento da autoria no ChatGPT diferente de Gallo (2023) que formula a noção de “hiperautor” para representar uma função vazia de sujeito. Por outro lado, entendo que o multiprocessamento do programa ao emular a produção de linguagem humana, produz um efeito de “simulacro” de autoria. Esse *simulacro de autoria* produz efeito de unidade e consistência ao texto como o faz a função-autor em um nível enunciativo-discursivo (Gallo, 2001), por meio do predomínio da paráfrase. Entretanto, quanto à singularidade e criatividade, que surgem apenas com a polissemia (Fernandes, 2017) no nível discursivo (Gallo, 2001), o ChatGPT-3.5 apresenta dificuldades por não representar um lugar de sujeito no discurso, vacilando quanto ao forjamento de condições de produção, mostrando inconsistência, assim, na mobilização de arquivos de leitura, da memória discursiva e do contexto sócio-histórico. Sua produção é uma repetição parafrástica de textos que



estão em sua memória de dados, portanto, não há uma tomada de posição consistente que produza o efeito de um gesto de singularidade que possa revelar uma ascendência à posição-autor como podemos observar na produção em SD1 analisada em contraste com a produção da IA em SD2.

Por fim, observo que, no estado atual de desenvolvimento da arquitetura GPT, a autoria possível para o ChatGPT-3.5 é a de um *simulacro de autoria* que não apresenta um grau elevado em que se possa observar uma tomada de posição-autor, mas pode se colocar em um grau intermediário como faria um sujeito-aluno, por exemplo, ao reproduzir o discurso pedagógico na posição-escrevente.

REFERÊNCIAS

DELEUZE, Gilles. **Diferença e repetição**. Rio de Janeiro: Graal, 1988.

FERNANDES, Carolina. **O visível e o invisível da imagem**: uma análise discursiva da leitura e da escrita de livros de imagens. Campinas, SP: Mercado de Letras, 2017.

FERNANDES, Carolina. **A autoria em textos produzidos por inteligência artificial e por alunos em uma perspectiva discursiva**. No prelo. (Preprint em: <https://mail.google.com/mail/u/1/#inbox/FMfcgzGwJcdVCFcLrzsJfKGktkMFwTlf>).

GALLO, Solange M. L. Autoria: questão enunciativa ou discursiva? **Linguagem em (Dis)curso**, Tubarão, v.1, n. 2, p. 61-70, jan./jun 2001.

GALLO, Solange, M. L. ChatGPT: hiperautor ou não autor? **Traços de Linguagem**, v. 7, n. 1, p. 84-95, 2023.

ORLANDI, E. P. **Discurso e texto**: formulação e circulação dos sentidos. 3. ed. Campinas, SP: Pontes, 2008.

ORLANDI, E. P. **Discurso e leitura**. 9. ed. São Paulo: Cortez; Campinas, SP: Unicamp, 2012.

PACÍFICO, S. M. R.; ROMÃO, L. M. S. (org.). **A pulsação pela palavra**. Liceu Albert Sabin, 1999.

PACÍFICO, S. M. R.; ROMÃO, L. M. S. **Era uma vez uma outra história**: leitura e interpretação na sala de aula. São Paulo: DCL, 2006.

PAVEAU, M.-A. **Análise do discurso digital**: dicionário das formas e das práticas. COSTA, J. L.; BARONAS, R. L. (org). Campinas, SP: Editora Pontes, 2021.

PÊCHEUX, M.; FUCHS, C. [1975]. Por uma análise automática do discurso; atualização e perspectivas. *In*: GADET, F.; HAK, T. (org.). **Por uma Análise Automática do Discurso**: Uma introdução à obra de Michel Pêcheux. 3. ed. Campinas: Editora da UNICAMP, 1997. p.163-187.