

MARCAS DISCURSIVAS EM VÍDEOS DE *DEEP FAKES*: VENDENDO CLOROQUINA E A RELAÇÃO DO ATOR HUMANO E DO ATOR MAQUÍNICO NA ENUNCIÇÃO AUDIOVISUAL

HENRIQUE DA SILVA PEREIRA*

Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), Programa de Pós-Graduação em Comunicação (PPGCom), Bauru, SP, Brasil.


Recebido em: 23 jan. 2024. Aprovado em: 5 abr. 2024.

Como citar este artigo: PEREIRA, H. da S. Marcas discursivas em vídeos de *deep fakes*: Vendendo cloroquina e a relação do ator humano e do ator maquínico na enunciação audiovisual. *Cadernos de Pós-Graduação em Letras*, v. 24, n. 1, p. 193-210, jan./abr. 2024. DOI 10.5935/cadernosletras.v24n1p193-210

Resumo

Este artigo examina as marcas discursivas no vídeo *deep fake Vendendo cloroquina*, de Bruno Sartori, explorando o desdobramento e a articulação dos atores humano e maquínico na enunciação audiovisual. O vídeo, uma crítica satírica à resposta de Jair Bolsonaro à pandemia da Covid-19, emprega tecnologia *deep fake* para transformar o rosto do ex-presidente em um personagem de um programa de TV que vende suplementos alimentares. Este artigo explora como a análise das marcas discursivas e o contrato de veridicção contribuem para uma melhor compreensão desse fenômeno e suas implicações no âmbito do discurso

* E-mail: henrique.pereira@unesp.br

 <https://orcid.org/0000-0002-1328-8694>

audiovisual, enfatizando a natureza das competências necessárias para enunciatóres e enunciatários.

Palavras-chave

Inteligência artificial. *Deep fakes*. Marcas discursivas.

No cenário contemporâneo, a evolução tecnológica tem proporcionado transformações significativas na produção e disseminação de conteúdo audiovisual, desencadeando um fenômeno desafiador aos destinatários da comunicação: os *deep fakes*. Até recentemente, a criação de vídeos com efeitos especiais estava restrita ao domínio de produtoras profissionais. Contudo, a democratização da capacidade de manipulação digital é evidenciada pela disponibilidade acessível de computadores e componentes de *hardware* para usuários domésticos, permitindo a exploração de efeitos visuais, gráficos avançados e processamento de dados para algoritmos de inteligência artificial, em especial o aprendizado de máquinas. Os *deep fakes* representam uma manifestação dessa convergência, pois utilizam algoritmos de inteligência artificial, como Redes Generativas Adversárias (*Generative Adversarial Networks* – GANs), para sintetizar vídeos, nos quais faces e vozes podem ser simuladas de maneira convincente. Nesse contexto, emergem os *deep fakes*, uma expressão da interseção entre avanços tecnológicos e a manipulação sofisticada de imagens, com implicações que vão além do entretenimento, adentrando os domínios da desinformação, política e ética. O presente artigo explora, a partir do vídeo crítico intitulado *Vendendo cloroquina*, do *youtuber* brasileiro Bruno Sartori, como a análise das marcas discursivas dos *deep fakes* pode contribuir para melhor compreender o fenômeno.

VERIDICÇÃO E DEEP FAKES

Há pouco tempo, a possibilidade de criar vídeos com efeitos especiais capazes de “enganar” o espectador era exclusividade de produtoras audiovisuais profissionais, quase sempre atreladas aos estúdios de Hollywood. Dessas, destacam-se a Industrial Light & Magic, do cineasta George Lucas, fundada em 1975 para realizar os efeitos do primeiro filme da saga *Star Wars*, pioneira

a utilizar o *software* Photoshop; e a Pixar Animation Studios, fundada pelo empresário Steve Jobs em 1986. As duas, posteriormente, foram adquiridas pela The Walt Disney Company, em 2012 e 2006, respectivamente. Atualmente, em compensação, como apontado por Oliveira (2020, p. 37-40), usuários domésticos encontram disponíveis e com preços relativamente populares computadores e componentes de *hardware* que permitem o desenvolvimento de efeitos visuais, gráficos e, principalmente, o processamento do volume de dados que possibilitam a execução de algoritmos de aprendizado maquínico e demais metodologias de inteligência artificial.

Nesse contexto surgem os *deep fakes*. A partir de 2017, usuários anônimos de fóruns da internet, notadamente o *Reddit*, começaram a disponibilizar de forma pública vídeos pornográficos que supostamente representariam atrizes famosas (como Daisy Ridley, Gal Gadot ou Emma Watson) em momentos íntimos. Pouco tempo depois, a imprensa começou a registrar essas ocorrências e a determinar o caráter adulterado dos vídeos. Uma das primeiras reportagens a tratar sobre o tema foi da *BBC*, em fevereiro de 2018, em que o jornalista David Lee (2018, tradução nossa) apontava:

À medida que essas ferramentas se tornaram mais poderosas e fáceis de usar, elas permitiram a transferência de fantasias sexuais da imaginação das pessoas para a internet. Ele ultrapassa não apenas os limites da decência humana, mas também nosso senso de acreditar no que vemos e ouvimos¹.

Lee indicou que os enunciados desenvolvidos a partir dessa tecnologia emergente podem ser lidos no escopo da crença, uma vez que tratam de imaginários, portanto desejados pelos destinatários, e que são presentificados e disponibilizados de forma *on-line*. É relevante apontar isso, uma vez que a crença implica um contrato tácito entre os polos da comunicação. O destinador e o destinatário acordam em um *crer-verdadeiro* instituído pelo enunciado. Afastamos a verdade (do campo do ser do ser) da análise semiótica, mas observamos as estratégias discursivas implicadas na construção de um efeito de *parecer-real*. Greimas e Courtés (2016, p. 530, grifos nossos) explicam:

1 “As these tools have become more powerful and easier to use, it has enabled the transfer of sexual fantasies from people’s imaginations to the internet. It flies past not only the boundaries of human decency, but also our sense of believing what we see and hear.”

O crer-verdadeiro do enunciador não basta, supomos, à transmissão da verdade: o enunciador pode dizer quanto quiser, a respeito do objeto de saber que está comunicando, que “sabe”, que está “seguro”, que é “evidente”; nem por isso pode ele assegurar-se de ser acreditado pelo enunciatário: um crer-verdadeiro deve ser instalado nas duas extremidades do canal da comunicação, e é este equilíbrio, *mais ou menos estável*, esse entendimento tácito entre dois cúmplices *mais ou menos conscientes* que nós denominamos contrato de veridicção (ou contrato enunciativo).

Interessa-nos pensar na expressão “mais ou menos estável e conscientes” utilizada pelos autores nessa citação. O contrato de veridicção, como caracterizado por Greimas e Courtés (2016), baseia-se em uma relação de querer de ambos os sujeitos da enunciação. Esse revestimento modal implica um querer do sujeito 1 em fazer com que o sujeito 2 acredite no que lhe é comunicado; da mesma forma que o sujeito 2 quer (ou deve) assumir o objeto de saber comunicado. Os autores explicam: “o contrato aparece como uma organização de atividades cognitivas recíprocas que provocam a transformação da competência modal dos sujeitos em presença” (Greimas; Courtés, 2016, p. 100). Retornaremos a essa questão em seguida.

Voltando à reportagem da BBC, o autor levantava a hipótese do uso da tecnologia do *deep fake* para usos de difamação política e adulteração de discursos públicos. Não tardou muito para que esse cenário se concretizasse. Na internet, em um contexto de *fake news*, surgiram vídeos *fakes* em plataformas de compartilhamento social, como YouTube e outras redes, que tinham por finalidade manipular os destinatários da comunicação em um crer em algo mentiroso. Nesse cenário, políticos e pessoas públicas que eram representados em vídeos disponibilizados *on-line* realizando veridicamente ações reprováveis começaram a utilizar-se da desculpa de que esses vídeos divulgados não passavam de *deep fakes*. Isto é, os *deep fakes* alcançaram um poder de persuasão discursiva suficientemente eficaz a ponto de não conseguirmos determinar se um vídeo foi adulterado antes de disponibilizá-lo publicamente. É o caso, por exemplo, do ex-governador do estado de São Paulo, João Dória, que teve um vídeo íntimo divulgado em 2018 e, após o ocorrido, foi a público se defender declarando o vídeo como falso. Um laudo da Polícia Federal emitido em 2022, em compensação, apontou que o vídeo era verdadeiro e não teria sido adulterado.²

2 Matéria reportando o caso disponível em: <https://www1.folha.uol.com.br/poder/2022/03/laudo-da-pf-diz-nao-ter-encontrado-sinal-de-adulteracao-em-video-de-orgia-associado-a-doria.shtml>. Acesso em: 3 fev. 2022.

Nessa esteira, observamos que muitas das discussões presentes na mídia sobre *deep fakes* tratam de um olhar jornalístico sobre o registro audiovisual. Isto é, a partir de agências de checagem de informações, busca-se validar ou rechaçar o vídeo enquanto discurso “verdadeiro” ou “falso”. Em voga, começam a ser debatidas estratégias e estruturas tecnológicas para verificar uma suposta “adulteração” de um determinado produto virtual, como a utilização das ferramentas de assinatura digital conhecidas como *blockchain*, em que os arquivos deixam lastros que registram todas as ações realizadas em seu interior. Entendemos, entretanto, a partir do olhar teórico-metodológico da semiótica discursiva, que todo texto, digital ou analógico, não passa de um discurso que quer ser interpretado como veridictório, buscando, portanto, a persuasão.

Não há, a rigor, no campo semiótico textos verdadeiros – todos, sem exceção, são construtos discursivos; todos são produzidos e revestidos semântica e sintaticamente de dispositivos voltados para a verificação. Barros (2011, p. 52) esclarece: “os dispositivos empregados na produção do discurso servem também de meios de persuasão, utilizados pelo enunciador para convencer o enunciatário da ‘verdade’ do seu texto”. Dessa forma, um registro em lastro que informa a veracidade de um determinado discurso, como ocorre nos *blockchains*, sob o olhar epistemológico da semiótica, não implica uma verdade discursiva. Os textos, inclusive aqueles não adulterados (no sentido da edição do conteúdo do vídeo), são discursos marcados por intencionalidades. Ao considerarmos, para exemplificar, uma câmera de vídeo de segurança fixada no teto de um determinado local, o vídeo produzido por esse registro é também um texto marcado por uma construção de intencionalidade: o enquadramento, a escolha do plano, a movimentação angular da câmera... todos esses elementos da linguagem audiovisual são manipuláveis, ou melhor, são marcas discursivas deixadas no enunciado pelo enunciador, permitindo-nos reconstruir, a partir dessas marcas deixadas no discurso, o ato enunciativo. Entendemos, assim, que, ao pensar nos vídeos de *deep fakes*, faz-se necessário reconstruir a enunciação a partir das marcas deixadas de forma implícita e explícita no enunciado. Dessa forma, acreditamos que o discurso é colocado em um contexto geral de interdiscursividades em que podemos verificar não a verdade em si do vídeo, mas a circulação de significados projetados por esse texto.

Vale retomar que no discurso não há um eu ontológico que fala para um tu ontológico. Existe um enunciador que tem competência, no sentido semiótico,

para performar um discurso que abrange tanto o (1) colocar em discurso (passar do virtual ao manifestado), o (2) domínio técnico das linguagens, suportes e discursos prévios e (3) um saber sobre o enunciatário pressuposto – corroborando, portanto, um fazer-ser, um significar com intencionalidades, entre as quais parecer verdadeiro. Considerando os *deep fakes*, entendemos necessário destacar as competências tanto relativas ao domínio técnico como também a projeção do enunciatário no discurso – permitindo-nos, ao mesmo tempo, verificar as estratégias discursivas voltadas para a persuasão e a imagem pressuposta do enunciatário. Nesse escopo, podemos considerar que essas duas competências estão englobadas, necessariamente, a uma competência anterior: o domínio da linguagem audiovisual.

Fiorin (2015), considerando as línguas naturais, elenca possíveis competências que o sujeito enunciadador precisa performar para enunciar: competência linguística, referente ao conhecimento dos sistemas gramáticos e léxicos, a fim de produzir enunciados aceitáveis; competência discursiva, que implica a produção de enunciados concretos, figurativizados, tematizados, e situados em uma lógica spatiotemporal, além dos elementos argumentativos; competência textual, a fim de produzir os enunciados e dominar os suportes e as tecnicidades da produção de cada linguagem; competência interdiscursiva, referente às inter-relações heterogêneas que constituem o discurso; competência intertextual, relacionada às relações que o texto mantém com outros; competência pragmática, que diz respeito aos valores ilocutórios presentes no enunciado; e competência situacional, permitindo a compreensão da situação e do parceiro de comunicação.

Ao considerarmos uma competência discursiva de produção de textos audiovisuais, devemos levar em conta seu caráter sincrético, isto é, a manifestação de múltiplas linguagens no plano de expressão, que serão percebidas, a partir do plano do conteúdo, de forma global e homogênea pelo enunciatário. Essa, por sua vez, insere-se em uma lógica que recupera a matriz da perspectiva artificial inaugurada no *Quattrocento* para figuração de suas imagens. Isso implica, portanto, um efeito de fazer parecer mais real que outros aparatos de figuração anteriores. O aumento do grau de iconização das imagens a partir do Renascimento determina uma certa filosofia visual que indicia uma forma de ver o mundo figurativizado. Greimas (2004, p. 81), em “Semiótica plástica e semiótica figurativa”, aponta:

É evidente, porém, que a figuratividade, entendida como um certo modo de leitura – e um modo de produção – das “superfícies construídas”, não se encontra necessariamente ligada a normalidade qualquer e que ela pode dar lugar a excessos e insuficiências: o desejo de fazer-parecido – de fazer-criar – manifestado por este ou aquele pintor, por esta ou aquela escola, por esta ou aquela época leva, mediante a adjunção e sobrecarga de traços visuais, à iconização da pintura; já o despojamento das figuras com a finalidade de tornar mais difícil o procedimento de reconhecimento, não deixando transparecer – como acontece com a Improvisação de Kandinsky – senão “objetos virtuais”, dá lugar, ao contrário, à abstração. Iconização e abstração não são, pois, quanto à sua natureza, duas maneiras de pintar diferente da pintura figurativa; constituem antes graus variáveis da figuratividade, já que esta é determinada – repitamos – por um certo modo de leitura – frequentemente, mas não necessário – dos objetos planares construído.

Dessa forma, a iconização, como efeito de sentido da figuratividade, permanece com a introdução da imagem em movimento, com as diversas experiências cinematográficas – destacando-se o cinematógrafo dos Lumière – e posteriormente com a invenção da imagem em movimento eletrônica e digital. O parecer real imbricado na imagem fotográfica, cinematográfica e videográfica decorre de um parecer real que se desenvolve em uma suposta continuidade, isto é, um efeito de sentido de movimento (Médola; Pereira, 2023). É suposta porque tanto os meios analógicos quanto os eletrônicos e digitais realizam a continuidade a partir de estruturas técnicas descontínuas: os fotogramas e as linhas de varredura. Esse efeito de sentido do *continuum* consegue representar (e toda representação é uma redução) o sentir sinestésico do viver: a ação, o movimento, o tempo. Essa lógica de figuração baseada primeiramente em fotogramas no cinema e, posteriormente, em linhas de varreduras no vídeo refere-se a modos de colocar em discurso um enunciado que tem por intencionalidade parecer real, que busca representar no discurso, logo, descontínuo, a continuidade do mundo natural.

A descontinuidade discursiva implica um domínio das técnicas de manipulabilidade dos suportes e da linguagem. Estas vão coagir o enunciador determinando modos de enunciar implicados ao suporte e aos modos de fruição. É nesse contexto de discursivização que os *deep fakes* se inserem. A veridicção projetada no discurso ocorre na continuidade a partir de um efeito de sentido de movimento. Dessa forma, o enunciador do *deep fake*, resultante do desdobramento humano e maquínico, precisa ter competência para manipular o discurso sem que ocorra um rompimento do contrato enunciativo.

Nessa esteira, observa-se na enunciação de *deep fakes* a incidência de um saber-fazer voltado à articulação depreendida pelo actante da enunciação. O enunciador pressuposto de todo enunciado é uno, entretanto, no caso de alguns objetos semióticos, a actorialização da enunciação ocorre a partir de fazeres articulados de diferentes atores projetados no enunciado. No caso do audiovisual, sabe-se que os fazeres dos atores figurativizados e tematizados como “diretor”, “diretor de fotografia”, “diretor de arte”, “continuista”, entre outros, são combinados para a produção de um todo de sentido. No caso dos *deep fakes*, além dessa própria atoralização do audiovisual, evidencia-se que o enunciador precedentemente se desdobra e se projeta no enunciado a partir das marcas de dois atores, ator humano e ator maquínico, que, articulados em uma lógica de interação discursiva, produzem sentido. Isso se dá porque a lógica enunciativa do *deep fake* é resultante de um processo imbricado em algoritmos de inteligência artificial, que, a partir de determinadas metodologias de programação, como Redes Generativas Adversárias, conseguem simular a imagem e a sonoridade de determinado actante da narrativa representada. Nesse contexto, a máquina procede a partir da análise e decomposição de um *big data*, banco de dados com milhares de exemplares, e, a partir dessa avaliação, atua na composição de figuras que buscam simular aquela figura original.

Consideramos que o ator maquínico é responsável pela geração da imagem da face humana. Ao ator humano atribui-se a criação do banco de dados com imagens da pessoa que se quer produzir *deep fakes*, esta, todavia, não é sua única ação. Cabe ao ator humano, também, a combinação (mistura) *manual* da face gerada pelo algoritmo com o corpo original. Essa ação se desenvolve a partir de *softwares* de manipulação gráfica como o Adobe After Effects. Antes disso, entretanto, e considerando a lógica de produção audiovisual, o ator humano desenvolveu um roteiro audiovisual para que o vídeo fosse desenvolvido. A partir desse *script*, o ator humano pode produzir os elementos necessários para a realização do vídeo: gravação da cena com o corpo original ou seleção de vídeos já disponíveis *on-line* para substituição do rosto humano. Posteriormente, o ator maquínico realiza a análise da geometria da face do sujeito original para servir de parâmetro para a criação das imagens de rosto adulterado. Tendo as imagens geradas pelo algoritmo, o ator humano as combinará no vídeo original aplicando os efeitos visuais necessários para fomentar um efeito de sentido de parecer-real.

Tratamos, assim, das competências voltadas ao colocar em discurso, dos domínios da técnica, suportes e linguagem. Agora, destacamos a competência

modal do saber desempenhada pelo enunciador voltada, portanto, à projeção do enunciatário no discurso. Sabemos que o enunciador é aquele que sabe e quer – por algum motivo – comunicar esse conhecimento. Além do saber sobre o objeto de comunicação, o enunciador também precisa saber sobre o enunciatário da enunciação. Greimas e Courtés (2016, p. 171) explicam:

[...] o enunciatário não é apenas destinatário da comunicação, mas também sujeito produtor do discurso, por ser a “leitura” um ato de linguagem (um ato de significar) da mesma maneira que a produção do discurso propriamente dito. O termo “sujeito da enunciação”, empregado frequentemente como sinônimo de enunciador, cobre de fato as duas posições actanciais de enunciador e de enunciatário.

Nesse sentido, podemos entender o porquê de o sujeito da enunciação ser um sujeito complexo, este é o desdobramento dos dois polos do enunciado: o seu destinador e destinatário. O enunciador, sincretizado como destinador da comunicação, para comunicar o saber precisa deter um saber sobre o enunciatário, compreendido como destinatário da comunicação. Assim, o discurso enunciado carrega tanto as marcas axiológicas do primeiro quanto do segundo. O enunciado, produto de uma colocação em discurso, é referente a uma leitura de mundo do enunciador que projeta um enunciatário. A comunicação, dessa forma, apenas ocorre quando o enunciatário percebe no texto enunciado um referente de sua leitura do mundo, isto é, quando o enunciador consegue desempenhar uma *performance* que representa, em alguma medida, as crenças do enunciatário. Essa lógica permite que o discurso enunciado se posicione como instância mediadora entre esses sujeitos discursivos, já que é um ato de linguagem que exprime a interpretação do sujeito complexo da enunciação. Nessa esteira, ao destacarmos a expressão “mais ou menos” utilizada por Greimas e Courtés (2016) na acepção do termo veridicção no *Dicionário de semiótica*, vista anteriormente, objetivamos demonstrar que a relação entre os sujeitos da enunciação e o enunciado não é totalmente estável e consciente, uma vez que a competência para enunciar passa por uma leitura de mundo e interpretação, para que assim possa ser realizada em discurso.

No caso do objeto desta pesquisa, a competência da interpretação de mundo passa por desempenhar um discurso que se insere em um modo de produção audiovisual que se desenvolve há mais de um século, em que se verificam uma gramática do audiovisual e determinados modos de fruição. O enunciador

e enunciatário compartilham desses saberes, entretanto a relação não é estável: o dinamismo das técnicas, das tecnologias e da linguagem determina novas estratégias discursivas que são colocadas em prática a cada novo enunciar, que por sua vez podem ser sancionadas, positiva ou negativamente, por aqueles que o fruem. Nesse cenário, o vídeo de *deep fake* se coloca como um desafio, sobretudo discursivo, para o sujeito da enunciação. Tanto o enunciador quanto o enunciatário demandam competências modais e pragmáticas para manipular, no caso do enunciador, e interpretar, no caso do enunciatário. O manipular exige um fazer interpretativo por parte do enunciador, já que este – conforme já discutido – interpreta o mundo no crivo de leitura do enunciatário para assim realizar o discurso.

ANÁLISE: VENDENDO CLOROQUINA

O vídeo *Vendendo cloroquina*, criado pelo *youtuber* Bruno Sartori, foi lançado em 9 de julho de 2020, quatro meses após o início do *lockdown* de isolamento decorrentes da pandemia de Covid-19. Sua distribuição ocorreu por meio das redes sociais YouTube e Twitter (atualmente X), além de disponibilizado para *download* direto do material. Atualmente, o vídeo não se encontra mais disponível no YouTube, tendo sido bloqueado pela rede social alegando infração das diretrizes de comunidade da plataforma de vídeos.³ Temos nesse vídeo uma crítica irônica sobre a atuação do então presidente da República Federativa do Brasil, Jair Messias Bolsonaro, à frente da crise sanitária de combate ao vírus letal.

Vendendo cloroquina é espacializado em um programa de televisão da emissora SBT, *Casos de família*. A premissa do programa é mediar problemas familiares com a presença de especialistas de diferentes áreas, como psicólogos, advogados, entre outros. Muito criticado por ser considerado sensacionalista, *Casos de família* foi um sucesso de audiência (alcançando média de 3 pontos de Ibope no período vespertino) e engajamento nas redes sociais (diversos vídeos no YouTube contam com milhões de visualizações). Um dos momentos mais

3 Não há uma justificativa por parte do YouTube dos motivos do bloqueio do vídeo. Provavelmente, o vídeo tenha sido removido da rede social devido à utilização de imagens com direito autoral da emissora de televisão SBT. Atualmente, pode-se acessar o vídeo no Twitter. Disponível em: <https://twitter.com/brunosarttori/status/1281207150111526919?s=20>. Acesso em: 14 jan. 2023.

icônicos do programa era o *merchandising* realizado pela TopTherm, empresa de suplementos alimentares de propriedade de Aracy Wolf, também propagandista dos produtos da marca.

No vídeo produzido por Sartori, vemos a apresentadora Christina Rocha ao lado de uma figura humana feminina que metamorfoseia Jair Bolsonaro e a vendedora Aracy Wolf. O vídeo situa-se no momento de anúncio publicitário do programa de televisão. Certamente, um momento nada convencional para um presidente, a venda de suplementos alimentares.

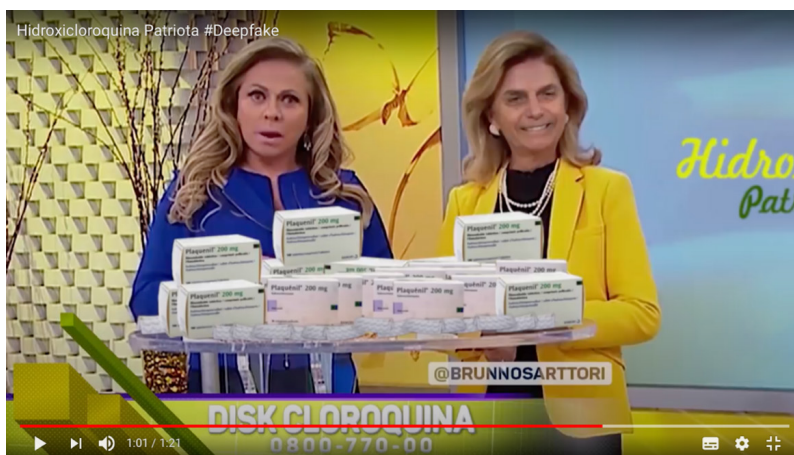


Figura 1 – Captura do vídeo de *deep fake* *Vendendo cloroquina #Deepfake*”

Fonte: Captura do vídeo realizada pelo autor.

Ao analisarmos atentamente, percebemos que a actorialização da figura metamorfoseada carrega marcas discursivas das duas pessoas que está representando: o timbre característico da voz da vendedora com a voz grave do presidente; o estilo do discurso de venda de Aracy (“Isso mesmo, Christina!”; “É bom para o coração, para a memória...”) com o universo semântico do presidente (hidroxicloroquina, Exército, população brasileira) e, o mais evidente, o corpo e as vestimentas da vendedora com o rosto de Jair Bolsonaro. Essas formas, sobrepostas e percebidas de forma global, são transformadas em uma nova figura: Aracy-Bolsonaro. Essa nova figura é marcada por referenciar dois macrodiscursos:

1. O discurso das constantes participações de Aracy em programas populares de auditório, apresentando e vendendo os suplementos alimentares de sua

- marca como produtos que operam em antienvelhecimento e melhora do sistema imune, atrelando valores como melhor qualidade de vida e saúde.
2. O discurso do ex-presidente Jair Bolsonaro em que, mesmo contrariando todas as recomendações científicas da área da saúde, decidiu produzir em larguíssima escala o composto hidroxicloroquina como remédio que supostamente agiria de forma eficaz contra a Covid-19.

Dessa forma, graças à metamorfose de *deep fake*, o vídeo equipara o discurso do líder do Poder Executivo de um país e suas decisões na área de saúde com uma vendedora de suplementos alimentares ineficazes contra qualquer doença. Entendemos que, para o enunciatário compreender o sentido crítico, irônico e cômico do vídeo, é necessário que este tenha uma competência interdiscursiva dos dois macrodiscursos resumidos já mencionados. Para assimilar o sentido do texto, o enunciatário relaciona o discurso de Aracy, o discurso de Jair Bolsonaro e o vídeo de *deep fake*.

Considerando as modalidades veridictórias, o enunciado com a figura metamorfoseada de Aracy-Bolsonaro não carrega marcas de um parecer-verdadeiro. O enunciador age considerando que o enunciatário tem competência semântica sócio-histórica para compreender que o vídeo produzido é um vídeo editado, manipulado. Percebe-se isso pelas marcas deixadas no enunciado, o título do vídeo tem a marcação “#deepfake” servindo como alerta caso alguém não seja competente nas interdiscursividades requeridas no texto, e, além disso, o vídeo é assinado “brunosarttori”, reafirmando seu caráter manipulado. Assim, o enunciador desenvolve um enunciado em que o não parecer é instaurado pelos recursos tecnológicos dos algoritmos de inteligência artificial. Já o enunciatário, munido das competências anteriormente descritas, exerce um fazer interpretativo que compreende o tom jocoso do vídeo e atesta o estatuto de não ser, modalizando, portanto, um vídeo em que a falsidade é construída.

Podemos considerar, ainda, uma segunda leitura. O enunciador e o enunciatário tacitamente concordam que o vídeo retrata uma realidade falsa, entretanto – ponderando o texto em seu sentido mais abstrato, em que temos uma venda de um composto químico que não atua de forma eficaz e não corrobora uma melhor qualidade de vida –, percebe-se que essa abstração diz respeito a episódios constantes no discurso público do presidente Jair Bolsonaro. O discurso, presente na mídia jornalística ao longo de 2020 e 2021 e em suas próprias redes sociais, foi marcado pela insistência dele em relação ao uso do

medicamento não reconhecido pela comunidade científica. Mesmo que o vídeo se comporte enquanto uma falsidade, o seu subtexto, em um movimento de abstração, pode ser compreendido como inspirado em um discurso verdadeiro, isto é, inspirado em um parecer que encontra eco no ser da vendedora de suplementos. E é nessa mudança de isotopia, ou seja, de leitura, que o humor acontece. Em perceber que mesmo o discurso falso carrega uma verdade profunda que articula o *Zeitgeist*, diversos outros discursos e axiologias próprias.

Podemos identificar no vídeo determinados elementos figurativos que se relacionam com a temática apresentada na sátira do *deep fake* e que corroboram, seguindo o estilo estético da TV, o fato de o discurso ser interpretado como parecer-verdadeiro. Em outras palavras: podem-se verificar elementos da linguagem televisual que foram adicionados ao vídeo e que nos permitem perceber que o enunciador de *Vendendo cloroquina* se municia de competência discursiva para fazer-ser o texto, em que o enunciatário foi interpretado como um conhecedor dos interdiscursos dispostos no enunciado.

Observa-se o uso dos seguintes elementos adicionados pelo enunciador, além da figura metamorfoseada de Aracy-Bolsonaro, que compõem o discurso de *Vendendo cloroquina*:

1. Uma base de gerador de caracteres (GC) com a legenda: “Disk Cloroquina 0800-770-00”.
2. Uma bancada com caixas do composto hidroxiclороquina, da marca “Plaquenil”.
3. Um telão no cenário em que está escrito “Hidroxiclороquina Patriota” nas cores amarela e verde, indiciando a bandeira do Brasil.
4. Ao longo do vídeo, observa-se que há cortes na sequência que produzem o efeito de divisão de duas cenas: o estúdio com a apresentadora Christina Rocha e a metamorfose Aracy-Bolsonaro, e *inserts* em que aparecem apenas caixas do fármaco Plaquenil.

Voltando o olhar para o desempenho do ator maquínico na enunciação, é relevante identificar dois pontos: as variações de resolução ao longo do vídeo e a figurativização dos elementos que compõem a face humana. No tocante ao primeiro, vê-se que *Vendendo cloroquina* se utiliza de variações de resolução dos elementos que compõem a imagem do vídeo. Primeiramente, observemos a feição da figura Aracy-Bolsonaro (Figura 2).



Figura 2 – Detalhe da feição de Aracy-Bolsonaro do *deep fake* *Vendendo cloroquina #Deepfake*

Fonte: Montagem realizada pelo autor.

Conseguimos identificar duas variações de resolução na composição da figura metamorfoseada de Aracy-Bolsonaro. Quando se observa a segunda imagem (estrutura do busto), percebe-se que está com qualidade inferior se comparada à terceira imagem (face de Bolsonaro). Isso fica evidente ao se atentar ao penteado de Aracy, já que na parte superior é possível reparar o cabelo com baixíssima resolução, ao passo que o rosto do político é dotado de maior qualidade visual. Além desse exemplo, podemos encontrar essa variação de resolução em outros momentos do vídeo:

1. Entre a resolução das caixas de Plaquenil inseridas na pós-produção e o vídeo original.
2. Entre o GC e o vídeo original.
3. Entre os *inserts* dos remédios e o vídeo original.

Vale pontuar, como discutido anteriormente, que esses três itens do vídeo dizem respeito às ações do ator humano – mas que são relevantes ao considerarmos as marcas figurativas presentes no enunciado do vídeo –, pois a baixa definição se traduz como uma marca de textualização que corrobora a leitura de que se trata de um não-parecer-verdadeiro.

Agora, considerando a figurativização dos elementos que compõem a face humana no *deep fake*, os olhos, ou melhor, os movimentos dos olhos, demonstram ser o mais complexo de se produzir artificialmente por uma inteligência artificial. Para melhor explicar esse ponto, desenvolvemos a seguinte montagem (Figura 3).

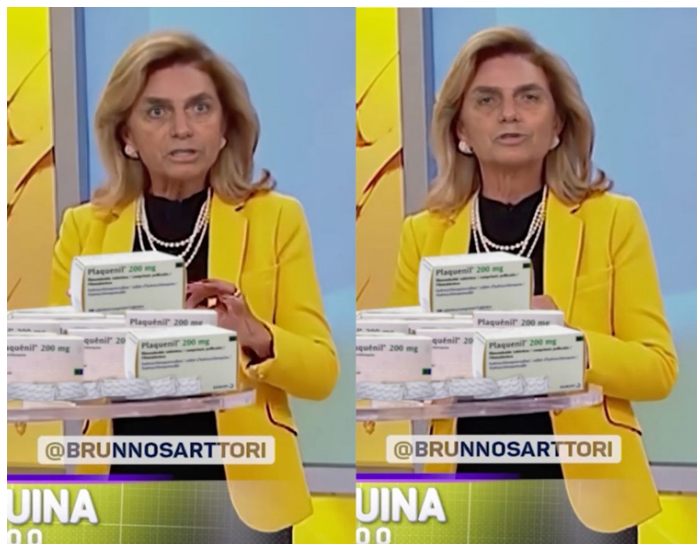


Figura 3 – Comparação movimento ocular de Aracy-Bolsonaro do *deep fake* *Vendendo cloroquina* #Deepfake

Fonte: Montagem realizada pelo autor.

Os movimentos oculares de Aracy-Bolsonaro não ocorrem de forma natural, isto é, não aparentam ser verdadeiros. Em poucos *frames* de distância, podemos observar que os olhos alteram seu foco. Na primeira imagem (aproximadamente em 00:35), vemos os olhos direcionados para o lado direito do quadro (esquerdo da personagem); já o segundo (aproximadamente em 00:36) volta-se para o centro do quadro, observando o enunciatário do vídeo – típico da linguagem televisual, em que apresentadores fitam a câmera criando um efeito de sentido de “conversar” com o espectador. Mais do que isso, sabe-se que a utilização de *teleprompter*⁴ os obriga a olhar diretamente para a câmera. Na seguinte imagem, demarcamos com um traço o sentido de observação de Aracy-Bolsonaro (Figura 4).

4 Equipamento para leitura de roteiro em programas não ficcionais.

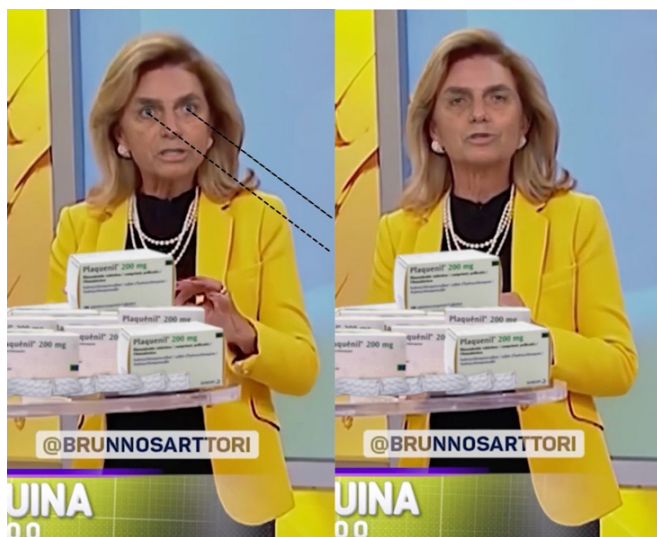


Figura 4 – Trajetória do movimento ocular de Aracy-Bolsonaro do *deep fake Vendendo cloroquina* #Deepfake

Fonte: Montagem realizada pelo autor.

CONSIDERAÇÕES FINAIS

A análise do vídeo *Vendendo cloroquina* à luz das marcas discursivas presentes no enunciado e a análise das competências modais, pragmáticas e interdiscursivas dos atores da enunciação nos permitem compreender a complexidade desse fenômeno discursivo. Os *deep fakes*, enquanto manifestação da convergência entre avanços tecnológicos e a manipulação sofisticada de imagens, desafiam não apenas a análise do contrato veridictório dos discursos, mas também as competências interpretativas dos enunciatários.

Comprova-se que o ator humano desempenha ações na enunciação que vão além da utilização da inteligência artificial. Assim, ao enunciar esses elementos figurativos que compõem o discurso do vídeo, o enunciador gera um enunciado que sustenta a geração de sentido desenvolvida nos níveis anteriores: o narrativo e o fundamental. Isto é, os elementos figurativos e temáticos analisados anteriormente revestem semanticamente a sintaxe que organiza o sentido do texto. Dessa forma, podemos afirmar que a vetorização e manutenção de sentido são desempenhadas pelo ator humano.

Além de os movimentos dos olhos da figura não conseguirem se assemelhar com o parecer-verdadeiro, pelo crivo de leitura do produto, vídeo cômico, o enunciatário pode aceitar determinadas imperfeições, uma vez que sabe que o vídeo não é verdadeiro. Outro ponto importante no que diz respeito à manutenção da veridicção do *deep fake* é a sincronia entre as falas das personagens e o movimento labial. Nesse vídeo, é possível verificar diversos momentos de assincronia – mas, novamente, por estar inserido em uma lógica jocosa, assume-se tal condição do fruir estético. A tecnologia de *deep fake*, em compensação, avança para aperfeiçoar tais problemas do campo da figurativização da figura humana no texto audiovisual.

Por fim, a análise do vídeo em questão destaca a interação entre ator humano e ator maquínico na produção de *deep fakes*. A competência técnica exigida para criar e combinar as imagens geradas pelo algoritmo com elementos humanos, como no caso de Aracy-Bolsonaro, ressalta a necessidade de compreendermos não apenas a linguagem audiovisual, mas também a interação entre humanos e máquinas na produção de discursos contemporâneos. A emergência dos *deep fakes* como fenômeno discursivo demanda uma reflexão contínua sobre as implicações éticas, políticas e culturais dessas práticas, incentivando uma abordagem crítica e informada por parte dos enunciatários e produtores de conteúdo.

Discursive marks in deep fake videos: *Selling chloroquine* and the relationship between the human actor and the machine actor in audiovisual enunciation

Abstract

This article examines the discursive marks in Bruno Sartori's deep fake video *Selling chloroquine*, exploring the deployment and articulation of human and machine actors in audiovisual enunciation. The video, a satirical critique of Jair Bolsonaro's response to the Covid-19 pandemic, uses deep fake technology to transform the former president's face into a character from a TV show selling food supplements. This article explores how the analysis of discursive marks and the contract of veridiction contributes to a better understanding of this phenomenon and its implications in the context of audiovisual discourse, emphasizing the nature of the skills required for enunciators and enunciated.

Keywords

Artificial intelligence. Deep fakes. Discursive marks.

REFERÊNCIAS

BARROS, D. L. P. *Teoria semiótica do texto*. São Paulo: Ática, 2011.

FIORIN, J. L. *Introdução à linguística I: objetos teóricos*. 6. ed. São Paulo: Contexto, 2015.

GREIMAS, A. J. Semiótica plástica e semiótica figurativa. In: OLIVEIRA, A. C. de (org.). *Semiótica plástica*. São Paulo: Hacker Editores, 2004. p. 75-96.

GREIMAS, A. J.; COURTÉS, J. *Dicionário de semiótica*. Vários tradutores. 2 ed. São Paulo: Contexto, 2016.

LEE, D. Deepfakes porn has serious consequences. *BBC*. 2018. Disponível em: <https://www.bbc.com/news/technology-42912529>. Acesso em: 3 jan. 2022.

MÉDOLA, A. S. L. D.; PEREIRA, H. da S. O efeito de sentido do movimento no audiovisual: um estudo sobre o cinetismo como formante da expressão a partir da vinheta Intolerância, da Globo News. *Estudos Semióticos*, v. 19, n. 2, p. 249-265, 2023.

OLIVEIRA, B. J. de. *Comunicação e sentido no audiovisual interativo para os meios digitais: estratégias enunciativas na construção de espaços, tempos e atores do discurso*. 2020. Tese (Doutorado em Comunicação) – Universidade Estadual Paulista, Bauru, 2020.