

Inteligência artificial generativa, ontologia da linguagem e inovação educacional progressista: revisão sistemática sob a perspectiva da gramática de construções baseada no uso e da sociocognição

RESUMO

Rebeca Emerich Alvarez
rebecaealvarez@gmail.com
<https://orcid.org/0009-0003-4978-7174>
Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, Rio de Janeiro, Brasil

O objetivo deste artigo é examinar como a literatura científica indexada entre 2020 e 2026 mobiliza ou silencia perspectivas *usage-based*, construcionais e sociocognitivas na análise da inserção de modelos de linguagem de grande porte (LLMs) em práticas educacionais, problematizando a racionalidade inovação-positividade que estrutura o debate contemporâneo. Para isso, foi realizada uma revisão sistemática orientada pelo protocolo PRISMA 2020, conduzida em seis bases (Scopus, Web of Science, ERIC, SciELO, LILACS, Lens.org), articulada a matriz de codificação temática indutiva. O *corpus* consolidado de 75 estudos distribui-se em quatro categorias temáticas, com hegemonia da racionalidade instrumental (50,7%) e marginalidade do aporte linguístico-cognitivo crítico (6,7%). Os dados sugerem padrão de sub-representação acentuada das tradições funcional-cognitivas em contexto educacional. O quadro empírico sustenta a contribuição original do estudo: a proposição de pedagogia linguística crítica, em diálogo com a Linguística Aplicada Crítica e os Estudos de Letramento, que reposiciona os LLMs como objeto de letramento progressista, ancorada na confluência entre linguística cognitiva, sociocognição brasileira, crítica filosófico-epistemológica e leitura freireana da técnica.

PALAVRAS-CHAVE: Inteligência artificial generativa. Modelos de linguagem. Linguística cognitiva. Sociocognição. Inovação educacional progressista.

INTRODUÇÃO

A inovação educacional ascendeu, nas últimas décadas, à condição de significativo privilegiado nos discursos pedagógicos (Demo, 2010; Messina, 2001; Hofman *et al.*, 2017), operando sob lógica que Tavares (2019) denomina racionalidade *inovação-positividade*: operação discursiva que, ao naturalizar a inovação como pressuposto axiológico, oculta as racionalidades político-econômicas subjacentes. Kassar e Oliveira (2024) demonstram como práticas e políticas educacionais mobilizam o léxico inovador como invólucro retórico para projetos atravessados pelo neoliberalismo (Dardot; Laval, 2017), pelas didáticas neotecnistas (Gonzalez, 2024) e pelos usos acríticos da tecnologia (Maruyama, 2013).

A consolidação dos modelos de linguagem de grande porte (*Large Language Models*, doravante LLMs), deflagrada pelo lançamento do GPT-3 em 2020 e amplificada pela popularização do ChatGPT em 2022, configura caso paradigmático dessa racionalidade. Ferramentas como ChatGPT, Gemini, Claude e Llama foram incorporadas ao espaço escolar mediante discursos que celebram personalização, automação de *feedback* e democratização do acesso (Kasneci *et al.*, 2023; Yigci *et al.*, 2025). Revisões recentes (Wu; Dang; Li, 2025) identificam crescimento exponencial da produção sobre IA generativa em educação, com prevalência de abordagens instrumentais centradas em desempenho.

A ontologia da linguagem subjacente a esses artefatos permanece subteorizada no debate educacional. Bender e Koller (2020) sustentam que um sistema treinado exclusivamente sobre a forma carece, *a priori*, de mecanismo para aprender significado, definido como relação entre forma linguística e intenção comunicativa. Bender, Gebru, McMillan-Major e Shmitchell (2021) caracterizam os LLMs como sistemas que costuram sequências de formas conforme informação probabilística de coocorrência. A perspectiva sociocognitiva e *usage-based*, que concebe a linguagem como prática situada, corporificada e intersubjetivamente distribuída (Goldberg, 2006; Bybee, 2010; Tomasello, 2003; Salomão, 2009; Marcuschi, 2007), oferece arsenal conceitual para problematizar essa apropriação: articulação que permanece, contudo, marginal nos estudos brasileiros e internacionais.

Diante desse quadro, o artigo responde à seguinte pergunta condutora: de que modo a literatura científica, no recorte 2020-2026, mobiliza ou silencia perspectivas *usage-based*, construcionais e sociocognitivas na análise da inserção dos LLMs em práticas educacionais? O objetivo central consiste em mapear, mediante revisão sistemática orientada pelo protocolo PRISMA 2020 (Page *et al.*, 2021), a produção indexada em seis bases (Scopus, Web of Science, ERIC, SciELO, LILACS e Lens.org), identificando padrões epistemológicos, lacunas teóricas e inflexões críticas. A contribuição original radica-se na proposição de horizonte pedagógico progressista que, ao recusar simultaneamente o determinismo tecnológico e o pessimismo paralisante, reposiciona a IA generativa como objeto de letramento crítico ancorado em ontologia coerente com as tradições funcional-cognitivas.

APORTE TEÓRICO

LINGUAGEM COMO PRÁTICA CONSTRUCIONAL, SITUADA E SOCIALMENTE DISTRIBUÍDA

A tradição *usage-based*, consolidada nas últimas três décadas no horizonte da Linguística Cognitiva, rompe com a partição saussureana entre *langue* e *parole*, com a dicotomia chomskyana entre competência e desempenho e com o postulado da modularidade. Para Goldberg (1995, 2006), a unidade fundamental da gramática reside na construção: pareamento convencionalizado entre forma e função, distribuído ao longo de um *continuum* que se estende do morfema à estrutura argumental. Bybee (2010) demonstra que a frequência de uso esculpe, de modo cumulativo e probabilístico, a representação cognitiva da gramática, dispensando regras autônomas em favor de modelo exemplarista.

Tomasello (2003) evidencia que a construção emerge do entrelaçamento entre atenção conjunta, intenção compartilhada e leitura intersubjetiva, capacidades sociocognitivas anteriores ao domínio sintático. Diessel (2019) sistematiza esse aparato em rede dinâmica de nós construcionais conectados por relações verticais e horizontais, perpetuamente reconfiguradas pelo uso.

Esse aparato dialoga, em chave brasileira, com a tradição sociocognitiva inaugurada por Salomão (1999, 2009) na recepção crítica da *Frame Semantics* fillmoreana. Salomão concebe a significação como atividade interativa de produção referencial, ancorada em *frames*: estruturas conceituais sociocognitivamente partilhadas que configuram o solo experiencial da compreensão. Marcuschi (2007) problematiza a referência como construto colaborativo; Koch (2002) desloca o texto da condição de produto para a de processo cognitivo-discursivo em ato; Morato (2010) demonstra a fertilidade do conceito de *frame* tanto em discursos cotidianos quanto em interações clínicas marcadas por afasia.

A convergência entre os paradigmas *usage-based*, construcional e sociocognitivo sustenta uma ontologia da linguagem cujos predicados se tensionam em face da tradição formalista: a linguagem é prática situada, corporificada, distribuída, emergente. Aquino e Palumbro (2019), ao articularem essa perspectiva com o ensino de argumentação, evidenciam o potencial pedagógico de tal ancoragem: aprender uma língua significa tornar-se membro de comunidade de prática mediante apropriação progressiva de redes construcionais que carregam, em sua materialidade, sedimentações culturais, ideológicas e afetivas.

A ONTOLOGIA DESCORPORIFICADA SUBJACENTE AOS MODELOS DE LINGUAGEM DE GRANDE PORTE

Os LLMs contemporâneos operam sobre arquitetura *transformer* (Vaswani *et al.*, 2017), treinada em *corpora* textuais de escala exorbitante mediante predição de *tokens* mascarados. O desempenho desses modelos em *benchmarks*

sugere convergência com o paradigma *usage-based*: tanto a arquitetura computacional quanto a aquisição infantil partem do uso e da extração estatística de regularidades distribucionais. Manning *et al.* (2020) evidenciam que estruturas hierárquicas sintáticas emergem espontaneamente nos *transformers* via mecanismos de atenção. Weissweiler *et al.* (2023) e Tayyar Madabushi, Divjak e Milin (2022) inauguram o *constructional probing*, mostrando que informação construcional se torna acessível às representações vetoriais dos LLMs. Dunn (2022) demonstra que gramáticas construcionais aprendidas computacionalmente convergem entre registros mediante exposição ampliada.

Cumprir registrar o contra-argumento que circula no campo. Piantadosi (2024) sustenta que o sucesso empírico dos LLMs em capturar regularidades sintáticas e construcionais constituiria evidência contra a Gramática Universal chomskyana e a favor de modelos baseados em uso, sugerindo comensurabilidade entre LLMs e tradição *usage-based*. A tese encontra refutação em três frentes: a captura distribucional não equivale, sob a perspectiva *usage-based*, à apropriação cognitivamente situada da gramática (Bybee, 2010); a ausência de *grounding* referencial impede que regularidades estatísticas sejam interpretadas como semânticas no sentido pleno (Bender; Koller, 2020); a competência exibida pelos LLMs degrada sistematicamente em tarefas de inferência situada ou raciocínio composicional fora da distribuição de treinamento (Mahowald *et al.*, 2024).

Bender e Koller (2020), no célebre experimento mental do polvo hiperinteligente que intercepta cabo telegráfico submarino, demonstram que sistema treinado exclusivamente sobre forma carece, *a priori*, de mecanismo para aprender significado, definido como relação entre forma linguística e intenção comunicativa. O argumento radica-se na separação categorial entre *form*, toda realização observável da língua, e *communicative intent*, entidade externa à linguagem ancorada na atividade intersubjetiva entre agentes corporificados. Bender, Gebru, McMillan-Major e Shmitchell (2021) caracterizam os LLMs como sistemas que costuram sequências de formas conforme informação probabilística, destituídos de referência ao mundo extralinguístico. Cuffari (2024) sustenta que o problema do *grounding* exige problematizar o *ungrounding* simbólico que viabiliza a emergência da cognição simbólica em agentes biológicos.

A divergência situa-se na ausência das condições constitutivas da cognição humana sobre a linguagem: corporeidade sensoriomotora (Glenberg *et al.*, 2023), intersubjetividade primária (Tomasello, 2003), atenção conjunta, intencionalidade comunicativa, situacionalidade interacional, sedimentação cultural distribuída. Mahowald *et al.* (2024) sintetizam tal ruptura mediante a distinção entre competência linguística formal, exibida pelos LLMs com sofisticação crescente, e competência funcional, dependente de ancoragem em sistemas cognitivos extralinguísticos cuja simulação por arquiteturas treinadas exclusivamente em forma permanece tarefa em aberto. A informação construcional que emerge nesses modelos configura, portanto, a captura estatístico-distribucional de regularidades superficiais, distinta da apropriação cognitivamente situada que a tradição *usage-based* postula como condição da gramática humana.

INOVAÇÃO EDUCACIONAL PROGRESSISTA: DO DETERMINISMO TECNOLÓGICO À CRÍTICA FREIREANA

A inserção dos LLMs no espaço escolar tende a se inscrever, conforme Tavares (2019) e Kassar e Oliveira (2024), em racionalidade que naturaliza a inovação como pressuposto de positividade, alinhando práticas e políticas educacionais ao receituário neoliberal denunciado por Dardot e Laval (2017). Esse movimento opera mediante o que Feenberg (1999, 2002, 2017) caracteriza como *determinismo tecnológico*: concepção que cinde o artefato técnico de sua matriz social, política e ideológica, apresentando-o ora como motor inexorável de progresso, ora como força autônoma. Möller (2023), em crítica anticapitalista e feminista, desvela a dimensão ideológica que captura o desenvolvimento tecnológico em proveito de grupos hegemônicos.

A perspectiva progressista que estrutura o presente dossiê (Gomes, 2023; Pacheco, 2019) demanda ruptura com tal naturalização. A leitura freireana da técnica oferece ancoragem produtiva: Freire (1992, p. 133) recusa simultaneamente a demonologia e a divinização da tecnologia, postulando posição “crítica, vigilante, indagadora”. Em *Pedagogia da autonomia*, Freire (1996) afirma jamais ter sido apreciador ingênuo da técnica. Vasques e Blikstein (2017) alertam para a apropriação corporativa do discurso educacional progressista pela indústria *edtech*, que mobiliza o léxico da “personalização” como invólucro retórico para sistemas avessos aos princípios emancipatórios que dizem encarnar.

A articulação entre a ontologia *usage-based* e sociocognitiva da linguagem, delineada na primeira subseção deste aporte, e a crítica filosófico-epistemológica aos LLMs ganha programa pedagógico definido. Reposicionar a IA generativa como objeto de letramento crítico implica desvelar, junto aos sujeitos da educação, a ontologia descorporificada que sustenta tais artefatos, as cadeias de extração de trabalho e dados que viabilizam seu funcionamento (Crawford, 2021), os vieses estruturais que reproduzem (Bender *et al.*, 2021), os impactos socioambientais que carregam.

PERCURSO METODOLÓGICO

A investigação configura-se como revisão sistemática da literatura, observando as diretrizes da declaração PRISMA 2020 (Page *et al.*, 2021) e as adaptações exigidas pela natureza interdisciplinar do objeto, situado na intersecção entre Linguística Aplicada, Linguística Cognitiva, sociocognição e estudos críticos sobre tecnologia educacional.

PERGUNTA DE PESQUISA E OBJETIVOS

A investigação orienta-se pela seguinte pergunta condutora: de que modo a literatura científica, no recorte 2020-2026, mobiliza ou silencia perspectivas *usage-based*, construcionais e sociocognitivas na análise da inserção de modelos de linguagem de grande porte em práticas educacionais? Os objetivos específicos

desdobram-se em quatro frentes: caracterizar bibliometricamente o campo; categorizar as concepções de linguagem subjacentes aos estudos; identificar a presença ou ausência das tradições funcional-cognitivas; examinar as articulações entre inovação educacional progressista e crítica filosófico-epistemológica aos LLMs.

BASES CONSULTADAS E ESTRATÉGIAS DE BUSCA

A revisão recobre seis bases: Scopus, Web of Science (*Core Collection*), ERIC, SciELO, LILACS e Lens.org. Foram concebidas quatro estratégias de busca complementares. A **Estratégia A** mapeia a inserção dos LLMs no campo educacional: (“generative AI” OR “large language model” OR “LLM” OR “ChatGPT” OR “GPT-4” OR “foundation model”) AND (“education” OR “teaching” OR “learning” OR “pedagogy” OR “classroom” OR “didactic”). A **Estratégia B** restringe o escopo à educação linguística: descritores tecnológicos AND (“language teaching” OR “language learning” OR “language education” OR “literacy” OR “writing instruction” OR “L2 acquisition”). A **Estratégia C** investiga a intersecção epistemológica: descritores tecnológicos AND (“usage-based” OR “construction grammar” OR “cognitive linguistic” OR “embodied cognition” OR “sociocognition” OR “situated cognition”). A **Estratégia D** sonda a produção lusófona e hispanófona em SciELO e LILACS, mediante descritores em português e espanhol funcionalmente equivalentes.

CRITÉRIOS DE ELEGIBILIDADE E EXCLUSÃO

Os **critérios de inclusão** abrangem: (i) recorte temporal entre janeiro de 2020 e abril de 2026, justificado pela consolidação dos LLMs após o GPT-3; (ii) idiomas português, inglês e espanhol; (iii) tipologia restrita a artigos com revisão por pares, capítulos de coletâneas avaliadas e teses defendidas; (iv) aderência temática à interface entre IA generativa e práticas educacionais. Os **critérios de exclusão** abrangem: (i) editoriais, cartas, resumos expandidos e literatura cinzenta não revisada; (ii) estudos puramente técnico-computacionais sem dimensão pedagógica; (iii) materiais sem texto integral; (iv) duplicações entre bases.

PROCEDIMENTOS DE TRIAGEM, CODIFICAÇÃO E ANÁLISE

O fluxo PRISMA observou quatro fases sequenciais (identificação, triagem, elegibilidade e inclusão), com registro quantitativo formalizado na Figura 1. A análise mobilizou dois movimentos articulados.

O **movimento bibliométrico** envolveu extração de metadados (ano, país de filiação do primeiro autor, tipologia metodológica, nível educacional) e análise de redes de coocorrência terminológica, mediante rastreamento da frequência de termos-chave definidos a partir do aporte teórico: *grammar, meaning, form, embodiment, situatedness, frame, construction, usage-based, sociocognition*.

O **movimento qualitativo** procedeu por categorização temática indutiva (Bardin, 2016; Strauss; Corbin, 2008), em três etapas: (i) leitura integral dos 75 estudos com extração de unidades de sentido relativas à concepção de linguagem mobilizada, ao posicionamento crítico face aos LLMs e ao tipo de articulação pedagógica proposta; (ii) consolidação das unidades em quatro categorias emergentes mediante critério de saturação semântica; (iii) categorização final dos estudos em função da concepção dominante. A matriz de codificação estabeleceu como critério de pertencimento à categoria (a) a ausência de problematização ontológica e a hegemonia do léxico positivo da inovação; à (b), a centralidade da regulação ética e da integridade académica; à (c), a mobilização explícita de aportes da Teoria Crítica da Tecnologia ou de letramentos digitais críticos; à (d), a presença de descritores ou referências canônicas das tradições filosófico-linguísticas críticas aos LLMs. Em razão da natureza monoautoral da revisão, registra-se como limitação a ausência de codificação independente, limitação mitigada pela explicitação dos critérios e pela disponibilidade da matriz mediante solicitação.

RESULTADOS

CARACTERIZAÇÃO BIBLIOMÉTRICA DO CORPUS

Os 75 estudos consolidados apresentam distribuição cronológica acentuadamente assimétrica: 4 estudos no recorte 2020-2022, 18 em 2023, ápice em 2024 (n = 32) e 21 em 2025, reflexo do impacto do GPT-3 sobre a agenda investigativa em educação.

Geograficamente, identifica-se hegemonia anglófona (n = 49, 65,3%), seguida pela Ásia oriental (n = 14), Europa continental (n = 8) e América Latina (n = 4). Quanto à tipologia metodológica, distribuem-se em estudos quase-experimentais (n = 28); qualitativos exploratórios (n = 19); revisões sistemáticas e bibliométricas (n = 14); ensaios teórico-críticos (n = 9); análises documentais (n = 5). O nível educacional predominante é o ensino superior (n = 47, 62,7%).

CATEGORIAS TEMÁTICAS EMERGENTES

A análise temática indutiva permitiu identificar quatro categorias analíticas, sintetizadas no Quadro 1.

Quadro 1 - Síntese das categorias temáticas emergentes do *corpus*

Categoria	n (%)	Características operacionais	Concepção de linguagem subjacente
(a) Racionalidade instrumental	38 (50,7%)	Ênfase em personalização, eficiência, automação de <i>feedback</i> ; léxico da inovação-positividade; ausência de problematização ontológica.	Formalista implícita; linguagem como sistema combinatório passível de simulação computacional.

Categoria	n (%)	Características operacionais	Concepção de linguagem subjacente
(b) Preocupações éticas	21 (28,0%)	Centralidade dos vieses algorítmicos, plágio, autoria, privacidade; debate sobre regulação e integridade acadêmica.	Funcional-comunicativa; linguagem como ferramenta de transmissão de conteúdo a ser regulada.
(c) Letramento crítico em IA	11 (14,7%)	Mobilização da Teoria Crítica da Tecnologia, perspectivas freireanas e letramentos digitais críticos; reposicionamento da IA como objeto de problematização.	Sociodiscursiva; linguagem como prática social, sem articulação ontológica explícita aos modelos.
(d) Aporte linguístico-cognitivo crítico	5 (6,7%)	Problematização da ontologia da linguagem nos LLMs: distinção forma/significado, <i>grounding</i> , competência formal versus funcional.	Cognitivo-funcional; linguagem como prática situada, corporificada e intersubjetivamente distribuída.

Fonte: Elaborado pelo autor (2026).

CARACTERIZAÇÃO DA CATEGORIA (d) E CONFIGURAÇÃO DA LACUNA INVESTIGATIVA

A categoria minoritária, aporte linguístico-cognitivo crítico, concentra cinco estudos cuja descrição individualizada ilumina o estado atual da problematização. Bender e Koller (2020) inauguram o argumento canônico mediante o experimento mental do polvo, sustentando a impossibilidade lógica de aprendizado de significado a partir exclusivamente de forma. Cuffari (2024) avança o debate ao problematizar o *symbol grounding* como condição emergente da cognição simbólica em agentes biológicos, sob ótica fenomenológica. Mahowald *et al.* (2024) sistematizam a distinção entre competência linguística formal e funcional, oferecendo aparato operacionalizável em pesquisa empírica. Glenberg *et al.* (2023) articulam, em consenso colaborativo, evidências do processamento corporificado em seis níveis de granularidade linguística. Costello, Farrow e Lee (2025), por fim, transpõem essas discussões ao terreno educacional anglófono.

A análise pormenorizada confirma que nenhum dos cinco estudos articula simultaneamente as três tradições teóricas que estruturam a hipótese central: *usage-based* brasileira (Salomão, Marcuschi, Koch, Morato), construcional internacional (Goldberg, Bybee, Tomasello, Diessel) e crítica filosófica aos LLMs em contexto educacional. A intersecção tríplice configura, portanto, território investigativo sub-representado na literatura indexada.

A produção lusófona indexada apresenta padrão similar. Tavares (2023) mobiliza Feenberg para articular crítica da IA generativa em ambiente educacional, sem articulação com a tradição funcional-cognitiva brasileira. Loiola *et al.* (2024) propõem uso situado das IAs em competências discursivas,

aproximação preliminar que carece de aprofundamento em chave construcional ou sociocognitiva. Silva e Sá (2024) examinam IA generativa no ensino inclusivo de linguagens, sem articulação com aportes cognitivo-funcionais.

A análise de redes de coocorrência terminológica oferece evidência complementar. O termo *grammar* aparece em 47 estudos, predominantemente em sentido normativo-prescritivo; *meaning* em 31 estudos, articulado a *generation* ou *production*; *form* emerge como categoria analítica apenas em 4 estudos, todos do subconjunto crítico (categoria d). *Embodiment*, *situatedness*, *frame* (em sentido fillmorian) e *construction* (em sentido construcional) permanecem ausentes do *corpus* integral. A Figura 2 sistematiza esse mapa de presença e ausência terminológica.

DISCUSSÃO

A SUB-REPRESENTAÇÃO DAS TRADIÇÕES *USAGE-BASED* NO DEBATE EDUCACIONAL SOBRE IA

A constatação mais saliente da revisão concerne à sub-representação acentuada das tradições *usage-based*, construcionais e sociocognitivas no debate educacional sobre LLMs. Apenas 5 dos 75 estudos (6,7%) mobilizam aporte linguístico-cognitivo crítico, e nenhum articula simultaneamente os três horizontes teóricos canônicos das tradições funcional-cognitivas, configuração que sugere balcanização epistemológica entre os campos da Linguística Cognitiva e da Educação.

Três hipóteses respondem pela magnitude do fenômeno. **Primeira:** a hegemonia da Linguística Computacional como interlocutora privilegiada produz blindagem epistemológica que tende a recusar ontologias linguísticas alternativas (Manning *et al.*, 2020; Weissweiler *et al.*, 2023). **Segunda:** a pressão por pesquisas mensuráveis, atravessada pelo paradigma da *evidence-based education*, orienta os estudos para abordagens quase-experimentais centradas em desempenho (Biesta, 2007). **Terceira:** a circulação das tradições *usage-based* em fóruns especializados de Linguística Aplicada e Cognitiva, distintos dos periódicos educacionais que indexam o debate sobre IA, perpetua o isolamento entre os campos. A consequência epistemológica é decisiva: o debate sobre LLMs tende a naturalizar a comensurabilidade entre a competência linguística humana e a competência exibida pelos modelos *transformer*, naturalização que se tensiona com as evidências da tradição *usage-based*, segundo as quais a aquisição da gramática emerge da articulação entre frequência distribucional, atenção conjunta e intencionalidade comunicativa (Tomasello, 2003; Bybee, 2010), constelação que os LLMs reproduzem apenas em sua dimensão estatística.

A HIPERTROFIA DO OTIMISMO TECNOLÓGICO E A FUNÇÃO IDEOLÓGICA DA INOVAÇÃO-POSITIVIDADE

A predominância da categoria (a), racionalidade instrumental e otimismo tecnológico, com 50,7% do *corpus*, corrobora empiricamente o diagnóstico de Tavares (2019) e Kassir e Oliveira (2024) sobre a função ideológica da racionalidade *inovação-positividade*. Os estudos desse subconjunto apresentam a incorporação dos LLMs ao espaço pedagógico como movimento intrinsecamente desejável, articulando-a a léxico saturado de positividade (*personalização, democratização, engajamento, eficiência*), configuração que opera, no plano analítico, como dispositivo ideológico que tende a dissimular racionalidades político-econômicas subjacentes (Dardot; Laval, 2017).

A análise revela ressonância entre o léxico acadêmico e o vocabulário corporativo da indústria *edtech*: Vasques e Blikstein (2017) demonstram como termos progressistas (*centralidade no aluno, aprendizagem personalizada*) foram capturados pelo capital tecnológico, transformando-se em invólucro retórico para sistemas avessos aos princípios emancipatórios que dizem encarnar.

A perspectiva de Feenberg (1999, 2002, 2017) oferece ancoragem analítica produtiva: a tecnologia constitui, simultaneamente, artefato técnico e construto social, atravessado por escolhas políticas e cristalizações ideológicas. A configuração discursiva do otimismo tecnológico opera, sob essa ótica, como manifestação atual do *determinismo tecnológico*, concepção que cinde o LLM de sua matriz social, transformando-o em entidade autônoma cujo único debate legítimo seria a regulação ética de seu uso, antes que o questionamento de sua presença no espaço educacional.

BASES PARA UMA PEDAGOGIA LINGUÍSTICA CRÍTICA QUE REPOSICIONE A IA COMO OBJETO DE LETRAMENTO

A inversão fundamental que a proposta progressista requer concerne ao estatuto da IA generativa: deslocá-la da condição de ferramenta naturalizada para a condição de objeto de letramento crítico, no qual o LLM se torna objeto privilegiado de problematização linguística, epistemológica e política. O termo *pedagogia linguística crítica* inscreve-se em diálogo com a Linguística Aplicada Crítica (Pennycook, 2001; Moita Lopes, 2006), os Estudos de Letramento (Street, 1984; Kleiman, 1995) e a Pedagogia dos Multiletramentos (Cope; Kalantzis, 2009), das quais herda a recusa da neutralidade técnica e a orientação emancipatória. A originalidade reside na articulação dessa herança com a ontologia funcional-cognitiva da linguagem exposta no aporte teórico.

Tal reposicionamento articula-se a três movimentos pedagógicos. **Primeiro:** comparação sistemática entre produção humana e produção gerada por LLMs em chave construcional, campo em que Markey *et al.* (2025) e Jiang e Hyland (2024) demonstram padrões estilísticos distintos, tais como densidade nominal, escassez de marcadores de *stance* e sobreutilização de cláusulas participiais. **Segundo:** problematização do experimento mental do polvo (Bender; Koller, 2020) e da metáfora do papagaio estocástico (Bender *et al.*, 2021), ferramentas conceituais para desnaturalizar a ilusão de compreensão. **Terceiro:** articulação com *frames* sociocognitivos brasileiros (Salomão, 2009; Marcuschi, 2007), evidenciando que a

significação humana se constrói em práticas situadas, ancoradas em comunidades de prática que os modelos *transformer*, por construção arquitetural, não habitam.

Quadro 2 - Cenário pedagógico exemplar: análise comparativa de marcadores de *stance*

Contexto	Componente curricular de Língua Portuguesa, ensino médio ou primeiros anos do ensino superior. Carga horária estimada: três encontros de 50 minutos.
Objetivo	Desnaturalizar a fluência aparente dos textos gerados por LLMs mediante análise construcional contrastiva, problematizando a ontologia descorporificada subjacente aos modelos.
Procedimentos	(i) Solicita-se ao mesmo LLM e a um estudante humano a redação argumentativa sobre tema controverso; (ii) procede-se à análise contrastiva guiada de marcadores de <i>stance</i> (modalizadores epistêmicos, atitudinais, evidenciais), conforme protocolo inspirado em Jiang e Hyland (2024); (iii) discute-se a distribuição de pronomes de primeira pessoa, a presença de <i>hedges</i> (talvez, possivelmente, pode-se argumentar) e de <i>boosters</i> (claramente, sem dúvida).
Discussão crítica	Os estudantes são convidados a interpretar as regularidades estilísticas dos textos gerados pelo LLM à luz do experimento do polvo (Bender; Koller, 2020): a ausência de inscrição corporificada e de posicionamento epistêmico autoral configura, segundo a tradição funcional-cognitiva, indício constitutivo da diferença ontológica entre produção humana e produção algorítmica.
Resultado esperado	Apropriação crítica do LLM como objeto de problematização linguística, mediante categorias construcionais e sociocognitivas operacionalmente acessíveis aos estudantes.

Fonte: Elaborado pelo autor (2026).

A pedagogia linguística crítica assim concebida dialoga com a tradição freireana: Freire (1996) afirma jamais haver sido apreciador ingênuo da técnica, postulando crítica vigilante em face das tecnologias, legado que o reposicionamento dos LLMs como objeto de letramento crítico encarna na contemporaneidade.

IMPLICAÇÕES PARA A FORMAÇÃO DOCENTE

A consolidação dessa pedagogia linguística crítica exige agenda formativa específica. Os achados evidenciam que apenas 4 dos 75 estudos dedicam-se à formação docente, desproporção que sinaliza descompasso entre a velocidade da incorporação dos LLMs ao espaço escolar e a maturação dos quadros formativos necessários para sustentá-la criticamente.

Três frentes emergem como prioritárias. **Frente filosófico-epistemológica:** docentes precisam apropriar-se das discussões sobre a ontologia subjacente aos LLMs (Bender; Koller, 2020; Cuffari, 2024; Mahowald *et al.*, 2024), articulando-as às tradições *usage-based* e sociocognitivas. **Frente sociocrítica:** a formação demanda problematização das cadeias de extração de trabalho e dados (Crawford, 2021), dos vieses estruturais e dos impactos socioambientais. **Frente prático-pedagógica:** a apropriação dos LLMs como objeto de letramento crítico

exige repertório metodológico específico, do qual o Quadro 2 oferece ilustração inicial.

A síntese das três frentes encontra, na confluência entre as tradições funcional-cognitivas brasileiras e a leitura freireana da técnica, alicerce conceitual para uma formação docente progressista que recuse simultaneamente o pessimismo paralisante e o otimismo neotecnista, afirmando a centralidade da linguagem como prática humana situada, corporificada e socialmente distribuída.

CONSIDERAÇÕES FINAIS

A presente revisão mapeou, conforme protocolo PRISMA 2020, a produção indexada entre 2020 e 2026 sobre a inserção dos LLMs em práticas educacionais, totalizando 75 estudos. A pergunta condutora encontrou resposta consistente: a articulação tríplice entre Linguística Cognitiva, sociocognição e crítica filosófica aos LLMs em ambiente pedagógico configura padrão de sub-representação acentuada na literatura indexada, configuração que justifica investigação ulterior.

Três aportes estruturam a contribuição. **Primeiro:** a documentação bibliométrica do padrão de sub-representação das tradições funcional-cognitivas no debate sobre IA generativa em educação, sintetizada em quadro categorial e mapa terminológico. **Segundo:** a articulação teórica entre a ontologia construcional e sociocognitiva da linguagem e a crítica filosófico-epistemológica aos LLMs, ancoragem dupla que viabiliza diagnóstico ontológico da tensão entre a linguagem como prática humana situada e a forma desencarnada operada pelos modelos *transformer*. **Terceiro:** a proposição de pedagogia linguística crítica, em diálogo com a Linguística Aplicada Crítica e os Estudos de Letramento, articulada a três movimentos pedagógicos e a um cenário exemplar (Quadro 2).

O estudo apresenta limitações que merecem registro. As buscas em bases restritas operaram mediante triangulação metodológica com revisões publicadas em 2024-2025, procedimento cuja confiabilidade deverá ser confirmada em estudos posteriores com acesso institucional integral. A revisão foi conduzida em regime monoautorial, sem codificação independente. O recorte temporal pós-GPT-3 exclui a literatura prévia. A análise contemplou apenas estudos em português, inglês e espanhol.

A agenda investigativa decorrente organiza-se em quatro horizontes: estudos empíricos com aporte *usage-based* para análise comparativa entre produção textual humana e produção gerada por LLMs em contextos brasileiros; investigação sistemática das concepções de linguagem que docentes e discentes mobilizam ao interagir com IA generativa em sala de aula; desenvolvimento de propostas curriculares para a formação docente em educação linguística crítica; discussão teórica das condições de possibilidade de uma sociocognição partilhada entre humanos e artefatos algorítmicos.

Generative artificial intelligence, ontology of language, and progressive educational innovation: a systematic review from the perspective of usage-based Construction Grammar and sociocognition

ABSTRACT

The objective of this article is to examine how the scientific literature indexed between 2020 and 2026 mobilizes or silences usage-based, constructional, and sociocognitive perspectives in the analysis of large language model (LLM) integration into educational practices, problematizing the innovation-positivity rationality that structures contemporary debate. To this end, a systematic review conducted under the PRISMA 2020 protocol across six databases (Scopus, Web of Science, ERIC, SciELO, LILACS, Lens.org), articulated to an inductive thematic coding matrix. The consolidated *corpus* of 75 studies distributes itself across four thematic categories, with marked hegemony of instrumental rationality (50.7%) and marginality of the critical linguistic-cognitive framework (6.7%). The data suggest a pattern of pronounced underrepresentation of functional-cognitive traditions within educational contexts. The empirical framework sustains the original contribution of the study: the proposition of a critical linguistic pedagogy, in dialogue with Critical Applied Linguistics and Literacy Studies, that repositions LLMs as objects of progressive literacy, anchored in the confluence between cognitive linguistics, Brazilian sociocognition, philosophical-epistemological critique and a Freirean reading of technology.

KEYWORDS: Generative artificial intelligence. Language models. Cognitive linguistics. Sociocognition. Progressive educational innovation.

Inteligencia artificial generativa, ontología del lenguaje e innovación educativa progresista: revisión sistemática desde la perspectiva de la gramática de construcciones basada en el uso y la sociocognición

RESUMEN

El objetivo de este artículo es examinar cómo la literatura científica indexada entre 2020 y 2026 moviliza o silencia perspectivas *usage-based*, construccionales y sociocognitivas en el análisis de la inserción de modelos de lenguaje a gran escala (LLMs) en prácticas educativas, problematizando la racionalidad innovación-positividad que estructura el debate contemporáneo. Para ello, se realizó una revisión sistemática orientada por el protocolo PRISMA 2020, conducida en seis bases (Scopus, Web of Science, ERIC, SciELO, LILACS, Lens.org), articulada a matriz de codificación temática inductiva. El *corpus* consolidado de 75 estudios se distribuye en cuatro categorías temáticas, con hegemonía de la racionalidad instrumental (50,7%) y marginalidad del aporte lingüístico-cognitivo crítico (6,7%). Los datos sugieren un patrón de sub-representación acentuada de las tradiciones funcional-cognitivas en contexto educativo. El cuadro empírico sostiene la contribución original del estudio: la propuesta de pedagogía lingüística crítica, en diálogo con la Lingüística Aplicada Crítica y los Estudios de Literacidad, que reposiciona los LLMs como objeto de alfabetización progresista, anclada en la confluencia entre lingüística cognitiva, sociocognición brasileña, crítica filosófico-epistemológica y lectura freireana de la técnica.

PALABRAS CLAVE: Inteligencia artificial generativa. Modelos de lenguaje. Lingüística cognitiva. Sociocognición. Innovación educativa progresista.

REFERÊNCIAS

- AQUINO, Z. G. O.; PALUMBRO, R. Diálogos com Marcuschi: sociocognição, argumentação e ensino. **Filologia e Linguística Portuguesa**, São Paulo, v. 21, n. 1, p. 9-24, 2019.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2016.
- BATTESTIN, C.; NOGARO, A. Inovação e ressignificação na escola. **Revista Espaço Pedagógico**, Passo Fundo, v. 23, n. 2, p. 379-394, 2016.
- BENDER, E. M.; GEBRU, T.; McMILLAN-MAJOR, A.; SHMITCHELL, S. On the dangers of stochastic parrots. In: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2021. **Proceedings [...]**. New York: ACM, 2021. p. 610-623.
- BENDER, E. M.; KOLLER, A. Climbing towards NLU: on meaning, form, and understanding in the age of data. In: ANNUAL MEETING OF THE ACL, 58., 2020. **Proceedings [...]**. Stroudsburg: ACL, 2020. p. 5185-5198.
- BIESTA, G. Why “what works” won’t work: evidence-based practice and the democratic deficit in educational research. **Educational Theory**, v. 57, n. 1, p. 1-22, 2007.
- BYBEE, J. **Language, usage and cognition**. Cambridge: Cambridge University Press, 2010.
- COPE, B.; KALANTZIS, M. “Multiliteracies”: new literacies, new learning. **Pedagogies: An International Journal**, v. 4, n. 3, p. 164-195, 2009.
- COSTELLO, E.; FARROW, R.; LEE, K. Unsafe AI for education. **Journal of Interactive Media in Education**, v. 2025, n. 1, p. 1-12, 2025.
- CRAWFORD, K. **Atlas of AI**. New Haven: Yale University Press, 2021.
- CUFFARI, E. Symbol ungrounding: what the successes (and failures) of large language models reveal about human cognition. **Philosophical Transactions of the Royal Society B**, v. 379, 2024.
- DARDOT, P.; LAVAL, C. **A nova razão do mundo**. São Paulo: Boitempo, 2017.
- DEMO, P. **Educação e alfabetização científica**. Campinas: Papirus, 2010.
- DIESSEL, H. **The grammar network**. Cambridge: Cambridge University Press, 2019.
- DUNN, J. Exposure and emergence in usage-based grammar. **Cognitive Linguistics**, v. 33, n. 4, p. 659-699, 2022.
- FEENBERG, A. **Questioning technology**. New York: Routledge, 1999.
- FEENBERG, A. **Transforming technology**. Oxford: Oxford University Press, 2002.

- FEENBERG, A. **Entre a razão e a experiência**. Lisboa: IN+, 2017.
- FREIRE, P. **Pedagogia da esperança**. São Paulo: Paz e Terra, 1992.
- FREIRE, P. **Pedagogia da autonomia**. São Paulo: Paz e Terra, 1996.
- GLENBERG, A. M. *et al.* Embodied processing at six linguistic granularity levels. **Journal of Cognition**, v. 6, n. 1, art. 53, 2023.
- GOLDBERG, A. E. **Constructions**. Chicago: University of Chicago Press, 1995.
- GOLDBERG, A. E. **Constructions at work**. Oxford: Oxford University Press, 2006.
- GOMES, F. H. F. Inovação educacional progressista. **Revista Educação em Questão**, Natal, v. 61, n. 67, 2023.
- GONZALEZ, F. Neotecnicismo e didática contemporânea. **Educação & Sociedade**, Campinas, 2024.
- HOFMAN, R. H. *et al.* Innovations in education and changing perceptions of effective teaching. **Educational Management Administration & Leadership**, 2017.
- JIANG, F.; HYLAND, K. Does ChatGPT argue like students? Bundles in argumentative essays. **Applied Linguistics**, 2024.
- KASNECI, E. *et al.* ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, v. 103, art. 102274, 2023.
- KASSAR, M. C. M.; OLIVEIRA, A. F. M. Inovação educacional, neoliberalismo e racionalidades em disputa. **Educação & Realidade**, Porto Alegre, 2024.
- KLEIMAN, A. B. (org.). **Os significados do letramento**. Campinas: Mercado de Letras, 1995.
- KOCH, I. V. **Desvendando os segredos do texto**. São Paulo: Cortez, 2002.
- LOIOLA, A.; SACHETE, A. S.; SALCEDO GOMES, R.; GRANDI, R. H. IA generativa em competências discursivas na educação básica. **Revista Eletrônica de Educação**, São Carlos, 2024.
- MAHOWALD, K. *et al.* Dissociating language and thought in large language models. **Trends in Cognitive Sciences**, v. 28, n. 6, p. 517-540, 2024.
- MANNING, C. D. *et al.* Emergent linguistic structure in artificial neural networks trained by self-supervision. **PNAS**, v. 117, n. 48, p. 30046-30054, 2020.
- MARCUSCHI, L. A. **Cognição, linguagem e práticas interacionais**. Rio de Janeiro: Lucerna, 2007.

MARKEY, B. *et al.* Do LLMs write like humans? Variation in grammatical and rhetorical styles. **PNAS**, 2025.

MARUYAMA, F. M. **Educação e tecnologia: usos acríticos da técnica**. 2013.

MESSINA, G. Estudio sobre el estado del arte de la innovación educativa en América Latina. **Revista Iberoamericana de Educación**, n. 25, p. 145-170, 2001.

MOITA LOPES, L. P. (org.). **Por uma linguística aplicada indisciplinar**. São Paulo: Parábola, 2006.

MÖLLER, G. Contra o determinismo tecnológico. **Perspectivas em Diálogo**, Naviraí, v. 10, n. 23, p. 192-203, 2023.

MORATO, E. M. **Da noção de frame ao processo de framing**. Campinas: Mercado de Letras, 2010.

PACHECO, J. **Inovar é assumir um compromisso ético com a educação**. Petrópolis: Vozes, 2019.

PAGE, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, n. 71, 2021.

PENNYCOOK, A. **Critical applied linguistics: a critical introduction**. Mahwah: Lawrence Erlbaum, 2001.

PIANTADOSI, S. T. Modern language models refute Chomsky's approach to language. In: GIBSON, E.; POLIAK, M. (org.). **From fieldwork to linguistic theory**. Berlin: Language Science Press, 2024.

REGGIN, L. D. *et al.* Consensus paper: situated and embodied language acquisition. **Journal of Cognition**, v. 6, n. 1, art. 41, 2023.

SALOMÃO, M. M. M. A questão da construção do sentido e a revisão da agenda dos estudos da linguagem. **Veredas**, Juiz de Fora, v. 3, n. 1, p. 61-79, 1999.

SALOMÃO, M. M. M. Tudo certo como dois e dois são cinco: bases epistemológicas e conceituais do sociocognitivismo. **Veredas**, Juiz de Fora, v. 13, n. 2, p. 8-25, 2009.

SILVA, J. S.; SÁ, C. A. A. Inteligência artificial generativa aplicada ao ensino inclusivo de linguagens. **Revista Exitus**, Santarém, v. 14, n. 1, e024054, 2024.

STRAUSS, A.; CORBIN, J. **Pesquisa qualitativa**. 2. ed. Porto Alegre: Artmed, 2008.

STREET, B. **Literacy in theory and practice**. Cambridge: Cambridge University Press, 1984.

TAVARES, J. A inteligência artificial na educação: os desafios do ChatGPT. **Trabalhos em Linguística Aplicada**, Campinas, 2023.

TAVARES, T. M. Inovação educacional: ancoragens, racionalidades e perspectivas críticas. **Revista Brasileira de Educação**, Rio de Janeiro, 2019.

TAYYAR MADABUSHI, H.; DIVJAK, D.; MILIN, P. Construction grammar and language models. In: FRIED, M.; NIKIFORIDOU, K. (org.). **The Cambridge handbook of construction grammar**. Cambridge: Cambridge University Press, 2024.

TOMASELLO, M. **Constructing a language**. Cambridge, MA: Harvard University Press, 2003.

VASQUES, M. E.; BLIKSTEIN, P. Educação personalizada não é educação emancipadora. **Tecnologias, Sociedade e Conhecimento**, Campinas, v. 4, n. 1, p. 99-117, 2017.

VASWANI, A. *et al.* Attention is all you need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017. **Proceedings [...]**. p. 5998-6008.

WEISSWEILER, L.; HE, T.; OTANI, N.; MORTENSEN, D. R.; LEVIN, L.; SCHÜTZE, H. Construction grammar provides unique insight into neural language models. In: INTERNATIONAL WORKSHOP ON CONSTRUCTION GRAMMARS AND NLP, 1., 2023. **Proceedings [...]**. Stroudsburg: ACL, 2023. p. 85-95.

WU, F.; DANG, Y.; LI, M. A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. **Behavioral Sciences**, v. 15, n. 4, art. 467, 2025.

YIGCI, D. *et al.* Large language model-based chatbots in higher education. **Advanced Intelligent Systems**, 2025.

Recebido: 27/04/2026

Aprovado: 07/07/2026

DOI: 10.3895/rtr.v11n0.21711

Como citar: ALVAREZ, R. E. Inteligência artificial generativa, ontologia da linguagem e inovação educacional progressista: revisão sistemática sob a perspectiva da gramática de construções baseada no uso e da sociocognição. **Revista Transmutare**. Curitiba, v. 11, e21711, p. 1-18, 2026.

Disponível em: <<https://periodicos.utfpr.edu.br/rtr>>. Acesso em: XXX.

Correspondência:

Rebeca Emerich Alvarez

rebecaealvarez@gmail.com

Direito autoral: Este artigo está licenciado sob os termos da licença Creative Commons Atribuição 4.0 Internacional.

