



UNIVERSIDADE FEDERAL DE UBERLÂNDIA – UFU
INSTITUTO DE LETRAS E LINGUÍSTICA
CURSO DE GRADUAÇÃO EM LETRAS INGLÊS – LICENCIATURA, NA
MODALIDADE A DISTÂNCIA

BETÂNIA CRISTINA MARQUES

O USO DO CHATGPT NO ENSINO-APRENDIZAGEM DE LÍNGUA INGLESA:
UMA REVISÃO INTEGRATIVA DE ESTUDOS EMPÍRICOS

Orientador: Prof. Dr. Rafael Matielo

ITUIUTABA

2026

SUMÁRIO

1 INTRODUÇÃO.....	3
2 REFERENCIAL TEÓRICO	5
2.1 ChatGPT:	5
2.2 Inglês como L2	6
2.3 IA e o Desenvolvimento de Habilidades Linguísticas.....	7
3 METODOLOGIA.....	8
3.1 Delineamento do Estudo.....	8
3.2 Perguntas de Pesquisa.....	8
3.3 Critérios de Inclusão e Exclusão	9
3.4 Busca e Seleção dos Estudos	9
3.5 Caracterização do Corpus.....	10
3.6 Procedimentos de Análise e Interpretação dos Dados.....	11
4 RESULTADOS E DISCUSSÃO	12
4.1 (PP1) Distribuição dos Estudos por Habilidade Linguística	12
4.2 (PP2) Foco em Compreensão Leitora (<i>Reading</i>)	14
4.3 (PP3) Foco em Produção Escrita (<i>Writing</i>).....	16
4.4 (PP4) Foco em Compreensão Oral (<i>Listening</i>).....	21
4.5 (PP5) Foco em Produção Oral (<i>Speaking</i>).....	23
5 CONSIDERAÇÕES FINAIS.....	25
6 REFERÊNCIAS	27

O USO DO CHATGPT NO ENSINO-APRENDIZAGEM DE LÍNGUA INGLESA: UMA REVISÃO INTEGRATIVA DE ESTUDOS EMPÍRICOS

RESUMO

Este estudo investigou os efeitos do uso do ChatGPT no ensino-aprendizagem de inglês como segunda língua (L2) por meio de uma revisão integrativa da literatura. A busca bibliográfica foi conduzida nas bases de dados ERIC e SciELO em julho de 2025. A consulta à SciELO não retornou resultados, ao passo que a busca na ERIC recuperou 210 registros, dos quais 22 estudos empíricos publicados entre novembro de 2022 e junho de 2025 compuseram o corpus final. Os estudos foram analisados quanto às habilidades linguísticas investigadas, aos principais resultados reportados e às limitações identificadas. Os resultados evidenciaram distribuição assimétrica entre as habilidades: a produção escrita (*writing*) concentrou aproximadamente 86% dos estudos, enquanto a compreensão leitora (*reading*) foi abordada em dois estudos e a produção oral (*speaking*) em apenas um. A compreensão oral (*listening*) esteve completamente ausente do corpus, configurando a lacuna mais expressiva da literatura. No domínio da escrita, os estudos apontaram vantagens relacionadas ao *feedback* imediato, ao apoio ao processo de revisão textual e ao aumento da autonomia dos aprendizes, contrastadas por limitações associadas à integridade acadêmica, à dependência excessiva da ferramenta e à qualidade variável das respostas geradas. Conclui-se que o ChatGPT apresenta potencial pedagógico relevante, desde que integrado de forma crítica e pedagogicamente orientada ao contexto de ensino de L2.

Palavras-chave: ChatGPT; inteligência artificial; inglês como L2; habilidades linguísticas; revisão integrativa.

ABSTRACT

This study investigated the effects of ChatGPT use on English as a second language (L2) teaching and learning through an integrative literature review. A bibliographic search was conducted in the ERIC and SciELO databases in July 2025. The SciELO search returned no results, while the ERIC search retrieved 210 records, of which 22 empirical studies published between November 2022 and June 2025 composed the final corpus. The studies were analyzed in terms of the linguistic skills investigated, the main results reported, and the limitations identified. Findings revealed an asymmetrical distribution across skills: writing accounted for approximately 86% of the studies, while reading was addressed in two studies and speaking in only one. Listening was entirely absent from the corpus, representing the most significant gap in the literature. Within the domain of writing, studies highlighted advantages related to immediate feedback, support throughout the text revision process, and increased learner autonomy, counterbalanced by limitations associated with academic integrity concerns, excessive tool dependency, and the variable quality of AI-generated responses. ChatGPT holds relevant pedagogical potential for L2 teaching and learning, provided it is integrated critically and in a pedagogically informed manner as a complement to human instruction.

Keywords: ChatGPT; artificial intelligence; English as an L2; language skills; integrative review.

1 INTRODUÇÃO

Nos últimos dez anos, a inteligência artificial (IA) tem se destacado como uma das mais importantes inovações tecnológicas da sociedade contemporânea, impactando diversos setores, incluindo a educação. Entre as ferramentas baseadas em IA, o ChatGPT, desenvolvido pela *OpenAI* e lançado em novembro de 2022, ganhou destaque devido à sua capacidade de gerar textos em linguagem natural, responder perguntas, produzir explicações detalhadas e simular interações semelhantes às realizadas entre seres humanos. Segundo Baskara (2023), trata-se de um modelo avançado de linguagem treinado a partir de grandes volumes de dados textuais, capaz de realizar diferentes tarefas linguísticas. Para Hong (2023) e Kohnke *et al.* (2023), o ChatGPT representa uma inovação significativa para o ensino e a aprendizagem de línguas, ao oferecer suporte personalizado e interativo aos estudantes.

No contexto de língua inglesa como segunda língua (L2), o uso de ferramentas de IA tem despertado crescente interesse entre pesquisadores e profissionais da educação. O potencial dessas tecnologias está relacionado à possibilidade de oferecer *feedback* imediato, promover a aprendizagem autônoma, ampliar oportunidades de prática linguística e personalizar experiências educacionais. Além disso, o ChatGPT pode ser utilizado para auxiliar no desenvolvimento de diferentes habilidades linguísticas, incluindo compreensão leitora (*reading*), produção escrita (*writing*), compreensão oral (*listening*) e produção oral (*speaking*), contribuindo para a criação de ambientes de aprendizagem mais dinâmicos e acessíveis.

Apesar do crescente interesse acadêmico e educacional, muitos aspectos da utilização do ChatGPT ainda permanecem pouco compreendidos. A rápida disseminação do ChatGPT tem gerado debates sobre seus potenciais benefícios para o ensino de línguas, mas também tem levantado questionamentos relacionados à dependência tecnológica, à integridade acadêmica, à qualidade do *feedback* fornecido pela IA e ao papel do professor em ambientes educacionais cada vez mais mediados por tecnologias digitais. Nesse sentido, Han (2024) destaca a necessidade de investigações sistemáticas que permitam compreender de forma mais aprofundada os impactos da ferramenta no processo de aquisição de uma L2. De maneira semelhante, Warschauer *et al.* (2023) argumentam que os avanços recentes da IA exigem pesquisas que examinem tanto suas potencialidades quanto suas limitações em contextos educacionais reais.

A relevância deste estudo justifica-se pela necessidade de reunir e analisar evidências científicas produzidas desde o surgimento do ChatGPT, identificando quais habilidades linguísticas têm sido investigadas, quais resultados têm sido obtidos e quais lacunas

permanecem na literatura. A identificação de vantagens, limitações e desafios pode contribuir para orientar professores, pesquisadores e instituições quanto ao uso responsável e pedagogicamente fundamentado dessa ferramenta.

Considerando a necessidade de sistematizar esse campo emergente, optou-se pelo método de revisão integrativa da literatura, por ser o delineamento mais indicado para reunir e sintetizar evidências provenientes de estudos com diferentes desenhos metodológicos (Torraco, 2005; Snyder, 2019). Esse tipo de revisão permite não apenas mapear o estado do conhecimento, mas também identificar tendências, divergências e lacunas que orientem agendas futuras de pesquisa. Diante desse cenário, este estudo teve como objetivo investigar os efeitos do uso do ChatGPT no ensino-aprendizagem de língua inglesa como L2 por meio de uma revisão integrativa da literatura. Para alcançar esse objetivo, foram analisados 22 estudos empíricos¹ publicados após o lançamento da ferramenta, buscando compreender como o ChatGPT tem sido utilizado para o desenvolvimento das habilidades linguísticas e quais resultados têm sido reportados pela literatura científica.

Este artigo está organizado da seguinte forma: primeiramente, apresenta-se o referencial teórico, abordando o ChatGPT no contexto educacional, o ensino-aprendizagem de inglês como L2 e o uso da IA no desenvolvimento de habilidades linguísticas. Em seguida, descreve-se a metodologia adotada, fundamentada no método de revisão integrativa. Posteriormente, são apresentados e discutidos os resultados obtidos a partir da análise dos 22 estudos selecionados, organizados em torno das cinco perguntas de pesquisa que nortearam a investigação. Por fim, apresentam-se as considerações finais, destacando as principais contribuições do estudo, suas limitações e sugestões para pesquisas futuras.

2 REFERENCIAL TEÓRICO

2.1 ChatGPT

A IA generativa, categoria de sistemas computacionais capazes de produzir conteúdo original a partir de padrões aprendidos em grandes volumes de dados, tem promovido mudanças significativas na forma como as pessoas produzem, acessam e compartilham informações

¹ Estudos empíricos são investigações baseadas em dados observados ou experimentais, coletados por meio de métodos sistemáticos como experimentos, questionários, entrevistas ou análises de desempenho. Diferem dos estudos teóricos ou conceituais por se fundamentarem em evidências concretas, permitindo avaliar de forma objetiva os efeitos de uma intervenção ou fenômeno em contextos reais

(Baskara, 2023). Entre as ferramentas mais populares dessa nova geração tecnológica destaca-se o ChatGPT, um modelo de linguagem de grande escala (*large language model* — LLM) desenvolvido pela *OpenAI* e disponibilizado ao público em novembro de 2022. Baseado em técnicas avançadas de processamento de linguagem natural (*natural language processing* — NLP), o sistema foi treinado com grandes volumes de dados textuais e é capaz de compreender comandos em linguagem natural, gerar respostas contextualizadas e realizar diversas tarefas relacionadas à comunicação escrita.

O rápido crescimento da popularidade do ChatGPT despertou o interesse de pesquisadores da área educacional, especialmente no campo do ensino de línguas. Segundo Hong (2023), a ferramenta representa uma inovação disruptiva por possibilitar interações mais dinâmicas entre estudantes e sistemas computacionais, aproximando a comunicação homem-máquina de padrões conversacionais humanos. De forma semelhante, Kohnke *et al.* (2023) argumentam que o ChatGPT pode atuar como um recurso complementar ao trabalho docente, oferecendo oportunidades de prática linguística, esclarecimento de dúvidas e apoio ao desenvolvimento de habilidades acadêmicas.

No contexto educacional, o ChatGPT tem sido utilizado para diferentes finalidades, incluindo produção textual, revisão gramatical, geração de atividades, prática de conversação e fornecimento de *feedback* automatizado (Kohnke *et al.*, 2023; Han, 2024). Entretanto, sua utilização também levanta preocupações relacionadas à qualidade das informações geradas, à dependência tecnológica e à integridade acadêmica (Yan, 2023; Rudolph; Tan; Tan, 2023). Dessa forma, a literatura enfatiza que o potencial pedagógico da ferramenta depende de uma utilização crítica e pedagogicamente orientada, em que professores e estudantes compreendam tanto suas possibilidades quanto suas limitações, evitando tanto a adoção acrítica quanto a rejeição infundada da tecnologia.

2.2 Inglês como L2

O ensino-aprendizagem de inglês como L2 constitui um dos campos mais investigados na área dos Estudos Linguísticos e suas interfaces. Embora os termos *língua estrangeira* (LE) e *L2* carreguem implicações teóricas e sociolinguísticas distintas, sendo L2 frequentemente reservado a contextos em que a língua-alvo desempenha função social relevante no ambiente imediato do aprendiz, ao passo que LE designa contextos em que essa língua não possui presença funcional cotidiana (Krashen, 1982; Ellis, 2008), a literatura especializada tem progressivamente relativizado essa dicotomia.

Ellis (2008) argumenta que não há evidências empíricas suficientes para sustentar que os processos de aquisição diferem de forma substantiva entre esses dois contextos, razão pela qual propõe tratá-los sob um mesmo espectro analítico. De modo semelhante, Ortega (2013) adota o termo L2 como designação abrangente para o processo de aprendizagem de línguas adicionais, independentemente do contexto de exposição, por entender que essa escolha confere maior precisão terminológica sem incorrer em simplificações contextuais indevidas. No cenário brasileiro, Matielo (2016) endossa essa posição ao destacar que, embora as diferenças contextuais entre LE e L2 não devam ser desconsideradas, adotar uma separação rígida entre os dois termos pode obscurecer mais do que esclarecer os processos envolvidos na aprendizagem de línguas adicionais. Assim, neste estudo, o termo L2 é empregado como designação abrangente para o processo de ensino-aprendizagem de língua inglesa, independentemente das especificidades sociolinguísticas de cada contexto investigado.

2.3 IA e o Desenvolvimento de Habilidades Linguísticas

A incorporação de tecnologias digitais ao ensino de línguas não é um fenômeno recente. O campo da Aprendizagem de Línguas Assistida por Computador (*Computer-Assisted Language Learning* — CALL) já investigava, desde as décadas de 1980 e 1990, o potencial de ferramentas tecnológicas para ampliar oportunidades de prática linguística, oferecer *feedback* corretivo e promover a autonomia dos aprendizes (Chapelle, 2001). Os avanços recentes da IA, contudo, ampliaram significativamente o escopo dessas possibilidades, introduzindo sistemas capazes de gerar linguagem natural, adaptar respostas ao perfil do usuário e sustentar interações comunicativas de forma dinâmica e contextualizada.

Nesse cenário, Warschauer *et al.* (2023) argumentam que sistemas de IA generativa apresentam potencial para apoiar o desenvolvimento da escrita em L2 por meio da oferta de *feedback* imediato, sugestões de reformulação textual e apoio à organização de ideias. Além disso, essas ferramentas podem contribuir para a aprendizagem autônoma, permitindo que os estudantes pratiquem fora do ambiente formal de ensino e recebam orientações instantâneas sobre suas produções linguísticas, o que se alinha diretamente ao conceito de *input* compreensível proposto por Krashen (1982), ao ampliar o contato significativo com a língua-alvo.

Han (2024) enfatiza que o surgimento do ChatGPT representa um momento de inflexão para a pesquisa em aquisição de L2, uma vez que a ferramenta oferece novas possibilidades de interação linguística mediada por tecnologia em escala sem precedentes. Contudo, o autor

ressalta que ainda são necessárias investigações empíricas sistemáticas para compreender de forma mais precisa seus efeitos sobre a aprendizagem. Nessa direção, Chappelle (2024) destaca que a integração de sistemas de IA ao ensino de línguas deve ser avaliada não apenas em termos de eficiência instrucional, mas também em relação ao desenvolvimento da agência linguística do aprendiz. Em outras palavras, a questão central não é se a ferramenta produz *output* linguístico de qualidade, mas em que medida ela promove processos cognitivos e metacognitivos que contribuem para a aquisição de L2 de longo prazo.

Tseng e Warschauer (2023) reforçam essa perspectiva ao defender que o uso educacional da IA deve ser acompanhado do desenvolvimento de letramentos críticos relacionados à avaliação, interpretação e utilização ética das respostas geradas pelos sistemas. Essa dimensão crítica é especialmente relevante no contexto do ensino de inglês como L2, em que a qualidade do *input* e do *feedback* recebidos pelo aprendiz exerce papel central no processo de aquisição linguística.

3 METODOLOGIA

3.1 Delineamento do Estudo

A pesquisa foi desenvolvida por meio do método de revisão integrativa da literatura, que permite reunir, analisar, comparar e sintetizar resultados de pesquisas já publicadas sobre determinado fenômeno (Jackson, 1980; Torraco, 2005; Snyder, 2019; Toronto; Remington, 2020). Esse método foi considerado adequado por possibilitar a identificação de tendências, vantagens, limitações e lacunas relacionadas ao uso do ChatGPT no ensino-aprendizagem de inglês como L2, seguindo as etapas de definição do problema, estabelecimento de critérios de inclusão e exclusão, busca, seleção, análise e interpretação dos resultados.

3.2 Perguntas de Pesquisa

A análise dos estudos foi orientada pelas seguintes perguntas de pesquisa (PP), desdobradas a partir dos objetivos específicos do estudo:

(PP1) Como os estudos empíricos sobre o uso do ChatGPT no ensino-aprendizagem de inglês como L2 se distribuem em relação às habilidades linguísticas investigadas?

(PP2) Quais são os efeitos do uso do ChatGPT no desenvolvimento da compreensão leitora (*reading*) em inglês como L2, em termos de vantagens e limitações?

(PP3) Quais são os efeitos do uso do ChatGPT no desenvolvimento da produção escrita (*writing*) em inglês como L2, em termos de vantagens e limitações?

(PP4) Quais são os efeitos do uso do ChatGPT no desenvolvimento da compreensão oral (*listening*) em inglês como L2, em termos de vantagens e limitações?

(PP5) Quais são os efeitos do uso do ChatGPT no desenvolvimento da produção oral (*speaking*) em inglês como L2, em termos de vantagens e limitações?

3.3 Critérios de Inclusão e Exclusão

Foram considerados elegíveis estudos empíricos (experimentais, quase-experimentais ou com coleta de dados primários) que investigassem os efeitos do ChatGPT no desenvolvimento de habilidades linguísticas em inglês como L2, publicados entre novembro de 2022 e junho de 2025. A opção por restringir o corpus a estudos empíricos justifica-se pela natureza das perguntas de pesquisa, que demandam evidências derivadas de dados coletados junto a aprendizes reais; estudos exclusivamente teóricos não permitem avaliar os efeitos da ferramenta sobre o desempenho linguístico e foram, portanto, excluídos. Os estudos deveriam abordar pelo menos uma das seguintes habilidades: compreensão leitora (*reading*), compreensão oral (*listening*), produção escrita (*writing*) ou produção oral (*speaking*). Foram descartados artigos de opinião, editoriais, resenhas, comentários, publicações duplicadas, estudos sobre outras línguas que não o inglês, investigações restritas a percepções de professores ou gestores sem foco no desenvolvimento linguístico dos aprendizes, e publicações cujo texto completo não estava disponível para acesso.

3.4 Busca e Seleção dos Estudos

A busca bibliográfica foi realizada em data única (14 de julho de 2025) nas bases de dados ERIC e SciELO, contemplando publicações disponibilizadas entre novembro de 2022, período correspondente ao lançamento público do ChatGPT, e junho de 2025. Foi adotado um descritor único e intencional, “CHATGPT”, aplicado em ambas as bases sem restrições terminológicas adicionais. Essa estratégia, orientada pela maximização da recuperação de registros (*high recall*), teve como objetivo reduzir o risco de exclusão de estudos relevantes que

não utilizassem terminologia padronizada, transferindo o refinamento do corpus para as etapas subsequentes de triagem.

A busca na base ERIC resultou em 210 registros. A consulta à SciELO não retornou nenhum resultado para os descritores utilizados, o que constitui, por si só, um indicativo relevante da ausência de representação do tema nas bases de dados de língua portuguesa e de circulação latino-americana. O total de 210 registros recuperados foi submetido à leitura de títulos e resumos, etapa em que foram excluídas publicações sem relação direta com o desenvolvimento de habilidades linguísticas em inglês como L2, incluindo, entre outros, estudos que investigavam exclusivamente percepções de professores sobre a ferramenta, sem foco em aprendizagem linguística mensurável. Após essa triagem, 60 estudos foram considerados potencialmente elegíveis e submetidos à leitura integral, resultando em 23 com delineamento empírico e 37 de natureza teórica. Os 37 estudos teóricos foram excluídos por não atenderem ao critério de inclusão que exigia a presença de dados empíricos capazes de responder às perguntas de pesquisa estabelecidas. Dos 23 estudos empíricos remanescentes, um foi excluído por inacessibilidade ao texto completo (LI, 2024), resultando em um corpus final de 22 estudos, considerados adequados para os objetivos desta revisão integrativa.

3.5 Caracterização do Corpus

Os 22 estudos incluídos no corpus apresentaram diferentes contextos educacionais, metodologias e perfis de participantes. A maioria das pesquisas foi conduzida em contextos de países asiáticos e do Oriente Médio, incluindo China, Turquia, Indonésia, Coreia do Sul, Taiwan, Arábia Saudita e Egito. Predominaram estudos realizados com estudantes universitários, embora algumas pesquisas também tenham envolvido alunos do ensino médio e estudantes de pós-graduação. O Quadro 1 a seguir apresenta uma síntese dos estudos do corpus.

Quadro 1 – Síntese dos estudos incluídos no corpus

Autor/Ano	Habilidade	Participantes	Contexto
Abdallah (2024)	<i>Writing</i>	50 professores em formação na área de Educação, aprendizes de inglês	Egito
Abduljawad (2024)	<i>Writing</i>	130 aprendizes universitários de inglês	Arábia Saudita
Bai & Wei (2024)	<i>Writing</i>	34 aprendizes universitários de Letras-Inglês	China
Çelik <i>et al.</i> (2024)	<i>Reading</i>	105 aprendizes universitários de inglês	Turquia
Dillon (2024)	<i>Writing</i>	99 aprendizes universitários de inglês	Coreia do Sul
EIEbyary & Shabara (2024)	<i>Writing</i>	Não informado	Egito
Escalante <i>et al.</i> (2023)	<i>Writing</i>	91 aprendizes universitários de inglês	Ásia/Pacífico

Geçkin <i>et al.</i> (2023)	<i>Writing</i>	43 aprendizes universitários de inglês	Turquia
Gozali <i>et al.</i> (2024)	<i>Writing</i>	18 aprendizes universitários de Letras-Inglês	Indonésia
Guo & Wang (2024)	<i>Writing</i>	5 docentes de inglês	China
Sapan & Uzun (2024)	<i>Writing</i>	66 aprendizes de inglês do Ensino Médio	Turquia
Sawangwan (2024)	<i>Writing</i>	50 aprendizes universitários de inglês	Tailândia
Seo (2024)	<i>Writing</i>	44 aprendizes universitários de inglês	Coreia do Sul
Solovey (2024)	<i>Writing</i>	30 aprendizes universitários de inglês	Israel
Tsai <i>et al.</i> (2024)	<i>Writing</i>	44 aprendizes universitários de Letras-Inglês	Taiwan
Üstünbaş (2024)	<i>Speaking</i>	4 aprendizes universitários de inglês	Turquia
Wang <i>et al.</i> (2024)	<i>Reading</i>	42 aprendizes universitários de inglês	China
Werdiningsih <i>et al.</i> (2024)	<i>Writing</i>	16 aprendizes de inglês de pós-graduação	Indonésia
Woo <i>et al.</i> (2024)	<i>Writing</i>	21 aprendizes de inglês do Ensino Fundamental 2 e Ensino Médio	Hong Kong
Yan (2023)	<i>Writing</i>	8 aprendizes universitários de Letras-Inglês	China
Yuan <i>et al.</i> (2024)	<i>Writing</i>	7 aprendizes de escrita acadêmica em inglês	Malásia
Zou & Huang (2024)	<i>Writing</i>	215 aprendizes de inglês de doutorado	China

Fonte: Elaborado pela autora com base nos estudos analisados (2026).

3.6 Procedimentos de Análise e Interpretação dos Dados

A análise dos estudos selecionados foi conduzida em duas etapas sequenciais. Na primeira etapa, os 22 estudos incluídos no corpus foram organizados em uma planilha de classificação, na qual cada publicação foi categorizada conforme os seguintes critérios: habilidade linguística investigada (*reading*, *writing*, *listening* ou *speaking*), perfil dos participantes, contexto educacional, país de realização e principais resultados reportados. Essa etapa permitiu distinguir os estudos de natureza empírica, únicos considerados elegíveis para inclusão, daqueles de natureza exclusivamente teórica, que foram excluídos do corpus.

Na segunda etapa, os estudos foram agrupados conforme a habilidade linguística investigada e submetidos à análise qualitativa e temática. A partir da leitura integral de cada publicação, foram identificadas e categorizadas as principais vantagens, limitações e lacunas de pesquisa relatadas pelos autores. Os resultados foram então comparados entre os estudos de cada categoria, permitindo identificar padrões de convergência e divergência. Esse procedimento possibilitou responder às cinco perguntas de pesquisa estabelecidas e construir uma síntese interpretativa dos efeitos do uso do ChatGPT no ensino-aprendizagem de inglês como L2.

Cabe registrar que a análise foi conduzida pela autora, sem procedimento formal de codificação interavaliadores. Para mitigar os riscos associados à subjetividade analítica inerente a esse tipo de procedimento, todas as decisões de categorização foram posteriormente revisadas e validadas pelo orientador, garantindo uma camada adicional de controle sobre a consistência interpretativa do processo. O conjunto de decisões analíticas foi documentado sistematicamente na planilha de classificação, assegurando rastreabilidade ao longo de todas as etapas da revisão. Não obstante, a ausência de um terceiro avaliador independente e de teste formal de concordância interavaliadores constitui uma limitação do estudo.

4 RESULTADOS E DISCUSSÃO

4.1 (PP1) Distribuição dos Estudos por Habilidade Linguística

Tabela 1 - Distribuição dos resultados dos estudos por habilidade

Habilidade	Nº de Estudos	% do corpus
<i>Writing</i> (Produção Escrita)	19	86,36%
<i>Reading</i> (Compreensão Leitora)	2	9,09%
<i>Speaking</i> (Produção Oral)	1	4,55%
<i>Listening</i> (Compreensão Oral)	0	0%
Total	22	100%

Fonte: A autora (2026).

O levantamento dos 22 estudos empíricos que compõem esta revisão integrativa, publicados entre novembro de 2022 e junho de 2025, evidenciou uma distribuição bastante desigual entre as quatro habilidades linguísticas analisadas, conforme apresentado na Tabela 1. A produção escrita (*writing*) concentra a maior parte das pesquisas, com 19 estudos, correspondendo a 86,36% do corpus. Em contrapartida, a compreensão leitora (*reading*) é abordada em apenas 2 estudos (9,09%), e a produção oral (*speaking*) em apenas 1 estudo (4,55%). Registra-se ainda a ausência completa de estudos empíricos dedicados ao efeito do ChatGPT sobre a compreensão oral (*listening*) em inglês como L2, o que evidencia uma lacuna expressiva na literatura da área.

Dada a natureza da ferramenta, não é surpreendente que a maioria dos estudos se concentre na produção escrita. O ChatGPT caracteriza-se como um modelo de linguagem generativo que opera fundamentalmente por meio de texto escrito: recebe *prompts* e gera respostas escritas. Essa característica o qualifica de forma direta para tarefas relacionadas à

escrita, como a geração de ideias, a revisão gramatical, a sugestão de vocabulário e o fornecimento de *feedback* sobre textos. Segundo Kohnke *et al.* (2023), as funcionalidades do ChatGPT articulam-se diretamente com o processo de escrita em língua adicional, o que contribui para compreender por que essa habilidade tem despertado maior interesse investigativo em comparação às demais.

A ausência de pesquisas com foco na compreensão oral (*listening*) merece atenção especial, uma vez que essa habilidade desempenha papel central na aprendizagem de L2. O *listening* é amplamente considerado a habilidade mais recorrente na comunicação cotidiana e a base para o desenvolvimento das demais competências linguísticas (Rost, 2011). O desenvolvimento dessa habilidade exige exposição a recursos de áudio e fala natural, elementos ausentes nas versões iniciais do ChatGPT, concebido originalmente para interações exclusivamente textuais. Apenas recentemente a ferramenta passou a incorporar interfaces de voz em tempo real, e como essa funcionalidade ainda era relativamente nova à época dos estudos analisados, sua ausência nas investigações identificadas é compreensível, embora não deixe de representar uma lacuna relevante a ser preenchida por pesquisas futuras.

Os resultados desta revisão integrativa convergem com os de Lo *et al.* (2024), que, ao conduzirem uma revisão sistemática de 70 estudos empíricos sobre a utilização do ChatGPT no ensino de inglês como L2, também identificaram predominância de pesquisas voltadas para a escrita, seguidas por estudos sobre produção oral e leitura, e ausência completa de investigações sobre compreensão oral. Essa convergência indica que o padrão observado não se limita às bases consultadas nesta revisão, mas reflete uma tendência mais ampla na literatura internacional. Assim, torna-se evidente a necessidade de ampliar as investigações sobre habilidades menos estudadas, especialmente *listening* e *speaking*, a fim de construir uma compreensão mais abrangente e equilibrada dos efeitos do ChatGPT no processo de aprendizagem de inglês como L2.

4.2 (PP2) Foco em Compreensão Leitora (*Reading*)

Entre os 22 estudos que constituem o corpus desta revisão integrativa, somente dois (9,09%) examinaram especificamente o efeito do ChatGPT na compreensão leitora em inglês como L2: Çelik *et al.* (2024) e Wang *et al.* (2024). Apesar da base empírica ainda ser bastante limitada, os resultados dessas pesquisas fornecem evidências preliminares relevantes a respeito das vantagens e limitações do uso do ChatGPT como instrumento auxiliar no desenvolvimento dessa habilidade, conforme sintetizado no Quadro 2.

Quadro 2 – Comparação dos estudos sobre *reading*

Critério	Çelik et al. (2024)	Wang et al. (2024)
País	Turquia	China
Participantes	105 graduandos de inglês como L2, nível B1	42 graduandos de inglês como L2 (N experimental = 22; N controle = 20)
Uso do ChatGPT	ChatGPT 3.5 para simplificação de texto autêntico de blog	Sistema ChatPRCS para geração personalizada de questões de compreensão
Design	<i>Within-subjects</i> , sessão única	Quase-experimental, 4 semanas
Foco	Compreensão leitora, inferência e ansiedade de leitura	Desempenho, motivação e carga cognitiva
Vantagens	Melhora significativa na compreensão ($Z = -8,142$; $p < ,001$) e inferência ($Z = -3,266$; $p < ,001$)	Questões sem diferença significativa em relação às de especialistas ($t = 0,47$; $p > ,05$); ganhos em desempenho ($\eta^2 = 0,62$) e motivação
Limitações	Sem redução significativa da ansiedade ($Z = -1,265$; $p = ,206$); sessão única; amostra de contexto único	Amostra restrita; avaliação de qualidade predominantemente subjetiva; eficácia dependente do volume de dados históricos

Fonte: A autora (2026).

No que se refere aos benefícios, a pesquisa de Çelik *et al.* (2024) revelou que os 105 graduandos de inglês como L2, nível B1, apresentaram médias consideravelmente superiores no pós-teste de compreensão ($M = 60,00$; $DP = 24,69$) em relação ao pré-teste ($M = 42,67$; $DP = 26,10$), após lerem versões simplificadas pelo ChatGPT 3.5 de um texto autêntico de blog ($Z = -8,142$; $p < ,001$). O estudo adotou um desenho *within-subjects*, no qual cada participante serviu como seu próprio controle. Os autores interpretaram esse resultado como evidência de que o ChatGPT pode atuar como um andaime pedagógico eficaz para aprendizes de nível intermediário, auxiliando na compreensão de textos autênticos sem substituí-los completamente. Ganhos significativos foram observados também na capacidade de inferência dos participantes ($Z = -3,266$; $p < ,001$), sugerindo que a exposição prévia ao texto simplificado forneceu o conhecimento de fundo necessário para que os estudantes inferissem o significado de termos desconhecidos na versão original.

Wang et al. (2024) seguiram uma abordagem distinta ao desenvolver o ChatPRCS (*ChatGPT-based Personalized Reading Comprehension Support*), um sistema personalizado de apoio à compreensão leitora fundamentado na teoria da Zona de Desenvolvimento Proximal de Vygotsky. Em um estudo quase-experimental conduzido ao longo de quatro semanas com 42 graduandos em uma universidade chinesa, o sistema utilizou o ChatGPT para gerar questões de

compreensão com graus de dificuldade adaptados ao perfil de cada estudante, considerando seu histórico de desempenho e modelando o esquecimento por meio da curva de Ebbinghaus. Os resultados do Teste-T aplicado às avaliações de três especialistas independentes não indicaram diferença estatisticamente significativa entre as questões geradas pelo ChatPRCS e as elaboradas por especialistas humanos ($t = 0,47$; $p > ,05$), com médias de 82,73 e 81,66 respectivamente, além de ganhos significativos em desempenho e motivação no grupo experimental em relação ao grupo controle ($F = 64,15$; $p < ,001$; $\eta^2 = 0,62$).

Em relação às limitações, Çelik *et al.* (2024) constataram que a simplificação de textos pelo ChatGPT não produziu redução estatisticamente significativa nos níveis de ansiedade de leitura dos participantes ($Z = -1,265$; $p = ,206$). Esse resultado sugere que a ferramenta não atua diretamente sobre o componente afetivo da leitura em L2, que envolve insegurança diante de textos autênticos e receio de não compreender plenamente. A ansiedade de leitura é reconhecida na literatura como um fator que compromete significativamente a compreensão leitora em contextos de L2 (Saito *et al.*, 1999), e sua persistência mesmo após a intervenção indica que o ChatGPT, quando utilizado de forma isolada e em sessão única (*one-shot*), não é suficiente para lidar com essa dimensão emocional. Além disso, o estudo foi conduzido em uma única instituição turca com aprendizes de nível B1, o que restringe a generalização dos resultados para outros contextos culturais e níveis de proficiência.

Wang *et al.* (2024), por sua vez, apontam como limitações o tamanho reduzido da amostra, restrita a um único curso de uma universidade chinesa, e a dependência de avaliações subjetivas para mensurar a qualidade das questões geradas. Os autores também indicam que a eficácia do sistema pode variar conforme o volume de dados históricos disponíveis sobre cada aprendiz, comprometendo a acurácia das predições nas fases iniciais de uso.

Esses resultados são consistentes com os de Lo *et al.* (2024), cuja revisão sistemática de 70 estudos empíricos sobre o ChatGPT no ensino de inglês como L2 também identificou escassez de pesquisas voltadas para a compreensão leitora, sinalizando que essa habilidade permanece subrepresentada na literatura. Em síntese, os dois estudos indicam que o ChatGPT apresenta potencial como recurso de apoio ao desenvolvimento da compreensão leitora, especialmente quando empregado para tornar textos autênticos mais acessíveis a aprendizes de nível intermediário ou para gerar questões personalizadas fundamentadas na Zona de Desenvolvimento Proximal. Contudo, a base empírica ainda é restrita, o que demanda estudos de maior duração, amostras mais diversas e desenhos metodológicos capazes de capturar não apenas ganhos cognitivos, mas também os impactos afetivos associados à leitura em L2.

4.3 (PP3) Foco em Produção Escrita (*Writing*)

A produção escrita (*writing*) foi a habilidade mais investigada nos estudos que compõem o corpus desta revisão integrativa, presente em 19 dos 22 estudos selecionados, o que corresponde a 86,36% do total. Essa concentração evidencia tanto a natureza textual da ferramenta quanto a centralidade da escrita no ensino de inglês como L2. Os estudos abordam aspectos como *feedback* automatizado, apoio ao processo de revisão, comparação com o *feedback* docente e percepções dos usuários. Ainda, os 19 estudos incluídos nesta seção apresentam características relativamente homogêneas quanto ao contexto de aplicação, perfil dos participantes e delineamento metodológico. Observa-se predominância de pesquisas conduzidas em países asiáticos e do Oriente Médio, como China, Indonésia, Turquia, Arábia Saudita, Egito, Coreia do Sul e Taiwan. Em relação ao perfil dos participantes, a maioria das pesquisas foi conduzida com estudantes universitários de graduação. Apenas um estudo investigou estudantes do ensino médio (Sapan; Uzun, 2024), e um envolveu mestrandos (Werdiningsih *et al.*, 2024). O Quadro 3 apresenta uma síntese dos estudos incluídos nesta seção, organizados por autor, ano, principal vantagem e principal limitação reportadas.

Quadro 3 — Síntese dos Estudos de *Writing*

Autor(es)	Ano	Principal vantagem reportada	Principal limitação reportada
Abdallah	2024	Melhora na autoexpressão e escrita reflexiva com apoio de IA e aprendizagem autorregulada	Design sem grupo controle; efeitos de longo prazo não investigados
Abduljawad	2024	<i>Feedback</i> personalizado, expansão de vocabulário e maior autonomia dos aprendizes	Perda de individualidade e criatividade; preocupações éticas; <i>feedback</i> contextualmente inadequado
Bai; Wei	2024	Apoio ao processo de revisão textual; superação de bloqueios criativos	Dependência excessiva; aceitação acrítica das sugestões
Dillon	2024	Treinamento em <i>prompt literacy</i> favorece uso estratégico do ChatGPT para escrita; estratégias compensatórias e metacognitivas bem avaliadas	Ausência de grupo controle e pré/pós-teste; interação restrita ao formato textual; estratégias sociais pouco aproveitadas
EIEbyary; Shabara	2024	<i>Feedback</i> imediato valorizado pelos aprendizes; retorno claro sobre aspectos linguísticos	Sugestões frequentemente genéricas, inadequadas ou incorretas

Autor(es)	Ano	Principal vantagem reportada	Principal limitação reportada
Escalante <i>et al.</i>	2023	<i>Feedback</i> eficaz; preferência dos estudantes pelo retorno do ChatGPT	Qualidade variável do <i>feedback</i> ; questões de integridade acadêmica
Geçkin <i>et al.</i>	2023	Avaliação automatizada da escrita com resultados comparáveis aos de avaliadores humanos	Limitações na avaliação de aspectos subjetivos da escrita
Gozali <i>et al.</i>	2024	Desenvolvimento do letramento em <i>feedback</i> ; integração como ferramenta de avaliação automatizada	Amostra muito reduzida; generalização limitada
Guo; Wang	2024	Apoio ao <i>feedback</i> docente; redução da carga de trabalho do professor	<i>Feedback</i> nem sempre adequado ao nível do estudante; necessidade de validação humana
Sapan; Uzun	2024	Integração ao ensino de inglês como L2 com ganhos em vocabulário	Instrução tradicional produziu resultados equivalentes ou superiores em alguns aspectos
Sawangwan	2024	<i>Feedback</i> imediato; apoio ao processo de escrita	Desempenho do grupo com instrução tradicional equivalente ao do grupo com ChatGPT
Seo	2024	Potencial para escrita narrativa; melhora em aspectos estruturais	Curta duração; efeitos de longo prazo não avaliados
Solovey	2024	<i>Feedback</i> percebido favoravelmente em comparação ao de pares e do professor	Estudo de pequena escala; generalização limitada
Tsai <i>et al.</i>	2024	Melhora na qualidade textual; percepções positivas dos estudantes	Curta duração; desafios de integridade acadêmica
Werdiningsih <i>et al.</i>	2024	Uso estratégico em diferentes etapas da escrita acadêmica	Dependência excessiva; aceitação acrítica das sugestões
Woo <i>et al.</i>	2024	Aumento da motivação e satisfação com o processo de escrita; redução da carga cognitiva	Preocupações com integridade acadêmica; dependência da ferramenta
Yan	2023	Melhora na proficiência escrita; apoio ao desenvolvimento linguístico	Riscos para a integridade acadêmica e autoria; uso inadequado em atividades avaliativas
Yuan <i>et al.</i>	2024	Aumento da confiança e autonomia; apoio à escrita acadêmica	Dependência excessiva; aceitação acrítica das sugestões; preocupações éticas

Autor(es)	Ano	Principal vantagem reportada	Principal limitação reportada
Zou; Huang	2024	Apoio em múltiplas etapas do processo de escrita	Riscos de plágio; qualidade variável do <i>feedback</i> ; dependência da ferramenta

Fonte: A autora (2026).

Uma das vantagens mais recorrentes identificadas nos estudos analisados foi a capacidade do ChatGPT de fornecer *feedback* imediato e personalizado aos estudantes. Diferentemente dos processos tradicionais de correção, que dependem da disponibilidade do professor, a ferramenta oferece retorno instantâneo sobre aspectos linguísticos da escrita, como gramática, vocabulário, coesão e organização textual. Estudos como os de ElEbyary e Shabara (2024), Guo e Wang (2024), Solovey (2024), Escalante *et al.* (2023) e Gozali *et al.* (2024) demonstram que os aprendizes valorizam especialmente a rapidez e a clareza das sugestões fornecidas. Além disso, o *feedback* automatizado permite múltiplos ciclos de revisão, favorecendo a identificação de erros recorrentes e estimulando a reflexão sobre o próprio processo de escrita.

Outra vantagem frequentemente observada refere-se à melhora do desempenho escrito dos estudantes. Estudos que utilizaram pré e pós-testes verificaram avanços significativos em aspectos como correção gramatical, organização textual, clareza das ideias e ampliação do repertório lexical. Abdallah (2024), Tsai *et al.* (2024), Seo (2024) e Abduljawad (2024) encontraram evidências de que os participantes produziram textos mais bem estruturados após o uso do ChatGPT. Em muitos casos, os estudantes demonstraram redução de erros linguísticos e maior capacidade de desenvolver argumentos de forma organizada, sugerindo que a ferramenta pode contribuir positivamente para o aprimoramento da escrita em inglês como L2.

Os estudos também destacam o papel do ChatGPT como apoio em diferentes etapas do processo de escrita. A ferramenta auxilia na geração de ideias, no planejamento do texto, na organização dos argumentos e na revisão linguística. Bai e Wei (2024), Werdiningsih *et al.* (2024), Zou e Huang (2024) e Yuan *et al.* (2024) relatam que os estudantes utilizaram o ChatGPT para superar bloqueios criativos, estruturar parágrafos e aperfeiçoar aspectos formais da escrita. Dessa forma, a IA funciona não apenas como corretora de erros, mas como parceira durante o desenvolvimento textual, contribuindo para tornar o processo de produção escrita mais fluido e menos intimidante para os aprendizes.

O aumento da motivação e da autonomia dos estudantes constitui outro benefício relatado. Woo *et al.* (2024), Yuan *et al.* (2024) e Abduljawad (2024) observaram que os

aprendizes se sentiram mais confiantes ao escrever com o apoio da ferramenta. A possibilidade de obter respostas imediatas, experimentar diversas formas de expressão e revisar textos sem receio de julgamento contribuiu para criar um ambiente mais receptivo para a aprendizagem. Além disso, muitos participantes relataram maior autonomia para identificar e corrigir problemas em seus próprios textos, adotando uma postura mais proativa no processo de aprendizagem.

Além dos benefícios para os estudantes, alguns estudos apontam contribuições relevantes para o trabalho docente. Guo e Wang (2024) e Escalante *et al.* (2023) destacam que o ChatGPT pode auxiliar professores ao fornecer *feedback* preliminar e identificar problemas linguísticos básicos, reduzindo a carga de trabalho relacionada à correção de textos. Esse apoio permite que o professor concentre sua atenção em aspectos mais complexos da escrita, como argumentação, criatividade e pensamento crítico. Assim, a ferramenta pode ser compreendida como um recurso complementar ao ensino, e não como substituta da atuação docente.

Uma limitação mencionada nos estudos refere-se às preocupações relacionadas à integridade acadêmica. Yan (2023), Yuan *et al.* (2024), Zou e Huang (2024) e Woo *et al.* (2024) alertam para os riscos de plágio, autoria indefinida e uso inadequado da ferramenta em atividades avaliativas. A facilidade com que o ChatGPT produz textos completos pode incentivar alguns estudantes a utilizarem conteúdos gerados pela IA sem participação efetiva no processo de escrita, comprometendo objetivos pedagógicos relacionados ao desenvolvimento da autoria e do pensamento crítico.

Outra limitação amplamente discutida é a possibilidade de dependência excessiva da ferramenta. Werdiningsih *et al.* (2024), Zou e Huang (2024) e Yuan *et al.* (2024) observam que alguns estudantes tendem a aceitar automaticamente as sugestões fornecidas pelo ChatGPT, reduzindo sua participação ativa no processo de revisão textual. Quando utilizada sem orientação adequada, a ferramenta pode diminuir o esforço cognitivo necessário para a construção autônoma do conhecimento e dificultar o desenvolvimento da capacidade crítica do aprendiz.

Embora muitos estudos relatem benefícios do uso do ChatGPT, alguns resultados indicam que a ferramenta não é necessariamente superior às práticas tradicionais de ensino. Sapan e Uzun (2024) e Sawangwan (2024) identificaram situações em que estudantes submetidos à instrução tradicional obtiveram desempenho semelhante ou superior ao dos grupos que utilizaram o ChatGPT. Esses resultados demonstram que a simples adoção da tecnologia não garante melhores resultados educacionais, reforçando a importância da

mediação pedagógica e do planejamento didático como condições indispensáveis para o aproveitamento eficaz da ferramenta.

Apesar de sua rapidez, o *feedback* gerado pelo ChatGPT nem sempre apresenta qualidade adequada. Guo e Wang (2024) e EIEbyary e Shabara (2024) relatam casos em que as sugestões fornecidas pela ferramenta foram excessivamente genéricas, inadequadas ao contexto ou incorretas. Em algumas situações, o ChatGPT modificou o sentido original do texto ou apresentou recomendações incompatíveis com o nível de proficiência do estudante. Essas limitações evidenciam a necessidade de supervisão humana para validar e interpretar o *feedback* automatizado.

Uma limitação metodológica transversal aos estudos analisados diz respeito à sua breve duração. A maior parte das pesquisas foi realizada ao longo de poucas semanas, o que dificulta a análise dos efeitos a longo prazo do uso do ChatGPT sobre a produção escrita em L2. Ainda não há evidências suficientes para concluir se os benefícios observados se mantêm ao longo do tempo ou se os estudantes desenvolvem dependência progressiva da ferramenta. Esse padrão é consistente com o observado por Lo *et al.* (2024) em sua revisão sistemática de 70 pesquisas empíricas sobre o ChatGPT no ensino de inglês como L2, que também identificou a limitação temporal como uma das principais lacunas metodológicas da área.

Um achado singular no conjunto dos estudos analisados diz respeito à estruturação pedagógica do uso do ChatGPT para a produção escrita por meio do treinamento explícito em *prompt literacy*². Dillon (2024), em estudo de delineamento misto conduzido com 99 estudantes universitários sul-coreanos de inglês em nível A2, organizou sessões semanais supervisionadas ao longo de seis semanas com o objetivo de familiarizar os participantes com *prompts* específicos alinhados às categorias do *Strategy Inventory for Language Learning* (SILL) e do modelo *Strategic Self-Regulation* (S2R), ambos de Oxford (1990, 2016). As atividades de escrita envolveram produção de *scripts*, correção gramatical, geração de tópicos e sumarização de conversas, todas mediadas por interação textual com o ChatGPT. Os resultados indicam que as estratégias compensatórias, especialmente correção de erros ($M = 4,18$) e ajuste de dificuldade ($M = 4,12$), foram as mais bem avaliadas pelos participantes, seguidas pelas estratégias metacognitivas de planejamento e sumarização ($M = 4,11$ a $4,15$). O estudo demonstra que o desempenho dos participantes nas tarefas de escrita foi favorecido não pelo

² O termo *prompt literacy*, utilizado em estudos educacionais, guarda estreita relação com o que a área da informática denomina engenharia de prompts (*prompt engineering*). Enquanto este último enfatiza o aspecto técnico de elaboração de instruções para sistemas de IA, o primeiro destaca a dimensão pedagógica e formativa da competência de usuários em interagir estrategicamente com tais ferramentas.

acesso irrestrito à ferramenta, mas pela formação sistemática em como interagir com ela de maneira estratégica, respondendo diretamente à lacuna identificada por estudos anteriores do corpus que apontaram a ausência de treinamento formal como um dos principais obstáculos ao aproveitamento pedagógico efetivo do ChatGPT em contextos de ensino de L2 (Shao; Xia, 2023; Vo; Nguyen, 2024, apud Dillon, 2024). As estratégias sociais, contudo, obtiveram os escores mais baixos ($M = 3,96$ a $4,07$), resultado que o autor associa às limitações de autenticidade interacional do sistema, coerente com as críticas de artificialidade das respostas reportadas em outros estudos deste corpus.

Em síntese, os resultados dos 19 estudos examinados sugerem que o ChatGPT possui potencial considerável como instrumento de apoio à produção escrita em inglês como L2, especialmente ao oferecer *feedback* instantâneo, apoiar o processo de revisão textual e aumentar a motivação dos aprendizes. No entanto, os resultados também evidenciam limitações significativas, como riscos associados à integridade acadêmica, dependência excessiva da ferramenta, qualidade variável do *feedback* e indícios de que a instrução tradicional pode gerar resultados equivalentes ou superiores em determinados contextos. Esses achados indicam que o ChatGPT deve ser incorporado ao ensino de escrita de maneira crítica e pedagogicamente orientada, como recurso complementar, e não como substituto do trabalho docente (Guo; Wang, 2024; Escalante *et al.*, 2023; Sawangwan, 2024).

4.4 (PP4) Foco em Compreensão Oral (*Listening*)

A análise do corpus desta revisão integrativa revelou ausência completa de estudos empíricos dedicados ao impacto do ChatGPT no desenvolvimento da compreensão oral (*listening*) em inglês como L2. Esse resultado não decorre de uma limitação metodológica do presente levantamento, mas constitui, em si mesmo, um achado científico relevante: evidencia que uma das quatro habilidades linguísticas centrais no ensino de L2 permanece sistematicamente negligenciada na literatura sobre o uso de IA generativa no contexto educacional. A convergência com os resultados de Lo *et al.* (2024), cuja revisão sistemática de 70 estudos empíricos publicados entre dezembro de 2022 e maio de 2024 também não localizou nenhuma investigação dedicada especificamente ao *listening*, reforça a percepção de que essa ausência reflete uma tendência estrutural da área, e não uma contingência das bases ou do período consultados nesta revisão.

A explicação mais direta para essa lacuna reside nas características técnicas das versões iniciais do ChatGPT. Quando disponibilizado ao público em novembro de 2022, o sistema

operava exclusivamente por meio de interações textuais escritas, sem qualquer capacidade de processamento ou geração de áudio. Essa arquitetura restringia objetivamente as possibilidades de aplicação da ferramenta ao *listening*, habilidade que exige exposição a elementos da fala natural que transcendem a linguagem escrita, como prosódia, entonação, ritmo e fenômenos de redução fonológica (Rost, 2011). Pesquisadores direcionaram suas investigações para habilidades mais compatíveis com o formato textual, em especial a produção escrita, e soma-se a isso o fato de que a avaliação da compreensão oral demanda metodologias de maior complexidade operacional, o que pode ter funcionado como fator adicional de inibição da produção científica na área.

A relevância dessa lacuna torna-se ainda mais evidente quando se considera o papel central que a compreensão oral desempenha na aquisição de L2. De acordo com Rost (2011), o *listening* é a habilidade linguística mais recorrentemente utilizada na comunicação cotidiana e constitui a principal via de acesso ao *input* compreensível, conceito central na teoria de aquisição de línguas de Krashen (1982), que postula a exposição a insumo linguístico ligeiramente acima do nível atual do aprendiz como condição necessária para o desenvolvimento da interlíngua. O *listening* também desempenha função estruturante para as demais habilidades: a construção de vocabulário receptivo, a internalização de padrões gramaticais e a automatização de estruturas fonológicas dependem, em larga medida, da quantidade e qualidade de exposição oral à língua-alvo. A ausência de investigações sobre como ferramentas de IA podem mediar esse processo representa, portanto, uma lacuna não apenas quantitativa na literatura, mas qualitativamente significativa para a compreensão dos efeitos do ChatGPT sobre o desenvolvimento comunicativo em L2.

O cenário tecnológico, contudo, tem evoluído de forma acelerada. As versões mais recentes do ChatGPT, em especial as baseadas no GPT-4o, incorporaram recursos multimodais que incluem reconhecimento de voz em tempo real e geração de áudio com prosódia naturalística, expandindo consideravelmente as possibilidades pedagógicas da ferramenta para além do texto escrito. Essas funcionalidades criam condições técnicas para o desenvolvimento de aplicações voltadas ao *listening*, como exposição a diferentes padrões de pronúncia, atividades de compreensão oral adaptadas ao nível do aprendiz, *feedback* sobre acurácia de compreensão auditiva e simulações de interações orais autênticas em contextos variados. Pesquisas futuras que explorem essas possibilidades são não apenas viáveis, mas necessárias para preencher a lacuna identificada nesta revisão e para construir uma compreensão mais equilibrada e abrangente dos efeitos do ChatGPT sobre o desenvolvimento das quatro habilidades linguísticas em inglês como L2.

4.5 (PP5) Foco em Produção Oral (*Speaking*)

No que diz respeito à habilidade de produção oral (*speaking*), o corpus desta revisão integrativa identificou apenas um estudo dedicado especificamente ao uso do ChatGPT para o desenvolvimento dessa competência (Üstünbaşı, 2024). Esse resultado reforça um padrão recorrente na literatura recente sobre o uso de IA no ensino de línguas, caracterizado pela predominância de investigações voltadas à produção escrita e ao *feedback* corretivo, em detrimento das habilidades orais. Tal tendência também é confirmada por Lo *et al.* (2024), que, ao analisarem 70 estudos empíricos sobre o uso do ChatGPT no ensino de inglês como L2, localizaram apenas cinco trabalhos centrados no *speaking*, evidenciando uma lacuna persistente na agenda de pesquisa, o que limita a consolidação de evidências robustas sobre o potencial da ferramenta para o desenvolvimento da oralidade.

O estudo de Üstünbaşı (2024), de natureza qualitativa e delineado como estudo de caso, investigou estudantes universitários turcos de inglês como L2 por meio de entrevistas semiestruturadas, registros das interações com o sistema e sessões de *recall* estimulado, submetendo os dados à análise temática. Os resultados indicam que os participantes perceberam o ChatGPT como um parceiro de prática linguística continuamente disponível, especialmente relevante em contextos nos quais há escassez de interlocutores humanos para interação em inglês fora do ambiente formal de ensino. Essa disponibilidade constante foi associada ao aumento de oportunidades de exposição à língua e ao fortalecimento da autoconfiança no uso do idioma, particularmente em atividades de simulação de conversas cotidianas e tópicos acadêmicos.

Entre os principais benefícios identificados, destaca-se a redução da ansiedade comunicativa, uma vez que a interação mediada por IA era percebida pelos participantes como livre de julgamento social, o que favorecia um ambiente seguro para a experimentação linguística. Nesse contexto, os estudantes relataram maior disposição para testar estruturas linguísticas mais complexas e para se engajar em práticas de produção discursiva sem o receio de erros, o que contribuiu para o desenvolvimento da autonomia de aprendizagem. Além disso, o ChatGPT foi utilizado como ferramenta para a simulação de diferentes gêneros interacionais, como entrevistas, debates e conversas informais, que permitiam ajustes no nível de complexidade linguística e promoviam maior personalização do processo de aprendizagem, reforçando seu potencial como recurso complementar ao ensino tradicional de línguas.

Apesar dessas vantagens percebidas, o estudo também evidencia limitações estruturais relevantes. A principal delas decorre do fato de que, no momento da investigação, as interações

com o ChatGPT eram realizadas exclusivamente por meio de texto escrito, o que impedia a caracterização da atividade como *speaking* em sentido estrito. Embora os participantes simulassem situações conversacionais, aspectos fundamentais da produção oral, como pronúncia, entonação, ritmo, fluência e prosódia, não eram efetivamente mobilizados, o que restringe significativamente o alcance pedagógico da ferramenta para o desenvolvimento da oralidade. Soma-se a isso a ausência de *feedback* fonológico, dado que o sistema não oferecia avaliação ou correção da pronúncia, limitando a contribuição para dimensões centrais da competência oral em L2.

Adicionalmente, foram observadas limitações relacionadas à naturalidade da interação e à qualidade discursiva das respostas geradas, que em alguns casos se apresentavam excessivamente longas, distantes de padrões interacionais naturais ou desalinhadas com convenções conversacionais humanas. Também ocorreram falhas de compreensão de comandos, o que interrompia a fluidez da interação e comprometia a experiência comunicativa. Outro fator crítico refere-se à dependência do letramento digital dos usuários, uma vez que a eficácia da prática estava condicionada à capacidade dos estudantes de formular *prompts* adequados, o que introduz desigualdades no acesso aos benefícios da ferramenta.

Por fim, os resultados devem ser interpretados com cautela dado o caráter de estudo de caso restrito a um único contexto e a ausência de mensuração objetiva de ganhos em fluência ou competência comunicativa. Deve-se considerar também que versões mais recentes do ChatGPT com recursos de voz ampliam as possibilidades para pesquisas futuras sobre fluência, pronúncia e interação oral autêntica.

5 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo investigar os efeitos do uso do ChatGPT no ensino-aprendizagem de inglês como L2 por meio de uma revisão integrativa. A análise dos 22 estudos empíricos selecionados, publicados entre novembro de 2022 e junho de 2025, permitiu mapear tendências, identificar vantagens e limitações recorrentes e evidenciar lacunas que permanecem abertas na agenda de pesquisa da área.

Em resposta à PP1, os resultados demonstraram distribuição fortemente assimétrica entre as quatro habilidades linguísticas investigadas. A produção escrita (*writing*) concentrou 19 estudos (86,36% do corpus), seguida pela compreensão leitora (*reading*), com 2 estudos (9,09%), e pela produção oral (*speaking*), com 1 estudo (4,55%). A compreensão oral (*listening*) esteve completamente ausente, constituindo a lacuna mais expressiva desta revisão, padrão

convergente com o observado por Lo *et al.* (2024) em levantamento de 70 estudos empíricos, o que indica tratar-se de uma tendência estrutural da área e não de contingência das bases consultadas.

Em resposta à PP2, Çelik *et al.* (2024) demonstraram que a simplificação de textos autênticos pelo ChatGPT produziu ganhos significativos na compreensão e na inferência de aprendizes de nível B1, embora sem efeito sobre a ansiedade de leitura. Wang *et al.* (2024) desenvolveram um sistema de geração personalizada de questões cuja qualidade foi estatisticamente equivalente à de especialistas humanos, com ganhos adicionais em motivação e desempenho. Ambos os estudos foram conduzidos em amostras reduzidas e contextos únicos, o que limita a generalização dos resultados.

Em resposta à PP3, os 19 estudos sobre produção escrita constituíram a base empírica mais robusta desta revisão, convergindo em torno de três vantagens principais: *feedback* imediato e personalizado, apoio às diferentes etapas do processo de escrita e aumento da motivação e da autonomia dos aprendizes. O treinamento explícito em *prompt literacy*, investigado por Dillon (2024), demonstrou ser fator determinante para o aproveitamento pedagógico efetivo da ferramenta, sugerindo que o acesso irrestrito ao ChatGPT sem mediação pedagógica estruturada não garante resultados superiores à instrução tradicional, como atestaram Sapan e Uzun (2024) e Sawangwan (2024). As limitações mais recorrentes foram preocupações com integridade acadêmica, dependência excessiva da ferramenta e qualidade variável do *feedback* gerado.

Em resposta à PP4, a ausência de estudos sobre *listening* reflete as restrições técnicas das versões iniciais do ChatGPT, concebidas para interações exclusivamente textuais. As versões mais recentes, baseadas no GPT-4o, incorporaram recursos de voz em tempo real, abrindo perspectivas concretas para investigações futuras. Em resposta à PP5, Üstünbaş (2024) ofereceu evidências qualitativas sobre a percepção do ChatGPT como parceiro de prática linguística, associado à redução da ansiedade comunicativa e ao aumento da autonomia, embora o caráter textual das interações investigadas impeça sua caracterização como *speaking* em sentido estrito.

Considerados em conjunto, os resultados permitem afirmar que o ChatGPT apresenta potencial pedagógico relevante para o ensino-aprendizagem de inglês como L2, especialmente no domínio da produção escrita, mas que esse potencial depende de condições pedagógicas específicas: formação docente adequada, treinamento explícito dos aprendizes no uso estratégico da ferramenta, desenho cuidadoso das tarefas e supervisão humana do *feedback*

gerado. A simples disponibilização da ferramenta, sem mediação pedagógica, mostrou-se insuficiente para garantir resultados superiores à instrução tradicional.

Este estudo apresenta limitações que devem ser consideradas. A busca foi conduzida em apenas duas bases de dados com um único descritor, o que pode ter excluído estudos relevantes indexados em outras plataformas ou publicados em idiomas distintos do inglês. A análise foi conduzida por uma única pesquisadora, sem procedimento formal de concordância interavaliadores, embora todas as decisões tenham sido revisadas pelo orientador. Para pesquisas futuras, recomenda-se priorizar investigações sobre *listening* e *speaking* com recursos multimodais das versões mais recentes do ChatGPT, estudos longitudinais com grupos controle e amostras diversificadas, investigações sobre *prompt literacy* como variável pedagógica independente e, especialmente, pesquisas conduzidas em contextos latino-americanos e lusófonos, cuja ausência representa a lacuna geográfica mais relevante identificada por esta revisão.

Do ponto de vista geográfico, a concentração dos estudos em países asiáticos e do Oriente Médio, sobretudo China, Turquia, Indonésia, Coreia do Sul, Taiwan, Arábia Saudita e Egito, representa uma limitação importante para a generalização dos achados. Contextos de língua portuguesa e latino-americanos estiveram completamente ausentes do corpus, o que foi evidenciado pela ausência de resultados na SciELO e aponta para uma lacuna geográfica e linguística significativa. Os efeitos do ChatGPT no ensino de inglês como L2 podem variar consideravelmente em função de fatores culturais, institucionais e de infraestrutura tecnológica, e a literatura atual ainda não oferece evidências suficientes para avaliar se os padrões identificados são generalizáveis a contextos como o brasileiro.

6 REFERÊNCIAS

- ABDALLAH, Mahmoud M. S. Using self-regulated learning supported by artificial intelligence (AI) chatbots to develop EFL student teachers' self-expression and reflective writing skills. **Academic Journal of Faculty of Education, Assiut University**, v. 40, n. 9, p. 1-50, set. 2024.
- ABDULJAWAD, Samah Abdulhadi. Investigating the impact of ChatGPT as an AI tool on ESL writing: prospects and challenges in Saudi Arabian higher education. **International Journal of Computer-Assisted Language Learning and Teaching**, v. 14, n. 1, p. 1-19, jan./dez. 2024. DOI: 10.4018/IJCALLT.367276

- BAI, Lifang; WEI, Yijia. Exploring EFL Learners' Integration and Perceptions of ChatGPT's Text Revisions: A Three-Stage Writing Task Study. *IEEE Transactions on Learning Technologies*, v. 17, p. 2215-2227, 2024. DOI: <https://doi.org/10.1109/TLT.2024.3491854>.
- BASKARA, Fidelis Rego. Exploring the implications of ChatGPT for language learning and teaching. **Englisia: Journal of Language, Education, and Humanities**, v. 10, n. 2, p. 179-193, 2023.
- ÇELİK, Ferdi; YANGIN ERSANLI, Ceylan; ARSLANBAY, Goshnag. Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? A one-shot intervention in Turkish EFL context. **International Review of Research in Open and Distributed Learning**, v. 25, n. 3, p. 287-303, ago. 2024.
- CHAPELLE, Carol A. **Computer applications in second language acquisition**. Cambridge: Cambridge University Press, 2001.
- CHAPELLE, Carol A. Open generative AI changes a lot, but not everything. **The Modern Language Journal**, v. 108, n. 2, p. 534-540. Disponível em: maio 2024. DOI: <https://doi.org/10.1111/modl.12927>.
Disponível em: <https://onlinelibrary.wiley.com/journal/15404781>.
- DILLON, Thomas. Korean University Students' Prompt Literacy Training with ChatGPT: Investigating Language Learning Strategies. **English Teaching**, v. 79, n. 3, p. 123-157, 2024. DOI: <https://doi.org/10.15858/engtea.79.3.202409.123>.
- ELEBYARY, Khaled; SHABARA, Ramy. CHATGPT-GENERATED CORRECTIVE FEEDBACK: DOES IT DO WHAT IT SAYS ON THE TIN?. **Teaching English with Technology**, v. 24, n. 3, p. 68-89, 2024. DOI: <https://doi.org/10.56297/vaca6841//BFFO7057/MYEH4562>.
- ELLIS, Rod. **The study of second language acquisition**. 2. ed. Oxford: Oxford University Press, 2008.
- ESCALANTE, Juan; PACK, Austin; BARRETT, Alex. AI-generated feedback on writing: insights into efficacy and ENL student preference. **International Journal of Educational Technology in Higher Education**, v. 20, n. 57, p. 1-21, 2023. DOI: <https://doi.org/10.1186/s41239-023-00425-2>.
- GEÇKİN, Vasfiye; KIZILTAŞ, Ebru; ÇINAR, Çağatay. Assessing second-language academic writing: AI vs. Human raters. **Journal of Educational Technology & Online Learning**, v. 6, n. 4, p. 1096-1108, 2023.

- GOZALI, Imelda et al. ChatGPT as an automated writing evaluation (AWE) tool: feedback literacy development and AWE tools' integration framework. **The JALT CALL Journal**, v. 20, n. 1, p. 1-22, 2024. DOI: <https://doi.org/10.29140/jaltcall.v20n1.1200>.
- GUO, Kai; WANG, Deliang. To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. **Education and Information Technologies**, v. 29, p. 8435-8463, 2024. DOI: <https://doi.org/10.1007/s10639-023-12146-0>.
- HAN, Zhaohong. ChatGPT in and for second language acquisition: a call for systematic research. **Language Learning & Technology**, v. 28, n. 1, p. 1-6, 2024.
- HONG, Wei Chun Henny. The impact of ChatGPT on foreign language teaching and learning: opportunities in education and research. **Journal of Educational Technology and Innovation**, v. 5, n. 1, p. 37-45, 2023.
- JACKSON, Gordon B. Methods for integrative reviews. **Review of Educational Research**, v. 50, n. 3, p. 438-460, 1980.
- KOHNKE, Lucas; MOORHOUSE, Benjamin L.; ZOU, Di. ChatGPT for language teaching and learning. **RELC Journal**, v. 54, n. 2, p. 537-550, 2023. DOI: 10.1177/00336882231162868
- KRASHEN, Stephen D. **Principles and practice in second language acquisition**. Oxford: Pergamon Press, 1982.
- LI, Mimi. Leveraging ChatGPT for Second Language Writing Feedback and Assessment. **International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)**, v. 14, n. 1, p. 1-11, jan. 2024. DOI: <https://doi.org/10.4018/IJCALLT.360382>. Disponível em: <https://www.igi-global.com/article/leveraging-chatgpt-for-second-language-writing-feedback-and-assessment/360382>.
- LO, Chung Kwan; HEW, Khe Foon; JONG, Morris Siu-yung. The influence of ChatGPT on student engagement: A systematic review and future research agenda. **Computers & Education**, v. 219, art. 105100, p. 1-18, jun. 2024. DOI: <https://doi.org/10.1016/j.compedu.2024.105100>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S036013152400109X>.
- MATIELO, Rafael. **Intralingual subtitles, interlingual subtitles, and L2 vocabulary learning: an exploratory study with Brazilian EFL students**. 2016. Tese (Doutorado em Inglês: Estudos Linguísticos e Literários) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2016. Disponível em: <https://repositorio.ufsc.br/handle/123456789/168145>.
- ORTEGA, Lourdes. **Understanding second language acquisition**. London: Routledge, 2013.
- ROST, Michael. **Teaching and researching listening**. 2. ed. London: Longman, 2011.

- RUDOLPH, Jürgen; TAN, Samson; TAN, Shannon. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? **Journal of Applied Learning & Teaching**, v. 6, n. 1, p. 342-363, jan. 2023. DOI: <https://doi.org/10.37074/jalt.2023.6.1.9>. Disponível em: <https://journals.sfu.ca/jalt/index.php/jalt/article/view/689>.
- SAITO, Yoshiko; GARZA, Thomas J.; HORWITZ, Elaine K. Foreign language reading anxiety. **The Modern Language Journal**, v. 83, n. 2, p. 202-218, 1999. DOI: 10.1111/0026-7902.00016
- SAPAN, Melek; UZUN, Levent. The effect of ChatGPT-integrated English teaching on high school EFL learners' writing skills and vocabulary development. **International Journal of Education in Mathematics, Science and Technology (IJEMST)**, v. 12, n. 6, p. 1657-1677, 2024. Disponível em: <https://doi.org/10.46328/ijemst.4655>.
- Sawangwan, Sirin. (2024). ChatGPT vs Teacher Roles in Developing EFL Writing. *International Journal of Computer-Assisted Language Learning and Teaching*, 14(1), 1–21. Disponível em: <https://doi.org/10.4018/IJCALLT.361235>
- SEO, Ji-Young. Exploring the Educational Potential of ChatGPT: AI-Assisted Narrative Writing for EFL College Students. **Language Teaching Research Quarterly**, v. 43, p. 1-21, 2024. Disponível em: <https://doi.org/10.32038/ltrq.2024.43.01>.
- SNYDER, Hannah. Literature review as a research methodology: an overview and guidelines. **Journal of Business Research**, v. 104, p. 333-339, 2019.
- SOLOVEY, Orit Zeevy. Comparing Peer, ChatGPT, and Teacher Corrective Feedback in EFL Writing: Students' Perceptions and Preferences. **Technology in Language Teaching & Learning**, v. 6, n. 3, p. 1-23, 2024. Disponível em: <https://doi.org/10.29140/tltl.v6n3.1482>.
- TORONTO, Cynthia E.; REMINGTON, Ruth. **A practical guide to integrative reviews**. Cham: Springer, 2020.
- TORRACO, Richard J. Writing integrative literature reviews: guidelines and examples. **Human Resource Development Review**, v. 4, n. 3, p. 356-367, 2005.
- TSAI, Chung-You; LIN, Yi-Ti; BROWN, Iain Kelsall. Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. **Education and Information Technologies**, v. 29, p. 22427-22445, 2024. Disponível em: <https://doi.org/10.1007/s10639-024-12722-y>
- ÜSTÜNBAŞ, Ümran. Hey, GPT, Can We Have a Chat?: A Case Study on EFL Learners' AI Speaking Practice. **International Journal of Modern Education Studies**, v. 8, n. 1, p. 91-107, jun. 2024. Disponível em: <https://doi.org/10.51383/ijonmes.2024.318>.
- WANG, Xizhe et al. ChatPRCS: a personalized support system for English reading comprehension based on ChatGPT. **IEEE Transactions on Learning Technologies**, v. 17, p. 1722-1736, 2024. DOI: 10.1109/TLT.2024.3405747

WARSCHAUER, Mark *et al.* The affordances and contradictions of AI-generated text for writers of english as a second or foreign language. **Journal of Second Language Writing**, v. 62, p. 101071, dez. 2023. Disponível em: <https://doi.org/10.1016/j.jslw.2023.101071>. Acesso em: 25 jun. 2026.

WERDININGSIH, Indah *et al.* Revolutionizing EFL writing: unveiling the strategic use of ChatGPT by Indonesian master's students. **Cogent Education**, v. 11, n. 1, 2399431, 2024. Disponível em: <https://doi.org/10.1080/2331186X.2024.2399431>.

WOO, David James; WANG, Deliang; GUO, Kai; SUSANTO, Hengky. Teaching EFL students to write with ChatGPT: students' motivation to learn, cognitive load, and satisfaction with the learning process. **Education and Information Technologies**, v. 29, p. 24963-24990, 2024. DOI: 10.1007/s10639-024-12819-4

YAN, Di. Impact of ChatGPT on learners in a L2 writing practicum: an exploratory investigation. **Education and Information Technologies**, v. 28, n. 11, p. 13943-13967, 2023. DOI: 10.1007/s10639-023-11742-4

YUAN, Ye; LI, Huan; SAWAENGDIST, Anan. The impact of ChatGPT on learners in English academic writing: opportunities and challenges in education. **CercleS**, [S. l.], v. 14, n. 1, p. 41-56, 2024. Disponível em: <https://doi.org/10.1515/cercles-2023-0006>.

ZOU, Min; HUANG, Liang. The impact of ChatGPT on L2 writing and expected responses: Voice from doctoral students. **Education and Information Technologies**, [S. l.], v. 29, p. 13201-13219, 2024. Disponível em: <https://doi.org/10.1007/s10639-023-12397-x>.