



O uso da tecnologia de síntese de fala baseada em Inteligência Artificial (IA) em dicionários pedagógicos digitais: contribuições para o aprendizado da pronúncia de línguas estrangeiras

The use of speech synthesis technology based on Artificial Intelligence (AI) in digital educational dictionaries: contributions to foreign language pronunciation learning

Juan Prete TOJEIRA-RAMOS*

Regiani Aparecida Santos ZACARIAS**

RESUMO: Este artigo científico investiga o uso da tecnologia de síntese de fala baseada em Inteligência Artificial (IA) em dicionários pedagógicos digitais, com ênfase em suas contribuições para o aprendizado da pronúncia de línguas estrangeiras. Adota-se uma abordagem teórica, fundamentada em revisão bibliográfica, para analisar benefícios como praticidade, personalização, eficiência, padronização acústica e controle de variáveis fonético-prosódicas. Discute-se também vantagens relacionadas à privacidade, à atualização contínua de conteúdos lexicográficos e à promoção da diversidade linguística em contextos educacionais digitais. Argumenta-se que a síntese de fala fornece modelos de pronúncia consistentes e padronizados, favorecendo a aprendizagem da produção oral e reduzindo problemas técnicos frequentes em gravações de fala natural, como ruídos, variações idiossincráticas e inconsistências articulatórias. Além disso, o estudo evidencia que a tecnologia contribui para a otimização da elaboração de dicionários pedagógicos digitais, ao reduzir custos de produção, eliminar a necessidade de gravações com falantes humanos e permitir maior escalabilidade na inclusão de novos verbetes e exemplos de uso. No campo educacional, destaca-se que o acesso a arquivos de fala sintética favorece o desenvolvimento de habilidades auditivas e articulatórias, pois possibilita ao aprendiz a repetição de modelos sonoros claros, estáveis e facilmente acessíveis. Em contextos bilíngues, ressalta-se ainda a possibilidade de representar diferentes variedades linguísticas de forma mais eficiente, ampliando o contato do estudante com diferentes padrões de pronúncia. Conclui-se que a integração de sistemas de síntese de fala em dicionários pedagógicos digitais pode elevar a qualidade do ensino de pronúncia em línguas estrangeiras, ao oferecer suporte tecnológico confiável, acessível e pedagogicamente consistente. Por fim, o artigo sugere a ampliação do uso dessa tecnologia em ambientes educacionais e incentiva a realização de pesquisas

* Doutorando pelo Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras (FCL), da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp). Araraquara, SP – Brasil. juan.tojeira@unesp.br

** Professora Livre-Docente junto ao Departamento de Letras Modernas da Universidade Estadual Paulista “Júlio de Mesquita Filho” (Unesp). Assis, SP – Brasil. regiani.zacarias@unesp.br

empíricas futuras que investiguem sua eficácia no ensino de línguas e na lexicografia pedagógica contemporânea, especialmente no contexto da formação de competências comunicativas em língua estrangeira, bem como na consolidação de práticas inovadoras mediadas por tecnologias digitais no ensino-aprendizagem de línguas.

PALAVRAS-CHAVE: Síntese de fala. Dicionários pedagógicos digitais. Inteligência Artificial. Aprendizado de pronúncia. Ensino de línguas estrangeiras.

ABSTRACT: This scientific article investigates the use of Artificial Intelligence (AI)-based speech synthesis technology in digital pedagogical dictionaries, with emphasis on its contributions to foreign language pronunciation learning. It adopts a theoretical approach grounded in a literature review to analyze benefits such as practicality, personalization, efficiency, acoustic standardization, and control of phonetic-prosodic variables. It also discusses advantages related to privacy, continuous updating of lexicographic content, and the promotion of linguistic diversity in digital educational contexts. It is argued that speech synthesis provides consistent and standardized pronunciation models, supporting oral production learning while reducing common technical issues found in natural speech recordings, such as noise, idiosyncratic variation, and articulatory inconsistencies. Furthermore, the study highlights that this technology contributes to optimizing the development of digital pedagogical dictionaries by reducing production costs, eliminating the need for human speaker recordings, and enabling greater scalability in the inclusion of new lexical entries and usage examples. In the educational field, it is emphasized that access to synthetic speech files fosters the development of listening and articulation skills, as it allows learners to repeatedly access clear, stable, and easily available sound models. In bilingual contexts, the possibility of representing different linguistic varieties more efficiently is also highlighted, expanding students' exposure to diverse pronunciation patterns. It is concluded that the integration of speech synthesis systems into digital pedagogical dictionaries can enhance the quality of pronunciation teaching in foreign languages by offering reliable, accessible, and pedagogically consistent technological support. Finally, the article suggests expanding the use of this technology in educational environments and encourages future empirical research to investigate its effectiveness in language teaching and contemporary pedagogical lexicography, particularly in the context of developing communicative competence in foreign languages as well as consolidating innovative practices mediated by digital technologies in language teaching and learning.

KEYWORDS: Speech synthesis. Digital pedagogical dictionaries. Artificial intelligence. Pronunciation learning. Foreign language teaching.

Artigo recebido em: 12.05.2026

Artigo aprovado em: 21.09.2026

1 Introdução

O setor de tecnologia digital tem experimentado avanços significativos no desenvolvimento de sistemas computacionais voltados a diversos âmbitos, como o

industrial, o mercadológico e o político, tanto no Brasil quanto no exterior. Essas inovações abrangem aplicações em telefonia, segurança, acessibilidade e monitoramento. No entanto, a influência dessas tecnologias na educação, sobretudo no ensino de línguas estrangeiras, ainda merece atenção, pois o uso de recursos computacionais e eletrônicos em contextos pedagógicos requer mais pesquisas (Yanhua, 2020; Liu, 2023; Tobing *et al.*, 2023; Yang, 2024; Chetveryk; Veretiuk, 2024; Law, 2024; Kolegova; Levina, 2024).

Entre os avanços tecnológicos mais notáveis, os sistemas de Inteligência Artificial (IA), originados da Ciência da Computação, têm desempenhado um papel crucial no progresso acadêmico e científico. De forma geral, a IA refere-se ao “desenvolvimento de sistemas de computador com a capacidade de raciocinar e executar funções cognitivas, como resolução de problemas, reconhecimento de objetos e palavras e tomada de decisões” (Eglinton; Tranter-Entwistle; Connor, 2023, p. 11, tradução nossa). Nos últimos anos, a IA expandiu sua presença a áreas como a educação linguística, impulsionada por colaborações entre instituições educacionais e corporativas.

Apesar de promissora na sociedade contemporânea, a aplicação de sistemas baseados em IA no ambiente pedagógico enfrenta desafios consideráveis, como os custos elevados, a necessidade de adaptação curricular e a garantia de acessibilidade tecnológica, em especial em situações marcadas por desigualdade educacional. Esses aspectos evidenciam a importância de pesquisas focadas na implementação de tecnologias digitais que beneficiem o ensino e a aprendizagem, como o estudo da pronúncia em línguas estrangeiras.

No campo da Lexicografia, que integra as Ciências do Léxico juntamente com a Lexicologia e a Terminologia (Biderman, 1984, 1998, 2001), observa-se uma crescente interface com tecnologias de IA. A Lexicografia, enquanto uma ciência dedicada à organização e produção de dicionários (Krieger, 2003, p. 71), é representada, neste estudo, por um de seus ramos particulares: a Lexicografia Pedagógica. Essa disciplina

investiga os aspectos educacionais dos dicionários escolares destinados ao ensino de línguas, bem como seu potencial didático e a qualidade dos materiais empregados (Krieger, 2011, p. 103). Dessa forma, a Lexicografia preocupa-se em propor dicionários que forneçam as informações das quais os consulentes necessitam, com prioridade às estratégias de consulta por eles utilizadas (Zacarias, 2018, p. 492).

Pesquisas recentes (Schryver, 2023; Lew, 2024; Nam; Na; Jung, 2024; Arias-Arias; Vázquez; Riveiro, 2024; Fuertes-Olivera, 2024) têm explorado a interface entre a Lexicografia (em sentido amplo) e a IA, com foco na produção de dicionários em contextos de transformação digital, por meio de técnicas de Processamento de Linguagem Natural (PLN). Contudo, ainda há lacunas quanto ao uso de sistemas de síntese de fala – tecnologia que converte texto em fala (Dutoit, 1997; Taylor, 2009) – em dicionários pedagógicos. Esses sistemas podem não apenas auxiliar na compreensão da pronúncia das palavras, mas também fomentar o desenvolvimento da competência oral de estudantes e consulentes.

Embora abordagens tradicionais de síntese de fala frequentemente dependam de dicionários de pronúncia, estudos recentes, como o de Liu, Ling e Chen (2023), têm buscado reduzir ou mesmo eliminar essa dependência por meio de modelos baseados em representações fonéticas aprendidas automaticamente¹. Esses avanços tecnológicos, no entanto, ainda não se refletem de forma consistente na incorporação de princípios da Lexicografia Pedagógica, sobretudo no que diz respeito ao ensino de pronúncia em contextos educacionais. Nesse cenário, observa-se que, no Brasil, a pesquisa acadêmica sobre a integração da síntese de fala em dicionários pedagógicos ainda é incipiente, o que destaca a necessidade de estudos mais aprofundados sobre o tema.

¹ Liu, Ling e Chen (2023) propõem um método de síntese de fala multilíngue que combina diferentes representações fonéticas aprendidas automaticamente e integra informações extraídas do sinal acústico e de modelos de reconhecimento, a fim de melhorar a inteligibilidade e a naturalidade do áudio.

Diante do exposto, o objetivo geral deste artigo científico é discutir, a partir de uma perspectiva teórica, o uso da tecnologia de síntese de fala baseada em IA em dicionários pedagógicos digitais. Os objetivos específicos, por sua vez, são: (1) analisar os benefícios dessa tecnologia para a elaboração de dicionários pedagógicos digitais; e (2) apresentar suas contribuições para o aprendizado da pronúncia de línguas estrangeiras. Já a pergunta que norteia a pesquisa é: como a tecnologia de síntese de fala baseada em IA pode ser utilizada em dicionários pedagógicos digitais, a fim de contribuir para o aprendizado da pronúncia de uma língua estrangeira? Nesse sentido, a hipótese que subjaz à pesquisa é a de que o uso de arquivos de fala sintética em dicionários pedagógicos digitais apresenta o potencial de contribuir, de maneira eficiente, para o aprendizado das características sonoras da língua estrangeira estudada pelo consulente, de modo a favorecer o desempenho de suas habilidades auditivas e articulatórias.

Do ponto de vista da metodologia, a pesquisa adota uma abordagem qualitativa, com base em uma revisão de material bibliográfico e análise de conceitos relativos à Lexicografia Pedagógica e à tecnologia de síntese de fala. Dada a escassez de trabalhos específicos sobre o uso dessa tecnologia em dicionários pedagógicos digitais, o estudo utiliza referências de áreas relacionadas, com o intuito de propor, sob uma perspectiva teórica, a aplicação de dados de fala sintética nessas obras. A proposta busca promover o aprendizado da pronúncia de línguas estrangeiras e, ao mesmo tempo, formular uma base teórica capaz de identificar as possibilidades e potenciais contribuições dessa tecnologia, a fim de alinhar as demandas curriculares aos avanços digitais.

A relevância do estudo reside em diversos aspectos. A seguir, estão listados os motivos para a realização do trabalho.

Em primeiro lugar, a pesquisa dialoga com os interesses de linguistas brasileiros, como Tojeira-Ramos e Massini-Cagliari (2024) e Tojeira-Ramos (2025), no que diz respeito à exploração do componente sonoro da tecnologia de síntese de fala,

a fim de favorecer seu uso no ensino de línguas. Ademais, o estudo promove um diálogo entre a Lexicografia Pedagógica, a Ciência da Computação e a Linguística Aplicada, para a integração de abordagens pedagógicas inovadoras e assistidas por tecnologias digitais.

A implementação da síntese de fala em dicionários pedagógicos digitais é promissora na democratização do acesso ao conhecimento linguístico, na redução de desigualdades educacionais e na promoção da inclusão, especialmente para alunos com dificuldades de leitura ou deficiência visual. Além disso, pela autonomia que propicia no aprendizado da pronúncia, essa tecnologia oferece suporte a professores e estudantes na exploração de aspectos sonoros de línguas estrangeiras.

Por fim, em uma época de crescente digitalização, é essencial investigar como ferramentas tecnológicas podem ser aplicadas a materiais educacionais, como dicionários pedagógicos. A aplicação da síntese de fala em tais recursos possibilita o aprimoramento do ensino de aspectos sonoros, o estímulo a práticas de aprendizado autônomo e a dinamização da educação linguística.

A discussão proposta é respaldada por estudos lexicográficos pedagógicos, como os de Welker (2008) e Krieger (2012), que estabelecem referenciais importantes para o campo no Brasil. A integração da síntese de fala por IA objetiva expandir o debate acerca das possibilidades tecnológicas aplicadas ao ensino, de modo a complementar as abordagens fonético-fonológicas tradicionais, que subsidiam, há décadas, a aprendizagem de línguas estrangeiras.

O artigo é composto de quatro seções principais. Após a “Introdução”, a seção “Fundamentação teórica” explora os princípios da Lexicografia Pedagógica, além de abordar o funcionamento e as possibilidades de aplicação da tecnologia de síntese de fala. Em seguida, a seção “Resultados e Discussões” analisa, em nível teórico, os benefícios da utilização dessa tecnologia em dicionários pedagógicos digitais, com atenção especial ao aprendizado da pronúncia de línguas estrangeiras. Por fim, a seção

“Considerações finais” revisa os principais aspectos abordados no estudo e apresenta reflexões sobre a temática.

2 Fundamentação teórica

A seção “Fundamentação teórica” é dividida em duas partes. A primeira subseção define e descreve os princípios da Lexicografia Pedagógica, enquanto a segunda subseção trata do funcionamento e das aplicações da tecnologia de síntese da fala.

2.1 Lexicografia Pedagógica

Conforme descrito na seção anterior, o principal interesse da Lexicografia Pedagógica é a pesquisa sobre as diversas facetas que compõem e envolvem os dicionários escolares relacionados ao ensino de línguas (materna e estrangeira). A referida disciplina é motivada pelo potencial didático dessas obras e se preocupa com a adequação e a qualidade dos materiais utilizados no contexto educacional em questão (Krieger, 2011, p. 103). Conhecidos como dicionários pedagógicos, esses instrumentos, que podem ser produzidos em formato digital ou impresso, levam em conta as habilidades, as dificuldades e as necessidades de consulta dos aprendizes de línguas (Krieger, 2011, p. 105). Assim, pode-se dizer que a Lexicografia Pedagógica tem duas vertentes: a Lexicografia Pedagógica Prática dedica-se ao desenvolvimento de dicionários pedagógicos, enquanto a Lexicografia Pedagógica Teórica (ou Metalexigrafia Pedagógica) concentra-se no estudo desses dicionários (Welker, 2011, p. 104). Embora não seja uma disciplina antiga, a Lexicografia Pedagógica ganha destaque no nível acadêmico pela importância dos dicionários no aprendizado de línguas (Krieger, 2011, p. 103).

Os dicionários pedagógicos digitais (ou eletrônicos) e impressos diferem em certas características, que são apresentadas nas citações a seguir.

O dicionário eletrônico possui, em relação ao dicionário convencional, impresso em papel, algumas diferenças básicas. A principal delas é que o dicionário eletrônico é constituído de bits, minúsculos pulsos de luz praticamente sem características físicas palpáveis, facilmente transmitidas de um computador a outro por linhas telefônicas, ondas de rádio ou qualquer suporte magnético. Por ser um arquivo digital, o dicionário eletrônico é extremamente maleável: pode ser facilmente compactado, ampliado e atualizado, sem grandes custos de produção. Além de textos e imagens pode incluir também animação, som e vídeo. Tem finalmente a característica da invisibilidade, só aparecendo ao usuário quando solicitado e mesmo assim mostrando apenas o verbete ou o dado solicitado, ocultando todo o resto dentro do computador ou no suporte que o sustenta. É impossível perceber um dicionário eletrônico em toda sua extensão (Leffa, 2006, p. 323).

O dicionário convencional, impresso em papel, é o oposto de todas essas características. Não é constituído de bits, mas de átomos (para ver a diferença entre bits e átomos, veja Negroponte, 1995). Por ser átomo, tem peso e dimensões físicas facilmente percebidas. O papel em que é impresso não pode ser fisicamente compactado e nem teletransportado de um lugar a outro. Qualquer atualização que precisar ser feita implica uma nova impressão de todo o texto, com altos custos de produção. Não oferece a possibilidade de incluir animação, som ou vídeo. É visível em sua totalidade; mesmo que o leitor esteja interessado em apenas uma palavra, tem que manusear o volume inteiro (Leffa, 2006, p. 323-324).

Além da noção de dicionário pedagógico, um conceito importante na Lexicografia Pedagógica é o léxico. Abaixo, são apresentadas suas especificidades.

O Léxico de qualquer língua constitui um vasto universo de limites imprecisos e indefinidos. Abrange todo o universo conceptual dessa língua. Qualquer sistema léxico é a somatória de toda experiência acumulada de uma sociedade e do acervo de sua cultura através das idades. Os membros dessa sociedade funcionam como sujeitos-agentes, no processo de perpetuação e re-elaboração contínua do Léxico de sua língua. Nesse processo em desenvolvimento, o Léxico, se expande, se altera, e, às vezes, se contrai. As mudanças sociais e culturais acarretam alterações nos usos vocabulares; daí resulta que unidades ou setores completos do Léxico podem ser marginalizados, entrar em desuso e vir a desaparecer. Inversamente, porém, podem ser ressuscitados termos que voltam à circulação, geralmente com novas conotações. Enfim, novos vocábulos, ou novas significações de

vocábulos já existentes, surgem para enriquecer o Léxico (Biderman, 2001, p. 178).

Quanto aos princípios da Lexicografia Pedagógica, Krieger (2011, p. 106) afirma que eles estão intrinsecamente ligados aos seus objetivos, como a adequação dos dicionários ao público-alvo e o uso produtivo dessas obras em diferentes projetos de ensino e aprendizagem. Segundo a autora, é fundamental compreender que o dicionário consiste em um texto organizado por regras particulares, capaz de sistematizar informações linguísticas, culturais e pragmáticas. Esse aspecto ressalta o potencial pedagógico dos dicionários e reforça o princípio de que, assim como existem livros didáticos apropriados para diferentes níveis de ensino, os dicionários também devem ser selecionados de acordo com as necessidades dos alunos. Em outras palavras, ao preparar dicionários pedagógicos, é essencial considerar as reais necessidades e habilidades dos usuários, o que implica produzir dicionários diferentes para alunos com variados níveis de competência linguística (Welker, 2011, p. 109).

Ao discutir os princípios da Lexicografia Pedagógica, Welker (2011, p. 106) ressalta que a maioria dos artigos científicos sobre o assunto tem como foco os dicionários pedagógicos para o ensino de línguas estrangeiras. Krieger (2011, p. 106), por sua vez, enfatiza a diversidade de tipos de dicionários, como os monolíngues, bilíngues e terminológicos, entre outras opções, que compartilham características e funções diferentes.

Com relação ao aspecto didático dos dicionários pedagógicos, Krieger (2011) e Welker (2011) fazem contribuições importantes. Para Krieger (2011, p. 110), os dicionários, independentemente de seu tipo, são didáticos, porque contêm informações lexicais, linguísticas e culturais que auxiliam os alunos na leitura, na escrita e na expressão. Esses materiais oferecem informações sistematizadas sobre as propriedades pragmáticas, semânticas, gramaticais e históricas das palavras, e contribuem para a alfabetização e a competência leitora dos alunos de língua materna, além de auxiliar na pesquisa descritiva de outras línguas (Krieger, 2011, p. 110). Por

outro lado, Welker (2011, p. 112) ressalta que os dicionários convencionais geralmente têm a função de informar os leitores, mas de forma não didática, ao passo que os dicionários pedagógicos são desenvolvidos especificamente para atender às necessidades do contexto educacional.

Embora Krieger (2011, p. 110) enfatize a natureza didática de um dicionário, a Lexicografia Pedagógica também se preocupa com as particularidades estruturais dos dicionários, o perfil do aprendiz e os componentes básicos de uma produção lexicográfica pedagógica. Abaixo, são descritos resumidamente os principais componentes listados por Krieger (2011).

No que diz respeito à seleção de verbetes, ela é obrigatória nos dicionários destinados ao aprendizado escolar. No entanto, de acordo com Krieger (2011, p. 110-111), ainda não existe, no Brasil, um princípio explícito que garanta a adequação dos dicionários ao seu público-alvo. Estudos lexicométricos ou de *corpus* que garantam registros lexicais mínimos de uso para alunos com diferentes níveis de conhecimento cultural ainda são raros na literatura nacional (Krieger, 2011, p. 111).

Krieger (2011, p. 111-112) também faz observações sobre o nível de linguagem e a forma gráfica dos dicionários. Quanto ao nível de linguagem, aplica-se o princípio da adequação, com simplificações que não comprometam a compreensão dos itens lexicais. Em relação à forma gráfica, a autora enfatiza que os dicionários devem valorizar recursos gráficos, como divisões silábicas, acentuação, desenhos e figuras que auxiliem a compreensão e delimitem as informações dos verbetes, principalmente no caso de palavras polissêmicas.

Tradicionalmente, o ensino da pronúncia nos dicionários é auxiliado por transcrições fonéticas, como as do Alfabeto Fonético Internacional (IPA, do inglês *International Phonetic Alphabet*), um instrumento consolidado para a representação da materialidade sonora (Tench, 1992). A tecnologia de síntese de fala baseada em IA oferece uma interface complementar que possibilita a audição instantânea e contextualizada dos verbetes. Ao integrar a dimensão auditiva à aprendizagem, esse

recurso expande as possibilidades de consulta lexicográfica discutidas em obras de referência, como as de Krieger (2012) e Welker (2008).

A próxima subseção trata dos principais conceitos e aplicações da tecnologia de síntese de fala.

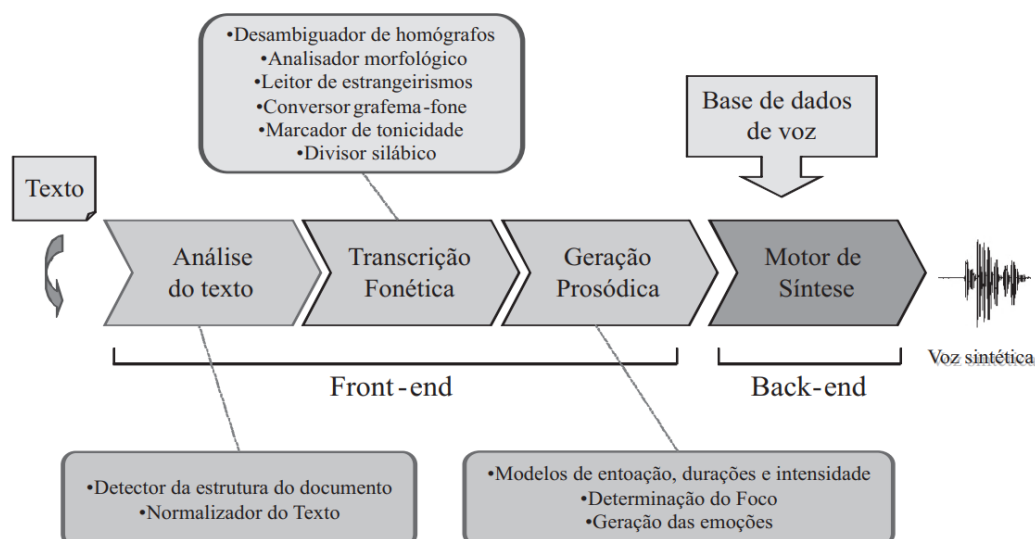
2.2 Tecnologia de síntese de fala

Entende-se por síntese de fala (TTS – Text-to-Speech), de um ponto de vista amplo, a “produção de fala por máquinas, por meio da fonetização automática das frases a serem pronunciadas” (Dutoit, 1997, p. 13, grifos do autor). Esse tipo de tecnologia de fala pode dar suporte a diferentes setores da sociedade, como o setor digital, em que o diálogo entre humanos e máquinas é realizado por meio de assistentes virtuais ou centrais de atendimento; o setor de acessibilidade e inclusão, em que pessoas com deficiência visual ou dificuldades de leitura podem contar com a ajuda de recursos para transformar texto escrito em fala audível; e o setor corporativo, em que as empresas costumam usar sistemas de síntese de fala para facilitar a comunicação entre clientes e funcionários.

Com o constante progresso da IA, a tecnologia de síntese de fala tornou-se mais aprimorada, a ponto de começar a se parecer com as características acústicas (físicas/ondulatórias) da voz humana. No entanto, ainda há desafios relacionados à expressividade da fala sintética, que exigem investimento contínuo em projetos de pesquisa compostos de profissionais de diferentes áreas do conhecimento, como Linguística, Ciência da Computação e Engenharia Eletrônica.

Apesar das possíveis diferenças de arquitetura entre os sistemas contemporâneos de síntese de fala, Braga (2007, p. 4) afirma que há, pelo menos, três blocos comuns a todos eles, conforme ilustrado na Figura 1: o pré-processamento de texto (*front-end*), o motor de síntese (*back-end*) e a base de dados de voz (*voice font*).

Figura 1 – Esquema da arquitetura de um sistema de conversão Texto-Fala.



Fonte: Braga (2007, p. 6).

O pré-processamento de texto (*front-end*), o motor de síntese (*back-end*) e a base de dados de voz (*voice font*) consideram informações linguísticas e computacionais para que a produção da fala sintética por máquinas ocorra de maneira satisfatória. As particularidades desses três blocos elementares de um sistema de síntese de fala são descritas a seguir.

A base de dados de voz é rigorosamente seleccionada, gravada e etiquetada, sendo as suas unidades seleccionadas por algoritmos de processamento de sinal que as concatenam e transformam em voz sintética. O *front-end* é composto por três componentes: 1) a análise de texto, do qual fazem parte o detector da estrutura do documento (responsável por identificar o tipo de texto, informação que pode ser útil para a geração dos modelos prosódicos) e o normalizador de texto (em que se incluem os sub-módulos de interpretação de pontuação, expansão de abreviaturas, leitura de siglas e de acrónimos e conversão de numerais, datas, números romanos e árabes, quantias em dinheiro, números ordinais); 2) a análise fonética, da qual fazem parte os módulos de leitura de estrangeirismos, de desambiguação de homógrafos heterófonos, de análise morfossintáctica e de conversão grafema-fone; 3) análise e geração prosódica, módulo que aproveita as informações anteriormente obtidas a partir do texto, como tipo de frase, divisão silábica, marcação de tónica, classificação morfossintáctica, para produzir modelos prosódicos. O *back-end* é composto pelo motor de síntese que interpreta a transcrição fonética

gerada pelo *front-end* e a transforma em fala sintética. Os critérios que subjazem à construção e etiquetagem de uma base de dados estão dependentes da técnica de síntese que se pretende utilizar e da qualidade que se pretende atingir (para um estudo comparativo das várias técnicas de síntese, veja-se Barros, 2002; Huang et al., 2001). Devem seleccionar-se textos que sejam suficientemente representativos da diversidade e riqueza fonética, sintáctica e prosódica da língua. Deve também ser tida em consideração a qualidade da gravação, que deve ser realizada com alta qualidade e em salas insonorizadas, com vista a obter-se a maior qualidade possível na síntese. A escolha do locutor é também fundamental, pois deve ter boa articulação e timbre de voz agradável. Após a recolha dos sons da futura base de dados, é necessário proceder-se à etiquetagem do sinal de fala com vista a servir de plataforma de trabalho para a programação de regras no sistema de conversão texto-fala, segundo critérios adequados à técnica usada (Braga, 2007, p. 4-5, grifos da autora).

Como se vê, os aspectos sonoros (segmentais e prosódicos) da língua em questão, bem como outros elementos gramaticais (como tópicos relacionados à morfologia, à sintaxe e à semântica), fazem parte da estruturação de um sistema de síntese de fala, de modo que o resultado acústico gerado pela máquina se assemelhe à voz natural. Na próxima seção, os benefícios da aplicação da tecnologia de síntese de fala aos dicionários pedagógicos digitais são discutidos a partir de uma perspectiva teórica, especialmente no que diz respeito ao aprendizado da pronúncia de uma língua estrangeira.

3 Resultados e discussões

Em concordância com Marques, Polizer e Zacarias (2020), que, por sua vez, baseiam-se em Nadin e Vargas (2016), os dicionários pedagógicos digitais, disponíveis *online* em *sites* ou aplicativos, são obras lexicográficas capazes de garantir resultados mais vantajosos para o consulente, em comparação com os dicionários pedagógicos impressos. Eles proporcionam aos alunos benefícios como a busca instantânea de palavras necessárias, o livre acesso ao material e a facilidade de pesquisar termos desconhecidos. Segundo os autores, esses dicionários estabelecem um diálogo com a

Base Nacional Comum Curricular (BNCC) e o Currículo Paulista, cujas diretrizes recomendam o uso de tecnologias digitais para o desenvolvimento de recursos pedagógicos contemporâneos. Assim como nos materiais impressos, os pesquisadores destacam que os dicionários pedagógicos digitais consideram as necessidades e prioridades linguísticas ao elaborar os verbetes, para que o aprendizado seja eficaz e os alunos não tenham dúvidas ou façam escolhas equivocadas na produção de textos.

Ao contrário das obras lexicográficas impressas, os recursos digitais possibilitam a integração de multimodalidades, como animações, vídeos e áudios (Leffa, 2006, p. 323). Entre esses recursos, os arquivos sonoros desempenham um papel crucial na promoção do aprendizado da pronúncia em línguas estrangeiras. Quando utilizados de maneira ética e estratégica, os recursos de síntese de fala ampliam a compreensão sonora dos alunos, além de favorecer o desenvolvimento das competências orais e auditivas dos consulentes. Ao ouvir a pronúncia de palavras em contextos discursivos ampliados, como frases ou enunciados completos, os estudantes têm a oportunidade de participar ativamente do processo de ensino e aprendizagem, o que lhes permite abranger as dimensões articulatória (produção sonora), acústica (transmissão sonora) e auditiva (percepção sonora). Esse processo possibilita o entendimento das características sonoras da língua e amplia a competência oral.

A gravação da voz de falantes humanos para cada verbete de um dicionário exige uma carga de trabalho considerável, pois são necessários instrumentos de gravação de áudio, cabines à prova de som para garantir um resultado acústico livre de ruídos, ferramentas de processamento de áudio por computador e protocolos burocráticos. Nesse contexto, a adoção de sistemas tecnológicos de síntese de fala baseados em IA pode ser uma opção eficaz, não apenas para promover um trabalho interativo, econômico e flexível, mas também para acelerar o processo de produção em larga escala, sem comprometer a qualidade do projeto (Dutoit, 1997; Morise; Yokomori; Ozawa, 2016; Han *et al.*, 2020; Huang *et al.*, 2020; Hu *et al.*, 2022; Yang *et al.*,

2023; Meyer *et al.*, 2023). Nas próximas subseções, são discutidos os benefícios da implementação da tecnologia de síntese de fala em dicionários pedagógicos digitais.

3.1 Logística, personalização e eficiência

Um benefício relevante da adoção de sistemas de síntese de fala em dicionários pedagógicos digitais diz respeito à logística, à personalização e à eficiência. Para atender às demandas educacionais de usuários no processo de ensino e aprendizagem de línguas, é necessário compilar um vasto repertório de palavras e contextualizá-las em termos discursivos e enunciativos, de modo que os alunos compreendam seu significado e saibam utilizá-las adequadamente na produção textual.

A utilização de arquivos de fala sintética, gerados automaticamente por sistemas computacionais acessíveis e frequentemente gratuitos, reduz os desafios associados à produção de gravações humanas. O lexicógrafo não precisa contratar locutores para gravar palavras isoladas ou inseridas em contextos mais amplos, ou se preocupar com a organização de ambientes de gravação controlados e isentos de ruídos. Ademais, elimina-se a necessidade de adquirir equipamentos técnicos especializados, como microfones e dispositivos de processamento de áudio de alta qualidade, que podem representar um alto investimento financeiro.

Outro aspecto relevante é o controle das variáveis associadas à fala natural, como idiossincrasias e variações de entoação, de intensidade e de taxa de elocução, que podem impactar a uniformidade do material acústico. A fala sintética permite mais padronização, o que facilita a gestão dessas variáveis por meio de técnicas linguístico-computacionais. Embora o resultado acústico de alguns sistemas tecnológicos ainda não alcance plenamente a expressividade da fala humana, avanços significativos têm sido feitos nesse campo. Empresas como Google, Microsoft e Amazon têm aprimorado suas tecnologias para oferecer reproduções vocais mais fiéis à fala natural, com articulação clara dos segmentos sonoros e precisão na posição do

acentos tônicos – parâmetros essenciais para o aprendizado das características sonoras de uma língua.

Dessa forma, a fala sintética não apenas oferece uma solução mais eficiente e econômica, mas também possibilita a produção em larga escala de dicionários pedagógicos digitais. Tal fato permite a disponibilização do material aos usuários sem o comprometimento de sua qualidade.

3.2 Privacidade

Um segundo benefício é a privacidade. Em um dicionário pedagógico digital, cujo objetivo é contribuir, dentre outros fatores, para o aprendizado da pronúncia de palavras de uma língua, é necessário que os alunos ouçam áudios elucidativos sobre os aspectos sonoros da língua em questão. Se forem utilizados áudios de fala natural, no entanto, surgem preocupações éticas relacionadas à privacidade das vozes dos participantes. Embora os participantes assinem um Termo de Consentimento Livre e Esclarecido (TCLE) e tenham acesso às condições éticas e legais do trabalho, suas vozes, ao serem inseridas em um ambiente eletrônico, ficam expostas a práticas ilegais e criminosas por parte de usuários mal-intencionados, que poderiam acessar essas gravações e realizar ações que infrinjam os valores morais da sociedade.

Com o uso de arquivos de fala sintética, essas preocupações são amenizadas, pois a voz à qual o consulente tem acesso é produzida por um dispositivo eletrônico, e não por um ser humano. Contudo, o uso de uma voz sintética não elimina totalmente o risco de uso ilegal desses dados, o que poderia prejudicar o trabalho da empresa responsável pelos serviços. Assim, as vozes sintéticas devem ser empregadas de maneira responsável e ética, para que o consulente tenha a oportunidade de compreender os aspectos sonoros da língua estudada.

3.3 Atualização

Um terceiro benefício do uso de dados de fala sintética é a possibilidade de atualização dos arquivos de som de um dicionário pedagógico digital. Quando o trabalho é atualizado na plataforma eletrônica com a inclusão de outras palavras, um sistema de síntese de fala pode ser utilizado para gerar áudios que exemplifiquem a pronúncia dos novos verbetes do dicionário, sem a necessidade de realizar outra sessão de gravação com falantes humanos. Esse método elimina a necessidade de custos e o trabalho logístico associado à seleção de falantes, caso eles tivessem sido previamente contratados para o projeto lexicográfico.

Além disso, se o resultado acústico gerado pelo sistema de síntese de fala apresentar inconsistências que prejudiquem o aprendizado da pronúncia, é possível ajustar os parâmetros linguístico-computacionais, como a variação entoacional e a velocidade de reprodução, para a produção de novos áudios. Por conseguinte, não seria necessária a regravação dos áudios nem o uso de recursos de edição de som, o que demandaria mais tempo. Quando um sistema de síntese de fala é aprimorado para fornecer resultados mais realistas, novos áudios podem ser facilmente gerados e incorporados ao dicionário pedagógico digital, o que beneficia o aprendizado da pronúncia e as habilidades auditivas dos estudantes.

3.4 Diversidade linguística

Um quarto benefício está relacionado à diversidade linguística. Em um dicionário pedagógico bilíngue, duas línguas são tratadas, o que exige a inclusão de arquivos de som para cada uma delas, especialmente se o objetivo do lexicógrafo for estabelecer uma comparação entre as línguas, como na articulação de certas vogais, nas características sonoras de consoantes específicas e na posição do acento tônico. Ao utilizar a tecnologia de fala sintética, o lexicógrafo pode gerar áudios de ambas as línguas em um curto período, sem a necessidade de gravar grupos adicionais de falantes proficientes para cada língua.

Nos sistemas tecnológicos contemporâneos, as variedades dialetais das línguas são levadas em consideração, o que permite que o lexicógrafo as inclua no dicionário pedagógico digital ou as sugira como conteúdo suplementar, a fim de enriquecer o processo de ensino e aprendizagem da pronúncia. No entanto, ao trabalhar com dados de fala natural, a questão da diversidade linguística pode ser negligenciada. O registro de uma quantidade suficiente de dados para ilustrar as diferenças dentro e entre as línguas exigiria um tempo considerável, o que pode ser um desafio para o profissional e sua equipe, sobretudo por causa da necessidade de produzir e publicar a obra lexicográfica em um período limitado para atender às demandas de um público específico.

A aplicação da tecnologia de síntese de fala requer a consideração das variedades sociolinguísticas. Para tanto, é imprescindível que o sistema reproduza usos linguísticos compatíveis com as situações de interação, em consonância com um conjunto de discussões da Sociolinguística e da Lexicografia Pedagógica, conforme os estudos de Bortoni-Ricardo (2004), Bagno (2007), Welker (2008, 2011), Krieger (2011, 2012) e Freitag (2024, 2025).

3.5 Qualidade

Um quinto benefício refere-se à qualidade do material disponível para os consulentes. Ao lidar com um grupo de pessoas, é provável que haja variações vocais entre elas, caso sejam consideradas variáveis sociolinguísticas, como sexo, idade, local de origem, entre outras. Ademais, durante as gravações, as pessoas podem estar sujeitas a condições ambientais que podem influenciar o resultado acústico do áudio. Por exemplo, mesmo que esse fator seja controlado antes do início das sessões de

gravação, se o locutor apresentar sinais de resfriado, é possível que o áudio disponha de alterações acústicas em função da condição clínica do falante².

Com os dados de fala sintética, o problema mencionado acima não ocorre. Essa situação se deve ao fato de que o lexicógrafo pode escolher a voz sintética desejada, e essa voz será responsável por gerar o áudio de todas as palavras e frases que farão parte do dicionário pedagógico digital, o que garante uniformidade e padronização. Ademais, não há erros humanos decorrentes de lapsos de memória, de cansaço, de desatenção ou de hesitações, como pausas imprevistas pela gramática da língua. Nos dados gerados por sistemas computacionais sofisticados, esses problemas são evitados, pois eles são programados para produzir áudios sem falhas técnicas que possam comprometer a compreensão do usuário, seja em termos de informação ou de pronúncia. Se o lexicógrafo quiser ajustar algum parâmetro relacionado, por exemplo, à entoação, à intensidade ou à velocidade de reprodução, poderá fazê-lo facilmente com as ferramentas adequadas.

3.6 Aprendizado de pronúncia

Os cinco critérios abordados nas subseções precedentes têm implicações substanciais para o aprendizado da pronúncia de uma língua estrangeira. Com foco na questão de saber se a fala sintética pode, de fato, contribuir para esse aprendizado, seguem algumas considerações.

Em primeiro lugar, uma tecnologia de fala sintética oferece aos usuários uma pronúncia “padrão”, de acordo com a tipologia linguística, não no sentido de ser a única ou a mais correta, mas a frequentemente utilizada. Esse recurso ajuda o usuário

² Embora as gravações humanas profissionais constituam um padrão de referência na Lexicografia, a escalabilidade e a necessidade de atualização constante dos dicionários digitais impõem desafios logísticos. Nesse contexto, a síntese de fala surge como um recurso de otimização. Quando devidamente parametrizada com base em princípios linguísticos e computacionais, ela pode assegurar uma consistência acústica voltada à aprendizagem do consulente. Essa solução também tem o objetivo de reduzir a variabilidade técnica presente nas gravações de fala humana, sem desconsiderar a legitimidade da variação linguística como um fenômeno constitutivo das línguas naturais (Labov, 1972; Tarallo, 1985).

a entender como articular e produzir os elementos sonoros das palavras em estudo, sem oferecer uma vasta gama de sutilezas ou possibilidades de pronúncia para uma língua específica, conforme diferentes variáveis sociolinguísticas.

Em segundo lugar, ao proporcionar uniformidade na pronúncia, os dados de fala sintética permitem que os alunos compreendam as características sonoras básicas da língua em questão, o que evita confusões com outras possibilidades de produção oral ou com influências de ruídos, interferências ou falhas técnicas que podem ocorrer em cabines de gravação de fala natural. Quanto à diversidade dialetal, é importante considerar que existem diferenças na produção de estruturas linguísticas específicas entre o português brasileiro e o português europeu, por exemplo, assim como entre o inglês americano e o inglês europeu. Um sistema de fala sintética pode auxiliar o lexicógrafo a incluir as pronúncias padrão dessas variedades, com vistas a potencializar o aprendizado linguístico.

Em terceiro lugar, nos sistemas de síntese de fala, alguns processos sonoros típicos da comunicação espontânea tendem a não ocorrer, a fim de possibilitar que o aluno aprenda a pronúncia das palavras estudadas, tanto isoladamente quanto em um contexto discursivo³. Todavia, é fundamental que o aluno tenha acesso à variação sonora da língua desde as etapas iniciais de aprendizagem por meio de recursos didáticos complementares. Tal conhecimento permitirá que, no futuro, ao se aprofundar na língua, o estudante compreenda as razões pelas quais esses fenômenos se manifestam na interação verbal. Ademais, por se tratar de um dicionário pedagógico, é crucial que a pronúncia das palavras seja clara, para garantir um aprendizado linguístico satisfatório.

Por fim, é importante observar que, nos dicionários pedagógicos, quando o aspecto da pronúncia não é negligenciado, há uma transcrição de como a palavra

³ Tais processos podem abranger, por exemplo, a omissão, a inserção e a coarticulação de segmentos sonoros. Embora ocorram em situações de fala espontânea entre seres humanos, esses fenômenos ainda não costumam ser observados, com frequência, em áudios sintéticos.

consultada costuma ser pronunciada. Além de ter acesso à representação da estrutura sonora da língua estrangeira estudada por meio da transcrição, o aluno pode ouvir como esse som é expresso ao acessar o arquivo de fala sintética. Os estudantes podem continuar a ouvir esses áudios sempre que desejarem, seja na escola ou em outros ambientes, o que lhes permite aprender de forma autêntica. Em outras palavras, ao ter acesso a um arquivo de fala sintética, o aprendiz é capaz de entender, de forma precisa e autônoma, como a palavra é pronunciada e de usá-la adequadamente nos contextos de interlocução oral necessários.

4 Considerações finais

O estudo teve como propósito discutir, a partir de uma perspectiva teórica, o uso da tecnologia de síntese de fala baseada em IA em dicionários pedagógicos digitais, com foco nas contribuições dessa abordagem para o aprendizado da pronúncia de línguas estrangeiras. A pesquisa buscou responder à seguinte pergunta norteadora: como a tecnologia de síntese de fala baseada em IA pode ser utilizada em dicionários pedagógicos digitais, a fim de contribuir para o aprendizado da pronúncia de uma língua estrangeira?

Em relação ao primeiro objetivo específico, o estudo identificou benefícios significativos decorrentes do uso dessa tecnologia. O uso de arquivos de fala sintética proporciona praticidade, personalização e eficiência, ao otimizar o trabalho dos lexicógrafos, reduzir custos e o tempo de produção e oferecer controle sobre variáveis acústicas. A padronização resultante torna a criação de dicionários pedagógicos digitais menos complexa e proporciona uma experiência de aprendizado mais consistente aos usuários.

Além disso, o uso da fala sintética traz vantagens no que diz respeito à privacidade, pois minimiza as preocupações éticas associadas ao uso de vozes humanas. Contudo, o trabalho destaca a importância de práticas responsáveis para evitar o uso indevido dessa tecnologia. Essa abordagem também se alinha à inovação

tecnológica e permite a rápida e precisa atualização dos dicionários, sem a necessidade de gravações humanas adicionais. Tal flexibilidade acelera a inclusão de novos verbetes e possibilita a correção de inconsistências que possam afetar negativamente o aprendizado da pronúncia.

A pesquisa também destacou o potencial dessa tecnologia para promover a diversidade linguístico-cultural, especialmente em dicionários bilíngues que incluem variedades dialetais, como o português brasileiro e o português europeu, ou o inglês americano e o inglês europeu. Embora a gravação e a exploração das variações dialetais demandem tempo, a síntese de fala possibilita enriquecer o ensino e a aprendizagem por meio de diversas representações linguísticas. A padronização do áudio elimina variações indesejadas, como inconsistências vocais ou erros humanos, e garante um resultado mais preciso e confiável.

Quanto ao segundo objetivo específico, o estudo demonstrou que a fala sintética é uma ferramenta valiosa para o aprendizado linguístico. Ela oferece aos usuários uma pronúncia padrão baseada na variedade dialetal mais comum, ajuda-os a produzir sons corretamente e permite que compreendam as características básicas da língua, sem interferências idiossincráticas ou desafios técnicos complexos.

Diante do exposto, o artigo respondeu à pergunta norteadora ao demonstrar que a tecnologia de síntese de fala baseada em IA contribui para o aprendizado da pronúncia por meio da oferta de uma voz padronizada, uniformidade sonora e suporte à diversidade dialetal. Esses elementos ajudam os alunos a desenvolver suas habilidades auditivas e articulatórias de maneira mais eficiente e confiável, o que comprova a hipótese do trabalho.

Conclui-se, portanto, que o artigo expôs as principais contribuições teóricas da implementação da fala sintética em dicionários pedagógicos digitais, a fim de beneficiar o ensino e a aprendizagem da pronúncia de línguas estrangeiras. Além disso, o estudo incentiva outros pesquisadores e educadores a explorarem o uso de tecnologias de síntese de fala em contextos educacionais, com a intenção de

desmistificar seu uso em sala de aula e promover investigações empíricas que validem e aprofundem as discussões teóricas e práticas sobre essa abordagem inovadora, com ênfase no contexto lexicográfico pedagógico.

Referências

ARIAS-ARIAS, I.; DOMÍNGUEZ VÁZQUEZ, M. J.; VALCÁRCEL RIVEIRO, C. Efficiency and intelligence in lexicography and artificial intelligence: can ChatGPT recreate the lexicographical text type? *Lexikos*, v. 34, n. 1, p. 51-76, 2024. DOI <https://doi.org/10.5788/34-1-1879>

BAGNO, M. **Nada na língua é por acaso**: por uma pedagogia da variação linguística. São Paulo: Parábola, 2007.

BARROS, M. J. A. de S. **Estudo comparativo e técnicas de geração de sinal para a síntese da fala**. 2002. Dissertação (Mestrado) – Faculdade de Engenharia, Universidade do Porto, Porto, 2002.

BIDERMAN, M. T. C. A ciência da lexicografia. *Alfa: Revista de Linguística*, São Paulo: UNESP, n. 28 (supl.), p. 1-26, 1984.

BIDERMAN, M. T. C. As ciências do léxico. In: OLIVEIRA, A. M. P. P. de; ISQUERDO, A. N. (org.). **As ciências do léxico**: lexicologia, lexicografia, terminologia. Campo Grande: UFMS, 1998. v. 1, p. 11-20.

BIDERMAN, M. T. C. **Teoria linguística**: teoria lexical e linguística computacional. 2. ed. São Paulo: Martins Fontes, 2001.

BORTONI-RICARDO, S. M. **Educação em língua materna**: a sociolinguística na sala de aula. São Paulo: Parábola, 2004.

BRAGA, D. Máquinas falantes: novos paradigmas da língua e da linguística. **Colóquio Política Linguística**, 2007.

BRASIL, Ministério da Educação. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais**. Brasília: MEC/SEF, 1998.

BRASIL, Ministério da Educação. Secretaria de Educação Básica. **Base Nacional Comum Curricular**. Brasília, DF: MEC, 2017.

CHETVERYK, V.; VERETIUK, T. The potential of artificial intelligence tools in foreign language teaching and learning. *In: Modern challenges and achievements of the scientific community of the 21st century: XLIII International Scientific and Practical Conference*, 2024, Narva, Estonia. p. 139-142.

DUTOIT, T. **High quality text-to-speech synthesis of the French language**. 1993. Tese (Doutorado) – Faculté Polytechnique de Mons, Mons, 1993.

DUTOIT, T. **An introduction to text-to-speech synthesis**. Kluwer Academic Publishers, 1997. DOI <https://doi.org/10.1007/978-94-011-5730-8>

EGLINTON, T.; TRANTER-ENTWISTLE, I.; CONNOR, S. Artificial intelligence in medicine: Promethean moment or Pandora's box? **New Zealand Medical Journal**, v. 136, p. 11-13, 2023. DOI <https://doi.org/10.26635/6965.6309>

FREITAG, R. M. K. Diversidade linguística e inclusão digital: desafios para uma IA brasileira. *In: Conferência Latino-Americana de Ética em Inteligência Artificial*, 1., 2024, Niterói. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 157-160. DOI <https://doi.org/10.5753/laai-ethics.2024.32476>

FREITAG, R. M. K. **Variação linguística: diversidade e cotidiano**. São Paulo: Contexto, 2025.

FUERTE-OLIVERA, P. A. Making lexicography sustainable: using ChatGPT and reusing data for lexicographic purposes. **Lexikos**, v. 34, n. 1, p. 123-140, 2024. DOI <https://doi.org/10.5788/34-1-1883>

HAN, Y.; LI, S.; CAO, Y.; MA, Q.; YOSHIKAWA, M. Voice-indistinguishability: protecting voiceprint in privacy-preserving speech data release. *In: IEEE International Conference on Multimedia and Expo*, 2020. IEEE, 2020. DOI <https://doi.org/10.1109/ICME46284.2020.9102875>

HU, Y.; LI, R.; WANG, S.; TAO, F.; SUN, Z. SpeechHide: a hybrid privacy-preserving mechanism for speech content and voiceprint in speech data sharing. *In: 2022 7th IEEE International Conference on Data Science in Cyberspace (DSC)*, 2022. IEEE, 2022. p. 345-352. DOI <https://doi.org/10.1109/DSC55868.2022.00054>

HUANG, X.; ACERO, A.; HON, H.-W.. **Spoken Language Processing: A guide to theory, algorithm and system development**. Upper Saddle River: Prentice Hall PTR, 2001.

HUANG, Y.; LI, J.; HE, L.; WEI, W.g; GALE, W.; GONG, Y. Using personalized speech synthesis and neural language generator for rapid speaker adaptation. *In: IEEE International Conference on Acoustics, Speech and Signal Processing*, 2020. p. 7399-7403. DOI <https://doi.org/10.1109/ICASSP40776.2020.9053104>

KOLEGOVA, I. A.; LEVINA, I A. Using artificial intelligence as a digital tool in foreign language teaching. **Bulletin of the South Ural State University. Series: Education. Educational Sciences**, v. 16, n. 1, p. 102-110, 2024. DOI <https://doi.org/10.14529/ped240110>

KRIEGER, M. da G. Dicionário de língua: um instrumento didático pouco explorado. *In: TOLDO, C. S. (org.). Questões de linguística*. Passo Fundo: UPF Editora, 2003. p. 70-87.

KRIEGER, M. da G.. Questões de lexicografia pedagógica. *In: XATARA, C.; BEVILACQUA, C. R.; HUMBLÉ, P. R. M. (org.). Dicionários na teoria e na prática: como e para quem são feitos*. São Paulo: Parábola Editorial, 2011. p. 103-113.

KRIEGER, M. da G. **Dicionário em sala de aula: guia de estudos e exercícios**. Rio de Janeiro: Lexikon, 2012.

KRIEGER, M. da G. O léxico com letra maiúscula: reflexo do trabalho de Maria Tereza Biderman. **Debate Terminológico**, v. 14, p. 89-91, 2015.

LABOV, W. **Sociolinguistic patterns**. Philadelphia: University of Pennsylvania Press, 1972.

LAW, L. Application of Generative Artificial Intelligence (GenAI) in language teaching and learning: a scoping literature review. **Computers and Education Open**, v. 6, p. 100174, 2024. DOI <https://doi.org/10.1016/j.caeo.2024.100174>

LEFFA, V. O dicionário eletrônico na construção do sentido em língua estrangeira. **Cadernos de Tradução**, Florianópolis, n. 18, p. 319-340, 2006.

LEW, R. Dictionaries and lexicography in the AI era. **Humanities and Social Sciences Communications**, v. 11, n. 1, p. 426, 2024. DOI <https://doi.org/10.1057/s41599-024-02889-7>

LIU, M. Y. Exploring the application of artificial intelligence in foreign language teaching: challenges and future development. *In: SHS Web of Conferences*, v. 168. EDP Sciences, 2023. DOI <https://doi.org/10.1051/shsconf/202316803025>

- LIU, C.; LING, Z.-H.; CHEN, L.-H. Pronunciation dictionary-free multilingual speech synthesis using learned phonetic representations. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, v. 31, p. 3706-3716, 2023. DOI <https://doi.org/10.1109/TASLP.2023.3313424>
- MARQUES, G. S. de O.; POLIZER, A. S.; ZACARIAS, R. A. S. Elaboração de e-verbetes português/inglês para dicionários pedagógicos. **Brazilian Journal of Development**, v. 6, n. 10, p. 78482-78490, 2020. DOI <https://doi.org/10.34117/bjdv6n10-325>
- MEYER, S.; TILLI, P.; DENISOV, P.; LUX, F.; KOCH, J.; VU, N. T. Anonymizing speech with generative adversarial networks to preserve speaker privacy. *In: IEEE Spoken Language Technology Workshop (SLT)*, 2023. IEEE, 2023. p. 912-919. DOI <https://doi.org/10.1109/SLT54892.2023.10022601>
- MORISE, M.; YOKOMORI, F.; OZAWA, K. WORLD: a vocoder-based high-quality speech synthesis system for real-time applications. **IEICE TRANSACTIONS on Information and Systems**, v. 99, n. 7, p. 1877-1884, 2016. DOI <https://doi.org/10.1587/transinf.2015EDP7457>
- NADIN, O. L.; VARGAS, M. D. Dicionário on-line de espanhol como língua estrangeira: panorama lexicográfico do ambiente virtual. **Trama**, Cascavel, v. 12, n. 24, p. 190-208, 2016.
- NAM, K.; AN, Y.; JUNG, H.-Y. Lexicography, artificial intelligence, and dictionary users. **Lexicography**, v. 11, n. 1, p. 1-6, 2024. DOI <https://doi.org/10.1558/lexi.30075>
- NEGROPONTE, N. **A vida digital**. São Paulo: Companhia das Letras, 1995.
- OPENAI. **ChatGPT (versão GPT-4)**: modelo de linguagem treinado pela OpenAI. Disponível em: <https://chat.openai.com/>. Acesso em: 2 de dezembro de 2024.
- SÃO PAULO (ESTADO). Secretaria da Educação. **Currículo Paulista**. São Paulo: SEE, 2019.
- SAEKI, T.; WANG, G.; MORIOKA, N.; ELIAS, I.; KASTNER, K.; BIADSY, F.; ROSENBERG, A.; RAMABHADHAN, B.; ZEN, H.; BEAUFAYS, F.; SHEMTOV, H. Extending multilingual speech synthesis to 100+ languages without transcribed data. *In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024. DOI <https://doi.org/10.1109/ICASSP48485.2024.10448074>

SCHRYVER, G.-M. de. Generative AI and Lexicography: The Current State of the Art Using ChatGPT. **International Journal of Lexicography**, v. 36, n. 4, 2023. DOI <https://doi.org/10.1093/ijl/ecad021>

TARALLO, F. **A pesquisa sociolinguística**. São Paulo: Ática, 1985.

TAYLOR, P. **Text-to-speech synthesis**. Cambridge: Cambridge University Press, 2009. DOI <https://doi.org/10.1017/CBO9780511816338>

TENCH, P. Phonetic symbols in the dictionary and in the classroom. In: BROWN, A. (org.). **Approaches to pronunciation teaching**. London: Macmillan, 1992. p. 90-102.

TOBING, M.; SNA, F. I.; HAYRUNNISA, N. R.; HASWURI, N. I. T.; SUTARSYAH, C.; MUNIFATULLAH, F. An exploration of artificial intelligence in English language teaching as a foreign language. **International Journal of Social Science and Human Research**, v. 6, n. 6, p. 3837-3843, 2023.

TOJEIRA-RAMOS, J. P. **Prosódia computacional do português brasileiro: a entoação declarativa neutra gerada por um sistema de síntese de fala baseado em Inteligência Artificial (IA)**. 2025. 148f. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2025.

TOJEIRA-RAMOS, J. P.; MASSINI-CAGLIARI, G. Associação tonal no domínio prosódico da palavra fonológica no português brasileiro: o caso de dados produzidos por uma tecnologia de síntese de fala baseada em Inteligência Artificial. In: Simpósio Mundial de Estudos de Língua Portuguesa (SIMELP), 9.; Congresso da Associação Internacional de Linguística do Português (AILP), 6., 2024, Recife. **Anais...** Recife: Universidade da Madeira, 2024. Textos selecionados.

WELKER, H. A. **Panorama geral da lexicografia pedagógica**. Brasília: Thesaurus, 2008.

WELKER, H. A. Questões de lexicografia pedagógica. In: XATARA, C.; BEVILACQUA, C. R.; HUMBLÉ, P. R. M. (org.). **Dicionários na teoria e na prática: como e para quem são feitos**. São Paulo: Parábola Editorial, 2011. p. 103-113.

YANG, A. Challenges and opportunities for foreign language teachers in the era of artificial intelligence. **International Journal of Education and Humanities**, v. 4, n. 1, p. 39-50, 2024. DOI [https://doi.org/10.58557/\(ijeh\).v4i1.202](https://doi.org/10.58557/(ijeh).v4i1.202)

YANG, K.; HU, T.-Y.; CHANG, J.-H. R.; KOPPULA, H. S.; TUZEL, O. Text is all you need: personalizing ASR models using controllable speech synthesis. In: **ICASSP 2023**

– 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023. p. 1-5. DOI <https://doi.org/10.1109/ICASSP49357.2023.10096971>

YANHUA, Z. The application of artificial intelligence in foreign language teaching. *In: International Conference on Artificial Intelligence and Education*, 2020. IEEE, p. 40-42. DOI <https://doi.org/10.1109/ICAIE50891.2020.00017>

ZACARIAS, R. A. S. Macro e microestruturas de dicionários escolares português-inglês/inglês-português. *Domínios de Lingu@gem*, v. 12, n. 1, p. 492-525, jan.-mar. 2018. DOI <https://doi.org/10.14393/DL33-v12n1a2018-17>