

# Recomendação de Palavras em Atividades de Leitura: Uma Revisão Sistemática

Geraldo Gomes da Cruz Júnior, Centro de Informática - UFPE, [ggc4@cin.ufpe.br](mailto:ggc4@cin.ufpe.br),  
<https://orcid.org/0000-0002-3856-5934>

Robson do Nascimento Fidalgo, Centro de Informática - UFPE, [rdnf@cin.ufpe.br](mailto:rdnf@cin.ufpe.br),  
<https://orcid.org/0000-0002-4714-2933>

**Resumo.** Aprender a ler não é um processo tão natural como o de aprender a falar. O treino da consciência fonológica a partir da recomendação de palavras apresenta bons resultados na aprendizagem da leitura, porém é um processo complexo. Este artigo apresenta uma revisão sistemática sobre estratégias computacionais para a recomendação de palavras para atividades de leitura. Foram analisados estudos sobre modelos LLMs, redes neurais, frameworks de PLN, métricas de avaliação e interpretabilidade. Os resultados indicam desafios quanto à disponibilidade limitada de dados de treinamento, a ineficiência de representações computacionais quanto a características complexas e ambíguas da língua natural e a ausência de modelos robustos para a língua portuguesa.

**Paravras-chave:** alfabetização. consciência fonológica. recomendação de palavras.

## *Word Recommendation in Reading Activities: A Systematic Review*

**Abstract.** Learning to read is not as natural a process as learning to speak. Training phonological awareness through word recommendation has shown positive results in reading acquisition; however, it is a complex process. This article presents a systematic review of computational strategies for recommending words in reading activities. The study analyzes research on LLMs, neural networks, NLP frameworks, evaluation metrics, and interpretability. The results highlight challenges such as the limited availability of training data, the inefficiency of computational representations in capturing complex and ambiguous features of natural language, and the lack of robust models for the Portuguese language.

**Keywords:** literacy. phonological awareness. word recommendation.

### 1. Introdução

A alfabetização é um processo complexo que envolve o domínio do princípio alfabético e a capacidade de formar palavras e frases (MORAIS, 2019). Para Freire (1983), alfabetizar vai além de decodificar símbolos: é possibilitar que o aluno estabeleça conexões com o mundo. No Brasil, os desafios são significativos, como aponta o Programa Internacional de Avaliação de Alunos (PISA) 2022: cerca de 50% dos jovens não atingiram o nível mínimo de proficiência em leitura (BRASIL, 2022). Além disso, segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD), 2.4 milhões de crianças são analfabetas (IBGE, 2022).

Diferente da oralidade, que se adquire naturalmente, a leitura e a escrita exigem competências cognitivas, perceptivas e emocionais (SIM-SIM *et al.*, 1997). Aprender a ler envolve compreender a relação entre sons e letras, sendo a consciência fonológica, habilidade

de perceber e manipular sons, um componente central nesse processo (TUNMER *et al.*, 1984; BRYANT e BRADLEY, 1985).

Estudos indicam que atividades pedagógicas baseadas na consciência fonológica, sobretudo aquelas que envolvem recomendação de palavras similares, contribuem para o avanço da leitura e escrita (JERKE, 2021; ADAMS, 2018; MOREIRA, 2016; MOTA e SILVA, 2007). O uso da similaridade fonológica favorece a identificação de padrões sonoros e fortalece a decodificação (SILVA e BRITTO, 2013; MECCA e LICHTIG, 2005). No entanto, modelar computacionalmente esses contextos é desafiador, devido à ambiguidade e variabilidade linguística (JERKE, 2021; GILLON, 2007).

Diante disso, este artigo apresenta uma Revisão Sistemática da Literatura (RSL) sobre estratégias computacionais para recomendação de palavras em atividades de leitura, com foco no treino da consciência fonológica. O objetivo é identificar modelos, algoritmos e *frameworks* aplicados à tarefa, buscando contribuir para o desenvolvimento de ferramentas mais eficazes, personalizadas e adequadas ao contexto da língua portuguesa.

## 2. Trabalhos Relacionados

A recomendação de palavras em atividades de leitura tem sido tema recorrente em pesquisas voltadas à promoção da fluência, compreensão e desenvolvimento lexical. Diversas abordagens computacionais vêm sendo propostas para apoiar esse processo, com destaque para o uso de *machine learning* e Processamento de Linguagem Natural (PLN).

Alenezi e Faisal (2020) analisaram o uso combinado de aprendizado de máquina e *crowdsourcing* para adaptar sugestões lexicais ao nível de proficiência dos estudantes, destacando benefícios em ambientes educacionais personalizados. Rivera *et al.* (2018) e Liu *et al.* (2022) reforçam essa abordagem ao evidenciar como sistemas baseados em PLN e *deep learning* viabilizam recomendações contínuas e adaptativas, especialmente no *e-learning*. Modelos linguísticos contextualizados também vêm ganhando espaço. De Gois *et al.* (2024) demonstraram, por meio de testes Cloze<sup>1</sup> com geração automática de respostas, o potencial de modelos de similaridade semântica na identificação de lacunas e na recomendação de vocabulário. Representações como *sense embeddings* e variantes do BERT<sup>2</sup> treinadas em português têm se destacado por capturar nuances semânticas e fonológicas mais complexas do que *embeddings* tradicionais, conforme sugerido por Abdullah *et al.* (2021).

Santana e Magalhães (2024), em revisão sistemática sobre o uso de PLN na educação, abordam técnicas como tokenização, análise semântica e cálculo de similaridade lexical. Destacam a crescente integração entre tecnologia e práticas pedagógicas, embora ainda haja lacunas no tratamento de aspectos fonológicos. Nesse escopo, De Sousa *et al.* (2021) propuseram classificadores para prever dislexia, evidenciando o potencial dos modelos computacionais no diagnóstico e apoio à aprendizagem da leitura. Já Meira *et al.* (2024) utilizaram métricas linguísticas automatizadas com o sistema NILC-Metrix para avaliar a complexidade textual em mais de dois mil trabalhos de conclusão de cursos, mostrando que indicadores como sofisticação lexical e coesão podem fortalecer sistemas de recomendação e avaliação adaptativa em contextos educacionais diversos.

Estudos sobre desenvolvimento da leitura também apontam a consciência fonológica como um dos principais preditores do sucesso na alfabetização (GILLON, 2007; ADAMS, 2018; MORAIS, 2019). No entanto, o uso explícito dessa dimensão como base para sistemas

---

<sup>1</sup> <https://www.scielo.br/j/prc/a/gtXWFDhqfQ4VQykPfNGFQqJ>

<sup>2</sup> <https://arxiv.org/abs/1810.04805>

de recomendação é escasso. Embora trabalhos como o de Rajaram (2018) investiguem os efeitos da similaridade fonológica na aquisição de palavras, sua aplicação em sistemas automatizados ainda é incipiente.

Pesquisas como a de Morciano *et al.* (2024) e Rauschenberger *et al.* (2022), voltadas à recomendação e detecção de dificuldades como dislexia, reforçam a necessidade de sistemas que considerem aspectos linguísticos finos, como ortografia e fonologia. Nesse mesmo eixo, Frances *et al.* (2021) e Zhao *et al.* (2017) investigam as representações cognitivas ortográficas e fonológicas, evidenciando como esses fatores impactam o reconhecimento de palavras em bilíngues e sua ativação em regiões específicas do córtex fusiforme. Adicionalmente, abordagens de avaliação e *feedback* automático com o uso de reconhecimento automático de fala (ASR), como as de Bai *et al.* (2023) e Molenaar *et al.* (2023), trazem contribuições importantes ao mostrar como a tecnologia pode apoiar diretamente o ensino da leitura, inclusive com métricas como acurácia e Cohen's Kappa para validação da leitura oral.

Apesar desses avanços, identificam-se lacunas relevantes: há poucos estudos voltados à recomendação lexical baseada especificamente em características fonológicas, sobretudo no contexto da língua portuguesa. Esta pesquisa busca preencher essa lacuna ao analisar modelos, algoritmos e frameworks computacionais voltados à recomendação de palavras em atividades fonológicas, considerando critérios como desempenho, interpretabilidade, personalização e aderência ao contexto educacional e linguístico do Brasil.

### 3. Metodologia

A RSL segue o protocolo de Kitchenham (2004), reconhecido na Engenharia de Software e adaptado para estudos interdisciplinares em tecnologia e educação, buscando reduzir vieses e garantir reprodutibilidade. O foco está em mapear modelos, técnicas, métricas e lacunas, como base para investigações e soluções futuras em Inteligência Artificial (IA) educacional.

Segundo Kitchenham (2004), uma RSL deve partir de motivações claras. Conforme discutido na Seção 1, embora a recomendação fonológica de palavras tenha mostrado bons resultados na alfabetização, sua aplicação ainda depende majoritariamente de especialistas, como educadores e fonoaudiólogos. Assim, a computação surge como aliada ao automatizar processos e possibilitar intervenções mais personalizadas. A RSL é, portanto, relevante para consolidar o conhecimento atual e evidenciar limitações e oportunidades nesse campo.

A formulação dos objetivos da pesquisa seguiu o paradigma GQM (*Goal-Question-Metric*), conforme Basili *et al.* (1994), e as questões foram estruturadas com base no modelo PICO (Huang *et al.*, 2006), adaptado ao contexto desta revisão, como detalhado na Tabela 1.

Tabela 1 - GQM da Revisão Sistemática.

| Analisar: Recomendação de palavras |                                                                       |
|------------------------------------|-----------------------------------------------------------------------|
| Com o propósito de                 | Identificar / Caracterizar / Avaliar                                  |
| Em relação a                       | Modelos, Algoritmos e Ferramentas                                     |
| Do ponto de vista das              | Pesquisas em modelagem e inteligência computacional                   |
| No contexto do                     | Ensino personalizado, Alfabetização, Leitura e Consciência fonológica |

#### 3.1. Questões de pesquisa

As questões de pesquisa deste trabalho visam entender como a tecnologia pode auxiliar na criação de atividades pedagógicas mais eficazes, com base em aspectos fonológicos. A seguir, são apresentadas as perguntas que norteiam esta revisão:

- **Q1.** Quais artefatos computacionais têm sido empregados para a recomendação de palavras em atividades de leitura?

- **Q2.** Quais variáveis são consideradas essenciais para o uso desses artefatos nos estudos analisados?
- **Q3.** Quais métodos de avaliação têm sido utilizados nas pesquisas identificadas?
- **Q4.** Quais os principais desafios e lacunas apontados nos estudos sobre essa temática?

### 3.2. Estratégia de busca

A busca sistemática foi realizada em cinco bases reconhecidas (Google Scholar, IEEE Xplore, SpringerLink, ScienceDirect e Scopus) considerando publicações de 2012 a 2024, em português ou inglês. As *strings* de busca incluíram sinônimos e variações relevantes ao tema, conforme ilustrado na Tabela 2.

Tabela 2 - *Strings* de busca.

| <i>String</i>                                                                                                                                                                          | Idioma    | Retorno          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------------------|
| ("Reading" OR "Literacy") AND ("Phonological Similarity" OR "Phonological Representation") AND ("recommendation system" OR "Artificial intelligence" OR "machine learning")            | Inglês    | 874<br>Trabalhos |
| ("Leitura" OR "Alfabetização") AND ("Similaridade Fonológica" OR "Representação Fonológica") AND ("Sistema de Recomendação" OR "Inteligência Artificial" OR "Aprendizagem de Máquina") | Português | 36<br>Trabalhos  |

### 3.3. Critérios de inclusão e exclusão

Com base nas diretrizes metodológicas propostas por Kitchenham (2004), foram definidos critérios de inclusão e exclusão para os estudos identificados. Os critérios para a inclusão de artigos são: I. Estudos primários revisados por pares; II. Estudos com a proposição de artefatos computacionais (e.g., modelos, ferramentas, técnicas, algoritmos ou *frameworks*) para o desenvolvimento da leitura.

Já os critérios para a exclusão de artigos são: I. Artigos não escritos na língua inglesa e portuguesa; II. Publicações não acessíveis na internet; III. Estudos no formato de pôsteres, *banners*, tutoriais ou palestras; IV. Estudos redundantes do mesmo autor. Nesse caso, apenas o artigo mais recente é mantido; V. Artigos anteriores a 2012 e posteriores a 2024; VI. Estudos que não apresentaram validação empírica, como testes com usuários, análises quantitativas ou avaliações experimentais.

### 3.4. Processos de seleção

Após definidos os critérios de inclusão e exclusão, a seleção dos estudos ocorreu em duas etapas. No Filtro I (Seleção Preliminar), foram analisados título e resumo, com base na aderência à questão de pesquisa. Cada decisão de inclusão ou exclusão foi justificada com base nos critérios estabelecidos. No Filtro II (Seleção Final), realizou-se a leitura integral dos estudos previamente selecionados, com o objetivo de aprofundar a análise e extrair dados mais precisos, seguindo os mesmos critérios de inclusão e exclusão.

Para garantir a confiabilidade da seleção, adotou-se o modelo de avaliação de qualidade proposto por Bibi *et al.* (2014), que considera aspectos como precisão e relevância dos estudos analisados, descrito a seguir:

- **Precisão:** Avaliada com base na descrição do ambiente de aplicação e coleta de dados. Classificada como forte (1), média (0,5) ou fraca (0).
- **Relevância:** Avaliada considerando a amostra, ambiente e método da pesquisa. Classificada como relevante (1) ou não relevante (0).

Por fim, observa-se o avanço das pesquisas em computação aplicadas à recomendação de palavras para o desenvolvimento da leitura. Uma linha do tempo evidencia as principais contribuições e métodos adotados, destacando lacunas ainda pouco exploradas e oportunidades futuras. A Figura 1 resume esse percurso e a quantidade de artigos em cada etapa da seleção.

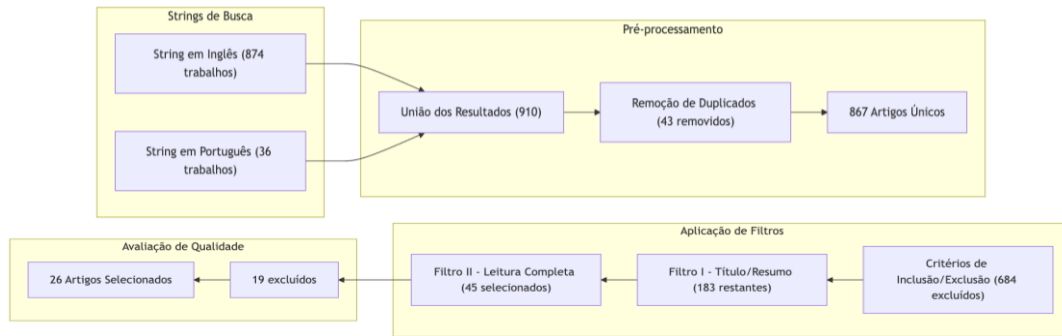


Figura 1 - Evolução das pesquisas e distribuição dos artigos por etapa da seleção.

#### 4. Resultados e Discussões

A partir do conjunto de 910 estudos analisados, 26 foram selecionados por serem os mais relacionados e relevantes ao tema pesquisado. Dentre os 910 estudos, muitos abordavam aspectos voltados para análises cognitivas e leitura do cérebro humano, estes trabalhos não foram considerados nesta pesquisa. A Tabela 3 ilustra a distribuição destas pesquisas ao longo dos anos, evidenciando um aumento gradual de produções, sobretudo a partir de 2020.

Tabela 3 - Distribuição de trabalhos por anos.  
Referências

| Ano  | Publicação | Referências                                                                                                                                                                                               |
|------|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2016 | 1          | [Fergadiotis <i>et al.</i> 2016]                                                                                                                                                                          |
| 2017 | 1          | [Zhao <i>et al.</i> 2017]                                                                                                                                                                                 |
| 2018 | 2          | [Yatsu and Araki 2018] e [Rajaram 2018]                                                                                                                                                                   |
| 2020 | 5          | [Low 2020], [Sen <i>et al.</i> 2020], [Modak <i>et al.</i> 2020], [Ferandes <i>et al.</i> 2020] e [Pazeto <i>et al.</i> 2020]                                                                             |
| 2021 | 7          | [Chernova e Podvigina 2021], [Shin e Park 2021], [Hasnine <i>et al.</i> 2021], [Frances <i>et al.</i> 2021], [Abdullah <i>et al.</i> 2021], [Sinclair <i>et al.</i> 2021] e [de Sousa <i>et al.</i> 2021] |
| 2022 | 4          | [Li 2022], [Weiss <i>et al.</i> 2022], [Rauschenberger <i>et al.</i> 2022] e [Al Madi 2022]                                                                                                               |
| 2023 | 3          | [Bai <i>et al.</i> 2023], [Molenaar <i>et al.</i> 2023], [Andrade <i>et al.</i> 2023]                                                                                                                     |
| 2024 | 3          | [Erbeli <i>et al.</i> 2024], [Alves <i>et al.</i> 2024], [Morciano <i>et al.</i> 2024]                                                                                                                    |

A RSL realizada oferece uma visão das estratégias computacionais aplicadas à recomendação de palavras em atividades de leitura, destacando técnicas, métodos, resultados e formas de avaliação. Identificam-se lacunas importantes, especialmente na personalização baseada em características fonológicas. Esses achados podem orientar futuras pesquisas e o desenvolvimento de novas ferramentas. As subseções a seguir apresentam as respostas às questões de pesquisa.

##### 4.1. Quais artefatos computacionais têm sido empregados para a recomendação de palavras em atividades de leitura?

Os estudos analisados destacam que a qualidade das recomendações depende diretamente da representatividade das características linguísticas modeladas (SHIN e PARK, 2021; ABDULLAH *et al.*, 2021). Observou-se o uso de representações fonéticas (via ARPabet<sup>3</sup>,

<sup>3</sup> <https://github.com/avlos/arpabet>

SAMPA<sup>4</sup>), semânticas e visuais das palavras, com menor exploração de níveis intrassilábicos ou silábicos (SINCLAIR *et al.*, 2021; FRANCES *et al.*, 2021).

As técnicas computacionais mais utilizadas envolvem algoritmos de classificação (SVM, árvores de decisão), filtragem colaborativa e *deep learning* (YATSU e ARAKI, 2018; LOW, 2020; LI, 2022). O surgimento dos Modelos de Linguagem de Grande Escala (LLMs), como o GPT-3<sup>5</sup>, aparece como um recurso promissor na representação contextualizada e semântica de palavras (HASNINE *et al.*, 2021; MORCIANO *et al.*, 2024), permitindo captar relações complexas e ajustadas ao contexto.

O uso de frameworks de PLN, como NLTK<sup>6</sup> e spaCy<sup>7</sup>, também foi frequente, especialmente em tarefas de análise morfossintática. Algoritmos de similaridade lexical mais tradicionais (como Levenshtein e Jaro-Winkler) foram observados em estudos mais antigos (FERGADIOTIS *et al.*, 2016), mas têm sido combinados com modelos modernos em abordagens híbridas (DE SOUSA *et al.*, 2021).

Outros estudos reforçam essa diversidade: Bai *et al.* (2023) e Molenaar *et al.* (2023) propõem sistemas de leitura automatizada baseados em reconhecimento de fala, apontando caminhos para integrar recomendação fonológica a tutores inteligentes. Morciano *et al.* (2024) explora métricas de similaridade vetorial voltadas a estudantes disléxicos, com implicações diretas na personalização. A Figura 2 apresenta uma visão geral das principais estratégias identificadas nos artigos.

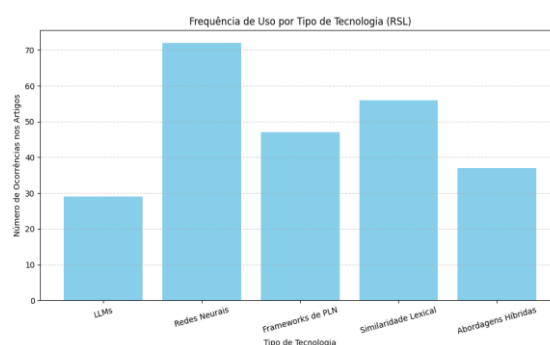


Figura 2 - Frequência das tecnologias encontradas na RSL.

#### 4.2. Quais variáveis são consideradas essenciais para o uso desses artefatos nos estudos analisados?

As variáveis mais comumente utilizadas envolvem aspectos fonéticos, semânticos e ortográficos das palavras, além de informações contextuais dos estudantes. Trabalhos como Al Madi (2022) e Zhao *et al.* (2017) destacam o uso de níveis de dificuldade, preferências de leitura e vocabulário adaptado. Nos LLMs, variáveis como corpus textual diversificado, vocabulário de alta cobertura e dados anotados foram essenciais para customizar o modelo ao contexto educativo (HASNINE *et al.*, 2021; MORCIANO *et al.*, 2024).

No caso de modelos neurais, a qualidade do conjunto de treinamento (em geral, textos paradidáticos ou corpus de alfabetização) é determinante. Andrade *et al.* (2023), por exemplo, destacam como a consciência fonológica e a memória fonológica são variáveis preditoras críticas no sucesso da leitura inicial. Já Alves *et al.* (2024) contribuem com um protocolo de avaliação que pode ser utilizado como fonte de variáveis para validação de recomendações computacionais. A necessidade de consideração do contexto, volume de dados e retroalimentação pedagógica aparece de forma recorrente.

#### 4.3. Quais métodos de avaliação têm sido utilizados nas pesquisas identificadas?

<sup>4</sup> <https://www.phon.ucl.ac.uk/home/sampa/>

<sup>5</sup> <https://openai.com/index/gpt-3-apps/>

<sup>6</sup> <https://www.nltk.org/>

<sup>7</sup> <https://spacy.io/>

A avaliação dos artefatos segue uma combinação de métricas quantitativas e análises qualitativas. Métricas como acurácia, precisão, *F1-score*, perplexidade e taxa de cliques são comuns em estudos focados em desempenho computacional (LI, 2022; AL MADI, 2022). Em contraste, estudos como Low (2020) e Sen *et al.* (2020) adotam entrevistas e feedbacks com professores para avaliar a relevância pedagógica das palavras recomendadas.

Bai *et al.* (2023) utilizam Cohen's Kappa e *F-measure* para comparar a classificação do sistema de *Automatic Speech Recognition* (ASR) com a de humanos, enquanto Molenaar *et al.* (2023) aplicam *Matthews Correlation Coefficient* (MCC) para validação da leitura oral. Esses métodos reforçam a importância da triangulação entre avaliação automatizada e julgamento humano. Em abordagens híbridas, a combinação dessas técnicas permite uma visão mais completa da efetividade dos sistemas.

#### 4.4. Quais os principais desafios e lacunas apontados nos estudos sobre essa temática?

Apesar dos avanços, permanecem desafios significativos. Há carência de modelos que capturem adequadamente aspectos fonológicos complexos, como sílabas complexas e regionalismos (SINCLAIR *et al.*, 2021). A maioria dos trabalhos utiliza representações fonéticas fixas, com pouca flexibilidade para personalização conforme *input* de especialistas. Ademais, a escassez de modelos voltados ao português brasileiro, dificultando aplicações em contextos locais.

Variações linguísticas, ambiguidade semântica e falta de explicação dos modelos são entraves recorrentes. Como observado em Erbeli *et al.* (2024), a dosagem ideal de intervenções fonológicas deve ser considerada para que sistemas de recomendação não apenas sugiram palavras, mas colaborem com a jornada pedagógica de maneira estruturada. A maioria dos sistemas identificados ainda carece de explicabilidade, o que limita a confiança de educadores.

Outro ponto crítico é a dependência de grandes volumes de dados para treinar modelos robustos. Isso não só dificulta aplicações em contextos de baixa infraestrutura como também limita a reprodutibilidade dos experimentos. Tais lacunas reforçam a necessidade de estudos futuros que combinem inteligência computacional com pragmatismo pedagógico, respeitando as especificidades do contexto linguístico e educacional brasileiro.

#### 4.5. Estratégias de Recomendação de Palavras em Atividades de Leitura

A revisão sistemática realizada revelou um conjunto diverso de estratégias computacionais aplicadas à recomendação de palavras em atividades voltadas ao desenvolvimento da leitura. Os principais artefatos identificados incluem modelos de LLMs, métodos baseados em redes neurais, *frameworks* de PLN, algoritmos de similaridade lexical e abordagens híbridas. Esses sistemas procuram representar palavras com base em aspectos fonológicos (sons), semânticos (significados) e visuais (forma escrita), com o objetivo de personalizar a experiência de leitura conforme o perfil e as necessidades dos estudantes. A Tabela 4 sintetiza as principais abordagens, relacionando cada técnica com suas principais variáveis de entrada e as métricas de avaliação comumente aplicadas.

Tabela 4 - Relação de técnicas, variáveis e métricas.

| Técnica                                | Variáveis                                                                                                        | Métricas                                                                                        |
|----------------------------------------|------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| LLMs                                   | Corpus textual diversificado; Vocabulário adaptado ao contexto escolar; Sentenças de treinamento representativas | Perplexidade; Avaliação por especialistas; Feedback de usuários (Morciano <i>et al.</i> , 2024) |
| Redes Neurais                          | Dados de treinamento anotados; <i>Word embeddings</i> pré-treinados; Corpus fonológico (Word2Vec, GloVe)         | Precisão, Revocação, <i>F1-score</i> , Acurácia (Li, 2022; Shin e Park, 2021)                   |
| <i>Frameworks</i> de PLN (NLTK, spaCy) | Dicionários linguísticos; Modelos <i>POS tagging</i> ; Lematizadores e Reconhecimento de Entidade Nomeada (NER)  | Precisão e revocação em tarefas de lematização, NER e <i>POS tagging</i>                        |

|                                    |                                                                                        |                                                                                                |
|------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|
| Algoritmos de Similaridade Lexical | Matrizes de similaridade (Levenshtein, Jaro-Winkler); Tabelas de frequência fonética   | Correlação com pares conhecidos; Análise de similaridade fonológica (Fergadiotis et al., 2016) |
| Abordagens Híbridas                | Combinação de <i>embeddings</i> semânticos e fonológicos; Pesos contextuais ajustáveis | Métricas combinadas (perplexidade + avaliação subjetiva)                                       |

## 5. Ameaças à Validade

Embora esta revisão sistemática tenha seguido protocolos consolidados, como os de Kitchenham (2004), algumas limitações devem ser consideradas. Restrições nas *strings* de busca, no idioma e na cobertura das bases podem ter excluído estudos relevantes, e a interpretação de dados com descrições limitadas pode ter introduzido vieses. Além disso, a heterogeneidade metodológica dos estudos dificulta comparações diretas, comprometendo a validade interna, enquanto a predominância de pesquisas internacionais limita a aplicabilidade dos achados ao contexto brasileiro. Reconhecer essas ameaças é fundamental para uma leitura crítica dos resultados e para orientar pesquisas futuras mais contextualizadas.

## 6. Considerações Finais

Esta RSL sintetizou o panorama atual das pesquisas que aplicam artefatos computacionais na recomendação de palavras para o desenvolvimento da leitura. Os estudos analisados demonstram um interesse crescente no uso de LLMs, redes neurais e frameworks de PLN, com resultados promissores no apoio a práticas educacionais.

Apesar dos avanços, desafios persistem, como a escassez de dados representativos, especialmente em português, e as limitações na personalização das recomendações. A complexidade da linguagem natural, com ambiguidade, polissemia e variações fonológicas, exige modelos mais refinados e sensíveis às nuances linguísticas e cognitivas dos usuários.

A avaliação dos modelos também se mostra crítica: embora métricas tradicionais sejam utilizadas, ainda há lacunas na medição da eficácia pedagógica. Futuras pesquisas devem focar na criação de corpora mais representativos, em mecanismos de recomendação adaptativos, em métricas que incorporem aspectos qualitativos da leitura e em estratégias que ampliem a interpretabilidade dos modelos, aspectos fundamentais para o avanço sustentável da área e sua aplicabilidade real no contexto educacional.

## Referências

- ABDULLAH, B. M.; MOSBACH, M.; ZAITOVA, I.; MÖBIUS, B.; KLAKEW, D. *Do acoustic word embeddings capture phonological similarity? An empirical study*. arXiv preprint arXiv:2106.08686, 2021.
- ADAMS, Marilyn J. *Consciência fonológica em crianças pequenas*. POA: Artmed, 2018.
- AL MADI, N. *Namesake: A checker of lexical similarity in identifier names*. In: 37th IEEE/ACM International Conference on Automated Software Engineering, 2022.
- ALENEZI, H. S.; FAISAL, M. H. *Utilizing crowdsourcing and machine learning in education: Literature review*. Education and Information Technologies, v. 25, n. 4, p. 2971–2986, 2020.
- ALVES, A. C.; CRUZ, J. F.; KICKHÖFEL, A. *Elaboração da bateria de avaliação metafonológica*. Anais do Congresso Internacional de Alfabetização, Porto Alegre: Mundo Alfal, 2024.



- ANDRADE, C. N.; OLIVEIRA, T. R.; LIMA, R. M. *Habilidades preditoras da leitura e escrita em escolares do 1.º e 2.º ano do Ensino Fundamental I*. Research, Society and Development, v. 12, n. 8, p. e51212833693, 2023.
- BAI, Y.; TEJEDOR-GARCÍA, C.; HUBERS, F.; CUCCHIARINI, C.; STRIK, H. *An ASR-based tutor for learning to read: how to optimize feedback to first graders*. arXiv preprint, 2023.
- BASILI, V. R.; CALDEIRA, G.; ROMBACH, D. H. *The Goal Question Metric Approach*. Encyclopedia of Software Engineering, v. 1, 1994.
- BIBI, S.; GEROGIANNIS, V. C.; KAKARONTZAS, G.; STAMELOS, I. *Ontology based Bayesian software process improvement*. In: 9th International Conference on Software Engineering and Applications (ICSOFT-EA). IEEE, 2014.
- BRASIL. *PISA 2022 – Relatório Nacional*. Brasília: INEP/MEC, 2022.
- BRYANT, P.; BRADLEY, L. *Phonetic analysis capacity and learning to read (reply)*. Nature, v. 313, n. 5997, p. 74, 1985.
- CHERNOVA, D.; PODVIGINA, D. *Phonological and orthographic representations in visual word recognition: ERP study of Russian homophones*. In: Advances in Cognitive Research, Artificial Intelligence and Neuroinformatics, Springer, 2021. p. 329–334.
- DE GOIS, T. S.; FREITAS, F. O.; TEJADA, J.; FREITAG, R. M. K. *NLP and education: Using semantic similarity to evaluate filled gaps in a large-scale Cloze test in the classroom*. The Mental Lexicon, p. 90–99, 2024.
- DE SOUSA, R. N.; DE SOUSA, R. N.; DE BRITO, R. X.; XIMENES, J. N. S. *Utilização de modelos computacionais baseados em classificadores para predição da dislexia em crianças*. In: XI Simpósio de Sistemas de Informação, 2021.
- ERBELI, L.; BLAIR, C.; REUTER, T. *A meta-analysis on the optimal cumulative dosage of early phonemic awareness instruction*. Scientific Studies of Reading, p. 3–23, 2024.
- FERGADIOTIS, G.; GORMAN, K.; BEDRICK, S. *Algorithmic classification of five characteristic types of paraphasias*. American Journal of Speech-Language, 2016.
- FERNANDES, A. C. P.; SANTOS, M. M.; ZUANETTI, P. A.; HAMAD, A. P. A. *Relação entre fatores do neurodesenvolvimento e a elaboração escrita em crianças do ensino fundamental*. Revista Família, Ciclos de Vida e Saúde no Contexto Social, v. 1, p. 493–502, 2020.
- FRANCES, C.; NAVARRA-BARINDELLI, E.; MARTIN, C. D. *Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals*. Scientific Reports, v. 11, n. 1, p. 12812, 2021.
- FREIRE, P. *Educação e mudança*. Rio de Janeiro: Paz e Terra, 1983.
- GILLON, G. T. *Phonological awareness: From research to practice*. New York: Guilford Press, 2007.
- HASNINE, M. N.; ISHIKAWA, M.; UEDA, H. *Broadening word learning scopes in informal learning using ubiquitous learning tools*. In: 2021 International Symposium on Educational Technology (ISET). IEEE, 2021. p. 232–234.
- HUANG, X.; LIN, J.; DEMNER-FUSHMAN, D. *Evaluation of PICO as a knowledge representation for clinical questions*. In: AMIA Annual Symposium Proceedings, 2006.
- IBGE. *Pesquisa Nacional por Amostra de Domicílios Contínua – PNAD Contínua*. Instituto Brasileiro de Geografia e Estatística, 2022.

- JERKE, R. B. *Método fônico de alfabetização: a nova indicação do MEC*. 2021.
- KITCHENHAM, B. *Procedures for performing systematic reviews*. Keele University, UK, 2004.
- LI, C. *English research learning and functional research based on constructivism theory and few-shot learning*. Computational Intelligence and Neuroscience, 2022.
- LIU, T.; WU, Q.; CHANG, L.; GU, T. *A review of deep learning-based recommender system in e-learning environments*. Artificial Intelligence Review, v. 55, n. 8, p. 5953–5980, 2022.
- LOW, S. *Applying machine learning to neuroimaging data to identify predictive models of reading disorder (RD)*. 2020.
- MECCA, F. F. D. N.; LICHTIG, I. *Investigação das funções da linguagem, da teoria da mente, do vocabulário e do desempenho escolar de alunos surdos*. 2005.
- MEIRA, R. R.; REATEGUI, E.; MOTZ, R. *Análise de trabalhos de conclusão de curso utilizando técnicas de processamento de linguagem natural*. Revista Novas Tecnologias na Educação, v. 22, n. 1, p. 456-465, 2024.
- MODAK, M.; WARADE, O.; SAIPRASAD, G.; SHEKHAR, S. *Machine learning based learning disability detection using LMS*. In: 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA). IEEE, 2020. p. 414–419.
- MOLENAAR, B.; TEJEDOR-GARCIA, C.; STRIK, H.; CUCCHIARINI, C. *Automatic assessment of oral reading accuracy for reading diagnostics*. arXiv preprint, 2023.
- MORCIANO, G.; LLERGO, J. M. A.; ZINGONI, A.; BOLIVAR, E. Y.; TABORRI, J.; CALABRÒ, G. *Use of recommendation models to provide support to dyslexic students*. Expert Systems with Applications, 2024.
- MORAIS, A. G. *Consciência fonológica na educação infantil e no ciclo de alfabetização*. Belo Horizonte: Autêntica, 2019.
- MOREIRA, J. S. *O valor preditivo da consciência fonológica e do reconhecimento de letras na aprendizagem da leitura*. Tese (Doutorado) – Universidade do Minho, 2016.
- MOTA, M. M. P. E.; SILVA, K. C. A. *Consciência morfológica e desenvolvimento ortográfico: Um estudo exploratório*. Revista Psicologia em Pesquisa, v. 1, n. 2, 2007.
- PAZETO, T. C. B.; DIAS, N. M.; GOMES, C. M. A.; SEABRA, A. G. *Prediction of reading and writing in elementary education through early childhood education*. Psicologia: Ciência e Profissão, v. 40, 2020.
- RAJARAM, M. C. *The influence of phonological similarity on multisyllabic word acquisition in children*. Tese (Doutorado) – The University of Texas, 2018.
- RAUSCHENBERGER, M.; BAEZA-YATES, R.; RELLO, L. *A universal screening tool for dyslexia by a web-game and machine learning*. Frontiers in Computer Science, v. 3, 2022.
- RIVERA, A. C.; TAPIA-LEON, M.; LUJAN-MORA, S. *Recommendation systems in education: A systematic mapping study*. In: International Conference on Information Technology & Systems (ICITS). Springer, 2018. p. 937–947.
- SANTANA, F. P.; MAGALHÃES, L. C.. *Aplicações do PLN no ambiente educacional: Uma revisão sistemática da literatura*. Revista Foco, 2024.
- SEN, A.; COX, C. R.; BORKENHAGEN, M. C.; SEIDENBERG, M. S.; ZHU, X. *Learning to read through machine teaching*. arXiv:2006.16470, 2020.
- SHIN, J.; PARK, J. *Pedagogical word recommendation: A novel task and dataset on personalized vocabulary acquisition for L2 learners*. arXiv preprint, 2021.

- SILVA, T. R.; BRITTO, D. B. O. *Variações semânticas nos enunciados de crianças em processo de desenvolvimento da linguagem oral: Estudo preliminar*. CEFAC, 2013.
- SIM-SIM, I.; DUARTE, I.; FERRAZ, M. J.; PORTUGAL. A língua materna na educação básica: Competências nucleares e níveis de desempenho. *Lisboa: Ministério da Educação de Portugal*, 1997.
- SINCLAIR, J.; JANG, E. E.; RUDZICZ, F. *Using machine learning to predict children's reading comprehension from linguistic features extracted from speech and writing*. *Journal of Educational Psychology*, v. 113, n. 6, p. 1088, 2021.
- TUNMER, W. E.; PRATT, C.; HERRIMAN, M. L. Metalinguistic awareness in children: Theory, research, and implications. *New York: Springer*, 1984.
- WEISS, B.; SZALMA, J.; VIDNYÁNSZKY, Z. *Data-driven detection of developmental dyslexia: A ML approach based on behavioral and eye-movement features*. 2022.
- YATSU, M.; ARAKI, K. *Comparison of pun detection methods using Japanese pun corpus*. In: Eleventh International Conference on Language Resources and Evaluation, 2018.
- ZHAO, L.; CHEN, C.; SHAO, L.; WANG, Y.; XIAO, X.; CHEN, C.; XUE, G. Orthographic and phonological representations in the fusiform cortex. *Cerebral Cortex*, v. 27, n. 11, p. 5197–5210, 2017.