



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
CIÊNCIA DA COMPUTAÇÃO

RAVI COSTA DE LUCENA

ANÁLISE DA CALIGRAFIA NA ESCOLA PRIMÁRIA VISANDO O MELHOR
RECONHECIMENTO AUTOMÁTICO

RECIFE

2026

Resumo

A escrita revolucionou a atividade humana, possibilitando a transmissão do conhecimento para gerações futuras.

A caligrafia, do grego kalos - belo e graphos - escrita, era até bem pouco tempo vista como extremamente relevante na educação.

A rápida disseminação da computação na sociedade moderna fez com que a caligrafia fosse negligenciada nas escolas, tendo sido vista como ultrapassada, já que os computadores são onipresentes.

Porém, a transcrição automática de textos manuscritos tornou-se extremamente relevante, uma vez que a avaliação manual de grandes massas de dados de redações, em concursos e vestibulares têm se tornado inviáveis.

Assim sendo, a caligrafia voltou a ser relevante, pois possibilita uma boa transcrição automática de textos manuscritos.

O presente trabalho visa analisar cópias de alunos da escola primária os orientando individualmente o que observar e corrigir de forma a possibilitar uma melhor transcrição automática do texto escrito.

1. Introdução

A escrita infantil constitui um elemento central no processo de alfabetização, desempenhando papel fundamental no desenvolvimento das competências linguísticas, cognitivas e motoras das crianças. Durante os primeiros anos de escolarização, a produção escrita não apenas reflete a apropriação do sistema alfabético, mas também revela aspectos da consciência fonológica, da coordenação e da construção progressiva das convenções ortográficas e textuais (Ferreiro; Teberosky, 1985).

Sob a perspectiva da psicologia e da pedagogia, a escrita manual ocupa um lugar privilegiado no desenvolvimento cognitivo infantil. Vygotsky (1987) demonstra que a produção escrita não se restringe à transcrição de símbolos, mas envolve a construção ativa do pensamento, a organização de ideias e a consolidação das relações entre linguagem e cognição. Nesse sentido, a escrita manuscrita constitui simultaneamente um instrumento de aprendizagem e um indicador do estágio de desenvolvimento linguístico da criança (Ferreiro, 1985).

Nesse contexto, a correção manual das produções escritas configura-se como uma atividade pedagógica essencial, porém é uma tarefa altamente demandante. No cenário educacional com turmas numerosas, a avaliação individualizada de textos manuscritos em larga escala torna-se um processo custoso em termos de tempo e esforço docente, o que frequentemente limita a frequência e a produtividade das devolutivas oferecidas aos alunos. Além disso, a ausência de feedback imediato compromete o ritmo da aprendizagem, dificulta a identificação precoce dos padrões recorrentes de erro e reduz as oportunidades de autorregulação do estudante.

Paralelamente, observa-se a escassez de bases de dados em língua portuguesa que contemplem escrita manuscrita infantil de forma estruturada e anotada. Embora existam conjuntos de dados amplamente utilizados para tarefas de reconhecimento óptico de caracteres (OCR), a maioria deles é composta por escrita adulta ou por palavras isoladas, não contemplando a complexidade gráfica e linguística das produções infantis. Essa lacuna limita o desenvolvimento, o treinamento e a avaliação de modelos computacionais voltados especificamente para o contexto de alfabetização em português.

Além disso, grande parte dos estudos existentes concentra-se em unidades lexicais isoladas ou em caracteres individuais, havendo ainda pouco foco em textos contínuos, como fábulas e narrativas curtas. Esse tipo de texto apresenta desafios adicionais, relacionados à segmentação de palavras, ao uso de pontuação, à marcação de discurso direto e à organização espacial do manuscrito, aspectos fundamentais tanto para a análise pedagógica quanto para o desempenho de sistemas de OCR.

Do ponto de vista computacional, a escrita manuscrita pode ser compreendida como um problema de reconhecimento de padrões, no qual sinais gráficos são convertidos em representações simbólicas. Técnicas de OCR e HWR, combinadas a etapas de pré-processamento, segmentação e aprendizado de máquina, permitem automatizar parte desse processo. No entanto, o desempenho desses sistemas está diretamente relacionado ao grau de legibilidade da escrita, uma vez que aspectos como regularidade do traçado, espaçamento entre letras e palavras, alinhamento e clareza dos caracteres influenciam de forma significativa a acurácia da transcrição automática. Em contextos de escrita infantil, nos quais a caligrafia ainda se encontra em processo de consolidação, variações acentuadas de forma, inclinação e organização espacial tendem a ampliar a ocorrência de ruídos e erros de reconhecimento,

impactando diretamente a qualidade dos resultados obtidos pelos sistemas de OCR. As principais contribuições do artigo são:

- (i) a coleta e organização de um conjunto de textos manuscritos infantis em português;
- (ii) a definição de categorias de erro em nível ortográfico e estrutural;
- (iii) uma análise quantitativa e qualitativa das divergências em relação ao texto original;
- (iv) a comparação entre diferentes ferramentas de OCR aplicadas à escrita infantil.

2. Trabalhos Relacionados

Estudos recentes destacam a importância da escrita manual como ferramenta de aprendizagem, evidenciando sua influência sobre aspectos cognitivos, motores e criativos. De modo geral, a literatura aponta que escrever à mão promove maior retenção de informações e organização do pensamento, sendo particularmente relevante durante a alfabetização, fase em que a coordenação visomotora e a consciência fonológica ainda estão em consolidação

Embora existam investigações sobre caligrafia e escrita manual, grande parte dos trabalhos apresenta limitações metodológicas, como amostras restritas, foco em adultos ou ausência de análises empíricas com textos contínuos infantis. Essas lacunas reforçam a relevância do presente estudo, que busca fornecer dados objetivos sobre a qualidade da escrita infantil e seus padrões de erro em contexto escolar

Pesquisas contemporâneas reforçam essa perspectiva ao comparar diretamente a escrita manual com métodos digitais. O estudo de *Ibaibarriga et al. (2025)* demonstra que crianças que praticam a escrita manual apresentam melhor desempenho na aprendizagem de letras e palavras quando comparadas àquelas que utilizam digitação, sugerindo que a experiência grafomotora contribui significativamente para o fortalecimento das representações ortográficas iniciais. Esses achados corroboram a chamada hipótese grafomotora, segundo a qual o ato físico de escrever desempenha papel fundamental no desenvolvimento da leitura e da escrita

O estudo de **Zhang et al. (s.d.)**, analisa a educação caligráfica em escolas primárias, O estudo mostra que a prática contínua da caligrafia melhora não apenas a legibilidade da escrita, mas também a concentração e os hábitos cognitivos positivos dos alunos, reforçando o vínculo entre habilidade motora e desempenho acadêmico. De modo complementar, **Smith (2025)** argumenta que a escrita manual estimula a criatividade e o pensamento crítico, enfatizando a importância de preservar essa prática em ambientes educacionais dominados por tecnologias digitais. A cópia manual, quando realizada de forma consciente e reflexiva, amplia a capacidade de observação e a compreensão textual — princípios diretamente aplicáveis ao ensino da escrita infantil, sobretudo na correção de erros e no aprimoramento da qualidade da produção textual. Por sua vez, **Li (2021)** destaca o papel da multimídia como complemento pedagógico, sugerindo um equilíbrio entre recursos tecnológicos e práticas manuais para manter o engajamento e o aprendizado significativo.

Apesar das contribuições teóricas e práticas, a maioria dos estudos analisados apresenta limitações metodológicas, como amostras restritas, ausência de comparações empíricas com métodos digitais e foco em públicos adultos. Essas lacunas reforçam a relevância do presente trabalho, que busca analisar a qualidade da escrita infantil e identificar erros no processo de aprendizagem, fornecendo dados objetivos sobre a influência da escrita manual na construção da competência linguística. Assim, este estudo não apenas complementa a literatura existente, mas também amplia sua aplicabilidade ao campo da educação básica, propondo uma abordagem empírica para compreender como aspectos motores, cognitivos e perceptivos interagem no desenvolvimento da escrita.

Diferentemente das ferramentas tradicionais de rastreamento, que operam majoritariamente sobre palavras isoladas ou textos curtos em língua inglesa, a abordagem proposta neste trabalho analisa produções manuscritas infantis em português a partir de textos narrativos completos, combinando medidas tradicionais de OCR (CER, WER e PMR) com análise linguística qualitativa, permitindo distinguir erros cognitivos de ruído de digitalização.

3. Metodologia

Este estudo adota uma abordagem experimental aplicada para a análise da escrita manuscrita infantil, integrando técnicas de OCR e procedimentos de anotação linguística. A metodologia foi estruturada da seguinte forma: definição do estímulo textual, caracterização dos participantes, coleta das produções manuscritas, preparação e digitalização dos documentos, transcrição automática e categorização dos erros em relação ao texto original. Esse processo permitiu a realização de análises quantitativas, por meio de métricas baseadas em edição, e qualitativas, a partir da identificação de padrões estruturais e cognitivos da escrita. A Figura 1 apresenta uma visão geral das etapas da pesquisa.

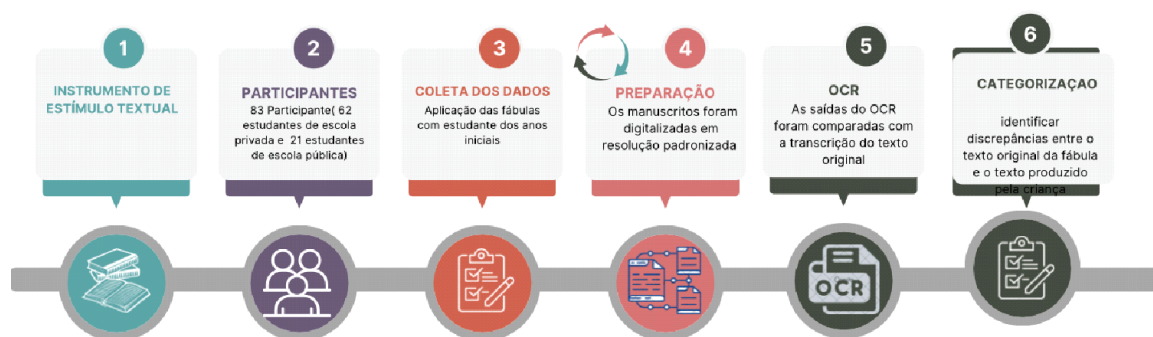


Figura 1: Etapas da pesquisa

3.1 Instrumento de estímulo textual

Como instrumento de coleta foi utilizada a fábula *O Urso e as Abelhas*, escolhida por sua adequação ao nível etário dos participantes e por permitir a produção de texto narrativo contínuo. Diferentemente de tarefas baseadas em palavras isoladas ou cópia mecânica de frases, o uso de um texto narrativo completo possibilita avaliar aspectos estruturais da escrita, como coerência global, segmentação de ideias e organização narrativa, além de características linguísticas locais, incluindo caligrafia, pontuação e flexão verbal.

A escolha dos textos também visa aproximar a atividade do contexto escolar real, uma vez que produções desse tipo são recorrentes nas práticas pedagógicas, permitindo observar o desempenho da criança em uma situação de escrita mais próxima daquela vivenciada no cotidiano.

3.2 Participantes

O estudo envolveu aproximadamente 83 estudantes do ensino fundamental, provenientes de duas instituições (uma pública e uma privada), com idades compatíveis com os anos iniciais da escolarização. A amostra incluiu alunos com diferentes níveis de proficiência em escrita, refletindo a heterogeneidade típica do ambiente escolar. Todos os participantes realizaram as atividades em contexto pedagógico regular, mediante autorização institucional.

3.3 Procedimento de coleta

Os estudantes produziram suas versões manuscritas da fábula em folhas pautadas, em ambiente escolar controlado, durante o horário regular de aula. Todos utilizaram os mesmos materiais (folha e instrumento de escrita) e receberam instruções padronizadas. Não houve limite rígido de tempo, permitindo que cada criança concluísse a tarefa conforme seu ritmo.

As produções manuscritas foram posteriormente digitalizadas em resolução padronizada e organizadas em relatórios, preservando-se a integridade dos textos originais.

3.4 Pré-processamento e transcrição

Após a digitalização, os textos foram submetidos ao processo de OCR. As saídas automáticas foram comparadas com transcrições de referência obtidas por revisão manual especializada, possibilitando a análise quantitativa da qualidade do reconhecimento e da escrita.

Foram calculadas três métricas principais: Character Error Rate (CER), Word Error Rate (WER) e Perfect Match Rate (PMR), que permitiram mensurar, respectivamente, a taxa de erro em nível de caractere, palavra e correspondência exata entre o texto transcrito automaticamente e o texto de referência.

3.5 Anotação linguística e categorização de erros

Cada produção textual foi analisada manualmente com o objetivo de identificar discrepâncias entre o texto original da fábula e o texto produzido pela criança. As divergências foram categorizadas em diferentes classes linguísticas, incluindo alterações de maiúsculas e minúsculas, omissões ou trocas de acentuação, erros ortográficos, problemas de pontuação, substituições lexicais, omissões de palavras ou trechos narrativos, bem como fragmentações estruturais.

Além da análise local, também foram observados aspectos globais do texto, como preservação da macroestrutura narrativa, presença ou ausência da Moral, continuidade semântica e organização dos parágrafos. Essa etapa permitiu distinguir erros predominantemente cognitivos, associados ao processo de alfabetização, de eventuais ruídos introduzidos pelo OCR.

3.6 Classificação dos perfis de escrita

Com base nas métricas quantitativas e na análise linguística qualitativa, os textos foram agrupados em cinco perfis cognitivos de escrita: cópia quase literal, alta fidelidade com fragilidade estrutural, reconstrução parcial, fragmentação severa e reescrita autoral. Essa classificação representa um continuum entre reprodução textual e produção autoral emergente, refletindo diferentes estágios do desenvolvimento da escrita infantil.

3.7 Análise comparativa entre escolas

A partir dessa categorização, realizou-se uma análise comparativa entre as duas instituições, considerando a frequência relativa de cada perfil de escrita. Essa abordagem possibilitou identificar diferenças nos padrões de apropriação textual entre os grupos, bem como níveis distintos de estabilidade ortográfica, estrutural e semântica. Esses resultados fornecem subsídios para compreender como fatores contextuais podem influenciar o desenvolvimento da escrita e reforçam o potencial da abordagem proposta como ferramenta de apoio à avaliação educacional.

3.8 Pipeline Computacional

O processamento dos dados manuscritos seguiu um pipeline computacional estruturado em etapas sequenciais, desde a digitalização das produções até o cálculo das métricas de avaliação. A Figura 8 apresenta uma visão geral do fluxo de processamento, detalhado a seguir:

(i) Digitalização: As produções manuscritas foram digitalizadas em formato de imagem (JPEG) com resolução de 300 DPI, garantindo qualidade suficiente para o processamento óptico. Todas as imagens foram armazenadas em repositório estruturado, preservando a identificação do participante, série escolar e instituição.

(ii) Pré-processamento: As imagens foram submetidas a etapas de normalização, incluindo conversão para escala de cinza, ajuste de contraste e correção de inclinação (deskewing), quando necessário. Não foram aplicadas técnicas de binarização agressivas, de forma a preservar as características caligráficas originais.

(iii) Transcrição de Referência: Uma transcrição manual de referência foi produzida por especialista, revisando cada produção caractere por caractere. Esta transcrição serviu como ground truth para o cálculo das métricas de erro.

(iv) Cálculo de Métricas: Foram calculadas as métricas CER (Character Error Rate), WER (Word Error Rate) e PMR (Perfect Match Rate) através da comparação automática entre cada transcrição OCR e a transcrição de referência, utilizando algoritmo de distância de edição (Levenshtein).

(v) Anotação Linguística: As divergências entre transcrições foram categorizadas manualmente em classes de erro (ortográficas, estruturais, de pontuação), permitindo análise qualitativa dos padrões de escrita.

4. Resultados do Experimento 1: Fábula O Urso e as Abelhas

4.1 Visão geral dos resultados

Nesta seção são apresentados e discutidos os resultados obtidos a partir da comparação entre o texto original da fábula *O Urso e as Abelhas* e as transcrições produzidas pelos alunos das duas instituições participantes. As divergências foram identificadas por inspeção manual e classificadas em categorias de erro, permitindo uma análise quantitativa, por meio da frequência e média de ocorrências, e qualitativa, a partir da identificação de padrões recorrentes de escrita e segmentação textual.

Para fins de análise, não foram contabilizados como erro os casos de separação silábica ao final da linha, por se tratar de um fenômeno gráfico de formatação, comum em textos manuscritos escolares, e que não caracteriza necessariamente uma transcrição inadequada.

De modo geral, observou-se ampla variabilidade entre os participantes, com produções que variaram desde cópias praticamente literais até textos fragmentados e, em alguns casos, reescritas autorais. Esse comportamento evidencia um continuum entre reprodução textual, reconstrução parcial e autoria emergente, refletindo diferentes estágios do desenvolvimento da escrita infantil.

4.2 Quantidade média de erros por aluno

A análise quantitativa revelou diferenças expressivas na quantidade média de erros entre os estudantes, bem como entre as duas escolas investigadas. Enquanto parte dos alunos apresentou baixo número de divergências, restritas principalmente a aspectos formais, como pontuação e uso de maiúsculas, outro grupo exibiu elevado volume de erros, incluindo omissões de palavras, substituições lexicais e fragmentação estrutural do texto.

A distribuição dos erros não foi homogênea, indicando a coexistência, em uma mesma turma, de níveis bastante distintos de proficiência escrita. Essa heterogeneidade foi mais pronunciada na Escola Estadual, onde se observou simultaneamente a presença de cópias de alta fidelidade e produções com perdas narrativas significativas.

4.3 Tipos de erros mais frequentes

Entre os tipos de erro mais recorrentes destacam-se alterações ortográficas, falhas de pontuação, problemas de espaçamento entre palavras e trocas de letras. Também foram observadas omissões de termos funcionais, como artigos e preposições, além de substituições verbais que, em alguns casos, alteraram o tempo ou o modo da frase.

Adicionalmente, verificou-se a ocorrência de erros que afetam a macroestrutura do texto, como a ausência da Moral da fábula ou a perda de trechos narrativos inteiros, especialmente em produções de menor fidelidade. Esses achados reforçam a importância de analisar não apenas erros locais, em nível de palavra, mas também aspectos globais da organização textual.

4.4 Problemas de espaçamento

Os problemas de espaçamento constituíram uma categoria relevante, manifestando-se tanto como junção indevida de palavras quanto como separação excessiva de sílabas ou caracteres. Em diversos casos, o espaçamento irregular dificultou a segmentação automática do texto pelo OCR, evidenciando a interação entre características da escrita manuscrita e o desempenho das ferramentas computacionais.

Do ponto de vista pedagógico, tais padrões sugerem instabilidade na consciência lexical e na delimitação de unidades linguísticas, aspectos frequentemente associados às fases iniciais da alfabetização.

4.5 Erros por categoria

4.5.1 Omissão de letras

As omissões ocorreram principalmente em sílabas átonas e em finais de palavras, afetando a integridade lexical e, ocasionalmente, comprometendo a compreensão do enunciado. Esse tipo de erro foi mais frequente em textos classificados como de reconstrução parcial ou fragmentação severa.

4.5.2 Troca de letras

As trocas envolveram, em sua maioria, fonemas próximos, indicando influência da percepção fonológica no processo de escrita. Exemplos incluem substituições vocálicas e consonantais que preservam parcialmente a sonoridade da palavra original, mas alteram sua grafia convencional.

4.5.3 Junção de palavras

A junção indevida de palavras foi observada sobretudo em sequências de alta frequência funcional, como artigos seguidos de substantivos. Esse fenômeno impacta diretamente métricas automáticas, como WER, e reflete dificuldades na segmentação textual.

4.5.4 Erros de pontuação

Os erros de pontuação incluíram ausência de pontos finais, uso inadequado de aspas e omissão de vírgulas explicativas. Embora muitas dessas ocorrências não comprometam gravemente o significado global, elas indicam domínio incipiente das convenções da escrita formal.

Resultados do texto *O Urso e as Abelhas*

A análise específica da fábula evidenciou que a maioria dos alunos conseguiu preservar a sequência básica dos eventos narrativos. No entanto, em produções de menor fidelidade, foram observadas alterações semânticas relevantes, incluindo inversões na Moral da história e reconstruções criativas do enredo. Esses casos indicam transição da simples cópia para processos de reinterpretação e autoria emergente.

4.6 Classificação dos perfis cognitivos da escrita :

A partir da análise combinada das medidas quantitativas (CER, WER e PMR) e da anotação linguística qualitativa, as produções foram agrupadas em quatro níveis principais de fidelidade textual, refletindo diferentes estratégias cognitivas empregadas pelos alunos durante a escrita.

O primeiro nível corresponde a produções de fidelidade muito alta, caracterizadas por cópias praticamente literais do texto-base, com divergências restritas a aspectos tipográficos, como uso de aspas ou pontuação final incompleta. Esses textos apresentaram PMR elevado e taxas mínimas de CER e WER, indicando reprodução textual precisa e baixo envolvimento reconstrutivo.

O segundo nível abrange produções de alta fidelidade com microdesvios linguísticos, incluindo perdas pontuais de acentuação, trocas verbais isoladas, omissões de vírgulas explicativas e pequenos erros de artigo ou plural. Nesses casos, a macroestrutura narrativa foi preservada, sugerindo escrita estável com interferência fonológica leve.

O terceiro nível corresponde a produções de fidelidade média, marcadas por substituições lexicais, alterações morfossintáticas, fragmentação estrutural e, em alguns casos, inversões semânticas da Moral da fábula. Esses textos apresentaram valores intermediários de WER e evidenciam um deslocamento da simples cópia para processos de reconstrução cognitiva do conteúdo.

Por fim, o quarto nível reúne produções de baixa fidelidade, caracterizadas por substituições fonéticas massivas, palavras sem correspondência lexical, omissão completa da Moral, ausência sistemática de pontuação e ruptura da continuidade narrativa. Esses textos exibiram PMR baixo e taxas elevadas de CER e WER, refletindo padrões típicos de estágios iniciais de alfabetização, com forte instabilidade grafêmica e semântica.

De modo geral, os resultados evidenciam um continuum progressivo entre cópia literal e reconstrução textual, no qual a maioria das divergências decorre de processos cognitivos próprios do desenvolvimento da escrita infantil, sendo mínima a interferência atribuível ao OCR.

4.7 Comparação institucional:

A análise comparativa entre as duas escolas revelou perfis distintos de apropriação da escrita, apesar do uso do mesmo instrumento textual. Na escola particular, observou-se predominância dos níveis de fidelidade muito alta e alta fidelidade com microdesvios linguísticos, com a maioria dos textos completos e erros concentrados em aspectos formais, como acentuação, pontuação e flexão verbal. Casos de reconstrução parcial foram menos frequentes, e praticamente não se observaram reescritas autorais.

Em contraste, a Escola Estadual apresentou maior heterogeneidade intra-grupo, com coexistência de cópias quase literais, fragmentações severas e produções autorais. Além disso, verificou-se maior incidência de omissões de blocos narrativos, quebras sintáticas, trocas fonológicas e abandono do texto-base, indicando uma transição mais dispersa entre reprodução textual, reconstrução parcial e autoria emergente.

A Tabela 1 sintetiza essa distribuição, evidenciando que, enquanto a escola particular concentrou-se majoritariamente nos níveis 1 e 2, a escola estadual apresentou maior frequência nos níveis 3 e 4, além de registros de reescrita criativa

Tabela 1: Distribuição dos perfis cognitivos de escrita por escola (Fábula *O Urso e as Abelhas*)

Categoria cognitiva	Descrição resumida	Escola Particular (São Francisco, ≈18)	Escola Estadual (Henrique Dias, 18)
---------------------	--------------------	--	-------------------------------------

Nível 1 – Cópia quase literal	Texto praticamente idêntico ao original; erros restritos a aspectos tipográficos (aspas, ponto final)	≈ 6	3
Nível 2 – Alta fidelidade com microdesvios linguísticos	Preservação da narrativa com perdas pontuais de acento, trocas verbais isoladas e pequenos erros gramaticais	≈ 5	3
Nível 3 – Reconstrução parcial	Substituições lexicais, alterações morfossintáticas, omissões de vírgulas e início de reconstrução semântica	≈ 4	5
Nível 4 – Fragmentação severa / omissão massiva	Frases quebradas, grandes trechos ausentes, Moral truncada ou perdida	≈ 2	5
Nível 5 – Reescrita autoral	Abandono do texto-base, nova narrativa e nova Moral	0–1	2

A Tabela 1 evidencia diferenças marcantes entre as instituições. Enquanto a escola particular concentra a maior parte das produções nos níveis de cópia quase literal e alta fidelidade, com erros predominantemente formais, a escola estadual apresenta maior dispersão entre as categorias, incluindo fragmentações severas e reescritas autorais. Esses resultados indicam uma heterogeneidade intra-grupo mais acentuada na escola estadual e reforçam a existência de um continuum entre reprodução textual, reconstrução parcial e autoria emergente.

4.8 Discussão: implicações dos padrões encontrados

Os resultados apontam que os erros mais recorrentes não se restringem a grafias incorretas, mas incluem falhas de segmentação e de convenções da escrita textual. Isso sugere que, ao analisar manuscritos infantis, torna-se necessário distinguir entre erros ortográficos em nível de palavra, erros estruturais relacionados à pontuação e organização narrativa, e erros potencialmente influenciados pela legibilidade e qualidade do manuscrito.

A coexistência de cópias literais, reconstruções parciais e reescritas autorais evidencia um continuum cognitivo da escrita, no qual diferentes estratégias são mobilizadas pelas crianças ao lidar com um texto-base. Tal comportamento reforça a importância de abordagens que considerem tanto métricas quantitativas, como CER, WER e PMR, quanto análises qualitativas da estrutura textual.

Consequentemente, recomenda-se que futuras análises tratem separadamente erros que alteram o significado, como substituições lexicais críticas e omissões narrativas, e erros predominantemente formais, como pontuação e espaçamento, dado que seus impactos são distintos do ponto de vista pedagógico e computacional. Enquanto os primeiros podem indicar dificuldades mais profundas de compreensão ou produção textual, os segundos tendem a refletir estágios iniciais de apropriação das normas da escrita.

Por fim, a comparação entre escolas evidencia que o mesmo instrumento textual pode revelar padrões bastante distintos de desenvolvimento da escrita, reforçando o potencial da abordagem proposta como ferramenta de apoio ao diagnóstico educacional baseado em textos contínuos manuscritos.

5. Resultados do Experimento 2: Fábula O Leão e o Javali

Neste tópico, são apresentados e discutidos os resultados obtidos a partir da comparação entre o texto original da fábula **O Leão e o Javali** e as transcrições produzidas pelos alunos do 3º e 5º ano. As divergências foram identificadas por inspeção manual e classificadas em categorias de erro, permitindo uma análise qualitativa e técnica a partir da identificação de padrões recorrentes de ortografia, pontuação e segmentação textual.

Observou-se que este fenômeno ocorreu de forma frequente e correta em diversas produções, como nas palavras "anima-is", "respon-deu" e "mai-ores", caracterizando-se como uma norma gráfica de formatação e não como uma falha de transcrição.

De modo geral, observou-se uma variabilidade significativa entre os participantes, conforme detalhado abaixo:

- Cópias Literais e Alta Precisão: Alguns alunos apresentaram produções com baixíssimo índice de erro (CER), demonstrando domínio da norma culta e fidelidade quase absoluta ao texto de referência.
- Reconstrução e Variação Linguística: Outras produções revelaram um processo de "reescrita mental", onde o aluno substitui termos por sinônimos ou variações verbais, como trocar "Olha" por "Olhe" ou "esperam" por "esperando", o que indica que a leitura para a cópia foi feita através do processamento do sentido e não apenas da decodificação mecânica.
- Textos Fragmentados e Desvios Ortográficos: Em casos específicos, como nas imagens de menor legibilidade, notou-se um padrão de escrita com simplificações fonéticas ("lião" por "leão", "inecia" por "iniciar") e falhas de segmentação ("adiscutir"), evidenciando alunos em estágios de consolidação da base alfabética e ortográfica.

Em suma, a variabilidade nos índices de legibilidade e acerto ortográfico reflete a coexistência de diferentes momentos do desenvolvimento da escrita dentro do mesmo grupo escolar. A distinção entre o **erro do aluno** (pedagógico) e o **ruído de OCR** (técnico) é crucial para compreender que as discrepâncias encontradas são, na verdade, marcas identitárias do percurso de aprendizagem de cada estudante.

Figura 2: Cópia de uma aluna do 5º ano escola particular (Fundamental 1/Anos iniciais)

O urso e as abelhas

"Um urso topou com uma árvore caída que queria de depósito de mel para um enxame de abelhas. Ele começou a farejar o tronco quando uma das abelhas voltou do campo de troncos. Ao perceber que o urso planejava roubar o mel, a abelha ferrou o animal e saiu.

O urso, chateado com a abelha, começou a arrastar o tronco com as garras. Então, o enxame inteiro de abelhas começou a sair do tronco atrás dele. O urso fugiu a toda velocidade e só se salvou porque mergulhou de cabeça num lago"

Moral: Aprenda a observar os sinais e não queria aquilo que não é seu

legendas dos erros: Primeira caixa vermelha identifica a troca de letra maiúscula por minúscula, a segunda caixa vermelha identifica a troca de letras(e/o , caixa azul identifica erro de divisão silábica incorreta e a última caixa vermelha identifica a troca da palavra queira por queria

Comparação entre séries/idades

A comparação entre as produções textuais de alunos do 3º e do 5º ano revela um amadurecimento que vai muito além da estética: trata-se de um salto na coordenação motora e na compreensão das regras silenciosas da escrita. Enquanto no 3º ano a criança ainda vive uma fase de transição em que o papel tenta acompanhar o ritmo da fala e as descobertas do som. O aluno do 5º ano já demonstra um domínio mais seguro da norma culta e da organização do texto no papel.

Ao analisarmos os manuscritos do 3º ano, percebemos o que os educadores chamam de autoria emergente. Nesse estágio, o ato de copiar não é passivo; o aluno projeta no papel suas próprias hipóteses sobre como as palavras deveriam ser escritas. Isso explica erros de omissão, como em "discuti", onde o "r" final é deixado de lado simplesmente porque não costuma ser pronunciado no dia a dia. A pontuação ainda aparece de forma tímida ou irregular, e o traçado das letras, com pressões variadas no lápis, denuncia alguém que ainda está aprendendo a dominar o próprio gesto.

Figura 3: Cópia de uma aluna do 3º ano (Fundamental 1/ Anos iniciais)

O leão e o javali

Um dia muito quente, um leão e um javali chegaram juntos a um poço. Os animais estavam com muita sede e começaram a discutir a ra ver quem beberia primeiro.

Nenhum cediu a verj ao outro. Teriam iniciar uma briga, mas o leão olhou para cima e viu vários urubus voando.

- Ahé lá! - disse o leão.

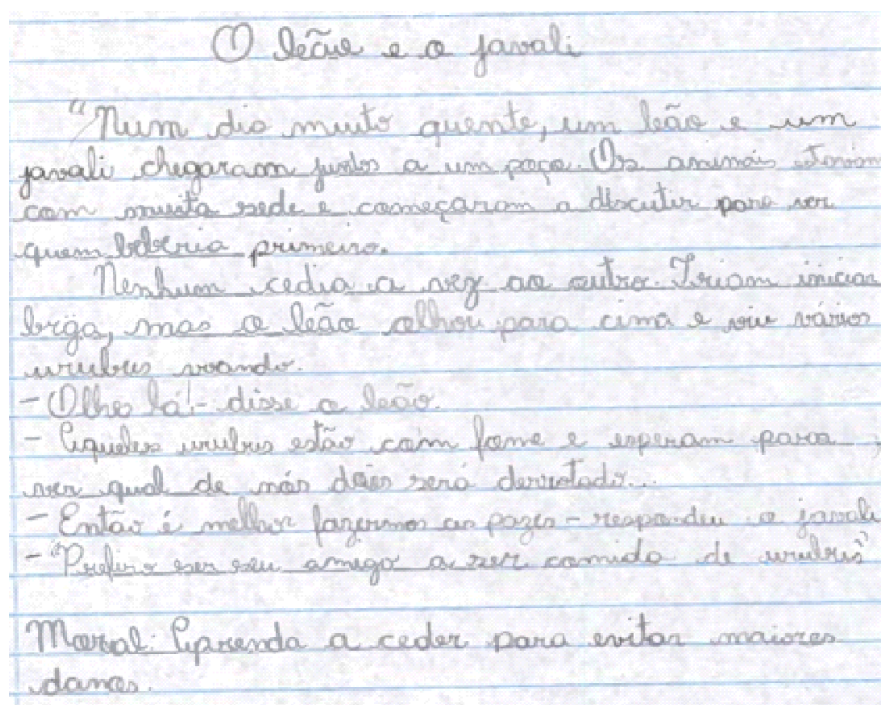
- Aquelos urubus estão com fome e estão para ver qual de nós dois será devorado...

- Então é melhor fazermos um acordo - restringem o javali

- Prefiro ser seu amigo a ser comida de urubus."

Em contrapartida, o texto do 5º ano nos apresenta uma **reprodução muito mais fiel e consciente**. A caligrafia cursiva ganha fluidez e proporção, sinalizando que a criança já não precisa concentrar tanto esforço apenas em "como desenhar a letra", podendo focar na precisão do conteúdo. A ortografia torna-se estável e até os erros mudam de natureza: se o aluno escreve "exame" em vez de "enxame", ele não está mais apenas tentando reproduzir um som, mas sim testando regras ortográficas mais complexas. A pontuação deixa de ser um desafio e passa a ser uma aliada, organizando o ritmo da história e dando clareza ao que é narrado.

Figura 4: Cópia aluna 5º ano



Esse progresso mostra que, entre os 8 e os 11 anos, ocorre uma mudança fundamental: a criança deixa de "escrever como fala" para começar a "escrever como se lê". A segmentação das palavras fica precisa, os parágrafos ganham seu devido espaço e a inclusão de conectivos torna a leitura mais agradável. O que antes era um esforço de decodificação no 3º ano, transforma-se, no 5º ano, em um exercício de rigor e autonomia, preparando o estudante para os desafios mais complexos que virão nas etapas seguintes da vida escolar.

O segundo experimento foi conduzido com uma amostra distinta de estudantes, utilizando a fábula O Leão e o Javali como instrumento textual. Este experimento complementa o primeiro, permitindo a comparação de resultados entre diferentes textos e grupos de participantes, além de possibilitar a análise da consistência metodológica entre as duas aplicações.

Comparação manual vs automático (OCR)

A análise da escrita manuscrita infantil pode ser conduzida tanto por procedimentos manuais, baseados na inspeção humana especializada, quanto por abordagens automatizadas, fundamentadas em sistemas de reconhecimento óptico de caracteres (OCR). Cada uma dessas estratégias apresenta vantagens e limitações específicas, especialmente quando aplicada a textos contínuos produzidos por crianças em fase de alfabetização.

A correção e análise manual constituem o método tradicionalmente adotado no contexto educacional. Esse procedimento permite uma interpretação contextualizada dos desvios encontrados, possibilitando distinguir erros ortográficos, estruturais e semânticos que decorrem de hipóteses cognitivas próprias do desenvolvimento da escrita infantil. Além disso, a análise humana é capaz de considerar aspectos discursivos e pedagógicos que extrapolam a dimensão gráfica do texto, como coerência narrativa, intenção comunicativa e autoria emergente. No entanto, esse tipo de análise apresenta limitações importantes em cenários de larga escala, uma vez que demanda elevado investimento de tempo, esforço cognitivo e padronização entre avaliadores, tornando-se pouco viável em turmas numerosas ou em avaliações sistemáticas frequentes.

Por outro lado, a abordagem automatizada por meio de OCR possibilita a transcrição rápida e sistemática de grandes volumes de textos manuscritos, viabilizando análises quantitativas objetivas baseadas em métricas como Character Error Rate (CER), Word Error Rate (WER) e Perfect Match Rate (PMR). Essas métricas permitem comparar produções individuais e coletivas de forma padronizada, além de facilitar a identificação de padrões recorrentes de erro ao longo de conjuntos extensos de dados. Contudo, sistemas de OCR apresentam limitações relevantes no contexto da escrita infantil, uma vez que variações caligráficas, instabilidade no espaçamento, fragmentação estrutural e traços incompletos podem gerar ruídos na transcrição automática, introduzindo erros que não correspondem necessariamente a falhas cognitivas do aluno.

Os resultados deste estudo indicam que, embora o OCR seja eficaz para apoiar a análise quantitativa e a organização dos dados, a interpretação exclusiva das saídas automatizadas pode levar a leituras equivocadas sobre o desempenho do estudante, sobretudo quando erros técnicos são confundidos com erros pedagógicos. Em textos com baixa legibilidade ou forte fragmentação estrutural, observou-se que parte das divergências identificadas automaticamente decorre de limitações do reconhecimento computacional, e não de dificuldades ortográficas ou semânticas intrínsecas à produção infantil.

Nesse sentido, a comparação entre os dois modos de análise evidencia que a abordagem manual oferece maior precisão interpretativa em nível cognitivo e pedagógico, enquanto a abordagem automatizada apresenta vantagens significativas em termos de escalabilidade, padronização e reprodutibilidade dos resultados. A integração entre ambas, conforme proposta neste trabalho, permite mitigar as limitações individuais de cada método: o OCR atua como ferramenta de apoio à triagem, quantificação e organização dos dados, enquanto a análise manual assegura a validação qualitativa e a correta interpretação dos padrões observados.

Assim, em vez de substituir a avaliação humana, os sistemas de OCR mostram-se mais adequados como instrumentos complementares ao processo de análise da escrita manuscrita infantil. Essa combinação possibilita ampliar o alcance das avaliações educacionais, reduzir o custo operacional da correção manual em larga escala e, ao mesmo tempo, preservar a sensibilidade pedagógica necessária para compreender os processos cognitivos subjacentes ao desenvolvimento da escrita.

Avaliação comparativa de sistemas de reconhecimento óptico de caracteres

Critérios de Avaliação

Os sistemas de OCR foram comparados com base nos seguintes critérios:

- (i) taxa de erro por caractere (CER);
- (ii) taxa de erro por palavra (WER);
- (iii) taxa de equivalência perfeita (PMR);
- (iv) número de quebras de linha inconsistentes;
- (v) ocorrência de inserção de texto inexistente no documento original;
- (vi) mistura indevida de idiomas.

As versões específicas das ferramentas utilizadas foram: Google Vision API (v1), Microsoft Azure OCR (v3.2), ChatGPT-5.2, Google Gemini 3 (modo Raciocínio) e Grok 3 (modo Especialista). As referências oficiais das APIs encontram-se disponíveis em: Google Cloud Vision (<https://cloud.google.com/vision>), Microsoft Azure Computer Vision (<https://azure.microsoft.com/services/cognitive-services/computer-vision/>), OpenAI ChatGPT (<https://openai.com/chatgpt>), Google Gemini (<https://gemini.google.com>) e xAI Grok (<https://x.ai>).

Tabela 2: Comparação entre sistemas de OCR (valores médios gerais) e desvio padrão

OCR	CER (%) + (±DP%)	WER (%) + (±DP%)	PMR (%) + (±DP%)	Nível de Inserções indevidas	Média de quebras de linha incorretas
Google Vision	43,5% (±44,55%)	54% (±50,91%)	40% (±49,50%)	Alto	20
Azure OCR	18% (±14,14%)	22,5% (±17,68%)	55% (±28,28%)	Muito baixo	0
ChatGPT	3,5% (±0,71%)	17,5% (±17,68%)	70% (±21,21%)	Alto	6
Gemini	7% (±4,24%)	11% (±5,66%)	81,5% (±9,19%)	Inexistente	0
Grok	54,5% (±57,28%)	66% (±59,40%)	38,5% (±51,62%)	Muito alto	15

Para os modelos baseados em LLM, foi utilizado o seguinte prompt:

You are an Optical Character Recognition (OCR) system specialized in transcribing handwritten children's texts from images. Your task is to extract the text exactly as it appears, character by character. Important rules: • Do not correct spelling, grammar, or punctuation. • Do not interpret or guess unclear words. • Preserve line breaks, capitalization, and any errors exactly as written. • Output only the transcription, nothing else.

O mesmo prompt foi utilizado para todos os modelos baseados em linguagem a fim de garantir isonomia na comparação.

Figura 5: Cópia de um aluno do 3º ano de escola particular utilizada para comparação (menor legibilidade):

O lão e o foali

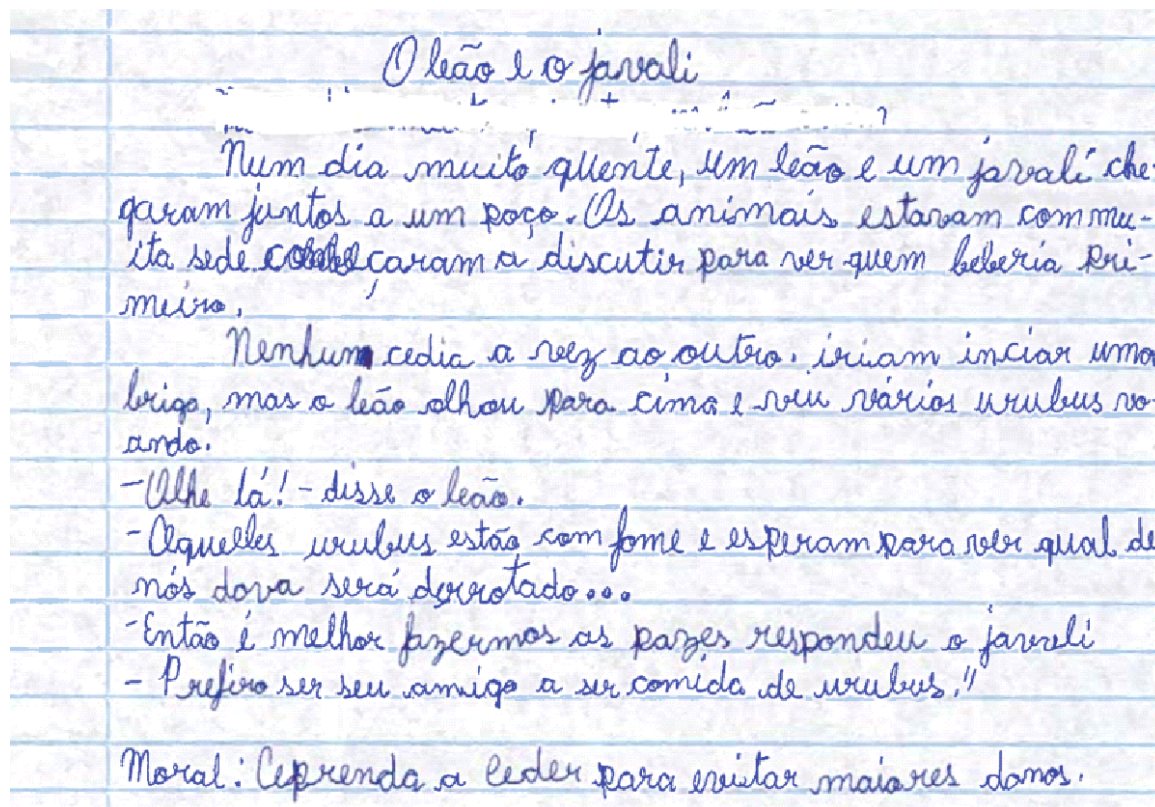
Num dia muito quente, um lão e um foali chegaram juntos a um peda. Os animais estavam com muita sede e começaram a discutir para ver quem bebia primeiro.

N mbum dia a vez de oter. Triam inicio uma briga, mas o lão abto para cima e sua varios urubus urubus.

- Lá - disse o lão.
- Pequitos urubus urubus com fome e esperavam para ver qual de nos do volta deotado...
- Então é melhor fazerem os potes - respondeu o foali.
- Prefiro ser ser amigo a ser comida de urubus!

metá: Espanta a ceder para evitar miolos domos.

Figura 6: Cópia de um aluno do 3º ano de escola particular utilizada para comparação (maior legibilidade):



Google Vision API - O serviço apresentou um desempenho satisfatório, apresentando um CER de 12%, WER de 18% e PMR de 75%, 7 quebras de linha inconsistentes e quase nenhuma inserção de palavras inexistentes ou mistura de idiomas, em uma redação com caligrafia mais nítida e regular. No entanto, mesmo quando a legibilidade ainda era considerada adequada, observou-se uma queda significativa na acurácia conforme a qualidade da escrita diminuía chegando a um CER de 75%, WER de 90% e PMR de 5%, 15 quebras de linha inconsistentes muitas inserções incorretas e mistura de idiomas bem presente em uma versão menos legível do mesmo texto. Em suma, verificou-se forte inconsistência na segmentação das linhas, bem como a introdução indevida de trechos em outros idiomas em áreas com menor clareza visual.

Microsoft Azure OCR 3.2 - O desempenho geral foi positivo. Em comparação ao Google Vision, o Azure demonstrou maior estabilidade, mantendo boa capacidade de reconhecimento mesmo quando a caligrafia se afastava de padrões mais formais, ao transcrever o mesmo texto menos legível obteve os seguintes resultados: CER de 28%, WER de 35% e PMR de 35%, e no texto mais legível chegando a resultados de: CER de 8%, WER de 10% e PMR de 75%. Destaca-se também a ausência quase total de quebras de linha injustificadas e mudança de idioma, também praticamente não há inserções indevidas, indicando maior coerência estrutural na transcrição.

ChatGPT-5.2 - O modelo apresentou elevada qualidade de transcrição, mantendo consistência ao lidar com os diferentes estilos de caligrafia, no texto menos legível obteve CER de 4%, WER de 30% e PMR de 55% com inserções e sem troca de idiomas. Esses números chamam a atenção pois o contraste do baixo erro de caracteres junto a quantidade de palavras diferente do original e a adição de trechos que não estão presentes no texto indicam que possivelmente ele está corrigindo o texto ao invés de seguir adequadamente o prompt. No texto com a caligrafia mais legível se obteve CER de 3%, WER de 5% e PMR de 85% sem inserções e sem troca de idioma, o que sugere um ótimo desempenho, o contraste entre os diferentes cenários indica que o modelo tem dificuldade em seguir o prompt ao analisar textos com a caligrafia menos legível.

Google Gemini 3 (Raciocínio) - Ao utilizar o Gemini como ferramenta de OCR direto, notou-se uma boa precisão ao seguir o prompt, realizando o OCR de maneira mais bruta como desejado, e não corrigindo palavras que realmente estão escrito errado no texto nem inserindo pedaços de texto inexistentes no texto original, até mesmo no texto menos legível obtendo os seguintes resultados: CER de 10%, WER de 15% e PMR de 75% sem inserções de palavras e sem troca de idiomas. No texto mais legível foi obtido o seguinte resultado: CER de 4%, WER de 7% e PMR de 88%, também sem inserção de palavras nem troca de idioma.

Grok 3 (Especialista) - Foi a pior das LLMs testada no navegador, não podendo ser considerada uma ferramenta de OCR válida, ao lidar com a redação menos legível, não apenas não conseguiu seguir o prompt como fez diversas inserções em nível de palavra, frase e narrativa, mudando completamente o contexto e adicionando elementos que não estavam nem próximos do texto original, na redação de menor legibilidade, obtendo os seguintes resultados: CER de 90%, WER de 98% e PMR de 2%. Em contrapartida, lidou bem com a redação com a caligrafia mais legível, obtendo bons resultados: CER de 9%, WER de 14% e PMR de 75% com pouca inserção e sem troca de idioma. Em suma, mesmo com o bom resultado na redação com melhor caligrafia, ainda não deve ser considerada uma opção pois o seu desempenho ao lidar com uma caligrafia um pouco pior foi muito insatisfatório.

Figura 7: Transcrição do Grok ao ler a redação com a caligrafia menos legível

O Leão e o Rato.

Num dia muito quente, um Leão, um Javali Dragão
juntos a um Rio. Os animais bebiam com muita sede e
conversavam animados. Pólo veio quem. Tinha primeiro.

Nenhum queria a vez do outro. Todos insistiam uma
solução. Mas o Leão disse: Para cima e os seus vorazes unhas aran
tr.

- Da! - gritou o Leão.
- Líquidos vindos de cima. Toda a água tem uma origem. Ota que qual é
mão de nossa dentada.....
- Então é melhor fazermos os poços - sugeriu um deles o Javali.
- Professor não seja amigo a só em comida de vultures!"

Não. e. pergunta a Ceder, pois autor mistério dome.

6. Ameaças à validade

Embora o estudo apresenta resultados relevantes para a análise da escrita infantil por meio de técnicas de OCR e métodos de análise estrutural, algumas ameaças à validade devem ser consideradas para contextualizar a interpretação dos achados.

6.1 Validade Interna

A validade interna refere-se ao grau em que os resultados obtidos podem ser atribuídos às técnicas e procedimentos adotados, e não a fatores externos não controlados. Neste estudo, variáveis como variações naturais na caligrafia infantil (pressão da escrita, velocidade, fadiga ou motivação no momento da atividade) podem ter impactado a quantidade e o tipo de erros observados. Além disso, pequenas diferenças no ambiente de coleta, como iluminação, ruídos ou postura das crianças, também podem ter influenciado a qualidade da escrita ou gerado ruídos na digitalização. A etapa de digitalização, apesar de padronizada, está sujeita a imperfeições no escaneamento (sombras, desalinhamentos, papel amassado), que podem afetar o desempenho do OCR e resultar em erros não atribuíveis à escrita infantil em si.

6.2 Validade Externa

A validade externa diz respeito à capacidade de generalização dos resultados para outros contextos. Neste estudo, a amostra é restrita a crianças do 3º, 4º e 5º ano de escolas específicas, o que limita a extrapolação dos achados para outras regiões, redes de ensino ou faixas etárias. Além disso, o estudo utiliza exclusivamente fábulas como gênero textual, o que pode influenciar os tipos de erros observados, dada a presença de elementos narrativos e estruturas específicas desse gênero. A escrita espontânea em outros contextos — como bilhetes, produções autorais ou atividades de avaliação — pode apresentar padrões de erro distintos, reduzindo a generalização dos resultados.

6.3 Validade de Construto

A validade de construto refere-se ao quanto às métricas e instrumentos utilizados representam adequadamente os conceitos teóricos investigados. Embora métricas como CER, WER e categorias de erros ortográficos permitam uma análise objetiva, nem todos os aspectos da qualidade da escrita infantil são capturados por elas. O OCR pode produzir falsos positivos ou falsos negativos, confundindo ruído com caracteres ou interpretando traços incompletos como omissões, o que pode alterar artificialmente a taxa de erros. Além disso, a análise estrutural, embora eficaz na identificação de problemas de espaçamento e alinhamento, pode não captar nuances pedagógicas relevantes, como coerência, intenção comunicativa ou complexidade sintática das produções.

6.4 Validade Estatística

A validade estatística diz respeito ao rigor das análises quantitativas realizadas. O tamanho da amostra e a distribuição desigual de participantes entre as séries podem impactar a comparabilidade dos resultados. Diferenças significativas no número de alunos por ano escolar podem aumentar ou reduzir artificialmente a variabilidade interna de cada grupo. Além disso, a comparação entre métodos de OCR pode ser influenciada pelo tipo de fonte manuscrita ou pela presença de ruídos específicos em algumas digitalizações. A ausência de testes estatísticos complementares para verificar significância ou variabilidade pode limitar a força das inferências.

7. Considerações finais e trabalhos futuros

Este trabalho investigou a escrita manuscrita infantil em língua portuguesa a partir de uma perspectiva interdisciplinar, integrando métodos de reconhecimento óptico de caracteres (OCR), métricas quantitativas de erro e análise linguística qualitativa aplicada a textos narrativos contínuos. Diferentemente de abordagens tradicionais que se concentram em palavras isoladas ou escrita adulta, a proposta aqui apresentada analisou produções reais de alunos do Ensino Fundamental, preservando a complexidade gráfica, estrutural e cognitiva do processo de alfabetização.

Os resultados obtidos evidenciam que a escrita infantil não pode ser adequadamente avaliada apenas por métricas automáticas de acurácia. Embora medidas como CER, WER e PMR sejam eficazes para quantificar divergências em relação ao texto-base, grande parte dos desvios observados decorre de processos cognitivos legítimos do desenvolvimento da escrita, como reconstrução semântica, autoria emergente e instabilidade na segmentação lexical. Nesse sentido, a análise estrutural e a anotação linguística mostraram-se fundamentais para distinguir erros pedagógicos genuínos de ruídos introduzidos por limitações técnicas dos sistemas de OCR.

A comparação entre ferramentas comerciais de OCR e modelos baseados em LLM revelou diferenças significativas não apenas em termos de taxa de erro, mas também no comportamento frente a manuscritos de menor legibilidade. Enquanto alguns sistemas apresentaram maior estabilidade estrutural e fidelidade literal, outros demonstraram tendência à correção implícita do texto, inserções indevidas ou reconstruções semânticas, contrariando o objetivo de uma transcrição fiel. Esses achados reforçam que modelos de linguagem, embora apresentem elevado potencial, ainda não podem ser utilizados como OCR “bruto” sem mecanismos adicionais de controle, especialmente em contextos educacionais sensíveis.

A análise comparativa entre escolas e séries evidenciou a coexistência de múltiplos perfis cognitivos de escrita dentro de um mesmo grupo, confirmando a existência de um continuum entre cópia literal, reconstrução parcial e reescrita autoral. Essa heterogeneidade reforça a inadequação de abordagens avaliativas binárias (certo/errado) e aponta para a necessidade de ferramentas computacionais que apoiem diagnósticos mais finos e contextualizados do desenvolvimento da escrita infantil.

Do ponto de vista da Ciência da Computação, este trabalho contribui ao demonstrar que pipelines de OCR aplicados a manuscritos infantis exigem não apenas técnicas robustas de reconhecimento visual, mas também integração com camadas de análise linguística e validação humana. A proposta apresentada mostra-se viável como ferramenta de apoio à análise em larga escala, desde que utilizada de forma complementar à avaliação manual, e não como substituta.

Como trabalhos futuros, diversas extensões podem ser exploradas. Em primeiro lugar, pretende-se ampliar a base de dados, incorporando manuscritos de diferentes regiões, séries e gêneros textuais, como produções autorais livres e textos argumentativos, visando aumentar a validade externa dos resultados. Em segundo lugar, a inclusão de técnicas de aprendizado de máquina supervisionado ou semi-supervisionado pode permitir a classificação automática dos perfis cognitivos de escrita identificados neste estudo, reduzindo a dependência da anotação manual.

Outra linha promissora consiste na investigação de abordagens híbridas que combinem OCR tradicional com modelos de linguagem restritos por regras explícitas de fidelidade textual, mitigando problemas de inserção e correção indevida observados em LLMs. Além disso, a incorporação de métricas estruturais adicionais, como coerência narrativa, preservação da macroestrutura textual e análise de progressão semântica, pode enriquecer a avaliação automática da escrita infantil.

Por fim, destaca-se o potencial de aplicação prática do pipeline proposto como ferramenta de apoio ao diagnóstico educacional, permitindo que professores e instituições identifiquem padrões recorrentes de erro e acompanhem a evolução da escrita ao longo do tempo. Assim, este trabalho estabelece uma base computacional sólida para futuras pesquisas que busquem integrar tecnologias de reconhecimento de escrita com práticas pedagógicas sensíveis ao processo cognitivo de alfabetização.

REFERÊNCIAS BIBLIOGRÁFICAS

FERREIRO, Emilia; TEBEROSKY, Ana. *Psicogênese da língua escrita*. Porto Alegre: Artmed, 1985.

FERREIRO, Emilia. *Reflexões sobre alfabetização*. 24. ed. São Paulo: Cortez, 1985.

IBAIBARRIAGA, Gorka; ACHA, Joana; PEREA, Manuel. The impact of handwriting and typing practice in children's letter and word learning: Implications for literacy development. *Journal of Experimental Child Psychology*, v. 253, p. 106195, 2025.

LI, M.; LV, T.; CHEN, J.; CUI, L.; LU, Y.; FLORENCIO, D.; ZHANG, C.; LI, Z.; WEI, F. TrOCR: transformer-based optical character recognition with pre-trained models. In: *AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE*, 37., 2023. *Proceedings...* [S.l.]: AAAI Press, 2023. p. 13094–13102.

PENG, Meifen; ZHANG, Hongwei. New challenges and countermeasures of calligraphy education in colleges and universities in the new era. *Creative Education*, v. 13, n. 8, p. 2544-2552, 2022.

SMITH, R. An overview of the Tesseract OCR engine. In: *INTERNATIONAL CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION (ICDAR)*, 9., 2007. *Proceedings...* Washington, DC: IEEE Computer Society, 2007. p. 629–633. Disponível em: <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/33418.pdf>. Acesso em: dia mês ano.