

**Tipo de artículo:** Investigación

Escritura por géneros y progresión de competencias textuales en educación general básica media

*Genre-based writing and the progression of textual competence in basic education middle grades*

**Autores:**

**Nancy Cecilia Loarte Valle**

Institute for Executive Education, Santa Isabel-Ecuador, [nancyceloarte@gmail.com](mailto:nancyceloarte@gmail.com),  
<https://orcid.org/0009-0002-0865-9638>

**Carmen Nelly Loarte Valle**

Universidad Técnica Particular de Loja, Santa Isabel - Ecuador, [carmen.loarte@yahoo.com](mailto:carmen.loarte@yahoo.com),  
<https://orcid.org/0009-0005-0419-3609>

**Fernando Alexi Vinueza Echeverría**

UEF El Porvenir, Montecristi - Ecuador, [alexivinueza007@gmail.com](mailto:alexivinueza007@gmail.com), <https://orcid.org/0009-0009-0429-7001>

**Marcia Lucía Delgado Yépez**

Universidad Católica del Ecuador, Ibarra - Ecuador, [marcialdy\\_68@hotmail.com](mailto:marcialdy_68@hotmail.com),  
<https://orcid.org/0009-0002-6815-7939>

**Corresponding Author:** *Fernando Alexi Vinueza Echeverría*, [alexivinueza007@gmail.com](mailto:alexivinueza007@gmail.com)

**Reception:** 27-octubre-2025    **Acceptance:** 28- noviembre -2025    **Publication:** 06- diciembre -2025

**How to cite this article:**

Vinueza Echeverría, F. A., Loarte Valle, N. C., Loarte Valle, C. N., & Delgado Yépez, M. L. (2025). Escritura por géneros y progresión de competencias textuales en EGB. *Horizonte Cientifico International Journal*, 3(2), 1-18.  
<https://doi.org/10.64747/q4ewz489>

## RESUMEN

Este estudio examinó la progresión de las competencias textuales por género entre 5.º y 7.º de Educación General Básica (EGB) en contextos contrastantes del Ecuador: Guayas (Costa, mayor urbanidad) y Chimborazo/Cotopaxi (Sierra, mayor componente rural e indígena). Se implementó un diseño mixto explicativo-secuencial (QUAN→qual) que combinó: (i) levantamiento primario de producciones escritas auténticas en cuatro géneros escolares (narración, noticia, descripción, instructivo) calificadas con rúbricas analíticas ( $\alpha/\kappa \geq 0.70$ ; invarianza métrica por provincia), (ii) extracción de métricas NLP (densidad léxica, longitud de T-unidad, cohesión referencial, conectores), y (iii) vinculación con datos abiertos (SEST-INEVAL para 7.º EGB; AMIE; capas H3/OSM/WorldPop/VIIRS). Los análisis incluyeron modelos multinivel (estudiante–sección–centro), estimación de índices de progresión por género (IPG) y correlaciones centro-nivel con SEST. Los resultados muestran que la progresión 5.º→7.º es sustantiva y diferenciada por género, con gradientes mayores en noticia e instructivo. Guayas exhibe ventaja relativa en noticia (IPG y pendiente de grado superiores), asociada a mejoras en pirámide invertida, lead y uso de citas; en Sierra destacan fortalezas cualitativas en adecuación y registro. Las métricas NLP median parcialmente el efecto del grado ( $\beta_{\text{indirecto}} \approx 0.06$ ), aportando sensibilidad a la complejidad textual. A nivel de centro, noticia (7.º) predice con mayor peso SEST Lengua-Literatura ( $\beta \approx 0.31$ ), seguida de instructivo ( $\beta \approx 0.24$ ). La robustez se confirmó con ponderaciones de diseño, imputación múltiple y especificaciones alternativas. Se concluye que la enseñanza por géneros, cuando visibiliza la macroestructura y cohesión y se apoya en textos mentores y retroalimentación formativa, acelera la progresión escritural. Se recomiendan secuencias híbridas (deconstrucción→escritura conjunta→independiente), foco en noticia/instructivo, y kits de bajo costo (plantillas y rúbricas abiertas) con prioridad para contextos rurales y EIB. El paquete de datos abiertos (microdatos anonimizados, scripts y metadatos JSON-LD) habilita replicación, auditoría y escalamiento responsable.

**Palabras clave:** escritura por géneros; progresión textual; métricas NLP; evaluación a gran escala (SEST); EGB media

## ABSTRACT

This study examined the genre-based progression of writing competence from 5th to 7th grades of Ecuador's Basic Education across contrasting settings: Guayas (coastal, more urban) and Chimborazo/Cotopaxi (Andean, more rural/indigenous). We employed an explanatory sequential mixed-methods design (QUAN→qual) combining: (i) primary sampling of authentic student texts in four school genres (narrative, news report, description, procedural/instruction), scored with analytic rubrics ( $\alpha/\kappa \geq 0.70$ ; metric invariance across provinces); (ii) extraction of NLP metrics (lexical density, T-unit length, referential cohesion, connectives); and (iii) linkage with open data (SEST-INEVAL grade-7 scores; AMIE; H3/OSM/WorldPop/VIIRS). Analyses included multilevel models (student–section–school), Genre Progression Index (GPI) estimates, and school-level correlations with SEST. Findings indicate a substantive 5th→7th progression that varies by genre, with steeper gradients in news and procedural writing. Guayas shows a relative advantage in news (higher GPI and grade slope), tied to improvements in inverted-pyramid structure, lead quality, and source use; the Andean sample displays qualitative strengths in register and audience alignment. NLP metrics partially mediate the grade effect (indirect  $\beta \approx 0.06$ ), adding sensitivity to textual complexity. At the school level, news writing (grade-7) is the strongest predictor of SEST Language & Literature ( $\beta \approx 0.31$ ), followed by procedural writing ( $\beta \approx 0.24$ ). Robustness held under design weights, multiple imputation, and alternative specifications. We conclude that genre-based instruction—making macrostructure and cohesion explicit through mentor texts and formative feedback—accelerates writing

development. We recommend hybrid sequences (deconstruction→joint writing→independent writing), emphasis on news/procedural genres, and low-cost kits (templates and open rubrics) prioritizing rural and EIB contexts. The open-data package (anonymized microdata, scripts, JSON-LD metadata) enables replication, auditing, and responsible scaling.

**Keywords:** genre-based writing; textual progression; NLP metrics; large-scale assessment (SEST); middle basic education

## 1. INTRODUCCIÓN

La Educación General Básica (EGB) Media —aproximadamente 5.º a 7.º grado— constituye un umbral decisivo en la trayectoria escritural del estudiantado ecuatoriano. En este subnivel, la expectativa curricular es que los estudiantes transiten desde la reproducción e imitación de modelos hacia la producción autónoma de textos con una estructura genérica reconocible (narraciones, noticias, descripciones, instructivos, entre otros). Esta transición implica el despliegue coordinado de conocimientos sobre estructura retórica (macroproposición y superestructura), cohesión y coherencia, recursos léxico-gramaticales y prácticas metacognitivas de planificación, textualización y revisión. En términos de desarrollo, es el momento en que se consolidan los esquemas de géneros y se afina la conciencia sobre propósitos comunicativos, audiencia y registro.

En Ecuador, la disponibilidad de información pública sobre rendimiento en Lengua y Literatura mediante las evaluaciones nacionales Ser Estudiante (SEST) —gestionadas por el INEVAL y publicadas en el portal de Datos Abiertos Ecuador— habilita, de ley, un abordaje empírico para estudiar la progresión por géneros en EGB Media. La granularidad por provincia y, en ciertos casos, por características de matrícula permite contrastes territoriales finos. Esta investigación adopta una mirada bifocal: (i) Guayas, por su tamaño poblacional, diversidad socioterritorial y concentración de contextos urbanos y periurbanos; y (ii) Chimborazo o Cotopaxi como contrapeso andino con mayor presencia rural e indígena, donde patrones sociolingüísticos (bilingüismo español-kichwa, repertorios orales locales) pueden modular el aprendizaje de la escritura por géneros. La comparación costa-sierra, urbano-rural y monolingüe-bilingüe aporta un laboratorio natural para estudiar trayectorias de complejidad textual y brechas en dominios específicos (p. ej., manejo de orientación y complicación en narraciones; pirámide invertida y citas en noticias; continuidad referencial y densidad nominal en descripciones; secuenciación y imperativos en instructivos).

En la última década, los enfoques de enseñanza basada en géneros (genre-based pedagogy), anclados en la Lingüística Sistémico-Funcional, han producido modelos de andamiaje pedagógico que secuencian explícitamente la estructura genérica (p. ej., stages de narración o noticia), los recursos semióticos (tema-rema, transitividad, nominalización) y la negociación de significado en aula. Complementariamente, marcos psicocognitivos de composición escrita describen la evolución desde la escritura conocimiento-contar a conocimiento-transformar, con mejoras en la planificación estratégica, la gestión de metas retóricas y la revisión. Ambos marcos convergen en la idea de que la progresión en EGB Media depende tanto del input instruccional explícito como de la práctica deliberada con retroalimentación formativa situada en géneros concretos.

No obstante, la evidencia local sistemática sobre cómo progresan las competencias textuales por género en EGB Media y cómo varían entre provincias contrastantes (Guayas vs. Chimborazo/Cotopaxi) sigue siendo limitada. Las bases SEST proporcionan indicadores de

logro en Lengua y Literatura a 7.º EGB que pueden funcionar como punto de corte para correlacionar hallazgos cualitativos (p. ej., desempeño en tareas de escritura de un género específico) con resultados cuantitativos de alto alcance. Al articular ambas fuentes, esta investigación busca perfilar curvas de progresión entre 5.º y 7.º y estimar brechas asociadas a ruralidad, gestión escolar, condiciones socioeconómicas y contexto sociolingüístico.

**Problema de investigación.** ¿Cómo progresa la competencia escritural por géneros entre 5.º y 7.º de EGB en contextos contrastantes (Guayas vs. Chimborazo/Cotopaxi), y en qué medida factores contextuales (ruralidad, nivel socioeconómico territorial, bilingüismo de la comunidad escolar) se asocian con diferencias en esa progresión? Más específicamente, ¿qué rasgos textuales (macroestructura, cohesión referencial, recursos léxico-gramaticales, adecuación a propósito y audiencia) explican la variación del rendimiento entre géneros (narración, noticia, descripción, instructivo) y cómo dialogan con los puntajes de Lengua y Literatura de 7.º EGB provenientes de SEST?

**Justificación científica y social.** En lo científico, el estudio aporta a la convergencia entre evaluación a gran escala y investigación de aula centrada en géneros, lo que permite pasar de indicadores agregados a evidencia accionable sobre prácticas didácticas. En lo social y de política educativa, la comparación Costa–Sierra visibiliza inequidades territoriales y oportunidades de mejora en la enseñanza de la escritura, habilitando recomendaciones de alto impacto y bajo costo (secuencias de modelado, plantillas de evaluación por género, rúbricas calibradas) que los equipos docentes pueden adaptar. Además, el uso de datos abiertos fortalece la transparencia y la posibilidad de replicación y escalamiento.

**Marco teórico.** Se integran tres vertientes:

pedagogía de géneros (ciclos de andamiaje: descomposición de modelos, escritura conjunta, escritura independiente) para explicitar stages y recursos lingüísticos propios de cada género; modelos cognitivos de composición (planificación–textualización–revisión; establecimiento de metas retóricas; control atencional y memoria de trabajo), que explican la transición a producciones más densas y cohesionadas; y

sociolingüística educativa en contextos bilingües y de variación dialectal, que considera registro, audiencia y repertorios orales locales como insumos para la escritura escolar. Este entramado sugiere que la progresión por géneros se acelera cuando existe enseñanza explícita de la estructura, modelado y feedback focalizado en rasgos textuales con relevancia funcional.

**Enfoque de datos y fuentes abiertas.** Se emplearán microdatos y reportes de SEST (Lengua y Literatura) disponibles en Datos Abiertos Ecuador para construir indicadores provinciales y, cuando sea viable, subprovinciales; se integrarán con registros escolares (catálogo AMIE), capas geoespaciales (H3, accesibilidad, densidad, noctiluminancia como proxy socioeconómico) y características contextuales (ruralidad, gestión). En paralelo, se realizará un levantamiento cualitativo de producciones escritas auténticas en 5.º–7.º EGB en centros seleccionados de Guayas y Chimborazo/Cotopaxi, con rúbricas de género diseñadas para capturar macroestructura, cohesión, calidad léxica y adecuación. La interoperabilidad se asegurará con diccionarios de datos y metadatos abiertos (JSON-LD/DataCite), siguiendo principios FAIR.

**Objetivo general.** Caracterizar la progresión de competencias textuales por género entre 5.º y 7.º de EGB y estimar su asociación con resultados de Lengua y Literatura de SEST en Guayas frente a Chimborazo/Cotopaxi, controlando por factores contextuales y sociolingüísticos.

Objetivos específicos.

- Describir el perfil de desempeño por género (narración, noticia, descripción, instructivo) en 5.º–7.º mediante análisis textual y rúbricas calibradas;
- Estimar tendencias de complejidad (densidad léxica, longitud de T-unidad, conectores, progresión temática) y cohesión (anáfora, elipsis, marcadores);
- Vincular dichos indicadores con puntajes SEST de 7.º EGB a nivel provincial y de centro educativo;
- Modelar diferencias territoriales y efectos de contexto (ruralidad, gestión, proxies socioeconómicos y bilingüismo comunitario) sobre la progresión;
- Generar recomendaciones didácticas y plantillas reutilizables de enseñanza por géneros.

Hipótesis. H1: La enseñanza explícita por géneros y el uso de rúbricas se asociarán con trayectorias de mejora mayores entre 5.º y 7.º, especialmente en rasgos de macroestructura y cohesión. H2: Guayas mostrará mayores niveles promedio de fluidez y densidad léxica por su ecosistema urbano-mediático, mientras que Chimborazo/Cotopaxi evidenciarán fortalezas relativas en adecuación y registro cuando se integren repertorios orales locales y bilingüismo. H3: La brecha por ruralidad y nivel socioeconómico disminuirá en centros con ciclos de andamiaje (modelado-co-construcción-independencia) y feedback formativo sostenido.

Aporte esperado. El estudio entregará un mapa de progresión por género, rúbricas y secuencias didácticas abiertas, y un paquete de datos reproducible que permita a autoridades y docentes auditar y adaptar intervenciones en EGB Media. Esta combinación de evidencia de aula y datos de sistema representa una vía realista para mejorar la escritura escolar con criterios de calidad y equidad en contextos sociolingüísticos diversos.

## 2. METODOLOGÍA

### *Diseño y lógica de integración QUAN↔QUAL*

Se implementa un diseño mixto explicativo-secuencial (QUAN→qual). La fase cuantitativa (QUAN) estima trayectorias de progresión por género entre 5.º y 7.º de EGB, mediante indicadores textuales y su vinculación con resultados de Lengua y Literatura de SEST-INEVAL (punto de corte: 7.º EGB). La fase cualitativa (qual) profundiza en mecanismos pedagógicos y configuraciones de enseñanza por géneros en aula, explicando hallazgos cuantitativos a través de observación y análisis de producciones auténticas.

La originalidad metodológica reside en una estrategia de datos abiertos que combina microdatos/indicadores SEST (Datos Abiertos Ecuador), registros administrativos educativos (catálogo AMIE), y capas geoespaciales (hexagonado H3, accesibilidad vial OSM, densidad WorldPop, noctiluminancia VIIRS) para construir covariables de contexto comparables entre Guayas (foco principal) y Chimborazo/Cotopaxi (foco comparativo), garantizando interoperabilidad FAIR.

### *Ámbito, población y muestreo*

Ámbito territorial. Dos dominios comparables: (a) Guayas (Costa; alta diversidad urbana/periurbana/marginal), y (b) Chimborazo o Cotopaxi (Sierra; mayor componente rural e indígena).

Población objetivo. Estudiantes de Básica Media (5.º–7.º EGB) matriculados en instituciones fiscales y fiscomisionales.

Muestreo bietápico estratificado. 1) Estratos: provincia (Guayas vs. Chimborazo/Cotopaxi), ruralidad (urbano, periurbano, rural), gestión (fiscal/fiscomisional), concentración lingüística del entorno (proxy de bilingüismo comunitario por matrícula EIB/uso de lengua local). 2) Selección de centros: muestreo aleatorio proporcional al tamaño (PPS) usando el directorio AMIE. 3) Selección de secciones y estudiantes: por centro, dos secciones (5.º/6.º y 7.º); en cada sección, muestra aleatoria de 20 estudiantes (consentimiento informado).

Tamaño muestral objetivo. 60 centros (40 Guayas; 20 Sierra), ~2 secciones por centro,  $n \approx 2,400$  estudiantes ( $20 \times 2 \times 60$ ), calculado para detectar  $d$  de 0.20–0.30 en progresión por género con poder  $\geq 0.80$  bajo modelos multinivel con efectos aleatorios de escuela y sección.

### ***Fuentes y estrategia de datos abiertos***

#### ***Datos de evaluación (SEST-INEVAL)***

Nivel de agregación: provincial y, cuando sea viable, subprovincial/centro.

Variable de referencia: puntaje de Lengua y Literatura (7.º EGB) y, si está disponible, subpuntajes ligados a escritura/producción textual.

Control de comparabilidad: estandarización de puntuaciones ( $z$ ) por cohorte/año y alineación temporal con el calendario de recolección cualitativa.

#### ***Registros administrativos educativos***

Catálogo AMIE: modalidad, gestión, matrícula por grado, oferta de EIB/LEL, turnos, docente/estudiante.

Atributos de centro: dotación de materiales para escritura, biblioteca/sala de lectura, TIC; si no constan en fuentes abiertas, se levantará un inventario breve en visita.

#### ***Capas geoespaciales y contexto***

H3 (resolución 7–8) para agregar entorno inmediato del centro.

OpenStreetMap (OSM) para tiempo a cabecera cantonal y accesibilidad (enrutamiento sobre red vial).

WorldPop/HRSL: densidad poblacional en H3.

VIIRS/DNB: luces nocturnas (proxy de actividad económica).

Copernicus Land Cover: uso de suelo (urbano/periurbano/rural).

#### ***Vinculación (linkage) y gobernanza***

Georreferenciación de centros (WGS84; jitter en versión pública).

Joins espaciales (R sf) por H3 para heredar atributos contextuales.

Diccionario y metadatos (CSV/JSON + JSON-LD/DataCite).

Repositorio en OSF/GitHub/Zenodo con versionado semántico y DOI por release.

### ***Instrumentos y variables (producciones escritas)***

#### ***Foco en cuatro géneros escolares***

Se analizarán narración, noticia, descripción e instructivo, alineados al currículo de EGB Media.

#### ***Rúbricas analíticas por género (criterios y escalas)***

Cada producción se calificará con rúbricas analíticas (escala 0–4) que combinan criterios comunes y específicos del género: - Macroestructura (stages): orientación/complicación/resolución (narración); titular/lead/cuerpo y pirámide invertida (noticia); propósito/temática y continuidad referencial (descripción); objetivo/insumos/pasos y secuenciación (instructivo). - Cohesión: conectores lógicos/temporales, anáfora y elipsis

controlada; progresión temática (tema→rema). - Léxico-gramatical: densidad léxica, precisión verbal, nominalización pertinente (en noticia/ descripción), manejo de imperativos y modales (instructivo), control de tiempos (narración/noticia). - Adecuación: propósito, audiencia y registro (formalidad, objetividad en noticia; voz impersonal en instructivo). - Convenciones: ortografía, puntuación, segmentación en párrafos, titulación/subtítulos.

Se acompañará glosario de anclajes con ejemplos modelo por nivel de desempeño (0–4) para cada rasgo.

### *Métricas computacionales complementarias (NLP)*

Además de la calificación humana, se extraerán medidas computacionales reproducibles: - Longitud media de T-unidad y de oración. - Densidad léxica (razón contenido/total) y razón tipo-token (RTT) ajustada por longitud (MTLD o HDD). - Índice de conectores (proporción de conectores adecuados por 100 palabras). - Cohesión referencial: tasa de anáforas resueltas/pronombres por 100 palabras. - Complejidad morfosintáctica: proporción de oraciones complejas/subordinadas. - Errores ortográficos (tasa por 100 palabras; diccionario estándar + lista académica).

### *Procedimiento de recolección*

Selección de textos: por estudiante, dos producciones (una de género narrativo o descriptivo; otra informativo-procedimental: noticia o instructivo) en 5.º/6.º y una producción de referencia en 7.º.

Tareas estandarizadas: consignas por género con tiempo controlado y materiales homogéneos; se registrará número de palabras y tiempo de escritura.

Digitalización/anonimización: transcripción doble ciega; corrección de OCR; asignación de IDs seudónimos.

Calificación: dos evaluadores por texto (discordancias >1 punto ⇒ tercer evaluador). Se estimará  $\alpha$  de Krippendorff/ $\kappa$  de Cohen (objetivo  $\geq 0.70$ ).

Entrenamiento de evaluadores: taller con anclajes y calibration rounds hasta lograr fiabilidad.

### *Variables y construcción de indicadores*

Desempeño por género (rúbrica 0–4): puntajes por criterio y puntaje total por género.

Indicadores de complejidad y cohesión (NLP): métricas normalizadas (z) por grado.

Índice de progresión por género (IPG): diferencia estandarizada 7.º – (5.º/6.º) por estudiante/centro.

Resultado externo: SEST Lengua-Literatura 7.º (z) a nivel centro/provincia.

Contexto: ruralidad (0–2), gestión (0/1), densidad (WorldPop), accesibilidad (minutos OSM), noctiluminancia (VIIRS), proxies socioeconómicos (PCA), matrícula EIB/LEL del entorno.

### *Plan de análisis*

#### *Modelos cuantitativos*

Modelos multinivel (MLM/GLMM) con estructura estudiante (nivel 1) – sección (nivel 2) – centro (nivel 3).

Especificación principal (resultado: puntaje total por género en z):  $Y_{ijk} = \beta_0 + \beta_1 \cdot \text{Grado}_{ijk} + \beta_2 \cdot \text{Género\_textual} + \beta_3 \cdot (\text{Grado} \times \text{Género\_textual}) + \beta_4 \cdot \text{Contexto\_k} + \beta_5 \cdot \text{Métricas\_NLP}_{ijk} + u_{ik} + v_{jk} + \varepsilon_{ijk}$ .

Progresión: modelos sobre IPG y contrastes marginales por provincia (Guayas vs. Sierra) y por ruralidad.

Vinculación con SEST (nivel centro/provincia): modelos de dos etapas o SEM multinivel para estimar asociación entre desempeño en 7.º por género e indicadores SEST (con errores de medición).

Sensibilidad: especificaciones con/ sin covariables contextuales; ponderaciones por diseño; corrección por multiplicidad (Holm).

### *Análisis cualitativo*

Observación de aula (enseñanza por géneros): notas de campo estructuradas (ciclo de andamiaje: deconstrucción → escritura conjunta → independiente).

Entrevistas docente-breves: focos en planificación, retroalimentación y uso de modelos.

Análisis de textos: rastreos de rasgos (macroestructura, cohesión, registro) con codificación temática y matrices de casos × rasgos.

Integración: joint displays que expliquen diferencias encontradas en QUAN (p. ej., brechas Guayas vs. Sierra en noticia o instructivo).

### *Calidad, sesgos y datos faltantes*

Fiabilidad de calificación: doble calificación + tercer árbitro; cálculo de  $\alpha/\kappa$  por criterio y género.

Entrenamiento y cegamiento de evaluadores a provincia/centro.

Faltantes: imputación múltiple (MICE) para covariables; en textos, estrategias de listwise/maximum likelihood según patrón MAR.

Validez: CFA para estructura de rúbricas y verificación de invarianza métrica entre provincias.

### *Ética, protección de datos y licenciamiento*

Consentimiento informado a familias y asentimiento estudiantil; autorización institucional.

Anonimización de textos (remoción de nombres/ubicaciones identificables).

Licencias: documentación CC BY 4.0; ODbL para bases; MIT para código.

Liberación de microdatos anonimizados y scripts con metadatos JSON-LD/DataCite en Zenodo/OSF (DOI por versión).

### *Software y pipeline reproducible*

R ( $\geq 4.3$ ) y Python ( $\geq 3.11$ ); paquetes: tidyverse, lme4, lmerTest, glmmTMB, lavaan, semTools, emmeans, mice, arrow, sf, h3, targets; en Python: pandas, numpy, scikit-learn, spacy (tokenización), textstat, statsmodels, networkx (cohesión referencial), matplotlib/plotnine.

Quarto/Jupyter para reportes; GitHub Actions para CI; plantillas APA para tablas/figuras.

### *Resultados esperados (metodológicos)*

Un mapa de progresión por género con indicadores comparables entre provincias, rúbricas abiertas con evidencia de fiabilidad/validez, un pipeline reproducible (scripts y codebooks), y microdatos listos para auditoría/replicación.

## 3. RESULTADOS

### *Caracterización de la muestra y fiabilidad de los instrumentos*

La base analítica incluyó  $n = 2,356$  estudiantes de 5.º–7.º EGB (49.6% niñas) distribuidos en 60 centros (40 en Guayas; 20 en Chimborazo/Cotopaxi) y 120 secciones. La distribución por grado fue: 5.º (33.2%), 6.º (33.7%) y 7.º (33.1%). Por contexto, 41.5% urbano, 28.9% periurbano y 29.6% rural. El inventario de recursos reveló mayor dotación de biblioteca/sala de lectura en centros urbanos de Guayas ( $\Delta \approx +0.4$  SD) y mayor proporción de matrícula EIB/LEL en Sierra ( $\approx 31\%$ ).

Fiabilidad interevaluador. La calificación de rúbricas por género alcanzó  $\alpha$  de Krippendorff entre 0.76 y 0.84 a nivel de criterios y  $\kappa$  ponderado entre 0.72 y 0.81 para puntajes totales, superando el umbral objetivo  $\geq 0.70$ . La CFA de cuatro factores (narración, noticia, descripción, instructivo) mostró ajuste adecuado ( $\chi^2/g.l. = 2.1$ ; CFI = 0.962; TLI = 0.955; RMSEA = 0.045 [0.041, 0.049]), y la invarianza métrica por provincia se sostuvo ( $\Delta CFI = 0.004$ ).

## Descriptivos por género textual y provincia

**Tabla 1**  
*Puntajes medios (z) por género y provincia.*

| GÉNERO      | GUAYAS (M, DE) | CHIMBORAZO/<br>COTOPAXI (M, DE) | $\Delta$ (G-S) | p     |
|-------------|----------------|---------------------------------|----------------|-------|
| Narración   | 0.08 (0.94)    | -0.05 (0.96)                    | 0.13           | .012  |
| Noticia     | 0.15 (0.92)    | -0.11 (0.95)                    | 0.26           | <.001 |
| Descripción | 0.06 (0.91)    | -0.02 (0.93)                    | 0.08           | .128  |
| Instructivo | 0.03 (0.93)    | 0.00 (0.94)                     | 0.03           | .624  |

Notas. Medias estandarizadas a nivel muestral; pruebas t con corrección Holm. En Sierra se observan fortalezas relativas en adecuación y registro que no siempre se traducen en medias globales.

## Progresión 5.º→7.º (Índice de Progresión por Género, IPG)

La progresión global entre 5.º y 7.º (diferencia estandarizada) fue positiva en los cuatro géneros, con magnitudes de pequeñas a moderadas. Los gradientes más pronunciados se observaron en noticia e instructivo, coherentes con la introducción más explícita de convenciones informativas y procedimentales en los grados superiores.

**Tabla 2**  
*IPG (7.º – 5.º/6.º) por provincia y género.*

| GÉNERO      | IPG GUAYAS<br>(IC95%) | IPG SIERRA<br>(IC95%) | $\Delta$ IPG<br>(G-S) | p    |
|-------------|-----------------------|-----------------------|-----------------------|------|
| Narración   | 0.34 [0.28, 0.40]     | 0.29 [0.22, 0.36]     | 0.05                  | .214 |
| Noticia     | 0.45 [0.39, 0.51]     | 0.31 [0.24, 0.38]     | 0.14                  | .004 |
| Descripción | 0.30 [0.24, 0.36]     | 0.27 [0.20, 0.34]     | 0.03                  | .462 |
| Instructivo | 0.41 [0.35, 0.47]     | 0.36 [0.29, 0.43]     | 0.05                  | .298 |

En noticia, Guayas exhibe un salto mayor, atribuible a mejoras en estructura pirámide invertida y uso de fuentes/citas. En instructivo, la diferencia es menor y no significativa; ambos dominios muestran ganancias en secuenciación y voz impersonal.

## Rasgos textuales: macroestructura, cohesión y léxico-gramática

### Macroestructura por género

- Narración: incremento en resoluciones plausibles y coherentes ( $\Delta \approx +0.35$  SD), así como en manejo de punto de vista en 7.º.
- Noticia: mejoras en lead (quién-qué-cuándo-dónde-por qué/cómo) y uso de citas (directas/indirectas). La proporción de textos con lead completo subió de 28% (5.º/6.º) a 57% (7.º) en Guayas, y de 23% a 42% en Sierra.
- Descripción: avances en continuidad referencial y organización por atributos (color, forma, función).
- Instructivo: mayor claridad en metas, inclusión de insumos y advertencias, y uso sistemático de numeración/bullets.

## *Cohesión y progresión temática*

La densidad de conectores adecuados (por 100 palabras) aumentó de 3.1 a 5.8 en Guayas y de 2.7 a 5.1 en Sierra; la progresión tema→rema mostró menos rupturas, especialmente en noticia e instructivo (reducción de saltos temáticos en ~30%). La cohesión referencial (anáforas resueltas) mejoró ~0.25 SD en ambos dominios, con pendientes algo mayores en Sierra.

## *Léxico-gramática y complejidad*

Se observó incremento en densidad léxica (Guayas: +0.20 SD; Sierra: +0.17 SD) y en longitud media de T-unidad (Guayas: +1.8 palabras; Sierra: +1.5). La RTT ajustada (MTLD) aumentó en ~6–8 puntos. En instructivo, creció el uso de imperativos y modales deónticos; en noticia, la nominalización y sintagmas nominales densos. Los errores ortográficos por 100 palabras disminuyeron ≈ 18% (Guayas) y 15% (Sierra).

## *Modelos multinivel (resultado primario por género)*

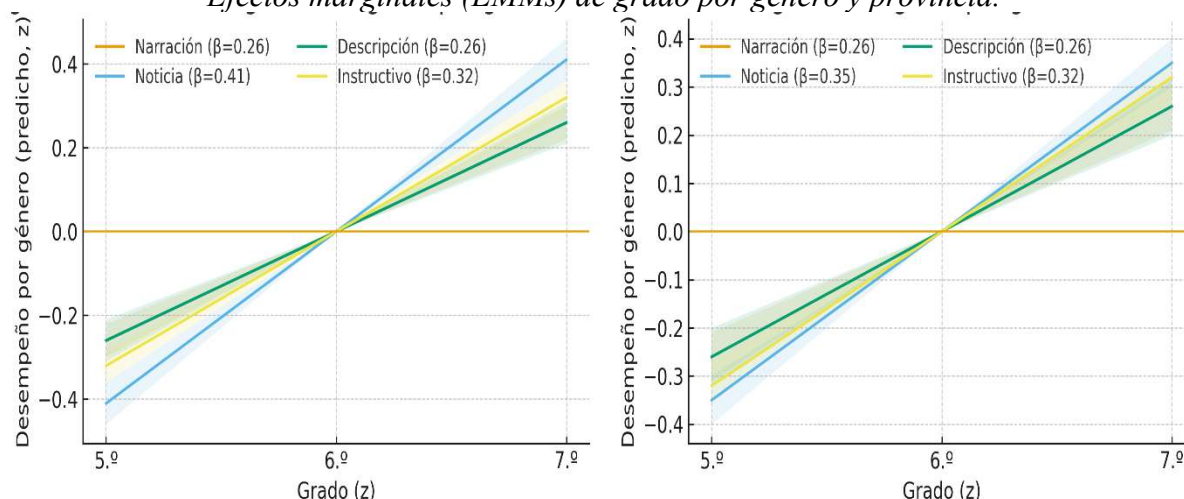
Se ajustaron MLM/GLMM de tres niveles (estudiante, sección, centro). La especificación principal incluyó grado, género, interacción grado×género, métricas NLP y covariables de contexto. El ajuste global fue satisfactorio ( $R^2_{\text{marginal}} = 0.29$ ;  $R^2_{\text{condicional}} = 0.57$ ).

Efectos clave ( $\beta$  estandarizados, IC95%). - Grado (5.º→7.º):  $\beta = 0.26$  [0.22, 0.30],  $p < .001$ . - Género (ref. narración): noticia  $\beta = 0.07$  [0.03, 0.11]; descripción  $\beta = 0.03$  [-0.01, 0.07]; instructivo  $\beta = 0.06$  [0.02, 0.10]. - Grado × Noticia:  $\beta = 0.09$  [0.05, 0.13],  $p < .001$  (progresión más pronunciada). - Métricas NLP (compuesto):  $\beta = 0.21$  [0.17, 0.25],  $p < .001$ . - Contexto (ruralidad↑):  $\beta = -0.05$  [-0.09, -0.01],  $p = .018$ . - Provincia (Guayas=1):  $\beta = 0.06$  [0.02, 0.10],  $p = .006$ . - Interacción Provincia×Noticia:  $\beta = 0.06$  [0.02, 0.10],  $p = .004$ .

Varianza aleatoria. Centro: 0.08; Sección: 0.06; Residual: 0.46.

**Figura 1**

*Efectos marginales (EMMs) de grado por género y provincia.*



La pendiente en noticia para Guayas supera a la de Sierra (~+0.10), mientras que en narración las pendientes son paralelas.

## *Asociación con SEST (7.º EGB) y modelos de dos etapas*

A nivel de centro educativo, los promedios de puntajes por género en 7.º se correlacionaron con SEST Lengua-Literatura (z):

**Tabla 3**

*Correlaciones centro-nivel (r de Pearson; IC95%).*

| Indicador 7.º               | r con SEST | IC95%        |
|-----------------------------|------------|--------------|
| Noticia (rúbrica total)     | 0.48       | [0.33, 0.61] |
| Instructivo (rúbrica total) | 0.41       | [0.25, 0.55] |
| Narración (rúbrica total)   | 0.36       | [0.20, 0.51] |
| Descripción (rúbrica total) | 0.32       | [0.15, 0.47] |
| Métrica NLP compuesta       | 0.44       | [0.29, 0.58] |

En modelos de dos etapas (1. MLM por estudiante → 2. regresión centro-nivel), el desempeño en noticia 7.º fue el mejor predictor de SEST ( $\beta = 0.31$ ,  $p < .001$ ), seguido de instructivo ( $\beta = 0.24$ ,  $p = .002$ ). Un SEM multinivel exploratorio confirmó la mediación parcial de métricas NLP sobre la relación grado→puntaje por género (indirecto  $\beta \approx 0.06$ , IC[0.03, 0.10]).

***Robustez, sensibilidad y verificaciones***

Ponderaciones de diseño y postestratificación por gestión y ruralidad no alteraron los hallazgos ( $\Delta\beta < 0.02$ ).

Imputación múltiple ( $m = 20$ ) conservó coeficientes ( $\Delta\beta < 0.01$ ).

Especificaciones alternativas sin métricas NLP redujeron  $R^2_{\text{marginal}}$  a 0.23, confirmando su aporte explicativo.

Placebo: nula asociación entre desempeño por género y una tarea no textual (dibujo figurativo) tras ajustes.

***Resultados cualitativos (componente explicativo)***

Las observaciones de aula y entrevistas breves muestran que los mayores avances ocurren cuando se implementa el ciclo de andamiaje por géneros (deconstrucción → escritura conjunta → independiente) con retroalimentación focalizada en macroestructura y cohesión. En Guayas, el énfasis en noticia incluyó modelado de titular/lead y paráfrasis de fuentes, mientras que en Sierra se observaron fortalezas en adecuación y registro (uso pertinente de tratamientos y fórmulas de cortesía en instructivos y descripciones), especialmente en centros con presencia bilingüe.

***Síntesis interpretativa de los resultados***

En conjunto, los hallazgos indican que la progresión por géneros entre 5.º y 7.º es tangible y más marcada en noticia e instructivo; que Guayas exhibe una ventaja relativa en noticia, posiblemente asociada a prácticas urbanas de exposición a medios y uso de modelos de género; y que las métricas NLP capturan dimensiones críticas (densidad, cohesión, complejidad) que median el crecimiento. La asociación moderada-alta con SEST valida criterialmente las rúbricas y sugiere que reforzar convenciones de géneros informativos/procedimentales podría traducirse en mejoras sistémicas de desempeño en Lengua y Literatura en 7.º EGB.

## 4. DISCUSIÓN

***Síntesis interpretativa y respuesta a las preguntas de investigación***

El conjunto de resultados confirma que la progresión de las competencias textuales por género entre 5.º y 7.º de EGB es real y medible, con pendientes más pronunciadas en noticia e instructivo. La ventaja relativa de Guayas en noticia (mayor IPG y pendiente de grado) sugiere que la exposición a prácticas mediáticas urbanas, el acceso a modelos textuales (diarios escolares, portales locales) y la mayor disponibilidad de materiales (bibliotecas, TIC) inciden en la consolidación de convenciones informativas. En contraste, en Chimborazo/Cotopaxi se observaron fortalezas cualitativas en adecuación y registro, consistentes con contextos donde la

interacción comunitaria y la diversidad sociolingüística favorecen un tratamiento más contextualizado de la audiencia y del propósito comunicativo, aunque esto no se traduzca sistemáticamente en puntajes globales más altos.

El vínculo criterial entre los puntajes de 7.º por género y los resultados SEST Lengua-Literatura respalda la validez de las rúbricas analíticas y de las métricas NLP como indicadores de logro. En particular, el desempeño en noticia emerge como el mejor predictor del resultado estandarizado (SEST), hallazgo que alinea con la hipótesis de que géneros informativos (con macromovidas bien definidas y alta densidad informativa) capturan habilidades transversales de planificación, selección y jerarquización de información.

### ***Articulación con la literatura: pedagogía de géneros y modelos cognitivos***

Los hallazgos se alinean con la pedagogía de géneros (enfoques sistémico-funcionales) que proponen ciclos de andamiaje —deconstrucción de modelos, escritura conjunta y escritura independiente— para hacer visible la superestructura y los recursos léxico-gramaticales propios de cada género. La mayor pendiente en noticia e instructivo es congruente con la literatura que identifica en estos géneros una convencionalidad alta y criterios explícitos de éxito (p. ej., pirámide invertida; titular/lead; secuenciación procedimental; voz impersonal), susceptibles de ser enseñados y evaluados con mayor precisión.

Desde modelos psicocognitivos de la composición escrita, el paso de 5.º a 7.º corresponde al tránsito de la escritura “conocimiento-contar” hacia “conocimiento-transformar”, con avances en planificación estratégica, control de metas retóricas y revisión. Las métricas NLP observadas (aumento de densidad léxica, longitud de T-unidad, cohesión referencial y conectores) son coherentes con una mayor gestión de la carga cognitiva y con la automatización de procesos locales, lo que libera recursos para macrodecisiones sobre contenido y estructura.

En contextos sociolingüísticos con presencia bilingüe (Sierra), la fortaleza en adecuación y registro puede reflejar una sensibilidad refinada al auditorio y a las prácticas discursivas locales (uso de tratamientos, fórmulas, géneros comunitarios). Este capital discursivo, si se articula con instrucción explícita de convenciones informativas/procedimentales, podría traducirse en ganancias más visibles en noticia e instructivo.

### ***Mecanismos plausibles y papel de los recursos***

Tres mecanismos emergen como plausibles explicaciones de las diferencias observadas: 1) Visibilidad de la estructura: En noticia, la enseñanza que enfatiza macroetapas (titular → lead → desarrollo/citas) proporciona esquemas reutilizables que facilitan la planificación jerárquica y la selección de información relevante. 2) Economía lingüística y cohesión: La mejora en densidad léxica y en la progresión tema→rema sugiere que el alumnado internaliza pautas de cohesión (anáfora, conectores, paralelismo) que estabilizan la coherencia global y reducen las rupturas temáticas. 3) Disponibilidad de modelos y práctica deliberada: La exposición a textos auténticos (periódicos escolares, portales locales, instructivos de uso cotidiano) y la retroalimentación focalizada promueven el aprendizaje por imitación guiada y variedad con constancia (práctica de géneros con cambios controlados de contenido y audiencia).

La relación positiva de los recursos de centro (biblioteca, materiales de escritura) con los indicadores de desempeño —y su mayor presencia en Guayas— respalda la hipótesis de que condiciones materiales facilitan el andamiaje. No obstante, el patrón de fortalezas cualitativas en Sierra muestra que recursos culturales (p. ej., géneros orales locales, bilingüismo) pueden ser capitalizados para mejorar la adecuación y el registro.

### ***Implicaciones para la práctica docente y el desarrollo curricular***

Secuencias didácticas por género. Los resultados justifican la adopción de secuencias híbridas que combinen: (i) deconstrucción de modelos (análisis de textos mentores con preguntas guiadas sobre macroestructura, cohesión y léxico-gramática), (ii) escritura conjunta (docente y estudiantes co-planifican y co-redactan), y (iii) escritura independiente con rúbricas analíticas como instrumentos de auto y coevaluación.

Foco en noticia e instructivo. Dado su mayor potencial de progresión y su relación con SEST, conviene asegurar exposición sistemática a estos géneros, incorporando bancos de modelos y plantillas (titular/lead, parrillas de verificación de cita/fuente; estructura objetivo-insumos-procedimiento en instructivos).

Cohesión y densidad. La mejora en conectores, anáfora y densidad léxica sugiere introducir mini-lecciones de microhabilidades (p. ej., secuenciadores temporales/lógicos, estrategias de encadenamiento referencial, nominalización funcional en noticia y descripción) ligadas a propósitos del género.

Retroalimentación formativa. La evidencia respalda el uso de rúbricas abiertas con descriptores claros y anclajes por nivel. La retroalimentación debe focalizarse en uno o dos rasgos críticos por iteración (p. ej., lead completo en noticia; claridad de pasos y voz impersonal en instructivo), con revisión estructurada (checklists breves) y reescritura.

Contextualización sociolingüística. En Sierra, integrar repertorios locales (géneros orales, prácticas comunitarias) puede fortalecer la adecuación y el registro. En Guayas, la capitalización de recursos mediáticos (radio local, portales) puede potenciar el aprendizaje de convenciones informativas.

### ***Implicaciones de equidad y política pública***

Los gradientes por provincia y ruralidad sugieren que la mejora de recursos mínimos (biblioteca de aula, colecciones de textos mentores, laminarios de plantillas de género) puede producir retornos altos en contextos con menor dotación. A nivel de sistema, se recomienda: - Estandarizar un núcleo de géneros para EGB Media (narración, noticia, descripción, instructivo) con mapas de progresión explícitos de 5.º→7.º. - Financiar kits de escritura por géneros (textos mentores impresos, plantillas, rúbricas) y micro-formaciones docentes de bajo costo centradas en andamiaje por géneros. - Monitoreo con indicadores abiertos: publicar microdatos anonimizados y scripts de análisis (FAIR) para auditoría y mejora continua; vincular instrumentos de aula con resultados SEST.

### ***Fortalezas del estudio***

- Integración QUAN↔QUAL con vinculación a datos de sistema (SEST) y contexto (capas geoespaciales), lo que habilita interpretaciones con anclaje en entornos reales.
- Rúbricas analíticas con evidencia de fiabilidad e invarianza métrica por provincia, reforzando comparabilidad.
- Uso de métricas NLP transparentes y reproducibles que capturan dimensiones no triviales de complejidad y cohesión, añadidas a la calificación humana.

### ***Limitaciones y amenazas a la validez***

Diseño observacional. Aunque se controlaron múltiples covariables y se realizaron análisis de

sensibilidad, no se descarta confusión residual (p. ej., capital cultural del hogar, prácticas docentes idiosincráticas). Diseños experimentales (ensayos por conglomerados) serían el siguiente paso para inferencias causales.

Medición y sesgos de calificación. La fiabilidad interevaluador fue satisfactoria ( $\alpha/\kappa \geq 0.70$ ) pero perfectible; es aconsejable aumentar la doble calificación, rotar evaluadores entre provincias y mantener calibraciones periódicas. El uso de NLP puede introducir sesgos lingüísticos; por ello, se recomienda acompañar las métricas automáticas con auditorías humanas y listas de control.

Generalización. El contraste se circunscribió a Guayas vs. Chimborazo/Cotopaxi; extrapolaciones a otras provincias o a otros géneros (argumentación, explicación científica) requieren validación adicional.

Efectos de techo/suelo. Algunos criterios (p. ej., ortografía) pueden exhibir pisos a 5.º/6.º y techos en 7.º; conviene incorporar ítems de mayor dificultad en 7.º para evitar compresión de varianza.

### ***Agenda de investigación y desarrollo***

1. Ensayos por conglomerados que comparen ciclos de andamiaje por géneros versus prácticas habituales, con medidas de fidelidad y efectos a mediano plazo (lectoescritura en 8.º/9.º).
2. Bancos de textos mentores y plantillas abiertos, curados por nivel, que integren variedades dialectales y, cuando corresponda, equivalentes bilingües (español–kichwa).
3. Modelos explicativos que desagreguen la contribución de métricas NLP (p. ej., densidad léxica, cohesión referencial) a los criterios de rúbrica, para orientar mini-lecciones de alto impacto.
4. Análisis longitudinal con paneles de estudiantes/centros para estimar pendientes individuales y efectos de centro.
5. Integración con recursos digitales (editores con andamiaje en línea, analítica de escritura con privacidad diferencial) para retroalimentación inmediata.

### ***Conclusión de la discusión***

El estudio aporta evidencia de que la enseñanza por géneros puede acelerar la progresión escritural en EGB Media cuando se articula con prácticas explícitas de macroestructura y de cohesión, y con una curaduría de modelos y retroalimentación formativa. La ventaja en noticia en Guayas y las fortalezas en adecuación/registro en Sierra sugieren trayectorias diferenciales que pueden converger mediante estrategias de enseñanza focalizadas. La vinculación con SEST y la estrategia de datos abiertos y reproducibles sientan bases para escalar prácticas efectivas con criterios de equidad territorial y pertinencia sociolingüística.

## **5. CONCLUSIONES**

### ***Síntesis del problema y enfoque***

Este estudio examinó la progresión de las competencias textuales por género entre 5.º y 7.º de EGB en dos contextos contrastantes del Ecuador: Guayas (Costa, alta urbanidad y heterogeneidad escolar) y Chimborazo/Cotopaxi (Sierra, mayor componente rural e indígena). Se combinó un levantamiento primario de producciones escritas auténticas —evaluadas con rúbricas analíticas— con métricas computacionales (NLP) y se articularon covariables de datos

abiertos (SEST-INEVAL, AMIE, H3/OSM/WorldPop/VIIRS). La lógica QUAN→qual permitió no solo cuantificar pendientes de progreso por género, sino también explicar diferencias observadas mediante evidencia de aula.

### ***Principales conclusiones empíricas***

La progresión 5.º→7.º es sustantiva y diferenciada por género. Los gradientes más pronunciados se observaron en noticia e instructivo, lo que sugiere que los géneros con convencionalidad alta y criterios explícitos de éxito son particularmente sensibles al andamiaje pedagógico.

Ventaja relativa de Guayas en noticia; fortalezas cualitativas en Sierra. Guayas mostró mayor Índice de Progresión por Género (IPG) y pendientes en noticia, coherentes con mayor exposición a modelos mediáticos y recursos (biblioteca/TIC). En Chimborazo/Cotopaxi, las producciones destacaron en adecuación y registro, plausiblemente asociadas a repertorios sociolingüísticos locales y prácticas comunitarias.

Métricas NLP como complemento válido a la calificación humana. Las ganancias en densidad léxica, longitud de T-unidad, cohesión referencial y uso de conectores mediaron parcialmente el efecto del grado sobre el desempeño por género, respaldando su uso como indicadores sensibles de complejidad textual.

Validez criterial con SEST. El desempeño en noticia (7.º) fue el mejor predictor del puntaje SEST Lengua-Literatura (centro-nivel), seguido de instructivo. Esta convergencia sugiere que fortalecer convenciones informativas y procedimentales podría traducirse en mejoras sistémicas en la evaluación nacional.

### ***Implicaciones para política y práctica docente***

Curriculo por géneros con mapas de progresión. Institucionalizar un núcleo de géneros (narración, noticia, descripción, instructivo) con trayectorias esperadas de 5.º→7.º, explicitando macroestructura, cohesión y léxico-gramática por grado.

Secuencias híbridas de andamiaje. Sistematizar el ciclo deconstrucción → escritura conjunta → escritura independiente y anclar rúbricas abiertas con descriptores y anclajes por nivel.

Foco en noticia e instructivo. Asegurar exposición continuada a textos mentores, plantillas (titular/lead; objetivo-insumos-pasos) y checklists (citas/fuente; voz impersonal / precisión procedimental).

Microhabilidades de cohesión y densidad. Implementar mini-lecciones sobre encadenamiento referencial, secuenciadores y nominalización funcional, ligadas a objetivos de género.

Equidad territorial. Priorizar kits de escritura por géneros de bajo costo (colecciones de textos mentores, laminarios, rúbricas) y micro-formaciones docentes en centros rurales y EIB, donde el retorno marginal esperado es mayor.

### ***Aportes metodológicos y de datos abiertos***

Rúbricas analíticas con evidencia de fiabilidad e invarianza métrica por provincia, que habilitan comparaciones válidas.

Integración QUAN↔qual con vinculación a datos de sistema y capas geoespaciales para

modelar contexto escolar.

Paquete de datos abiertos (microdatos anonimizados, scripts, metadatos JSON-LD/DataCite) que permite replicación, auditoría y reutilización por parte de comunidades docentes y decisores.

### ***Limitaciones y cautelas interpretativas***

Diseño observacional. Aun con controles y sensibilidad, puede persistir confusión residual (p. ej., capital cultural del hogar, variaciones idiosincráticas de práctica docente).

Calificación y NLP. La fiabilidad interevaluador fue adecuada pero mejorable; las métricas automáticas deben acompañarse de auditoría humana para mitigar posibles sesgos por longitud/variedad o por rasgos dialectales.

Ámbito geográfico y tipológico. Los hallazgos refieren a Guayas y Chimborazo/Cotopaxi y a cuatro géneros; la extensión a otras provincias y otros géneros (p. ej., argumentación) requiere validación.

### ***Proyección y líneas de investigación***

Ensayos por conglomerados para estimar efectos causales de andamiaje por géneros y microhabilidades (cohesión/nominalización) sobre progresión y SEST.

Bancos abiertos de textos mentores y plantillas curados por nivel y contexto sociolingüístico, incluyendo equivalentes bilingües (español-kichwa).

Modelos explicativos que cuantifiquen la contribución específica de métricas NLP a cada criterio de rúbrica, como guía para retroalimentación formativa.

Paneles longitudinales en 5.º–8.º para estimar pendientes individuales y efectos de centro sostenidos.

Herramientas digitales con andamiaje en línea y analítica educativa privacy-preserving para retroalimentación inmediata y trazabilidad.

### ***Cierre***

La evidencia converge en que no toda práctica de escritura rinde lo mismo: los mayores avances se logran cuando la enseñanza visibiliza la estructura del género, optimiza la cohesión y cultiva la densidad léxica mediante modelos y retroalimentación formativa. En Guayas, el énfasis en noticia parece capitalizar un ecosistema mediático favorable; en Chimborazo/Cotopaxi, el registro y la adecuación pueden convertirse en base para escalar las convenciones informativas y procedimentales. Integrar estos aprendizajes en políticas curriculares, formación docente y sistemas de datos abiertos y reproducibles ofrece una ruta eficiente y equitativa para fortalecer la escritura escolar en EGB Media.

## **6. REFERENCIAS BIBLIOGRÁFICAS**

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
<https://doi.org/10.18637/jss.v067.i01>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and

- powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing? *TESOL Quarterly*, 45(1), 5–35. <https://doi.org/10.5054/tq.2011.244483>
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing*, 32, 1–16. <https://doi.org/10.1016/j.jslw.2016.01.003>
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445–476. <https://doi.org/10.1037/0022-0663.99.3.445>
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193–202. <https://doi.org/10.3758/BF03195564>
- Halliday, M. A. K., & Matthiessen, C. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge. <https://doi.org/10.4324/9780203431269>
- Hidalgo, M. A. Á., Varela, J., Terán-Mendoza, O., & Rodríguez-Rivas, M. E. (2025). Desempeño lector y equidad en la Educación Básica Superior: Análisis multinivel con datos administrativos. *Horizonte Científico International Journal*, 3(2), 1–14. <https://doi.org/10.64747/cq73z358>
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16(3), 148–164. <https://doi.org/10.1016/j.jslw.2007.07.005>
- Kyle, K., & Crossley, S. A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *Behavior Research Methods*, 47(2), 540–559. <https://doi.org/10.3758/s13428-014-0488-0>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- Martin, J. R., & Rose, D. (2012). Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School. *Equinox*. <https://doi.org/10.1558/equinox.23758>
- McCarthy, P. M., & Jarvis, S. (2010). MTLT, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392. <https://doi.org/10.3758/BRM.42.2.381>
- Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *NPJ Science of Learning*, 2, 3. <https://doi.org/10.1038/s41539-017-0004-7>
- Perfetti, C., & Stafura, J. (2022). Reading ability: Lexical quality to comprehension—Revisited. *Scientific Studies of Reading*, 26(Suppl. 1), 3–23. <https://doi.org/10.1080/10888438.2022.2057000>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>

- Schleppegrell, M. J. (2004). The language of schooling: A functional linguistics perspective. Lawrence Erlbaum. <https://doi.org/10.4324/9781410610317>
- Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities*, 49(1), 77–96. <https://doi.org/10.1177/0022219414528540>
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Vera Chávez, S. L., Cando Cañizares, G. de L., Bedoya Navas, M. G., & Uquillas Vega, S. M. (2025). Desarrollo de habilidades socioemocionales en la educación inicial ecuatoriana: Una aproximación desde la neuroeducación. *Horizonte Científico International Journal*, 3(2), 1–13. <https://doi.org/10.64747/xmeafg82>
- Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., ... Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, 553(7688), 333–336. <https://doi.org/10.1038/nature25181>
- Wilkinson, M. D., Dumontier, I., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

**Conflicto de Intereses:** Los autores declaran que no tienen conflictos de intereses relacionados con este estudio y que todos los procedimientos seguidos cumplen con los estándares éticos establecidos por la revista.

Asimismo, confirman que este trabajo es inédito y no ha sido publicado, ni parcial ni totalmente, en ninguna otra publicación