enunciación

Publicación preprint

Este artículo fue aprobado para publicación en el Vol. 32 N.°1 de 2027 de la revista *Enunciación*, revista editada por la Universidad Distrital Francisco José de Caldas. Para su publicación, fueron tenidos en cuenta los conceptos de los pares evaluadores y los cambios realizados por los autores, para cumplir con la calidad académica establecida en nuestras pautas. Por lo tanto, se publica la versión preliminar del artículo para su consulta y citación. Esta versión puede consultarse, descargarse y citarse según se indica a continuación, pero debe recordarse que el documento final (PDF, HTML y XML) puede ser diferente.

Cómo citar: Jaramillo, L. (2027). Políticas para el Bilingüismo y prácticas evaluativas de la lengua en el Centro-Occidente Colombiano. *Enunciación, 32(1)*, e23957. https://doi.org/10.14483/22486798.23957

In prepress

The following article was approved for publication in Vol. 32 N.°1 (2027) in *Enunciación*, a journal published by Universidad Distrital Francisco José de Caldas. For its publication, it was subjected to an academic peer-review process and the authors incorporated their suggestions to comply with the academic quality established in our guidelines. Therefore, the preliminary version of the article is published for consultation and provisional citation. This version can be consulted, downloaded, and quoted as indicated below, but please consider that the final document (PDF, HTML, and XML) might be different.

Artículo de investigación. Lenguaje, sociedad y escuela.

Políticas para el bilingüismo y prácticas evaluativas de la lengua en el centro-occidente colombiano

Policies for Bilingualism and Language Assessment Practices in West-Center Colombia

Políticas para o bilinguismo e práticas de avaliação da linguagem no Centro-Oeste da Colômbia

Luis Felipe Jaramillo-Calderon[1]

VERSIÓN PREPRINT

Highlights

- El bilingüismo, más que un proyecto nacional, se ha convertido en un cliché en los discursos políticos.
- Las pruebas estatales se han vuelto políticas de facto que promueven prácticas descontextualizadas de enseñanza de inglés.
- Es posible que las pruebas estatales también carezcan de validez consecuencial.
- El enfoque heteroglósico permite la integración de prácticas multilingües como el cambio de código en las evaluaciones.
- Los docentes y académicos han abogado por una transformación en la evaluación de lenguas, aspecto prometedor para la evaluación bilingüe en el país.

Resumen

En los últimos años, las políticas lingüísticas colombianas han puesto en el centro del debate el papel del inglés en la educación y la sociedad. Sin embargo, poco se ha explorado la manera en que estas políticas inciden en las prácticas evaluativas en diferentes niveles del sistema educativo y cómo algunas evaluaciones en el contexto colombiano pueden llegar a ser, a su vez, políticas implícitas. Este artículo examina la interacción entre las políticas lingüísticas colombianas y los ejercicios evaluativos en la región centro-occidental del país, con especial atención al impacto de las políticas nacionales sobre los diversos actores del sistema de educación en lenguas. Adoptando un enfoque de casos múltiples, esta investigación se basa en entrevistas con formuladores de

[1] Magíster. Universidad Tecnológica de Pereira. Correo electrónico: luisfe00798@utp.edu.co. ORCID: https://orcid.org/0009-0002-8782-4047

Cómo citar: Jaramillo-Calderon, L. F. (2027). Políticas para el bilingüismo y prácticas evaluativas de la lengua en el centro-occidente colombiano. *Enunciación*, 32(1), e23957. https://doi.org/10.14483/22486798.23957

Fecha de postulación: 28 de julio de 2025; fecha de aceptación: 6 de febrero de 2026

políticas, desarrolladores de pruebas, administradores escolares y maestros, junto con un análisis de documentos de políticas relevantes. Los hallazgos indican que las evaluaciones estandarizadas de alto riesgo, en particular las Pruebas Saber 11 (exámenes nacionales para estudiantes de 11.° grado), han asumido el papel de política *de facto*, al dar forma y limitar los métodos de enseñanza y evaluación. Estos exámenes ejercen una presión significativa sobre las escuelas y los educadores, quienes con frecuencia priorizan la preparación para las pruebas a expensas del desarrollo lingüístico más amplio. Este artículo culmina con unas consideraciones sobre las implicaciones de estos hallazgos para el futuro de la educación hacia el bilingüismo en Colombia, abogando por prácticas de evaluación que sean equitativas, democráticas y alineadas con el desarrollo de la competencia multilingüe.

Palabras clave: educación de lenguas; evaluación de lenguas; política lingüística

Abstract

In recent years, Colombian language policies have placed the role of English in education and society at the center of debate. However, little has been explored about how these policies impact assessment practices at different levels of the education system and how some tests for the Colombian context, in turn, may become hidden policies. This article examines the complex relationship between Colombian language acquisition policies and assessment practices in the country's west-central region, emphasizing the impact of national policies on various stakeholders within the bilingual education system. Adopting a multi-case study approach, this research draws from interviews with policymakers, test developers, school administrators, and teachers, alongside an analysis of relevant policy documents. The findings indicate that high-stakes standardized assessments, notably the Pruebas Saber 11 (the state exams for 11th graders), have assumed the role of *de facto* policy, both shaping and constraining teaching and assessment methods. These exams exert significant pressure on schools and educators, who frequently prioritize test preparation at the expense of broader language development. This article ends with some considerations on the implications of these findings for the future of bilingual education in Colombia, advocating for assessment practices that are equitable, democratic, and aligned with the development of multilingual competence.

Keywords: language assessment; language education; language policies

VERSION PREPRINT

Resumo

Nos últimos anos, as políticas linguísticas colombianas têm colocado o papel do inglês na educação e na sociedade no centro do debate. No entanto, pouco se tem explorado sobre como essas políticas impactam as práticas de avaliação em diferentes níveis do sistema educacional e como algumas avaliações no contexto colombiano podem, por sua vez, se tornar políticas ocultas. Este artigo examina a complexa relação entre as políticas colombianas de aquisição da língua e as práticas de avaliação na região centro-oeste do país, com foco principal no impacto das políticas nacionais sobre os diversos atores do sistema educacional bilíngue. Adotando uma abordagem de estudo de caso múltiplo, a pesquisa baseia-se em entrevistas com formuladores de políticas, desenvolvedores de testes, administradores escolares e professores, além de uma análise de documentos de políticas relevantes. Os resultados indicam que avaliações padronizadas de alto risco, notadamente o Pruebas Saber 11 (exames nacionais para alunos do 11º ano), assumiram o papel de política de

fato, moldando e restringindo os métodos de ensino e avaliação. Esses exames exercem pressão significativa sobre escolas e educadores, que frequentemente priorizam a preparação para os testes em detrimento do desenvolvimento mais amplo da linguagem. Este artigo termina com algumas considerações sobre as implicações dessas descobertas para o futuro da educação bilíngue na Colômbia, defendendo práticas de avaliação que sejam equitativas, democráticas e alinhadas ao desenvolvimento da competência multilíngue.

Palavras-chave: avaliação de línguas; educação lingüística; políticas linguísticas

Introducción.

El bilingüismo, más que un proyecto nacional, se ha convertido en un cliché en los discursos educativos y políticos del país (Jaramillo-Calderon, 2023; Miranda y Fajardo, 2022). En los últimos años, el bilingüismo inglés-español se ha vuelto la meta o el objetivo de diversas políticas lingüísticas, en especial desde el Programa Nacional de Bilingüismo 2004-2019 (en adelante, PNB) (Ministerio de Educación Nacional [MEN], 2004), principalmente por visiones utilitarias de la prestigiosa lengua inglesa a nivel mundial. Se fundamenta esta elección en la competitividad económica que puede ofrecer un mayor dominio de esta lengua para la nación. A pesar de la importancia que el programa da al aprendizaje de lenguas extranjeras y al bilingüismo, el programa ha recibido numerosos cuestionamientos por parte de los académicos (Valencia y Vega, 2022). Por ejemplo, autores como Mackenzie (2020) y Bastidas y Jiménez-Salcedo (2021) critican la definición estrecha de bilingüismo, una definición que se centra principalmente en lograr el dominio del inglés como lengua prestigiosa y desestima otras lenguas extranjeras, étnicas y el español como L1 o L2. Además, se han planteado preocupaciones sobre la adopción acrítica de marcos extranjeros que no tienen en cuenta las realidades locales (Fandiño-Parra *et al*., 2012). Más aún, Correa y Usma (2013) señalan un refuerzo de las pruebas estandarizadas y las medidas de rendición de cuentas de los docentes dentro de la política, lo que podría descuidar otros aspectos valiosos de la educación en lenguas, como procesos de educación más holísticos y contextualizados, o competencias comunicativas o de mediación no capturadas por los instrumentos de medición empleados. Los autores señalan que este tipo de política lingüística suele recurrir a pruebas estandarizadas como instrumento principal para medir la efectividad de las políticas impuestas, lo cual no es solo una manera muy limitada de hacer seguimiento, sino que sirve como medida coercitiva para imponer las decisiones políticas, pues genera lógicas de competitividad entre docentes y entre instituciones y ansias por mejorar los resultados de las pruebas más que los propios procesos de enseñanza y aprendizaje.

VERSIÓN PREPRINT

Este tipo de políticas impulsadas por visiones estrechas del bilingüismo y la adopción de estándares extranjeros y otras tendencias políticas internacionales a menudo vienen acompañadas de pruebas estandarizadas que las refuerzan. Pruebas que en sí mismas han suscitado varias críticas, pero que pueden seguir dando forma a la manera en que se instruye y se evalúa la lengua inglesa en el país. Es por eso que este artículo explora la intrincada relación entre las políticas de lengua nacionales y los ejercicios de evaluación de lengua dentro y fuera del aula en la educación de lenguas del centro-oeste de Colombia (Jaramillo-Calderon, 2023). Esto puede contribuir no solo a comprender mejor la condición actual de cómo se evalúa el inglés como lengua adicional, sino también proporcionar información valiosa para una mejor formulación de políticas en el futuro.

Es importante aclarar que en este artículo se hace referencia al bilingüismo como un fenómeno social, interpersonal e individual de lenguas en contacto. En este sentido, aprendices de una segunda lengua, aunque sea en contexto de lengua extranjera, se configuran como bilingües emergentes (García *et al.*, 2008). Además, si bien históricamente se ha realizado una distinción entre la educación bilingüe y la educación tradicional en lenguas, en la que en la primera se integra la educación en lenguas con los contenidos mientras que en la segunda se mantienen aparte (García, 2009), la misma autora reconoce que estas distinciones son cada vez más difusas. Y así, recientemente autores como Turnbull (2017) invitan al reconocimiento de escenarios de enseñanza en lengua extranjera como instancias de educación bilingüe. Así mismo, Baker (2001) ha reconocido la educación tradicional en lenguas como un tipo de educación *para* el bilingüismo, aunque quizás no el más ideal.

Así pues, se puede afirmar que las políticas nacionales, si bien son ambiciosas, a menudo adoptan un enfoque descontextualizado *de arriba hacia abajo*, es decir impuesto burocráticamente por hacedores de políticas lingüísticas oficiales y agentes internacionales (Usma, 2015). Este enfoque de política lingüística no tiene verdaderamente en cuenta las diferentes realidades y necesidades lingüísticas de las regiones y las partes interesadas dentro del país, como docentes, estudiantes y padres de familia (Fandiño-Parra *et al*., 2012; Cely y Osorio, 2024). Esta tensión se complica aún más si agregamos a la ecuación el poder de las pruebas

VERSIÓN PREPRINT

estandarizadas como Pruebas Saber 11. Sobre estas pruebas, Lopez (2009), a partir de su análisis, ha generado unas críticas frente a las prácticas evaluativas estandarizadas que se realizan en la nación. Él ilustra cómo estas pruebas, aunque basadas en cierta medida en pruebas internacionales, tienden a carecer de autenticidad en las tareas que proponen e incluyen una limitada serie de habilidades. Esto podría devenir en una imagen incompleta del perfil bilingüe de los estudiantes que toman esta prueba, pues solo se les evalúa el vocabulario, la gramática y hasta cierto punto la comprensión lectora. Esto, según el autor, podría poner en cuestión la validez de las generalizaciones que se pretenden hacer sobre los tomadores de las pruebas a partir del limitado número de tareas.

Además, puesto que estas pruebas suelen tener un alto nivel de impacto, tanto para los estudiantes como para los docentes y las instituciones, las pruebas podrían incidir negativamente en la manera en que se llevan a cabo las clases de inglés *(negative washback)*, puesto que las clases se podrían orientar excesivamente hacia la preparación para los exámenes (*teaching to the test),* como lo señalan López *et al*. (2021) frente a la competencia de lectura. Por todo esto, y el hecho de que estas pruebas suelen ser usadas como mecanismos para forzar políticas lingüísticas y para realizar rendición de cuentas, es posible que estas pruebas también carezcan de validez consecuencial (Jaramillo-Calderon, 2023). Más aún en el contexto de un país altamente desigual, en el que muchos colegios y estudiantes marginados no tienen los recursos mínimos para alcanzar las metas esperadas en materia de bilingüismo, pero se exigen los mismos resultados y se utilizan los mismos instrumentos de evaluación.

Aunque estas pruebas puede que no evidencien de manera más amplia las dinámicas lingüísticas de bilingües o bilingües emergentes, sí han sido utilizadas para informar o justificar las nuevas políticas lingüísticas. Por tanto, esta investigación profundiza en cómo estas políticas lingüísticas, a su vez, dan forma e influyen en las prácticas de evaluación en varios niveles. Al examinar esta interacción entre la política en lengua y los ejercicios de evaluación, este estudio pretende obtener una comprensión integral de su interconexión dentro del sistema de educación bilingüe del centro-oeste de Colombia. Así, la pregunta que pretende contestar este estudio es: ¿Cuál es la influencia de las políticas lingüísticas colombianas en las prácticas de evaluación de los diferentes actores del sistema educativo bilingüe en el centro-oeste del país?

VERSIÓN PREPRINT

Marco teórico

Evaluación en lenguas

La evaluación de las competencias lingüísticas de aprendices de lengua puede consistir en muchas estrategias de recolección, análisis y juicio de lo que los estudiantes pueden hacer con sus recursos lingüísticos. La evaluación en lenguas (*language assessment)* puede ser un proceso continuo e informado por diversas técnicas, tanto formales como informales, formativas como sumativas, que den cuenta de las competencias lingüísticas de los usuarios en un momento dado (Brown y Abeywickrama, 2018). A menudo, se limita la evaluación a pruebas o exámenes de lengua, los cuales solo representan una de las posibilidades en las que se puede llevar a cabo este ejercicio (Chapelle y Plakans, 2012). El diseño de las evaluaciones lingüísticas se basa consciente o inconscientemente en constructos, los cuales suelen ser conceptos abstractos como la capacidad o el desempeño lingüístico y que responden a una serie de paradigmas sociales e ideológicos (Chapelle *et al*., 2020). Estos constructos determinan el tipo de evaluaciones y el tipo de tareas que los estudiantes deben afrontar, por ejemplo, una prueba orientada a la gramática o un ejercicio alternativo orientado al desempeño en una tarea que simule contextos auténticos de uso de la lengua.

Washback, impacto y validez consecuencial

Las evaluaciones lingüísticas suelen tener tanto efectos a nivel de enseñanza y aprendizaje como efectos a nivel personal, institucional y social (Wall, 1997). A estos primeros efectos se les denomina en la literatura *washback,* y pueden ser positivos en el sentido de que potencian los procesos de enseñanza y aprendizaje o negativos si por el contrario inhiben o entorpecen dichos procesos. El impacto, por su parte, se refiere a los efectos más amplios que las pruebas, en especial las de alta importancia, tienen sobre los individuos, las políticas educativas y la sociedad en general. Para ilustrar posibles tipos de efectos, las pruebas de alto impacto pueden representar, por ejemplo, el acceso o no a oportunidades de educación superior y por ende mayores oportunidades de desarrollo lingüístico tanto en inglés como en español. Las pruebas estandarizadas, debido a su alcance, pueden tener efectos tanto en un nivel micro como en uno macro nivel. A nivel micro, las pruebas de alto impacto pueden llevar a que los docentes e

VERSIÓN PREPRINT

instituciones se orienten excesivamente hacia la preparación específica para estas pruebas, dejando de lado alternativas instruccionales que puedan ser, potencialmente, más significativas. Además, pueden representar un factor excesivo de presión y jerarquización de la calidad de acuerdo con los indicadores de dichas pruebas.

Como indica Messick (1989), las pruebas pueden tener efectos positivos o negativos para el aprendizaje; en este caso, el efecto negativo puede ser más marcado si la misma prueba no representa dinámicas más auténticas del uso de la lengua objetivo. Por lo tanto, Messick (1981) defiende que las pruebas nunca son neutrales, ya que tienen valores inherentes y pueden producir consecuencias intencionales o no intencionales, tanto dentro como fuera del aula. En este sentido, el autor insiste en que los diseñadores y aplicadores de pruebas lingüísticas no deben preocuparse únicamente por la validez de constructo, es decir, que estas midan lo que pretenden medir, sino que amplíen su concepción de validez para tener en cuenta los posibles efectos de washback e impacto que pueda tener la aplicación de estas pruebas y así, procurar que sean lo más proporcionales y lo menos negativas posible.

Política lingüística

La política y planificación lingüística se configura tanto como una práctica social como un campo de estudio en sí mismo. Como práctica, tradicionalmente se refiere a los esfuerzos para estructurar, asignar funciones sociales y controlar la manera en la que los hablantes usan y aprenden una o más lenguas, frecuentemente, pero no siempre, a través de instituciones oficiales como ministerios de educación (Tollefson, 2006). Estos tipos de esfuerzos dan forma a las estrategias implementadas en el sector educativo para reforzar y/o desestimar la manera en la que ciertas lenguas son usadas o adquiridas. En este sentido, Spolsky (2004) expande un poco el espectro de la política lingüística y reconoce tres componentes: la manera real en la que los hablantes usan las lenguas, es decir las convenciones que de manera orgánica se generan; los valores simbólicos que se atribuyen a cada lengua, variedad o código, en gran medida informados por las ideologías lingüísticas de la comunidad; y, finalmente, los esfuerzos implícitos o explícitos para controlar y modificar la manera en la que las lenguas son usadas y adquiridas en una comunidad lingüística.

Como campo académico, desde los años sesenta se han desarrollado diferentes escuelas de pensamiento dependiendo de cómo los académicos y creadores de políticas han visto la

VERSIÓN PREPRINT

lengua, la educación y el ejercicio político en sí mismo. Tradicionalmente se ha conformado una visión racionalista que ve a las instituciones educativas como organizaciones burocráticas regidas por jerarquías, estándares y números, buscando maximizar la eficiencia y el control de lo que ocurre en las aulas (Perrow, 1986). Este enfoque ha priorizado la estandarización de los procesos educativos, la evaluación sumativa en lo posible orientada a la norma (*norm-oriented testing)*, y la toma de decisiones desde los roles de mando, enfatizando estándares internacionales y procesos de rendición de cuentas. Este tipo de política ha sido enfáticamente cuestionada por autores como Usma (2015), quien además señala que la manera en la que históricamente se ha abordado la política en Colombia guarda muchas similitudes con esta postura racionalista.

Es por esto que Usma (2015) insiste en un cambio de una política lingüística racionalista que limita las posibilidades de acción, a una visión más comparativa, crítica y sociocultural. Este cambio teórico y práctico permite comprender las políticas lingüísticas como fenómenos sociales en los que intervienen múltiples actores que, desde sus propios contextos, las interpretan, negocian y moldean, dando lugar a formas diversas de implementación. Este tipo de visión da también más voz y prioridad a los actores que se encuentran en los eslabones bajos e intermedios de la política educativa, desde los estudiantes y docentes hasta los administrativos locales, académicos y diseñadores de pruebas, por ejemplo.

VERSIÓN PREPRINT

Evaluación en lenguas como política lingüística

Las prácticas de evaluación y evaluación lingüística no son independientes de las políticas lingüísticas; más bien, están profundamente entrelazadas. Como lo indica Shohamy (2017), los exámenes de lengua no son únicamente instrumentos para medir competencias lingüísticas, sino que reflejan una ideología lingüística dominante, un uso aceptado y a menudo artificial de usar la lengua, con estrictas separaciones de códigos lingüísticos. Adicionalmente, cuando estas pruebas son estandarizadas y de gran escala, pueden incidir en la manera en la que los docentes llevan a cabo sus clases, en la definición de qué incluir y excluir de los currículos escolares, e incluso pueden influir en usos de la lengua más allá del aula de clase.

Estos efectos no ocurren de manera aislada; son tendencias globales que se posicionan mediante instituciones internacionales con gran influencia y poder. Entre estas tendencias se encuentra la estandarización de procesos educativos y la rendición de cuentas (Mølstad *et al*.,

2020; Guerrero-Nieto y Quintero-Polo, 2021). Una manera elemental en la que estas tendencias políticas son reforzadas es a través de las pruebas estandarizadas de alto alcance, pues estas pueden incidir, guiar y desestimar prácticas en las instituciones y aulas. Es por esto que Menken (2017) denomina a las pruebas estandarizadas políticas *de facto* que intencionalmente o no inciden lingüísticamente en todos los niveles del sistema educativo. De manera similar, McNamara (1998) las posiciona como un brazo o una extensión de las políticas lingüísticas en el panorama educativo, pues, debido a sus altas consecuencias, los actores tienen poca opción más que alinearse con la norma en busca del mejor resultado posible. En consecuencia, estas pruebas o políticas *de facto* resultan afectando de manera desproporcionada a sectores marginados de la sociedad y, paradójicamente, pueden desestimar, a largo plazo, procesos más orgánicos de desarrollo lingüístico.

Metodología

Este estudio cualitativo utilizó un enfoque de estudio de casos múltiples (Stake, 2013) para explorar las perspectivas e ideologías de diversas partes interesadas, como los encargados de formular políticas, los diseñadores de pruebas, los administradores de instituciones y los académicos, con respecto a las políticas lingüísticas y las prácticas de evaluación. La investigación incorporó marcos críticos y socioculturales (Levinson y Sutton, 2001) para examinar las experiencias de estas partes interesadas. Los hallazgos que se presentan en este artículo corresponden a una aproximación sintética de un estudio de mayor alcance (ver reconocimientos). En este se priorizan los hallazgos más directamente relacionados con la pregunta presentada, de acuerdo con las convenciones del formato de artículo académico.

Contexto y participantes

El estudio involucró a nueve actores de diversos contextos, incluidos formuladores de políticas, diseñadores de pruebas, académicos, administradores y docentes, todos involucrados en los ámbitos de la política y la evaluación lingüísticas. Estos participantes, algunos con múltiples roles, fueron seleccionados intencionalmente de varias ciudades para garantizar la representación en diferentes niveles (desde el nacional hasta el aula), sectores (públicos y privados) y niveles de

VERSIÓN PREPRINT

experiencia. Todos los participantes fueron informados sobre el propósito del estudio, el carácter voluntario de su participación, la preservación de su anonimato y la confidencialidad de la información. La participación se formalizó mediante la firma de un consentimiento informado, en concordancia con los principios éticos de la investigación académica. Además, como parte de la investigación, se analizaron documentos de políticas, incluida la Ley 1651 (Congreso de la República de Colombia, 2013) y el Plan Sectorial de Educación 2018-2022 (MEN, 2021). A continuación, se presenta mayor detalle de los casos de estudio.

Tabla 1. Caracterización de casos

P.	Código	Rol/Posición	Tipo de institución	Ciudad	Perfil y bagaje
1	PM-PB-DC	Responsable de políticas, líder del programa nacional	Sector público / política nacional	Bogotá D.C.	Diseña y promueve estrategias para el bilingüismo y el aprendizaje de idiomas; más de 15 años en liderazgo (e.g. Consejo Británico); se ha relacionado con instituciones nacionales y locales.
2	TD/Ter-PB-RIS	Diseñador de pruebas y profesor de idiomas	Universidad pública (instituto)	Eje Cafetero	Más de 13 años de docencia; diseña y administra pruebas de competencia; asesor académico; alinea las pruebas con los objetivos institucionales; participa en iniciativas locales.
3	Sch/TD-PV-DC	Académico y diseñador de pruebas	Universidad privada / servicios de evaluación	Bogotá D.C.	Doctorado en educación (bilingüismo/evaluación); creó un centro de evaluación; experiencia en escuelas bilingües y empresas de evaluación; investigación sobre evaluación bilingüe/multilingüe.
4	Ter-PV-RIS	Profesor de idiomas, líder de equipo de lenguas extranjeras	Colegio privado	Eje Cafetero	Profesor de inglés de escuela secundaria y grados décimo y once; más de 4 años de experiencia; se enfoca en la promoción del inglés y el desempeño en exámenes nacionales.

VERSIÓN PREPRINT

P.	Código	Rol/Posición	Tipo de institución	Ciudad	Perfil y bagaje
5	Sch/Adm-PB-ANT	Académico y administrador	Universidad pública (instituto)	Antioquia	Postura crítica sobre políticas nacionales; investigación en desarrollo profesional y políticas lingüísticas; promueve múltiples idiomas, incluidos los locales.
6	Sch/Adm-PB-CAL	Académico y director del programa	Universidad pública	Eje Cafetero	Administrativo de un posgrado en Educación en Idiomas; experto en alfabetización evaluativa; capacita extensamente a docentes en servicio y en formación; se enfoca en desarrollo profesional y evaluación de idiomas.
7	Ter-PB-RIS	Profesor de idiomas provisional	Escuela pública (rural)	Eje cafetero	Trabaja con estudiantes rurales de escuelas de grados sexto a noveno y de décimo a once; también enseña preparación para exámenes nacionales; destaca el bajo nivel de inglés inicial y la escasez de recursos.
8	Adm-PV-VAC	Administrador, director del instituto	Universidad privada	Valle del Cauca	Dirige un instituto de idiomas; roles anteriores de director en Eje Cafetero; la institución se centra en la internacionalización y el inglés es central en este proceso.
9	PM/Adm-PB-QUI	Responsable de políticas y administradores locales	Enlace universidad pública/gobierno	Eje Cafetero	Conecta universidades, gobiernos locales y nacionales y actores privados (por ejemplo, el British Council); promueve iniciativas de bilingüismo a nivel regional.

*Nota***:** Los códigos asignados a cada participante se utilizaron posteriormente para identificar los extractos de entrevista presentados en los resultados; estos códigos por sus siglas en inglés indican el(los) rol(es) principal(es) (profesor, administrador, desarrollador de pruebas, académico o hacedor de políticas), el tipo de institución (pública o privada) y la región (Bogotá, Antioquia, Caldas, Risaralda, Valle del Cauca y Quindío). Estos se emplearon con el fin de preservar el anonimato de los participantes y facilitar la trazabilidad analítica de los hallazgos.

Fuente: elaboración propia

VERSIÓN PREPRINT

Métodos de recolección de datos

Los datos de este proyecto consisten en entrevistas semiestructuradas con los participantes y el análisis de dos documentos de políticas. Se eligieron entrevistas semiestructuradas porque su naturaleza de cuestionarios verbales permite obtener información precisa y a profundidad de los participantes (Fraenkel, 2012). En total, se realizaron entrevistas semiestructuradas a la totalidad de los nueve participantes incluidos en el estudio. La mayoría de las entrevistas duraron entre 40 minutos y una hora y se realizaron en español, la lengua materna de los participantes, a excepción de una realizada en inglés, ajustándose a la preferencia del entrevistado. La estructura de cada entrevista semiestructurada, aunque se organizó a partir de ejes temáticos comunes para todos los participantes, relacionados con políticas lingüísticas, prácticas de evaluación y sus implicaciones, se ajustó al rol o roles que cada participante pudiera tener (docente, administrativo, diseñador de pruebas o gestor de políticas) siguiendo los principios de dicho instrumento. Las entrevistas se realizaron de manera individual, en un solo encuentro por participante. Todas las entrevistas fueron registradas en audio para luego ser transcritas y analizadas. Además, como afirma Bowen (2009), el análisis de documentos puede señalar situaciones que requieren atención especial; puede ayudar a enfocar la investigación y hacerla más crítica y exhaustiva. Por tanto, se analizaron dos documentos de políticas: la Ley 1651 (2013), que describe los mandatos nacionales de educación bilingüe, y el Plan Sectorial 2018-2022 (MEN, 2021), que detalla las políticas de desarrollo lingüístico durante ese período de gobierno.

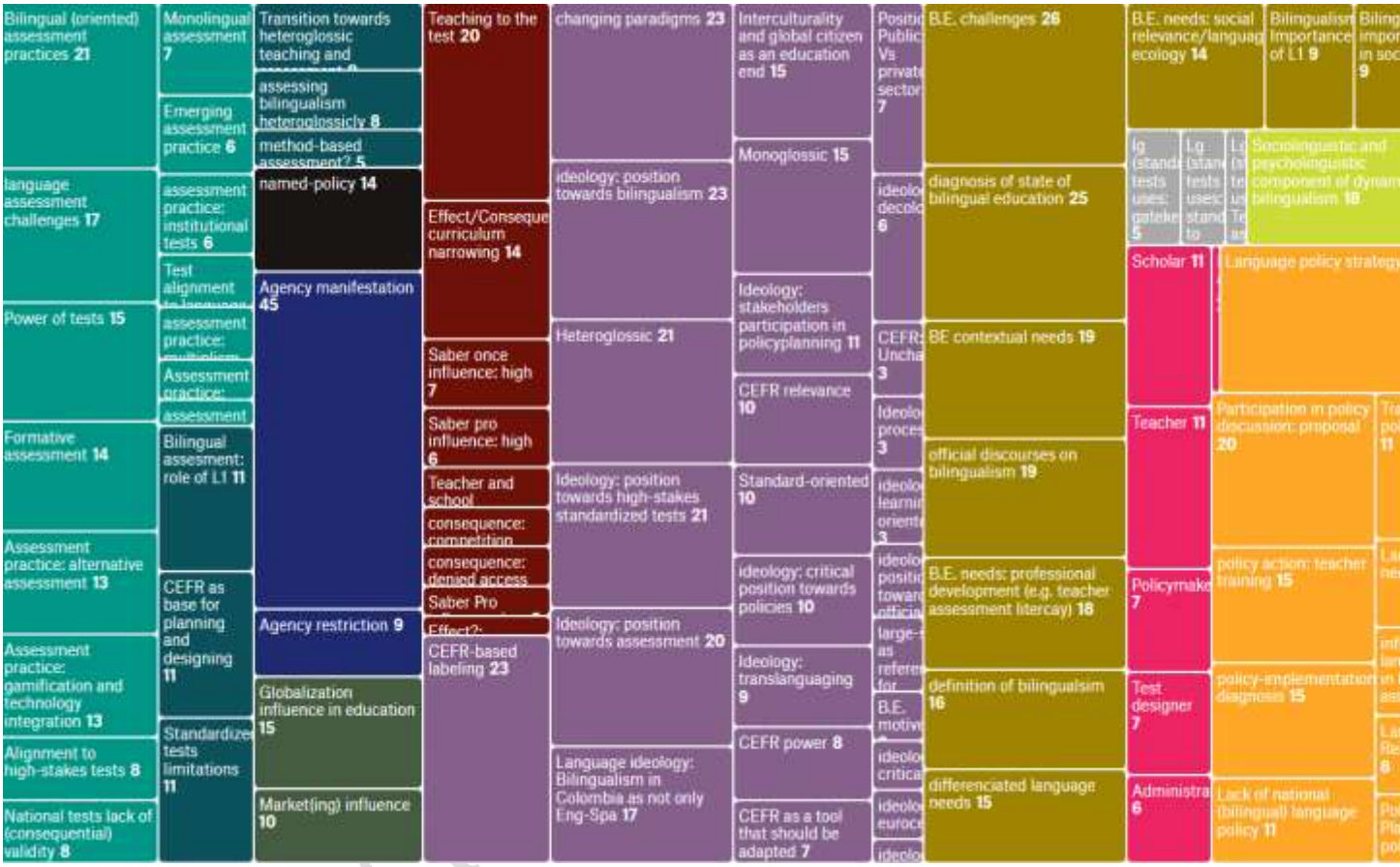
Análisis de datos

Para el análisis de los datos se implementó la técnica de análisis de contenido (Fraenkel *et al*., 2012). Esta técnica busca interpretar el contenido de los datos cualitativos para la identificación de temas y significados clave. Desde el análisis de contenido, se va más allá del contenido superficial para descubrir significados más profundos en los datos (Dörnyei, 2007). El análisis de este estudio siguió procedimientos híbridos, incluida la generación de códigos *a priori* y *ad hoc*, la codificación de los datos, la tematización de los códigos y datos a partir de los patrones y tendencias para así sustentar los diferentes resultados y dar respuesta a la pregunta planteada. Este proceso se llevó a cabo en Atlas.ti versión web. El diagrama de mapa de árbol

VERSIÓN PREPRINT

presentado en la Figura 1 ilustra de manera visual la codificación realizada a partir de las entrevistas y los documentos de políticas escritas.

Figura 1. Mapa de árbol, Atlas.ti web



Fuente: elaboración propia en Atlas.ti web.

El análisis abarcó la totalidad de los datos, los cuales comprenden las nueve transcripciones completas de las entrevistas y los documentos de política examinados. Cada uno de los casos resultó esencial para definir, codificar y construir los temas analíticos. Para la presentación de los hallazgos, y debido a los límites de extensión de un artículo, se optó por incluir solo ciertos fragmentos ilustrativos.

Resultados y análisis

Incidencia de las políticas en lengua en las prácticas evaluativas: factores limitantes

En cuanto al primer hallazgo, la incidencia de las políticas en lengua en los procesos de evaluación de la educación colombiana es significativa. Los datos muestran que las pruebas nacionales de inglés se han convertido en políticas *de facto*, que orientan y limitan la enseñanza y la evaluación. Como lo halló el estudio, estas pruebas tienen el potencial de perpetuar las desigualdades, limitar los currículos y pasar por alto otras lenguas y prácticas culturales. Las partes interesadas, como académicos, docentes y administrativos, reconocen la fuerte influencia de las políticas y las pruebas en el ámbito escolar, lo que conduce a la adaptación de las prácticas evaluativas para alinearse con estas pruebas y políticas. Los docentes muestran cómo sus escuelas, tanto públicas como privadas, dedican una cantidad significativa de tiempo, desde muy temprano en el proceso escolar, a la preparación de los estudiantes para las pruebas.

> **Extracto 1:** Sch/TD-PV-DC
> Se pone la prueba para que refuerce la política y mucha gente lo que hace es modificar lo que está haciendo en la práctica del aula de clase para que estén alineadas con la prueba. (...) Entonces, infortunadamente como te dije, se utiliza la evaluación como una herramienta para obligar a todo el mundo a reaccionar a esas políticas en vez de trabajar con los usuarios, los profesores, estudiantes, en cómo crear esas políticas y cuáles políticas serían ideales, entonces infortunadamente se recurrió a esa práctica que ha sido criticada por muchos años.

> **Extracto 2:** Sch/Adm-PB-ANT
> Entonces digamos que en los colegios eso [la influencia de las políticas] es muy fuerte, se ve tan fuerte que por ejemplo usted les dice a muchos maestros les dice, venga, vamos a enseñar de otra manera. Vamos a integrar las cuatro habilidades. Vamos a poner a los muchachos y las muchachas a que hablen más, hagamos esto un poco más comunicativo, no enseñamos tanta gramática vocabulario comprensión de lectura y le dicen a uno pero es que, si eso es lo que van a evaluar, yo, para qué les voy a enseñar lo otro y si yo no les enseño eso y les va mal por ponerme a hacer con ellos conversaciones y sale y les va mal y entonces a la persona que le van a echar la culpa es a mí acá en el colegio.

El análisis de las entrevistas con las partes interesadas reveló un reconocimiento compartido de la importante incidencia de las políticas y los exámenes lingüísticos, en particular

VERSIÓN PREPRINT

a nivel escolar. El extracto 1 sugiere que las políticas suelen ir acompañadas de exámenes correspondientes que las hacen cumplir. Los altos riesgos asociados con estos exámenes presionan a las escuelas y a los docentes a adaptar sus prácticas para alinearse con el examen y, en consecuencia, con la política en sí. Este actor también expresa su malestar con la forma en que se utilizan los exámenes nacionales como una herramienta para hacer cumplir la política, en lugar de trabajar en colaboración con los educadores para desarrollar políticas específicas para el contexto que aborden las necesidades reales. De manera similar, otro actor (extracto 2) destaca cómo las funciones de control y rendición de cuentas de los exámenes nacionales obligan a los docentes y las instituciones a limitar sus prácticas de enseñanza y evaluación para ajustarse al formato del examen, lo que limita los enfoques más amplios para el desarrollo del lenguaje. Las entrevistas posteriores con maestros de escuelas públicas y privadas confirman esta presión percibida para alinear las prácticas con los exámenes estandarizados, como lo muestran los siguientes extractos:

Extracto 3: Ter-PV-RIS

El colegio, como te conté, quedó en número uno en la ciudad de Dosquebradas, y ese es uno de los objetivos que siempre tienen los administradores y el rector. Entonces nosotros, al menos desde que estoy yo en la institución, se ha intentado que se empiece un proceso al menos desde temprana edad con los niños de que reconozcan cómo sería el formato de ese tipo de evaluaciones. La influencia es total, o sea, todo lo que te mencionaba ahorita sobre las pruebas bimestrales, aunque se hacen individuales, nos basamos en el formato del ICFES, sí. Nos esforzamos para que los niños desde muy pequeños tengan ese conocimiento [de toma de exámenes] y puedan presentarlas y sepan qué hacer en el momento de que tengan que presentar [el examen].

Extracto 4: Ter-PB-RIS

En el colegio damos cuatro horas de inglés y como es un colegio enfocado a las pruebas hay que ver una hora de Pruebas Saber 11 de inglés, ¿cierto? (...) En cuanto a la evaluación obviamente todo muy procedimental, sin embargo, obligatoriamente y por la modalidad del colegio, al final de cada periodo se tiene que dar un tipo de pruebas Saber,

VERSIÓN PREPRINT

o sea, una prueba con opción múltiple ¿cierto?, todos los niños desde sexto hasta 11 tienen ese tipo de prueba de evaluación.

Los maestros de escuelas públicas y privadas (extractos 3 y 4) describen la naturaleza de alto riesgo de las Pruebas Saber 11. Estos exámenes, explican, ejercen una influencia tan poderosa que el currículo, la enseñanza y las prácticas de evaluación se orientan a lograr buenos resultados en las pruebas. Un entrevistado (extracto 3) sugiere que las escuelas priorizan estos exámenes en la medida en que los estudiantes son entrenados en estrategias para tomar exámenes desde una edad muy temprana. De manera similar, el maestro de escuela pública describe no solo alinear las pruebas finales con el formato nacional sino también dedicar una clase semanal específicamente a capacitar a los estudiantes para las Pruebas Saber 11[2], comenzando en sexto grado (cuando el examen en sí se administra en undécimo grado). Este énfasis en la preparación para las pruebas resalta la función de las pruebas no solo como medidas de control y rendición de cuentas, sino también como un marcador de prestigio escolar y una fuente de competencia entre escuelas. Los puntajes de las pruebas, según este punto de vista, se consideran un indicador de la calidad general de la escuela.

Extracto 5: Sch/Adm-PB-CAL
Bueno uno es el efecto que tienen las pruebas en el currículo de las instituciones públicas en el país es que hay instituciones que encarrilan o enfocan su currículo a que los estudiantes se preparen para presentar una prueba y subir indicadores en la institución parte de esas consecuencias es que haya más prestigio para las instituciones lo cual digamos es un poco irónico porque sube el prestigio tomando como base los resultados de una prueba que es limitada porque es la naturaleza de las pruebas.

El extracto anterior refuerza el argumento de que las pruebas nacionales limitan el currículo y la evaluación, al tiempo que destaca la consecuencia no deseada de una mayor competencia entre las escuelas debido a su mal uso. El entrevistado señala la ironía de utilizar los puntajes de las pruebas como una medida del prestigio de la escuela, dadas las limitaciones

[2] Pruebas nacionales colombianas que se deben presentar en el último año escolar. Estas pruebas incluyen, entre otras materias, un componente de inglés.

VERSIÓN PREPRINT

inherentes de las pruebas estandarizadas para capturar de manera integral el conocimiento de los estudiantes. Esto es particularmente cierto, como se argumenta en este proyecto, en el caso de las Pruebas Saber, que se centran en un conjunto restringido de habilidades. Finalmente, el extracto sugiere que la influencia de las Pruebas Saber 11 puede extenderse más allá de las prácticas en el aula e incluso dar forma a las estrategias de políticas nacionales.

> **Extracto 6:** PM-PB-DC
> Be the One Challenge, que este sí me gustaría mostrártelo, para nosotros es muy importante porque en este momento ya más de 400.000 estudiantes la han descargado y la han utilizado. Estamos en la versión 3.0, es una versión de una aplicación que es gamificada, y si tú ves la app y la abres, es básicamente la Prueba Saber de Inglés. Hay cuatro niveles, el estudiante escoge en qué nivel están, los cuatro niveles que son pre A1, A1, A2, B1 y cuando los niños escogen cada una de esas misiones, dentro de cada misión tiene siete sub-misiones corresponden a la Prueba Saber en inglés.

El último extracto muestra a uno de los responsables de políticas gubernamentales mencionando la aplicación Be the One Challenge, una iniciativa política que gamifica las herramientas de evaluación lingüística. Si bien este enfoque representa una nueva dirección interesante en la evaluación de políticas lingüísticas, la política en sí está orientada a capacitar a los estudiantes para las Pruebas Saber 11. El participante explica que toda la aplicación refleja el formato de prueba nacional. Esto resalta el amplio alcance de la Prueba Saber 11, que influye no solo en las prácticas en el aula sino también en la configuración de las demás estrategias nacionales de desarrollo lingüístico.

Como afirma Wall (1997), el énfasis excesivo en la preparación para las pruebas, como priorizar la lectura (López *et al.*, 2021), puede llevar a descuidar el pensamiento crítico, la resolución de problemas y las experiencias auténticas de aprendizaje de idiomas. Esta orientación a la preparación para la prueba va más allá de los salones de clase y condiciona las estrategias de políticas nacionales y los esfuerzos gubernamentales de desarrollo lingüístico, como el desafío Be the One, que, a pesar de pretender ser una herramienta de evaluación formativa, está orientado a las pruebas. Alineándose de nuevo con un paradigma racionalista (Usma, 2015), este interés por

VERSIÓN PREPRINT

las pruebas sugiere una mayor preocupación por elevar los indicadores de las pruebas que por asegurar realmente procesos llamativos y significativos de aprendizaje de lengua fuera del aula.

Esto demuestra la alta influencia de las pruebas de Estado en la educación en lenguas adicionales en el contexto colombiano. Las pruebas, que originalmente eran utilizadas para reforzar políticas lingüísticas, se han vuelto a su vez políticas *de facto* que promueven aún prácticas descontextualizadas de enseñanza de inglés a pesar de que, en el papel, políticas lingüísticas más recientes traten de separarse de estos objetivos. Como lo afirma Menken (2017), estas implicaciones rara vez se cuestionan, pero tienen un efecto profundo no solo en las pruebas sino también en los educadores y estudiantes, y en los procesos de instrucción y aprendizaje de lenguas adicionales.

Como se presentó anteriormente, la validez de una prueba va más allá de su contenido o constructos a evaluar, y se debe extender hasta los usos y consecuencias que pueden presentar. Messick (1989) subraya la importancia de que las pruebas se utilicen debidamente y produzcan más consecuencias positivas que negativas, tanto a nivel de aula como social. En este sentido, se podría cuestionar la validez consecuencial de esta prueba. Esto se debe a que el mal uso de la prueba —como instrumento de rendición de cuentas, como regulador del acceso a mayores oportunidades de estudio, en especial para los estudiantes más marginados, o como mecanismo de estandarización de las prácticas educativas en la mayoría del territorio nacional— puede limitar el alcance de los currículos y las posibilidades de desarrollo global de lenguas adicionales.

Prácticas evaluativas heteroglósicas emergentes

A pesar de la prevalencia de prácticas monoglósicas, en este caso orientadas al inglés, reforzadas por las políticas *de facto* en el sistema educativo, los datos sugieren que está surgiendo entre los docentes, académicos y administrativos un cambio hacia prácticas de evaluación bilingües más heteroglósicas, es decir, visiones más inclusivas que abogan por la coexistencia de múltiples lenguas. En este caso, estas visiones propenden por reconocer la importancia de la primera lengua de los estudiantes y los usos dinámicos de ambos códigos en la comunicación. Este cambio abraza el papel del repertorio lingüístico completo de los estudiantes en su proceso de aprendizaje de la lengua y el contenido. Los siguientes extractos pueden servir como ilustración:

VERSIÓN PREPRINT

Extracto 7: Sch/TD-PV-DC

Como te digo, le damos mucha importancia cuando hablamos de bilingüismo a la lengua extranjera y no nos podemos olvidar de la lengua materna, sobre todo evaluarla. De ese proceso del uso del lenguaje muchos de esas habilidades o competencias se pueden transferir a la segunda lengua.

Extracto 8: Sch/Adm-PB-CAL

Pero digamos como la tendencia grande, que seguirá habiendo es el uso de la evaluación por tareas para activar la competencia de lengua de los estudiantes. O sea que no es evaluar el bilingüismo, sino lo que se puede hacer con el bilingüismo … Si el enfoque no es filosofía, sino inglés, como lengua extranjera ¿cómo puedo yo crear, pensar estrategias pedagógicas donde yo pueda evaluar en inglés como lengua extranjera activando el bilingüismo de mis estudiantes? Es decir, si yo quiero evaluar un aspecto la comprensión de escucha de mis estudiantes en inglés, ¿cómo puedo diseñar un instrumento de evaluación que tome digamos el español de mis estudiantes para poder yo mirar y evaluar su comprensión de escucha en inglés? Entonces digamos bilingüismo como un recurso en el área de aprendizaje de inglés como lengua extranjera.

Estos fragmentos anteriormente presentados son una muestra del creciente valor que están tomando las visiones heteroglósicas en las instancias evaluativas. Los participantes, como se muestra en los extractos 7 y 8, identifican la necesidad de incluir contextualmente la primera lengua de los estudiantes en los procesos instruccionales orientados a una lengua adicional. Por ejemplo, el participante 3 menciona la relevancia de priorizar la lengua objetivo sin dejar de lado los apoyos que una primera lengua puede ofrecer. Por su parte, el participante 6 habla del potencial que tiene la primera lengua, por ejemplo, para demostrar los entendimientos que tienen bilingües emergentes en un proceso receptivo, es decir, recurrir a la lengua materna como apoyo en el proceso institucional, y de tal manera involucrarse en prácticas más bilingües o heteroglósicas de la educación en lenguas adicionales.

Estos puntos de vista posicionan el repertorio del bilingüismo como un recurso valioso tanto en la adquisición de lenguas como en su evaluación. Este cambio se alinea con el giro

VERSIÓN PREPRINT

multilingüe más amplio en la lingüística aplicada acuñado por Conteh y Meier (2014), que enfatiza una visión holística del desarrollo del lenguaje. Esto también sugiere una posible influencia del discurso académico en las perspectivas de las partes interesadas. El multilingüismo promueve una visión más holística de la adquisición del lenguaje (Lafford, 2007), un concepto que puede ser relativamente nuevo en las prácticas de evaluación como lo señaló una parte interesada: "las grandes discusiones sobre el bilingüismo como un constructo en la evaluación de lenguas extranjeras comenzaron alrededor de 2017 o 2019, por lo que es relativamente nuevo" (Sch/Adm-PB-CAL).

El enfoque heteroglósico, como lo describen Lopez *et al*. (2017), permite la integración de prácticas multilingües como el cambio de código y el translenguaje en las evaluaciones. Esto contrasta con las evaluaciones tradicionales, en las que se espera que los estudiantes tengan un desempeño monolingüe, a pesar de que podrían tener un mejor desempeño si tuvieran acceso a su lengua materna. En este sentido, los actores colombianos pueden estar creando evaluaciones más inclusivas y efectivas que aprovechen los diversos repertorios lingüísticos de los aprendices en el proceso de instrucción y evaluación.

VERSIÓN PREPRINT

Conclusiones

Respecto a la pregunta planteada en este estudio, se puede concluir que, si bien los discursos recientes han comenzado a virar hacia paradigmas más alternativos y heteroglósicos, las políticas tradicionales dominantes (como las pruebas estandarizadas de alto riesgo) siguen ejerciendo una influencia significativa sobre las instituciones, en particular las escuelas. Estas pruebas, que sirven como políticas *de facto* debido a su uso con fines de rendición de cuentas y sus altas consecuencias sociales para los estudiantes (Menken, 2017), han llevado a las escuelas a enfocarse en gran medida en mejorar los resultados en las pruebas nacionales. Este enfoque orientado a los exámenes nacionales tiene una serie de desventajas a nivel educativo, como la limitación de un desarrollo lingüístico más amplio e integral en los estudiantes, reducción de prácticas de evaluación diversas, limitación de los planes de estudio y fortalecimiento de enfoques monoglósicos para el aprendizaje del inglés (Guerrero, 2008).

A pesar de estas fuertes influencias, los docentes, administrativos, académicos y diseñadores de pruebas se han apropiado de su proceso educativo y su rol político. De esta

manera, no solo han buscado resistir ciertas presiones políticas en sus propias prácticas, sino que también han participado activamente en los debates sobre políticas lingüísticas y su implementación y, en algunos casos, han influido mediante la presión en la propia formulación de políticas lingüísticas. Frente a las prácticas de resistencia en el aula, se identificaron prácticas emergentes manifestadas por los participantes, como métodos de evaluación alternativos, la integración de la tecnología de manera formativa y formulaciones e interpretaciones heteroglósicas de la evaluación lingüística. Aspectos como la integración tecnológica pueden tanto promover prácticas alternativas como reforzar prácticas tradicionales monoglósicas y orientadas a la supervisión, dependiendo de cómo se incorporen (Lopez *et al*., 2014). Aunque estas nuevas iniciativas evaluativas aún no son tan comunes, su priorización en los enfoques formativos de evaluación refleja la influencia de los docentes y académicos que han abogado por una transformación en la evaluación de lenguas, aspecto prometedor para la evaluación bilingüe en el país.

Reconocimientos

El presente artículo se deriva del trabajo de grado del autor para obtener el título de Magíster en Educación Bilingüe de la Universidad Tecnológica de Pereira.

Referencias

Baker, C. (2001). *Foundations of bilingual education and bilingualism (3.* ª ed.). Multilingual Matters.

Bastidas, J. G. y Jiménez-Salcedo, J. (2021). La política lingüística educativa en Colombia: análisis de la literatura académica sobre el Plan Nacional de Bilingüismo. *Tonos Digital,* 41,1-31.

Bowen, G.A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27-40. https://doi.org/10.3316/QRJ0902027

Brown, H. D. y Abeywickrama, P. (2018). Language assessment: Principles and classroom practices (3.ª ed.). Pearson Education.

VERSIÓN PREPRINT

Cely, B. L. y Osorio, M. (2024). Colombia towards bilingualism policies, achievements, failures, and challenges. *Kurdish Studies*, *12*(2), 1277-1294.

Chapelle, C.A. Kremmel, B. y Brindley, G. (2020). Assessment. En N. Schmitt (Ed.), *An introduction to applied linguistic (*pp 247-267). Hodder Education.

Chapelle, C.A. y Plakans, L. (2012). Assessment and testing: Overview. En C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0603

Congreso de la República de Colombia (12 de julio de 2013). Ley 1651, por medio de la cual se modifican los artículos 13, 20, 21, 22, 30 y 38 de la Ley 115 de 1994 y se dictan otras disposiciones-Ley de bilingüismo. (1994). Recuperado de: http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/2013/LEY%201651%20DEL%2012%20DE%20JULIO%20DE%202013.pdf

Conteh, J. y Meier, G. (2014). *The multilingual turn in language education: Opportunities and challenges*. Multilingual Matters.

Correa, D. y Usma, J. (2013). From a bureaucratic to a critical-sociocultural model of policymaking in Colombia. *HOW Journal*, *20*(1), 226-242. https://www.howjournalcolombia.org/index.php/how/article/view/3

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Fandiño-Parra, Y. J., Bermúdez-Jiménez, J. R. y Lugo-Vásquez, V. E. (2012). Retos del Programa Nacional de Bilingüismo. Colombia Bilingüe. *Educación y Educadores*, 15(3), 363-381.

Fraenkel, J. R., Wallen, N. E., y Hyun, H. H. (2012). *How to design and evaluate research in education* (8.ª ed.). McGraw-Hill.

VERSIÓN PREPRINT

García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Wiley.

García, O., Kleifgen, J. A. y Falchi, L. (2008). *From English language learners to emergent bilinguals*. Equity Matters: Research Review No. 1. A Research Initiative of the Campaign for Educational Equity.

Guerrero, C. H. (2008). Bilingual Colombia: What does it mean to be bilingual within the framework of the National Plan of Bilingualism. *Profile: Issues in Teachers' Professional Development*, 10, 27-45.

Guerrero-Nieto, C. H. y Quintero-Polo, Á. H. (2021). Emergence and development of a research area in language education policies: Our contribution to setting the grounds for a local perspective on policymaking. *HOW*, *28*(3), 119-133.

Jaramillo-Calderon, L. F. (2023). *The influence of education policies for bilingualism in stakeholder's language assessment practices and views in West-Center Colombia language education: A multicase study*. [Tesis de maestría, Universidad Tecnológica de Pereira] UTP Campus Repository. https://repositorio.utp.edu.co/items/738955dd-f46c-42c9-9a7b-db02d7e5541d

Lafford, B. (2007). Second language acquisition reconceptualized? The impact of Firth and Wagner. *The Modern Language Journal*, 91, 735-756.

Levinson, B. A. U. y Sutton, M. (2001). Introduction: Policy as/in practice— A sociocultural approach to the study of educational policy. En M. Sutton (Ed.), *Policy as practice: Towards a comparative sociocultural analysis of educational policy* (pp. 1-22). Ablex.

Lopez, A. (2009). *Impact of a policy-driven test effort on foreign language programs in Colombian universities*. [Ponencia] at Pensando Educación. Universidad de los Andes, Colombia.

VERSIÓN PREPRINT

Lopez, A. A., Guzman-Orth, D. y Turkan, S. (2014). *A study on the use of translanguaging to assess the content knowledge of emergent bilingual students*. [Ponencia] Invited Colloquium: Negotiating the Complexities of Multilingual Assessment, at the annual American Association for Applied Linguistics (AAAL) Conference.

Lopez, A. A., Turkan, S. y Guzman-Orth, D. (2017). *Conceptualizing the use of translanguaging in initial content assessments for newly arrived emergent bilingual students* (Research Report No. RR-17-07). Educational Testing Service. https://doi.org/10.1002/ets2.12140

López, M., Ramírez, O., y Arango, M. (2021). Enseñanza-aprendizaje del inglés bajo las políticas del Ministerio de Educación en Colombia: voces desde el aula de clase. *Revista Boletín Redipe*, *10*(12), 200-215.

Mackenzie, L. (2020). A critical evaluation of the social justice implications of the Colombian government's English-Spanish bilingualism policies. *Language Planning*. *21*(2), 117-134. https://doi.org/10.1080/14664208.2019.1615743

McNamara, T. (1998). Policy and social considerations in language assessment. *Annual Review* of *Applied Linguistics*, 18, 304-319.

Menken, K. (2017). *High-stakes tests as de facto language education policies*. En E. Shohamy, L. Or, y S. May (Eds.), Language testing and assessment. (3ª ed., pp. 185-196). Springer.

Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Research*, 10, 9-20.

Messick, S. (1989). Validity. En R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). American Council on Education: Macmillan.

Ministerio de Educación Nacional (MEN). (2004). *Programa Nacional de Bilingüismo. Colombia 2004- 2019*. https://www.mineducacion.gov.co/1621/articles-132560_recurso_pdf_programa_nacional_bilinguismo.pdf

VERSIÓN PREPRINT

Ministerio de Educación Nacional (MEN). (2021) *Plan sectorial 2018-2022: Pacto por la equidad, pacto por la educación* Recuperado de https://www.mineducacion.gov.co/1759/articles-349495_recurso_140.pdf

Miranda, I. R., y Fajardo, A. (2022). Políticas de bilingüismo: reconfiguración discursiva del concepto de nación y ciudadano bilingües. *Cuadernos de Lingüística Hispánica,* (39).

Mølstad, C., Prøitz, T. y Dieude, A. (2020). When assessment defines the content: Understanding goals in between teachers and policies. *The Curriculum Journal, 32*(2), 290-314.

Perrow, C. (1986). Complex organizations: A critical essay (3ª ed.). Random House.

Shohamy, E. (Ed.). (2017). *Language testing and assessment*. (3ª ed.). Springer.

Spolsky, B. (2004). *Language policy*. Cambridge University Press.

Stake, R. E. (2013). *Multiple case study analysis*. Guilford Publications.

Tollefson, J. W. (2006). Critical theory in language policy. En T. Ricento (Ed.), *An introduction to language policy: Theory and method*. Blackwell.

Turnbull, B. (2017). Towards new standards in foreign language assessment: Learning from bilingual education. *International Journal of Bilingual Education and Bilingualism*, 23(4), 488–498. https://doi.org/10.1080/13670050.2017.1375891

Usma, J. (2015). *From transnational language policy to local appropriation: The case of the National Bilingual Program in Medellín, Colombia*. Deep University Press.

Valencia, A. y Vega, L. A. (2022). Caracterización del programa nacional de bilingüismo en los últimos cinco años (un enfoque a las percepciones de los estudiantes). *Ingenio Libre*, *10*(20), 79-100. https://doi.org/10.18041/2322-8415/ingelibre.2022.v10n20.9365

Wall, D. (1997). Impact and washback in language testing. En C. Clapham y D. Corson (Eds.), *Encyclopedia of language and education* (pp. 291-302). Dordrecht: Kluwer Academic Publications.

VERSIÓN PREPRINT

Anexo

Autor(es) (Apellido y Nombre)	Año de publicación	Título del documento	Tipo de documento (artículo, libro, tesis doctoral, etc)	Fuente de acceso
PM-PB-DC	2023	Interview: PM-PB-DC	Transcripción entrevista	https://docs.google.com/document/d/1vszbgOLXqJaJnDGpJTMH6K27dqZ9GBcL/edit?usp=sharing&ouid=100168795783332986871&rtpof=true&sd=true
TD/Ter-PB-RIS	2023	Interview-TD/Ter-PB-RIS	Transcripción entrevista	https://docs.google.com/document/d/1_GnQENe4mUH4vjgmNoDrj1wryfZ3gGevLUvduhgLPuI/edit?usp=sharing
Sch/TD-PV-DC	2023	Interview: Sch/TD-PV-DC	Transcripción entrevista	https://docs.google.com/document/d/1OzjeiCI14XKwMZieXgptQTGVNYpJesh5i3urWJ9fw-Q/edit?usp=sharing
Ter-PV-RIS	2023	Interview: Ter-PV-RIS	Transcripción entrevista	https://docs.google.com/document/d/1wr4nvOslZ666h42fcDCFtiFKt4A2_ontB8rBchKikyA/edit?usp=sharing
Sch/Adm-PB-ANT	2023	Interview: Sch/Adm-PB-ANT	Transcripción entrevista	https://docs.google.com/document/d/1xF-7Qing_7XSQw4WUHdqZVwyHqluxJk4JprilYuYtao/edit?usp=sharing

Sch/Adm-PB-CAL	2023	Interview: Sch/Adm-PB-CAL	Transcripción entrevista	https://docs.google.com/document/d/1YRYV_WBWqw-UDaagmx78u55uBRPoLU1ll0Y-FB7roU4/edit?usp=sharing
Ter-PB-RIS	2023	Interview: Ter-PB-RIS	Transcripción entrevista	https://docs.google.com/document/d/1c0J45SSGjf8626hJ5f68hsz8NEsOECDZY1AeVrTyD2Q/edit?usp=sharing
Adm-PV-VAC	2023	Interview: Adm-PV-VAC	Transcripción entrevista	https://docs.google.com/document/d/1IDmz1HhRbsfSvbU_ciO0oCQK5A3I1YIr-q0VNYFgNWM/edit?usp=sharing
PM/Adm-PB-QUI	2023	Interview: PM/Adm-PB-QUI	Transcripción entrevista	https://docs.google.com/document/d/1yd3WvC8tOGO9Mco8G_-whjfkGMXxpseHFts7BxQgQCo/edit?usp=sharing
Congreso de la República de Colombia	2013	Ley 1651 de 2013	Ley	https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=53770
Ministerio de Educación	2022	PLAN SECTORIAL DE EDUCACIÓN 2018 – 2022	documento de política	https://www.mineducacion.gov.co/1780/articles-349495_recurso_141.pdf