

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE LETRAS E LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

BRUNO CHAVES BORJA

VOZES DO SUL: EXPLORANDO ESTUDOS SOBRE AVALIAÇÃO NO ENSINO DE
LÍNGUAS EM PERIÓDICOS LATINO-AMERICANOS

UBERLÂNDIA/MG

2025

BRUNO CHAVES BORJA

VOZES DO SUL: EXPLORANDO ESTUDOS SOBRE AVALIAÇÃO NO ENSINO DE
LÍNGUAS EM PERIÓDICOS LATINO-AMERICANOS

Tese apresentada ao Programa de Pós-graduação em Estudos Linguísticos do Instituto de Letras e Linguística da Universidade Federal de Uberlândia, como requisito para obtenção do título de Doutor em Estudos Linguísticos.

Área de Concentração: Estudos em Linguística e Linguística Aplicada.

Linha de pesquisa: Linguagem, ensino e sociedade.

Orientadora: Profa. Dra. Marileide Dias Esqueda.

UBERLÂNDIA/MG

2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

B734
2025 Borja, Bruno Chaves, 1991-
Vozes do Sul [recurso eletrônico] : explorando estudos sobre
avaliação no ensino de línguas em periódicos latino-americanos /
Bruno Chaves Borja. - 2025.

Orientadora: Marileide Dias Esqueda.

Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-
graduação em Estudos Linguísticos.

Modo de acesso: Internet.

DOI <http://doi.org/10.14393/ufu.te.2025.482>

Inclui bibliografia.

Inclui ilustrações.

1. Linguística. I. Esqueda, Marileide Dias, 1973-, (Orient.). II.
Universidade Federal de Uberlândia. Pós-graduação em Estudos
Linguísticos. III. Título.

CDU: 801

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Estudos Linguísticos

Av. João Naves de Ávila, nº 2121, Bloco 1G, Sala 1G256 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4102/4355 - www.ileel.ufu.br/ppgel - secppgel@ileel.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Estudos Linguísticos				
Defesa de:	Tese de Doutorado - PPGEL				
Data:	Vinte e cinco de junho de dois mil e vinte e cinco	Hora de início:	08:30	Hora de encerramento:	11:30
Matrícula do Discente:	12113ELI006				
Nome do Discente:	Bruno Chaves Borja				
Título do Trabalho:	Vozes do Sul: Estudos sobre avaliação no ensino de línguas em periódicos latinos-americanos				
Área de concentração:	Estudos em Linguística e Linguística Aplicada				
Linha de pesquisa:	Linguagem, ensino e sociedade				
Projeto de Pesquisa de vinculação:	Ensino e aprendizagem de tradução e interpretação: pressupostos teóricos, epistemológicos e metodológicos				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Estudos Linguísticos, assim composta pelos Doutores: Sandra Dias Loguercio - UFRGS; Márcio Issamu Yamamoto - UFJ; Letícia de Sousa Leite - UFU; Valeska Virgínia Soares Souza - UFU; Marileide Dias Esqueda - UFU, orientadora da Tese.

Iniciando os trabalhos a presidente da mesa, Dra. Marileide Dias Esqueda, apresentou a Comissão Examinadora e o candidato agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir, a senhora presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar, foram encerrados os trabalhos. Foi lavrada a presente ata que, após lida e achada conforme, foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Márcio Issamu Yamamoto, Usuário Externo**, em 25/06/2025, às 11:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Valeska Virginia Soares Souza, Professor(a) do Magistério Superior**, em 25/06/2025, às 12:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Letícia de Sousa Leite, Usuário Externo**, em 25/06/2025, às 12:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marileide Dias Esqueda, Professor(a) do Magistério Superior**, em 25/06/2025, às 12:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Sandra Dias Loguercio, Usuário Externo**, em 25/06/2025, às 12:24, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **6451097** e o código CRC **E4E1EE1C**.

AGRADECIMENTOS

Gostaria de expressar meu sincero agradecimento à minha orientadora, Marileide Dias Esqueda, pela orientação excepcional, paciência e acolhimento ao longo de nossa trajetória conjunta. Suas palavras de apoio, incentivo e orientação acolhedora tornaram minha jornada como pesquisador mais prazerosa e enriquecedora. Desde o início da nossa relação, em 2010, durante a graduação, temos construído uma parceria leve e profundamente significativa. Espero poder me inspirar na sua prática docente e transmitir todo o carinho dessa relação aos meus futuros alunos. Marileide é uma inspiração como ser humano. Minha mais profunda admiração e respeito!

Aos meus pais: sinto-me verdadeiramente privilegiado por ter pais que me incentivaram nos estudos desde a infância e que nunca mediram esforços para garantir que eu tivesse acesso aos melhores professores, escolas e cursos. Reconheço que todos os sacrifícios valeram a pena para que eu e minha irmã pudéssemos receber uma educação de qualidade. Obrigado, mãe, por nos ensinar, através do exemplo, a ter disciplina e compromisso. Pai, agradeço por nos ensinar o valor da leitura e da informação. Sou profundamente grato a vocês.

Ao Kélltony, agradeço por me proporcionar companheirismo e leveza ao longo de toda a minha rotina nos últimos meses. Obrigado pela escuta atenta quando mais precisei nos dias cansativos de estudos e trabalho.

Aos meus alunos, agradeço o companheirismo e o fato de me inspirarem a ser um profissional mais completo e preparado para vocês. Obrigado por confiarem no meu trabalho e serem tão fiéis. As nossas trocas são sempre lembradas com muito carinho.

Agradeço a todas as pessoas amadas, assim como aos profissionais da saúde que puderam me auxiliar em um momento delicado no Hospital das Clínicas da Universidade Federal de Uberlândia.

RESUMO

As práticas de avaliação impactam estudantes, pais, professores e diretores, embora a magnitude de suas consequências varie conforme o propósito e tipo de instrumento avaliativo utilizado. Ouvir as vozes dos estudantes submetidos a avaliações alinha-se com as teorias decoloniais, pois essas vozes raramente são ouvidas e dificilmente geram mudanças sistêmicas nas práticas avaliativas. Os avaliados são os mais afetados pelas repercussões negativas das avaliações, que afetam sua autoestima e limitam suas oportunidades. Grande parte da produção acadêmica do Sul Global se dirige ao Norte Global, destacando a necessidade de questionar a colonialidade do conhecimento e defender a inclusão de epistemologias do Sul. Este estudo descritivo utiliza indicadores bibliométricos para examinar pesquisas latino-americanas sobre avaliação no ensino de idiomas na base de dados da SCImago. Todas as plataformas acessíveis dos respectivos periódicos foram revisadas para identificar artigos relevantes sobre o tema. Dentre as principais conclusões do estudo, destacam-se, nesses artigos, o Letramento em avaliação por parte do corpo docente e uma mudança gradual rumo a práticas avaliativas mais diversas e qualitativas.

Palavras-chave: avaliação da aprendizagem; ensino de línguas; decolonização do conhecimento; vozes dos estudantes; instrumentos avaliativos; pesquisa bibliométrica.

RESUMEN

Las prácticas de evaluación impactan a estudiantes, padres, profesores y directores, aunque la magnitud de sus consecuencias varía según el propósito y el tipo de instrumento evaluativo utilizado. Escuchar las voces de los estudiantes sometidos a evaluaciones se alinea con las teorías decoloniales, ya que estas voces rara vez son escuchadas y difícilmente generan cambios sistémicos en las prácticas evaluativas. Los evaluados son los más afectados por las repercusiones negativas de las evaluaciones, que afectan su autoestima y limitan sus oportunidades. Gran parte de la producción académica del Sur Global se dirige hacia el Norte Global, destacando la necesidad de cuestionar la colonialidad del conocimiento y defender la inclusión de epistemologías del Sur. Este estudio descriptivo utiliza indicadores bibliométricos para examinar investigaciones latinoamericanas sobre evaluación en la enseñanza de idiomas en la base de datos de SCImago. Se revisaron todas las plataformas accesibles de los respectivos periódicos para identificar artículos relevantes sobre el tema. Entre las principales conclusiones del estudio, hubo, en dichos artículos, una preocupación especial respecto al alfabetismo en evaluación por parte del cuerpo docente, así como un cambio gradual hacia prácticas de evaluación más diversas y cualitativas.

Palabras-clave: evaluación educativa; enseñanza de lenguas; decolonización del conocimiento; voces de los estudiantes; instrumentos evaluativos; investigación bibliométrica.

ABSTRACT

Assessment practices impact students, parents, teachers, and principals, although the magnitude of their consequences varies depending on the purpose and type of assessment instrument applied. Listening to the voices of students subjected to evaluations aligns with decolonial theories, as these voices are rarely heard and seldom generate systemic changes in assessment practices. Those evaluated are the most affected by the negative repercussions of assessments, which impact their self-esteem and limit their opportunities. A significant portion of the academic output from the Global South is directed toward the Global North, highlighting the need to question the coloniality of knowledge and advocate for the inclusion of Southern epistemologies. This descriptive study uses bibliometric indicators to examine Latin American research on assessment in language teaching within the SCImago database. All accessible platforms of the respective journals were reviewed to identify relevant articles on the subject. Among the main conclusions of the study, there was a particular concern, in these articles, regarding assessment literacy among teachers, as well as a gradual shift toward more diverse and qualitative assessment practices.

Keywords: educational assessment; language teaching; decolonization of knowledge; student voices; assessment instruments; bibliometric research.

LISTA DE FIGURAS

Figura 1 – Perguntas sobre avaliação no <i>site</i> chileno	51
Figura 2 – Número de periódicos por país	61

LISTA DE GRÁFICOS

Gráfico 1 – Nomes dos periódicos e quantidade de artigos por publicação	62
Gráfico 2 – Publicações no período de 1994 a 2023	63
Gráfico 3 – Línguas nas quais os artigos foram escritos.....	64
Gráfico 4 – Citações nas publicações latino-americanas	65

LISTA DE QUADROS

Quadro 1 – Natureza e função dos tipos de instrumentos avaliativos	23
Quadro 2 – Periódicos da base de dados da SCImago	57
Quadro 3 – Peronard (1994)	65
Quadro 4 – Mourente Miguel (2004)	68
Quadro 5 – Ashraf Rizvi (2005)	68
Quadro 6 – Muñoz e Álvarez (2008)	71
Quadro 7 – Muñoz Marín (2009)	72
Quadro 8 – Benítez Figari (2009)	73
Quadro 9 – López Mendoza e Bernal Arandia (2009)	75
Quadro 10 – León Chica, D’Costa Martínez e Franco Jácome (2010)	77
Quadro 11 – Livingstone (2011)	78
Quadro 12 – Muñoz, Palacio e Escobar (2012)	79
Quadro 13 – Baleghizadeh e Zarghami (2012)	82
Quadro 14 – Baitman e Véliz Campos (2012)	83
Quadro 15 – Pineda (2014)	84
Quadro 16 – Neira Martínez, Reyes Reyes, Riffo Ocares (2014)	85
Quadro 17 – Riffo Ocares <i>et al.</i> (2014)	87
Quadro 18 – Furlan, Ferrero e Gallart (2014)	88
Quadro 19 – Vine Jara e Ferreira Cabrera (2015)	88
Quadro 20 – Palacio, Gaviria e Brown (2016)	89
Quadro 21 – Farías (2016)	92
Quadro 22 – Pastene Labrín, Díaz Chavarría e Molina Castillo (2016)	93
Quadro 23 – Guzmán, Véliz e Reyes (2016)	95
Quadro 24 – Abad e Alzate (2016)	96
Quadro 25 – Amengual-Pizarro e García-Laborda (2017)	98
Quadro 26 – Muñoz e Ferreira (2017)	98
Quadro 27 – Vega Rodríguez, Torres Rodríguez e Del Campo Rivas (2017)	99
Quadro 28 – Gutiérrez (2018)	101
Quadro 29 – Martín Alonso (2018)	102
Quadro 30 – Scliar-Cabral (2018)	103
Quadro 31 – Alfaro Chat, Gómez Álvarez e Saez Carrillo (2018)	105
Quadro 32 – Giraldo (2018)	107

Quadro 33 – Gallego Ortega, Figueroa Sepúlveda e Rodríguez Fuentes (2019)	107
Quadro 34 – Saavedra-Jeldres e Campos-Espinoza (2019)	108
Quadro 35 – Machado Key	109
Quadro 36 – Andiñ Herrero e Criado de Diego (2019)	110
Quadro 37 – Figueroa, Meneses e Chandía (2019)	111
Quadro 38 – Novoa Lagos (2019)	112
Quadro 39 – Giraldo (2020)	115
Quadro 40 – Duque-Aguilar (2020).....	115
Quadro 41 – Cárcamo (2020).....	116
Quadro 42 – Benavides (2020).....	117
Quadro 43 – Lizasoain e Toledo Veja (2020)	118
Quadro 44 – Muñoz Moreno e Mavrou (2020)	119
Quadro 45 – Giraldo (2021)	122
Quadro 46 – Cruz-Ramos e Herrera-Díaz (2021).....	123
Quadro 47 – Cadena-Aguilar e Álvarez-Ayure (2021).....	122
Quadro 48 – Vallejos Carrasco, Véliz de Vos e Sáez Carrillo (2021).....	125
Quadro 49 – Forero-Mondragón e Quintero-Polo (2021).....	126
Quadro 50 – González (2021)	127
Quadro 51 – Almeida, Castillo e Rosero (2021)	128
Quadro 52 – Nayernia, Nosrati e Mohebbi (2022)	129
Quadro 53 – Boudon <i>et al.</i> (2022)	130
Quadro 54 – Tajeddin, Saeedi e Panahzadeh (2022)	131
Quadro 55 – Martínez Giménez e Sellés Nohales (2022).....	133
Quadro 56 – Taboh <i>et al.</i> (2022)	134
Quadro 57 – Iacobuzio <i>et al.</i> (2022)	135
Quadro 58 – Valdés-Léon, González Riffo e Molina Olivares (2022).....	136
Quadro 59 – Valenzuela Muñoz (2022).....	137
Quadro 60 – Díaz Larenas <i>et al.</i> (2022).....	138
Quadro 61 – Ramos-Leiva <i>et al.</i> (2022).....	139
Quadro 62 – Espinosa Aguirre, Figueroa Miralles e Ávila Reyes (2022)	140
Quadro 63 – Hernández-Ocampo (2022).....	143
Quadro 64 – Sayyadi (2022).....	144
Quadro 65 – Ramírez Santana <i>et al.</i> (2023)	145
Quadro 66 – Özdemir-Yilmazer e Özkan (2023)	146

SUMÁRIO

1 MINHAS MOTIVAÇÕES	15
2 INTRODUÇÃO	18
3 FUNDAMENTAÇÃO TEÓRICA	23
3.1 Avaliação e ensino de línguas.....	25
3.2 Avaliação crítica	27
3.3 Letramento em avaliação	33
3.4 Decolonialidade	38
3.5 Mudanças no cenário da avaliação on-line e o ChatGPT	44
3.6 Marcos Curriculares	48
4 METODOLOGIA DA PESQUISA	54
5 RESULTADOS	58
5.1 Resultados bibliométricos: dados gerais	58
5.2 Quadros com resultados da análise de conteúdo.....	66
5.2.1 <i>La evaluación de la comprensión de textos escritos: el problema del resumen – Revista Lenguas Modernas – 1994.....</i>	66
5.2.2 <i>Teacher correction versus peer-marking – Revista Profile – 2004.....</i>	67
5.2.3 <i>Using students’ analysis in teaching public speaking for business – Revista Profile – 2005</i>	68
5.2.4 <i>Preliminary Evaluation of the Impact of a Writing Assessment System on Teaching and Learning – Revista Profile – 2008</i>	70
5.2.5 <i>Exploring teachers practices – Revista Profile – 2009.....</i>	71
5.2.6 <i>Análisis descriptivo de narraciones – Revista Literatura Y Linguística – 2009</i>	73
5.2.7 <i>Language testing in Colombia: a call for more teacher education and teacher training in language assessment – Revista Profile – 2009.....</i>	74
5.2.8 <i>Designing proficiency tests to accredit previous knowledge in American and British literature in a bilingual education program – Revista Profile – 2010.....</i>	77
5.2.9 <i>La evaluación de un modelo metodológico mixto para mejorar las habilidades productivas y receptivas en español como segunda lengua: un estudio cuasi- experimental – Revista Lengua y Habla – 2011</i>	79
5.2.10 <i>Teachers’ Beliefs about Assessment in an EFL Context in Colombia – Revista Profile – 2012</i>	80

5.2.11 The impact of conferencing assessment on EFL students' grammar learning – Revista Profile – 2012.....	81
5.2.12 A comparison of oral evaluation ratings by native English speaker teachers and non-native English speaker teachers – Revista Literatura y Lingüística – 2012	82
5.2.13 The feasibility of assessing teenagers' oral English language performance with a rubric – Revista Profile – 2014.....	83
5.2.14 Experiencia académica y estrategias de comprensión lectora en estudiantes universitarios de primer año – Literatura e Lingüística – 2014.....	85
5.2.15 Competencia léxica, comprensión lectora y rendimiento académico en estudiantes de enseñanza media – Literatura y Lingüística – 2014	86
5.2.16 Ansiedad Ante Los Exámenes, Procrastinación y Síntomas – Revista Argentina de Ciencias del Comportamiento – 2014	88
5.2.17 Propuesta de un modelo para una prueba con fines específicos académicos en ele – Revista Literatura Y Lingüística – 2015.....	89
5.2.18 Aligning English Language – Revista Profile – 2016.....	90
5.2.19 Evaluación de una intervención pedagógica de alfabetización visual en la comprensión lectora de textos multimodales en programas de formación inicial docente en Lenguaje y Comunicación – Revista Literatura y Lingüística – 2016	92
5.2.20 Producción del Ensayo Argumentativo en Estudiantes de dos Establecimientos de Educación Media Técnico-Profesional de la Provincia de Ñuble. Resultados de un Diagnóstico – Revista Literatura y Lingüística – 2016.....	93
5.2.21 Memoria operativa, comprensión lectora y rendimiento escolar – Revista Literatura y Lingüística – 2016.....	94
5.2.22 Strategies Instruction to Improve the Preparation for English Oral Exams – Revista Profile – 2016.....	96
5.2.23 Analysing Test Takers' Views On A Computer Based Speaking Test – Revista Profile – 2017	97
5.2.24 El feedback correctivo escrito indirecto en el aprendizaje de la forma comparativa de adjetivos en inglés – Revista Logos – 2017	99
5.2.25 Habilidades metamorfológicas y su incidencia en la comprensión lectora – Revista Signos – 2017	100
5.2.26 Habilidades favorecedoras del aprendizaje de la lectura en alumnos de 5 y 6 años Revista Signos – 2018.....	101
5.2.27 Validez y fiabilidad de una escala de evaluación oral en ele – 1 – Revista RLA – 2018	

.....	102
5.2.28 Evaluación de la escritura en niños de tercer grado en Brasil – Revista Boletín De Lingüística – 2018.....	103
5.2.29 Complejidad, precisión y fluidez en el desempeño oral de aprendices con distintos niveles de proficiencia en inglés como L2 – Revista Signos – 2018.....	104
5.2.30 Language Assessment Literacy and Teachers’ Professional Development: A Review of the Literature – Revista Profile – 2018.....	106
5.2.31 La comprensión lectora de escolares de educación básica – Revista Literatura y Lingüística – 2019.....	107
5.2.32 Chilean Pre Service Teachers – Revista Profile – 2019.....	108
5.2.33 La competencia intercultural en la clase de francés lengua extranjera: ¿qué y cómo evaluarla? – Revista Lengua y Habla – 2019.....	109
5.2.34 Consideraciones sobre la variación fonética en español y sus implicaciones para la evaluación internacional de la lengua como LE/L2: dele, siele y celu – Revista RLA – 2019.....	110
5.2.35 Desempeños en la calidad de explicaciones y argumentaciones en estudiantes chilenos de 8° básico – Revista Signos – 2019.....	111
5.2.36 Complejidad textual, vocabulario pasivo y comprensión lectora en escolares chilenos de segundo ciclo básico – Revista Literatura y Lingüística – 2019.....	112
5.2.37 A Post-Positivist and Interpretive Approach to Researching Teachers’ Language Assessment Literacy – Revista Profile – 2020.....	113
5.2.38 Teachers’ Assessment Approaches Regarding EFL Students’ Speaking Skill – Revista Profile – 2020.....	114
5.2.39 Classifying Written Corrective Feedback – Revista Profile – 2020.....	115
5.2.40 Level of English in Colombian Higher Education: A Decade of Stagnation – Revista Profile – 2020.....	116
5.2.41 Evaluación de la producción escrita de estudiantes secundarios haitianos por parte de profesores de Lengua y Literatura en Chile – Revista Estudios Filológicos – 2020.....	117
5.2.42 La influencia del componente no verbal asociado al feedback correctivo en la producción oral en ELE – Revista Signos – 2020.....	119
5.2.43 A Reflection on Initiatives for Teachers’ Professional Development Through Language Assessment Literacy – Revista Profile – 2021.....	120
5.2.44 Assessment of Students’ Oral Communicative Competence in English through a Web Conferencing Platform – Revista Profile – 2021.....	121

5.2.45 Self and Peer Assessment of Student Generated Podcasts – Revista Profile – 2021..	123
5.2.46 Estructura hipertextual y habilidad lectora: Análisis de la producción digital de textos argumentativos en estudiantes de 4° Año Medio – Revista Signos – 2021.....	124
5.2.47 Disciplinary Power Lying Behind the Requisite of English Language Mastery in International Scholarships – Revista Profile – 2021	125
5.2.48 The Impact of Assessment Training – Revista Profile – 2021	126
5.2.49 Assessment practices in some language centers in public Universities of Ecuador and student proficiency outcomes – Revista Literatura y Linguística – 2021.....	128
5.2.50 The Factors Contributing to Language Teachers’ Effectiveness in an EFL Learning Context: A Questionnaire Validation Study – Revista Profile – 2022	129
5.2.51 Mapudunguken: cuestionario para evaluar competencia oral en mapudungun en niños mapuche – Revista Literatura y Lingüística – 2022.....	130
5.2.52 English Language Teachers Perceived Classroom – Revista Profile – 2022	131
5.2.53 El desarrollo del conocimiento de los componentes y funciones del lenguaje escrito en estudiantes prelectores – Revista Signos – 2022	132
5.2.54 Evaluación del lenguaje oral en niños y niñas con hipoacusia: los tests estandarizados y la edad auditiva – Revista Signos – 2022	133
5.2.55 Estrategias y desafíos en la creación de un instrumento de evaluación lingüística infantil – Revista Signos – 2022.....	135
5.2.56 Competencia comunicativa en educación superior: hacia una evaluación diagnóstica integral – Revista Logos – 2022.....	136
5.2.57 Desempeño escrito de los estudiantes de secundaria en Chile: una evaluación directa e indirecta de la escritura – Revista Logos – 2022.....	137
5.2.58 EFL In-Service Teachers’ Self-Reported Perspectives on English Language Assessment – Nueva Revista Del Pacifico – 2022	138
5.2.59 La ansiedad hacia la evaluación en el aprendizaje del inglés como lengua extranjera en estudiantes de enseñanza media – Revista Lenguas Modernas – 2022.....	139
5.2.60 Escribir en L2 en la escuela chilena: Una caracterización de la escritura en español de estudiantes de origen haitiano en 5° básico – Revista Signos – 2022.....	140
5.2.61 Colombian’ Scholars Discussion About Language Assessment: A Review of Five Journals – Revista Profile – 2022	141
5.2.62 In-Service University Level Efl Instructors Language Assessment Literacy and Training Needs Ali Sayyadi – Revista Profile – 2022	143
5.2.63 Intervención en Lectura Temprana en Alumnado con Trastorno del Desarrollo del	

<i>Lenguaje – Revista Signos – 2023.....</i>	<i>145</i>
<i>5.2.64 Challenges and Complexities in Teacher Conceptualization of Dynamic Assessment in the L2 Classroom: A Case Study in Türkiye – Revista Profile – 2023.....</i>	<i>146</i>
6 DISCUSSÃO	148
6.1 Tipos de Avaliação	149
<i>6.1.1 Testes padronizados</i>	<i>149</i>
<i>6.1.2 Rubricas e critérios avaliativos</i>	<i>149</i>
<i>6.1.3 Provas de compreensão leitora e escrita</i>	<i>149</i>
<i>6.1.4 Questionários e escalas</i>	<i>149</i>
6.2 População avaliada.....	150
6.3 Competências avaliadas	156
<i>6.3.1 Ranking de competências mais e menos comuns.....</i>	<i>157</i>
<i>6.3.2 Análise comparativa das competências nos 64 artigos sobre avaliação no ensino de línguas.....</i>	<i>157</i>
6.4 Métodos de instrumentos utilizados	159
6.5 Resultados e conclusões presentes nos 64 artigos sobre avaliação	161
<i>6.5.1 Pontos em comum</i>	<i>161</i>
<i>6.5.2 Diferenças.....</i>	<i>162</i>
<i>6.5.3 Tendências mais comuns</i>	<i>163</i>
<i>6.5.4 Tendências menos comuns.....</i>	<i>163</i>
6.6 Análise quantitativa das tendências (1 a 64)	164
<i>6.6.1 Tendências gerais.....</i>	<i>164</i>
<i>6.6.2 Fatores específicos.....</i>	<i>165</i>
<i>6.6.3 Métodos de avaliação.....</i>	<i>165</i>
6.7 Desafios, limitações e recomendações	165
<i>6.7.1 Semelhanças identificadas nos 64 artigos.....</i>	<i>165</i>
<i>6.7.2 Diferenças destacadas.....</i>	<i>166</i>
<i>6.7.3 Ranking de tendências identificadas</i>	<i>168</i>
<i>6.7.4 Semelhanças nas recomendações</i>	<i>168</i>
6.8 Pontos semelhantes e divergentes entre os artigos	170
<i>6.8.1 Pontos semelhantes</i>	<i>170</i>
<i>6.8.2 Pontos diferentes.....</i>	<i>171</i>
<i>6.8.3 Pontos interessantes</i>	<i>172</i>
6.9 Comparações entre os resultados desta tese e Borja (2020a).....	172

7 CONSIDERAÇÕES FINAIS.....	178
REFERÊNCIAS.....	183
REFERÊNCIAS DO <i>CORPUS</i>.....	186

1 MINHAS MOTIVAÇÕES

Sempre fui afetado negativamente por determinados instrumentos avaliativos ao longo de minha trajetória como estudante e acho que isso nunca foi exclusividade minha. Ao longo de meu percurso escolar, apesar da afinidade com determinadas disciplinas e embora sentisse prazer ao estudá-las, sentia que o interesse por alguns conteúdos diminuía após ser submetido a testes. Os conteúdos que eu havia entendido perfeitamente durante a aula pareciam ser acrescidos de um grau de dificuldade consideravelmente maior ao realizar uma prova. Lembro de sentir indignação pelo fato de algumas frases ou exercícios serem desenhados de determinada maneira, com o intuito de confundir os alunos, e, mesmo quando me sentia seguro quanto ao conteúdo em sala, o resultado em notas não era necessariamente satisfatório nos testes. Como a maioria das crianças, no entanto, não questionaria algo, pois, além de tímido, acreditava que alunos apenas obedeciam.

Tive minha autoestima abalada inúmeras vezes por resultados numéricos de testes, até que me autointitulei inferior e segui com essa crença limitante por anos. Estudar e me esforçar nem sempre era traduzido em notas boas, e aquilo foi frustrante, mas estranhamente libertador. Saía-me bem nos chamados “trabalhos”, tendo em vista que tínhamos tempo para pesquisar, refazer, discutir e simplesmente desfrutar do processo com menos ansiedade e pressão.

Durante os meus anos de graduação, tive um pouco mais de voz para perguntar e questionar determinados instrumentos avaliativos e seus resultados, mas ainda de forma tímida e imatura (por vezes reativo e defensivo). No ensino superior, por estudar disciplinas que realmente me interessavam, ser avaliado parecia mais leve, apesar de ainda atrelado a ansiedade. Quando, anos mais tarde, optei por fazer uma especialização em Formação de Professores, deparei-me com uma disciplina exclusivamente dedicada à Avaliação, que me abriu os olhos e trouxe conforto. Lembro que, a cada leitura, me sentia abraçado por uma mensagem de “o problema não está em você, há vários instrumentos avaliativos possíveis”. Foi gratificante reconquistar a minha autoconfiança pouco a pouco, e me pergunto quantas pessoas nunca mais se sentiram confiantes em suas trajetórias intelectuais, por terem sido submetidas a instrumentos avaliativos pouco diversos e excludentes. Confesso que, após a disciplina sobre Avaliação, foi difícil me interessar por outros textos e debates, pois queria seguir aprofundando meus estudos sobre Avaliação. Foi a disciplina em questão que me despertou a curiosidade de pesquisar sobre o tema em um possível mestrado.

Após muito pesquisar em editais de mestrado, percebi que poucos professores atualmente dedicavam seus estudos ou parte deles à Avaliação, o que, a princípio, me

surpreendeu. Uma vez que tive uma conversa com a minha orientadora e confirmei o seu interesse pela área, dediquei-me aos estudos de artigos que haviam sido publicados em território nacional sobre avaliação no ensino de línguas estrangeiras, com foco em periódicos brasileiros da área de Linguística e Linguística Aplicada com Qualis A1. Os resultados, por mais que abrangessem um período longo de publicações, ainda pareciam escassos. O volume de artigos sobre a área era reduzido, e a diversidade de temas também, muitas vezes enfatizando que o Quadro Europeu Comum deveria ser considerado nas avaliações, revelando um ideal de falante nativo.

Durante o mestrado, certifiquei-me de que cursaria uma disciplina sobre avaliação, assim como, anos mais tarde, no doutorado. Após novas leituras, decidi expandir a pesquisa, já em nível de doutorado, para o território latino-americano.

Acredito que as novas leituras sobre a Avaliação no Ensino de Línguas trazem um panorama ainda mais amplo sobre o tema, para enriquecer a pesquisa antes iniciada nos cursos de especialização e mestrado (Borja, 2020a). Associadas às leituras dos artigos de diversos países latino-americanos, as leituras que realizei sobre Avaliação Crítica, assim como aquelas sobre os Estudos Decoloniais, contribuíram fortemente para que eu entendesse ainda mais a voz dos avaliados. Embora os dados coletados não sejam exaustivos, mas possivelmente impressionísticos, sobretudo em razão dos contextos em que se inserem, espero, com esta tese de doutorado, não só enriquecer a minha prática com uma visão mais acolhedora na sala de aula, como também (com esperança) despertar o interesse da discussão sobre Avaliação, tendo em vista que ela é uma realidade em todas as áreas e em todos os contextos educacionais.

Assim, concebo a avaliação em sala de aula de ensino de línguas como um processo de acompanhamento bastante amplo, realizado ao longo de semestres e anos, no dia a dia da sala de aula, levando em consideração as tentativas, o esforço e os desafios enfrentados pelo aluno.

Além disso, vale ressaltar que esta tese, atrelada à linha 3 de pesquisa em Linguagem, Ensino e Sociedade do PPGEL/UFU (Programa de Pós-graduação em Estudos Linguísticos da Universidade Federal de Uberlândia), adota a concepção de linguagem de Bakhtin (1992), que a considera um processo de interação. O linguista russo, que tem uma importância fundamental nos estudos sobre o discurso, esclarece que a interação verbal constitui, assim, a realidade fundamental da linguagem. Dessa forma, a linguagem, além de exteriorizar os pensamentos e ser veículo de comunicação, é o lugar de interação humana, pela produção de efeitos de sentido entre interlocutores, em uma dada situação de comunicação e em um contexto sócio-histórico-ideológico. Por fim, vale mencionar que esta tese foi pensada e construída junto às discussões

implementadas no Grupo de Estudos e Pesquisa em Tradução, Tecnologias, Ensino e Cienciometria (GETTEC) ao qual pertenço e do qual minha orientadora é líder.

2 INTRODUÇÃO

Independentemente do contexto escolar em questão, a avaliação sempre se faz presente e é acompanhada de um peso considerável em contextos educacionais. É comum que ementas, materiais didáticos e decisões em sala de aula sejam fortemente voltados para a realização de avaliações em momentos específicos do semestre, que geralmente levam em conta os conteúdos contemplados e avaliados nos exames nacionais de maior impacto, a saber, os vestibulares e o Enem no contexto brasileiro, sobretudo porque são utilizados como parâmetros futuros de formação profissional. Segundo dados do IBGE (2011),

O Enem busca oferecer uma referência para autoavaliação com vistas a auxiliar nas escolhas futuras dos cidadãos, tanto com relação à continuidade dos estudos quanto à sua inclusão no mundo do trabalho. A avaliação pode servir como complemento do currículo para a seleção de emprego.

Exames nacionais como o Enem e os vestibulares exercem forte influência na maneira como as aulas de línguas (e também aquelas de outras áreas do conhecimento) são conduzidas nas escolas, fazendo com que exercícios, livros, correções, calendário e até mesmo punições sejam direta ou indiretamente determinados por eles, fenômeno conhecido na literatura como “efeito retroativo”, conforme pude constatar em minha dissertação de mestrado (Borja, 2020a).

O peso dessas avaliações é vultoso, por se tratar de instrumentos que permitirão ou não o acesso de alunos às principais universidades ou aos melhores empregos no país. Vale ressaltar que o desempenho dos alunos nos exames, muitas vezes, é um “termômetro para medir” o desempenho de professores e escolas também, acarretando diversas consequências aos envolvidos. Para Scaramucci (2011), é de suma importância que uma análise seja feita com base nos impactos sociais, intencionais ou não, advindos dos exames nacionais de maior prestígio, algo que nem sempre os professores têm autoridade para questionar. Tamanha expectativa, associada a momentos (socialmente) específicos de avaliação, pode desviar o foco para a implementação de uma avaliação processual, que pode ser feita ao longo do semestre, no dia a dia da sala de aula, levando-se em consideração as tentativas, o esforço e os desafios enfrentados pelos alunos. Também vale ressaltar que a correção, com base em resultados numéricos, tem fortes consequências na motivação e na autoestima do aluno, que tem seu esforço constatado com um número, e nem sempre intervenções são feitas para que ele evolua a partir daquele resultado. Para Méndez (2002), o caráter formativo da avaliação justifica-se por formar o aluno intelectual e humanamente.

Em se tratando do ensino e aprendizagem de línguas adicionais, a polêmica não é diferente. A avaliação nessa área no Brasil é fortemente influenciada não apenas pelos exames nacionais de grande escala, mas também pelos modelos de exames de proficiência internacionais, trazendo para as experiências formativas do aluno uma ansiedade atrelada à imagem idealizada do falante nativo. Na visão de Scaramucci (2000), o conceito de proficiência deve considerar a finalidade de uso da língua, e não um falante proficiente absoluto.

Além disso, o contexto escolar brasileiro enfrenta uma série de desafios políticos e econômicos e não deve ser comparado a modelos internacionais. Frequentemente, no ensino fundamental e médio de escolas públicas brasileiras, as aulas de língua adicional acontecem uma vez por semana e com turmas numerosas, o que configura um empecilho para o desenvolvimento das quatro habilidades que fazem parte do Quadro Europeu Comum de Referência para Línguas: produção oral, compreensão oral, produção escrita e compreensão escrita. A possível expectativa de utilizar o Quadro Comum Europeu como modelo para práticas de ensino e avaliação é pouco realista no contexto em questão. Mesmo as salas com número reduzido de alunos dificilmente são padronizadas e homogêneas, o que as torna ainda mais interessantes.

Considerando-se o desafio que envolve avaliar em sala de aula e a importância do ensino de línguas em um mundo globalizado, a avaliação não pode ser considerada concluída ou desvendada e deve ser pesquisada em suas diversas nuances. Apesar de presente no contexto escolar desde os primórdios, as avaliações são pouco questionadas, fazendo com que poucas pesquisas sejam desenvolvidas em nível nacional e tenham pouca diversidade de enfoque.

Diante da importância e da necessidade de se desconstruírem os hábitos arraigados e associados a experiências avaliativas e almejando descortinar quais experiências formativas a elas relacionadas poderiam ser mais motivadoras para os alunos de línguas, propus-me, nesta tese, a investigar, na literatura internacional, por meio de técnicas bibliométricas, e em comparação com os resultados de minha dissertação de mestrado, o tema avaliação no ensino de línguas.

As pesquisas nacionais sobre avaliação selecionadas em minha dissertação de mestrado foram provenientes de um *corpus* que envolveu sete periódicos Qualis A1 da área de Linguística e Literatura encontrados na Plataforma Sucupira da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) (Borja, 2020a, 2020b). A Capes, fundação do Ministério da Educação, desempenha importante função no acesso às pesquisas científicas em nível de pós-graduação (mestrado e doutorado) de diversas regiões do Brasil, e os periódicos Qualis A1 têm

(ou tiveram, já que este parâmetro está em vias de reformulação) o maior impacto e uma produção científica constante.

Assim, o objetivo geral desta tese é o de investigar a frequência e o foco das principais pesquisas acerca de avaliação no ensino de línguas em periódicos internacionais latino-americanos, bem como compreender as principais práticas avaliativas que emergem de tais artigos. No que concerne aos objetivos específicos, esta tese

- i. identifica possíveis tendências e diferenças de foco de pesquisas acerca do tema avaliação no ensino de línguas em periódicos internacionais;
- ii. problematiza possíveis diferenças de foco em pesquisas sobre avaliação; e
- iii. compara certo volume de dados mais proeminentes com aqueles já coletados em Borja (2020a).

A busca por artigos em periódicos internacionais foi o primeiro passo para identificar, bibliometricamente, possíveis diferenças de foco e frequência nas pesquisas da área de Avaliação em ensino de línguas. Interessei-me por fazer um apanhado dos principais periódicos internacionais latino-americanos em Linguística Aplicada, pois, além de expandir os resultados da minha pesquisa de mestrado, há, a meu ver, uma necessidade de trazer visibilidade aos pesquisadores sul-americanos e latinos como um todo, por vezes apagados ou invisíveis diante da hegemonia científica dos países norte-americanos e europeus.

Sulear os discursos acadêmicos é, portanto, necessário. Para Matos e Santos (2022), sulear é escutar a voz dos saberes relegados, como os dos povos indígenas, afrodescendentes, LGBTQIA+, imigrantes e outros. Portanto, trabalhar com os saberes suleados é sinônimo de resistência e luta, visando a incluir saberes anteriormente desvalorizados.

Um número considerável de pesquisas do *corpus* de artigos brasileiros tinha como foco, conforme revelou minha dissertação de mestrado (Borja, 2020a), as contribuições do Quadro Europeu Comum de Referência para as pesquisas sobre avaliação, o que revelou um eurocentrismo hegemônico nas investigações da área de Avaliação. Interessei-me em saber se as pesquisas dos periódicos internacionais latino-americanos traziam essas contribuições suleadas para a área, além de diferentes discussões dentro da própria área de Avaliação, como a diversidade de temas.

O ambiente acadêmico frequentemente prioriza um discurso monocultural. No contexto de ensino e aprendizagem de línguas adicionais, livros, autores e manifestações relacionadas ao contexto europeu são comumente priorizados. Assim, a corrente decolonial nasce com o intuito de questionar a hierarquização de certos saberes tidos como legítimos e romper com o

silenciamento de contribuições, como as oriundas da América Latina. Decolonizar, portanto, é um ato de resistência diante dos anos de apagamento.

A base de dados do SCImago (SCImagoJR, 2025), que atua como medidor de influência científica de periódicos acadêmicos de diversos países, foi utilizada como forma de auxílio na seleção dos principais periódicos da área, sobretudo de países latino-americanos. O SCImago Journal & Country Rank, portal público que inclui os periódicos e indicadores científicos, foi criado e é ainda mantido por parceiros das universidades de Granada, de Extremadura e de Carlos III, na Espanha. Os periódicos podem ser comparados ou analisados separadamente e agrupados por área temática (27 grandes áreas temáticas), por categoria temática (309 categorias temáticas específicas) ou por país. Os dados de citação são extraídos de mais de 34.100 títulos de mais de 5.000 editores internacionais e métricas de desempenho de 239 países.

Essa base de dados leva o nome do indicador SCImago Journal Rank (SJR), desenvolvido pela SCImago, com base no amplamente conhecido algoritmo Google PageRank. Esse indicador mostra a visibilidade dos periódicos contidos na base de dados da Scopus desde 1996.

Dentre os periódicos internacionais latino-americanos de Linguística e Linguagem indexado na base de dados gerada pela SCImago, constam periódicos colombianos, chilenos, mexicanos, argentinos, venezuelanos, peruanos e uruguaios. Todos os periódicos cujos *sites* eram de livre acesso foram selecionados, com o intuito de se fazer uma busca por artigos relacionados à avaliação no ensino de línguas, que posteriormente foram comparados aos resultados oriundos da pesquisa antes feita com artigos brasileiros.

Trata-se, assim, de uma pesquisa descritiva que utilizou indicadores bibliométricos para verificar o que tem sido estudado no cenário internacional sobre avaliação no ensino e aprendizagem de línguas. Dada a relativa escassez de investigações sobre o tema em cenário brasileiro (Borja, 2020a), levantei a hipótese de que a avaliação também não é um tema recorrente nos principais periódicos latino-americanos atrelados às áreas de Linguística e Linguística Aplicada e as discussões nele fomentadas dão ênfase aos testes externos como forma de balizar as avaliações e os testes em sala de aula de ensino de línguas.

O foco principal foi identificar, por um lado, até que ponto questões políticas e sociais influenciam as pesquisas sobre avaliação no ensino de línguas em diferentes países e, por outro, as principais tendências investigativas nos contextos em questão.

Esperava-se que o número de artigos internacionais sobre avaliação fosse superior ao do *corpus* de artigos nacionais já investigados, que se mostrou reduzido e pouco diverso em relação aos temas. O uso de tecnologias digitais em sala de aula (presencial e *online*) e,

consequentemente, na avaliação, também foram foco da investigação, bem como possíveis pareceres oficiais dos principais órgãos educacionais sobre mudanças acarretadas pela pandemia da covid-19.

Por décadas, profissionais da educação foram resistentes a mudanças relacionadas à avaliação, e a pandemia da covid-19 trouxe uma urgência de adaptação e atualização nunca antes vista por parte de professores e instituições com mais ou menos experiência. Em um momento tão crucial, minha tese buscou contribuir com pesquisas internacionais acerca da avaliação, trazendo as principais descobertas e inovações praticadas internacionalmente para o cenário nacional.

Esta tese contém as seguintes seções: Fundamentação Teórica, Metodologia de Pesquisa, Resultados, Discussão e Considerações Finais. Na Fundamentação Teórica, reuni contribuições teóricas da literatura da área de Avaliação, assim como da Avaliação no Ensino de Línguas, reflexões da área de Avaliação Crítica, a relevância do Letramento em Avaliação, contribuições sobre Decolonialidade, a mudança de paradigma impulsionada pelo agente conversacional de inteligência artificial ChatGPT e a análise de documentos/pareceres educacionais sobre avaliação, através dos Marcos Curriculares. Na Metodologia de Pesquisa, abordei as contribuições da Bibliometria, bem como da Análise de Conteúdos para auxílio na análise do *corpus* coletado com base no *ranking* da SCImago. Por último, analisei o *corpus* de artigos sobre avaliação no ensino de línguas, através de critérios como: Tipo de avaliação; População avaliada; Competências avaliadas; Métodos e instrumentos utilizados; Resultados e conclusões; Desafios e limitações; e Recomendações futuras. Cada artigo foi analisado separadamente, seguido de análise conjunta posterior, com as principais tendências observadas, culminando nos resultados, nas discussões e nas considerações finais.

3 FUNDAMENTAÇÃO TEÓRICA

avaliação (substantivo).

uma abordagem sistemática para coletar informações e fazer inferências a respeito da habilidade de um estudante ou da qualidade ou sucesso de um percurso de ensino com base em várias fontes de evidência. A avaliação pode ser feita por meio de testes, entrevistas, questionários, observações etc. Por exemplo, a avaliação da habilidade de compreensão de um estudante imigrante pode ser importante para descobrir se o aluno está apto a seguir o curso de estudo em uma escola ou se aulas extras do idioma serão necessárias. Estudantes podem ser testados no começo e novamente ao fim de um curso para avaliar a qualidade do ensino no curso. O termo “teste” é frequentemente associado a grandes testes padronizados, ao passo que o termo “avaliação” é usado em um sentido muito mais amplo, se referindo a uma variedade de abordagens em testes e avaliações (Assessment..., 2002, p. 35, tradução nossa).

Sobre a natureza e a função dos tipos de instrumentos avaliativos mais comumente praticados, Rodrigues, Rocha e Gonçalves (2011) apresentam um quadro com base em Herculano (2001):

Quadro 1 – Natureza e função dos tipos de instrumentos avaliativos

Tipos de instrumentos de avaliação	Natureza	Função
Prova escrita objetiva	Questões diretas que possuem apenas uma resposta correta	Avaliar o quanto o aluno apreendeu dos dados singulares e específicos do conteúdo.
Prova escrita discursiva	Questões indiretas com uma maior flexibilidade nas respostas	Avaliar a capacidade de analisar o problema central, abstrair fatos, formular ideias e redigi-las.
Prova escrita mesclada (objetiva e discursiva)	Questões diretas e indiretas (já descritas anteriormente)	Avaliar tanto a aquisição de dados singulares quanto a capacidade de formular ideias e redigi-las.
Prova oral	Questões diretas e/ou indiretas feitas oralmente	Verificar a capacidade de compreensão e produção oral.
Ditado	Questões diretas	Verificar a compreensão oral e ortografia.
Produção textual	Questões indiretas	Avaliar a capacidade de expressão, interpretação e organização na produção

		textual.
Trabalho individual	Atividades de naturezas diversas realizadas individualmente	Avaliar a capacidade individual do aluno de compreensão, elaboração e organização de ideias.
Trabalho em grupo	Atividades de naturezas diversas realizadas coletivamente	Avaliar a capacidade de compreensão, elaboração e organização de ideias por meio da socialização.
Seminário para um determinado público	Exposição oral para um público utilizando a fala e possíveis materiais de apoio para explanação do assunto	Avaliar a capacidade de expor oralmente um conteúdo/assunto
Participação/conceito	Análise feita pelo professor da participação do aluno nas atividades propostas	Avaliar a participação do aluno nas atividades propostas.
Autoavaliação	Análise oral ou escrita que o aluno faz do seu próprio processo de aprendizagem	Permitir ao aluno se autoavaliar julgando seu processo de aprendizagem e percebendo os pontos positivos e negativos.

Fonte: Rodrigues, Rocha e Gonçalves (2011, p. 12-13).

Conforme o Quadro 1, Rodrigues, Rocha e Gonçalves (2011) apresentam uma variedade de instrumentos de avaliação, cada um com sua natureza e função específicas. As provas escritas, objetivas e discursivas avaliam desde a apreensão de dados específicos até a capacidade de análise e formulação de ideias. A prova oral verifica a compreensão e a produção oral, enquanto o ditado focaliza a compreensão oral e a ortografia. A produção textual e os trabalhos – individuais ou em grupo – avaliam a expressão, a interpretação e a organização de ideias. O seminário avalia a capacidade de exposição oral, e a participação/conceito e a autoavaliação permitem ao professor e ao aluno refletirem sobre o processo de aprendizagem.

A avaliação, portanto, como a conhecemos hoje, tornou-se parte indissociável do ensino com a escolaridade obrigatória, como forma de abarcar e obter registros do desempenho de todos os novos alunos matriculados. Junto a ela, preocupações mais estatísticas foram surgindo, em detrimento das pedagógicas.

Para Luckesi (2000), a avaliação da aprendizagem requer, antes de mais nada, uma disposição psicológica específica do avaliador: a de acolher e incluir em detrimento de classificar; de aceitar a situação inicial do aluno, satisfatória ou não, para intervir posteriormente, visando o seu crescimento. Se o aluno é prejudicado, o avaliador não está cumprindo o seu papel, pois é apenas após acolher que o avaliador fará tomadas de decisão. Apenas constatar um *status* não tirará o aluno do seu ponto de partida. Luckesi (2000) ainda

traz luz ao fato de que os instrumentos avaliativos não devem ser utilizados como instrumentos de punição e ameaça de reprovação, como frequentemente acontece, e acrescenta:

Quaisquer que sejam os instrumentos – prova, teste, redação, monografia, dramatização, exposição oral, arguição etc. – necessitam manifestar qualidade satisfatória como instrumento para ser utilizado na avaliação da aprendizagem escolar, sob pena de estarmos qualificando inadequadamente nossos educandos e, conseqüentemente, praticando injustiças. Muitas vezes, nossos educandos são competentes em suas habilidades, mas nossos instrumentos de coleta de dados são inadequados e, por isso, os julgamos, incorretamente, como incompetentes. Na verdade, o defeito está em nossos instrumentos, e não no seu desempenho. Bons instrumentos de avaliação da aprendizagem são condições de uma prática satisfatória de avaliação na escola (Luckesi, 2000, p. 5).

Méndez (2002) chama a atenção para o fato de que, antes de qualquer tentativa de mudança no sistema avaliativo das escolas, devemos considerar as características institucionais do contexto em questão e o fato de que todos os professores foram socializados de forma específica. Logo, todos trazem suas experiências como alunos para a sala de aula, o que também inclui formas de avaliar. A avaliação tornou-se um ritual tão padronizado, que frequentemente é assumida de forma acrítica, de modo a apenas aferir e pouco intervir. É comum que vejamos professores sendo considerados bons com base na quantidade de alunos que reprovam e no quão complexos os seus instrumentos avaliativos são. Tais conceitos são, no mínimo, contraditórios, tendo em vista que professores deveriam ser valorizados com base no auxílio que dão a seus alunos para a superação de obstáculos, evitando, assim, a evasão escolar e a mera memorização de informações.

3.1 Avaliação e ensino de línguas

Um dos primeiros encontros de que se tem registro entre ensino de línguas e avaliação no Brasil, conforme apontado por Day (2012), tem o Colégio Pedro II como cenário. No Brasil Império, o colégio em questão foi fundado em 1837, tornando-se referência nos estudos secundários. A chegada da Família Real trouxe mudanças em vários âmbitos, incluindo o educacional. As línguas modernas receberam, pela primeira vez no país, a mesma atenção que as línguas clássicas. Os alunos passaram a ter aulas de inglês, francês e alemão no colegial, que agora oferecia uma educação nos moldes europeus. O Colégio Pedro II foi pioneiro também na aplicação de exames que possibilitavam acesso ao ensino superior.

Para Harmer (2001), os tipos mais comuns de avaliação no ensino de línguas são:

- i. nivelamento: direciona os alunos para a turma mais adequada. Geralmente essa avaliação faz uso de conteúdos que os alunos utilizarão quando tiverem seus níveis definidos. Conhecimentos sobre vocabulário e gramática são comumente testados, bem como habilidades produtivas e receptivas. Algumas escolas pedem que os próprios alunos façam uma autoavaliação como parte do processo de nivelamento;
- ii. diagnóstico: focaliza as dificuldades, as lacunas e os pontos a serem trabalhados com os alunos durante o curso. Com a clareza de quais são os principais desafios enfrentados pelos estudantes, uma intervenção se torna mais fácil;
- iii. progresso: mede o progresso em relação ao cronograma em execução. Essa avaliação é, geralmente, elaborada pelos próprios professores e aplicada com maior frequência em uma perspectiva formativa. Vale ressaltar que, nesse tipo de instrumento, os elementos que o compõem devem ser familiares. Isso não implica repetir textos e exercícios, mas utilizar textos e exercícios em formatos semelhantes aos aplicados em sala normalmente. Considerando que o intuito é avaliar os conhecimentos do aluno em relação ao que ele vem estudando ao longo do semestre, formatos de avaliação diferentes seriam contraditórios e, mesmo que pudessem medir conhecimentos da língua de uma forma geral, não seriam a opção ideal. Os testes em questão não devem expor as fraquezas do aluno, mas guiar e auxiliar tomadas de decisão futuras e mudanças no cronograma, sempre visando à superação de obstáculos;
- iv. proficiência: permite uma visão geral dos conhecimentos do aluno (e não do progresso). Esse tipo de avaliação é comumente solicitado para a admissão em universidades internacionais, vagas de emprego e certificados. É comumente marcado pelo efeito retroativo que exerce e, apesar de injustamente, associa a reputação do colégio/professor ao êxito ou não do aluno nos testes de proficiência;
- v. portfólio: reúne as principais atividades do aluno ao longo do semestre. Os testes de proficiência e progresso avaliam as habilidades em um momento específico, dando aos alunos uma única chance de mostrar o que sabem. Por diversos fatores, como questões emocionais, avaliar os conhecimentos pontualmente, em detrimento do processo, pode ser injusto e causar ansiedade desnecessária no aluno. Pensando nisso, algumas instituições sugerem que os alunos montem portfólios. Com um caráter processual, já que avaliam as tentativas de um longo período, os portfólios permitem que os alunos sejam avaliados com base em três ou quatro de seus melhores trabalhos. Os efeitos são muito positivos, pois permitem que o estudante se torne mais

autônomo e tenha mais chances de se aperfeiçoar; em contrapartida, os portfólios podem demandar tempo excessivo em sua construção. Muitas instituições são contra a adoção de portfólios por não saberem se a identidade dos alunos será confirmada ou se receberam ajuda de terceiros para concluir as atividades.

Como todo e qualquer tipo de inovação em contextos marcados por tradições milenares, como o caso da avaliação nas escolas, as mudanças e inovações sempre enfrentam algum tipo de resistência, que deixa de abrir espaço para formas novas e mais justas de avaliar. É comum que grande parte das pesquisas acerca de avaliação no ensino de línguas, sobretudo em contexto nacional, não se aprofunde em questões relacionadas ao uso de tecnologias digitais na educação. Apesar de o ensino de línguas e a avaliação mediados por computador serem uma opção já há um tempo considerável, essas atividades sofreram certa resistência por parte de professores e alunos que desconheciam a dinâmica virtual envolvida e não foram apresentados às principais plataformas disponíveis de ensino e avaliação de línguas. Questões econômicas também influenciam a não implementação de tais formas de avaliação, tendo em vista que muitos alunos não têm acesso a computador e serviço de *internet* de qualidade nem condições financeiras para a manutenção das máquinas.

Recentemente, com as questões relacionadas à pandemia da covid-19, o cenário educacional viu-se obrigado a se adaptar às mudanças. Após muito debate entre os principais órgãos responsáveis e o adiamento da volta às aulas nas principais universidades e escolas, muitos professores, pais e alunos tiveram que se familiarizar com ferramentas de ensino e avaliação *on-line*, tendo em vista que, por questões de segurança, essa era a única opção viável.

Tal marco histórico não só vai mudar o cenário educacional, como já está mudando: ao se abrirem a novas possibilidades, os alunos verão os pontos positivos do uso de tecnologias digitais em sala de aula, e muitos profissionais da educação poderão, finalmente, implantar mudanças no cenário resistente e pouco questionado da avaliação. Os acontecimentos atuais também poderão fazer com que os programas de inclusão digital para alunos sejam tratados com mais urgência e seriedade por parte das autoridades, por meio de auxílio financeiro aos alunos na aquisição de dispositivos eletrônicos e pacotes de dados para acesso à *internet*, visando à inclusão e à educação de qualidade ao maior número possível de pessoas.

3.2 Avaliação crítica

Shohamy (2001, p. 7), em seu capítulo sobre as vozes dos avaliados, lembra-nos que é pouco provável que se encontre alguém que nunca tenha vivido a experiência de ser submetido a um teste ou que não tenha um relato de como um único instrumento avaliativo afetou a sua vida (positiva ou negativamente). Os efeitos dos testes ecoam em suas memórias por anos. Apesar disso, na literatura sobre avaliação, os que são submetidos a testes dificilmente têm suas vozes ouvidas, como se os criadores dos instrumentos avaliativos não estivessem interessados em ouvir tais vozes, o que poderia auxiliar em novas perspectivas sobre avaliação, caso fossem levadas em consideração.

O poema abaixo, *A Death Wish*, escrito por J. McGahern (1977 *apud* Shohamy, 2001, p. 7), ilustra o impacto que os testes têm em uma criança prestes a ser avaliada:

*Please God may I not fail
Please God may I get over sixty per cent
Please God may I get a high place
Please God may all those likely to beat me get killed in road accidents and
may they die roaring.*

O trágico poema em questão mostra a tensão e o medo relacionados à experiência avaliativa. O medo associado ao teste demanda orações, como se a criança fosse ser salva de um desastre prestes a acontecer. Além de gerar tensão, o teste também pode gerar competitividade com os colegas de sala, criando um clima tenso generalizado.

Shohamy (2001) fornece exemplos, em um capítulo dedicado às vozes e aos depoimentos dos avaliados, de depoimentos que relatam o estigma associado a desempenhos ruins em avaliações, sobretudo em se tratando de exames nacionais de peso. Em um artigo de 1991, publicado no jornal Haaretz, uma experiência de exame nacional de compreensão de leitura foi compartilhada. No dia do exame, que aconteceria na escola dos alunos em questão, alguns não tiveram permissão para fazê-lo. O motivo, segundo o diretor, é que a participação dos alunos em questão faria com que a média do resultado caísse, afetando o resultado/pontuação da escola. Tais alunos, proibidos de fazer o teste, foram para uma sala de filmes durante sua aplicação. Esse exemplo, com certeza, não acontece sem deixar marcas irreversíveis na autoestima dos envolvidos, que provavelmente se sentiram diminuídos e humilhados pela postura antiética e desumana do diretor (Shohamy, 2001, p. 10). Além de cruéis, os resultados desse exame não serão realistas e válidos, tendo em vista que não puderam capturar todos os perfis de estudantes da escola. Os exames nacionais são comumente temidos, tendo em vista o *status* envolvido em um resultado considerado positivo.

A validade também pode ser questionada de outras formas, como quanto à

impossibilidade de saber se os alunos, de fato, estudaram ou se memorizaram “truques” para que pudessem depositar suas respostas o mais rápido possível no instrumento, antes que se esquecessem de vez do conteúdo. Ademais, muitos apenas estudam “motivados” pela pressão e pelo medo dos resultados. Muitos alunos não sentem a confiança necessária para questionar respostas e fazer suas contribuições para além do roteiro esperado para determinada resposta, já que o controle está nas mãos dos avaliadores e qualquer tipo de questionamento teria consequências, limitando o perfil questionador de determinados estudantes (Shohamy, 2001, p. 13). Os alunos veem os testes como ferramentas poderosas, ao passo que se sentem sem poder nenhum quanto à sua realização, tendo em vista que têm pouco ou nenhum controle sobre eles e suas consequências. Assim sendo, cabe aos alunos renderem-se às regras (Shohamy, 2001, p. 14). Alguns dos termos utilizados por alunos para definir os testes foram:

- ameaçadores/assustadores;
- desconectados da realidade (pouco realista);
- levam a consequências prejudiciais;
- poderosos;
- geram competição;
- requerem estratégias de sobrevivência;
- têm efeitos a longo prazo;
- injustos;
- desanimadores;
- parciais;
- são indicação/prova do fracasso;
- dependem da sorte e de poderes sobrenaturais; e
- levam a comportamentos antiéticos.

Escutar as vozes dos alunos submetidos a testes vai ao encontro do defendido pelas teorias decoloniais, tendo em vista que são vozes raramente ouvidas e, se ouvidas, infelizmente não geram mudanças no sistema avaliativo como um todo. Sob essa mesma ótica, Freire (1970) chama atenção para o fato de que “o educador, que aliena a ignorância, se mantém em posições fixas, invariáveis. Será sempre o que sabe, enquanto os educandos serão sempre os que não sabem”. Os avaliados são os principais afetados pelas consequências negativas dos testes, seja na autoestima ou na privação de oportunidades. Ouvir seus pontos de vista e angústias é um meio de contribuir para as pesquisas e reformulação de instrumentos avaliativos, visando ao constante aperfeiçoamento dos instrumentos, para que sejam mais válidos, humanos e

acolhedores. A prática de ouvir os alunos a respeito dos testes deveria ser mais amplamente adotada por profissionais da educação, tendo em vista o peso que os instrumentos avaliativos têm em qualquer contexto educacional e o fato de que a aplicação de testes altera o comportamento/rotina daqueles que serão submetidos aos testes.

Os testes trazem consequências para alunos, pais, professores e diretores, mas o peso de suas consequências varia conforme a intenção e o tipo de instrumento. Um termo amplamente discutido na literatura de avaliação, o *efeito retroativo*, diz respeito à medida em que um teste, sobretudo exames nacionais de peso, tem a capacidade de mudar o formato de um curso, uma grade curricular etc. Exames como o Enem ou vestibular fazem com que todo o conteúdo estudado no Ensino Médio seja alterado, levando-se em consideração resultados satisfatórios no Enem. Como e o que será estudado e como os alunos serão avaliados são questões fortemente influenciadas por tais exames de peso, mesmo que os alunos não pretendam cursar nenhuma universidade. Tal possibilidade (a de não cursar uma universidade) muitas vezes nem é cogitada pelas escolas, como se apenas um caminho fosse possível. Logo, alunos que vão ou não prestar o Enem, por exemplo, são afetados por esse preparo para a prova, inevitável em suas escolas, quer isso acarrete consequências positivas ou negativas. A título de opinião, creio que as próximas gerações vão e já estão questionando os formatos e as trajetórias educacionais atuais.

Segundo Shohamy (2001, p. 105), os efeitos dos testes podem trazer mudanças aos comportamentos daqueles a eles submetidos, que terão que se “render” às regras e demandas. Suas consequências, no geral, estão fortemente relacionadas à privação de oportunidades, punição, sentimento de falha e reprovação. Os professores, por sua vez, alteram suas estratégias em sala de aula, já que o “sucesso” de seus alunos é interpretado como um reflexo de seu ensino. Os pais podem investir em aulas particulares, já que o sucesso de seus filhos pode ser lido como reflexo de sua criação. O “sucesso”, aqui interpretado vagamente como notas altas em tais testes, pode e deve ser questionável.

É comum que escolas sejam escolhidas pelos pais com base na quantidade de alunos aprovados em determinado teste nacional de peso ou em uma pontuação alta, o que faz com que muitos as considerem boas escolas. Isso reforça o peso dos resultados em testes na nossa cultura, afetando todos os envolvidos direta ou indiretamente (Shohamy, 2001, p. 106). Os resultados satisfatórios em testes podem ser associados a melhores oportunidades de trabalho, melhores salários e expansão de oportunidades econômicas como um todo. Também podem ser traduzidos em capital simbólico, como prestígio, reconhecimento, mesmo que isso não resulte em ganhos financeiros diretos. Assim sendo, os resultados seguem uma lógica econômica, já que o alto desempenho resulta em capital ou capital simbólico.

Os testes de maior peso também podem gerar ansiedade nos professores, já que existe uma pressão por resultados considerados satisfatórios e, muitas vezes, pouca autonomia para questionar o foco em tais testes, fazendo com que o professor tenha menos autonomia na sua prática e mais pressão por resultados. O teste, por fim, ganha mais autoridade do que o próprio professor.

Na área da Avaliação Crítica (Shohamy, 2001, p. 131), os testes são vistos como ferramentas poderosas inseridas em um contexto social e político e associadas a intenções e consequências. A Avaliação Crítica sugere a necessidade de desenvolvimento de estratégias críticas para examinar o uso e as consequências dos testes, monitorando o poder que exercem, a força prejudicial, revelando o seu possível uso inadequado e empoderando os alunos submetidos a testes. Dessa forma, os professores aplicadores de testes terão mais informações e poderão refletir melhor sobre o impacto dos instrumentos avaliativos.

A área de Avaliação Crítica argumenta que os testes de línguas não são neutros, e sim um produto cultural, social, político e ideológico que molda a vida dos envolvidos. Essa área defende que os alunos submetidos a testes devem ser encorajados a questionar os instrumentos avaliativos, desenvolvendo uma visão crítica sobre eles, já que são vistos como sujeitos políticos dentro de um contexto político. A Avaliação Crítica reconhece que os testes são ferramentas diretamente associadas ao sucesso e questiona quais intenções (e de quem) estão por trás de um teste. Questiona, inclusive, se os testes têm a intenção de avaliar conhecimento ou de ditar conhecimento. O que é considerado verdade nos testes de línguas são “verdades” fixas ou abertas à negociação? Os testes também consideram as vozes de alunos, pais de alunos, professores e autores de testes?

A Avaliação Crítica considera que o conhecimento de qualquer autoridade avaliativa é incompleto e deve ser enriquecido com os conhecimentos de outras fontes para obter uma interpretação mais precisa dos conhecimentos avaliados. Ela desafia a Psicometria, tendo em vista que não considera verdades absolutas, mas que diferentes interpretações são consideradas válidas no momento de aplicar uma pontuação. As definições de conhecimento por trás dos testes reafirmam o poder das autoridades ou consideram os diversos grupos da sociedade de forma democrática?

A Avaliação Crítica questiona o fato de o teste ser considerado o único instrumento válido para avaliar o conhecimento e propõe, alternativamente, múltiplas possibilidades que, aliadas, poderão oferecer uma melhor avaliação dos conhecimentos dos indivíduos avaliados (Shohamy, 2001, p. 133). Questiona também o que é considerado válido em testes, propondo que o que trouxe consequências negativas para a educação, ou usos inadequados dos testes,

não pode ser considerado válido atualmente.

Alguns questionamentos se mostram necessários. Quem são os examinadores? Quais são suas intenções? Quem são os avaliados? Qual o seu contexto? Qual o contexto do conteúdo sendo avaliado? Quem vai se beneficiar dos testes? Por qual motivo o teste está sendo aplicado? Como os seus resultados serão usados? O que está sendo avaliado e por quê? O que não está sendo avaliado e por quê? Quais os valores subjacentes por trás do teste? Quais são os métodos avaliativos? Quais evidências adicionais estão sendo coletadas sobre o conteúdo? Quais tipos de decisões são tomados com base nos testes? Além do examinador, quem é incluído na criação do teste e em sua aplicação? Qual ideologia é transmitida através do teste? Quais as mensagens sobre alunos, professores e sociedade que o teste presume? Qual tipo de *feedback* está sendo dado através do teste e para quem? O teste pode ser questionado? Quais são os seus usos intencionais e não intencionais? Qual o seu efeito retroativo? Quais são as formas que os avaliados usam para questionar o teste?

A Avaliação Crítica ainda reconhece que os testes podem ser facilmente usados de maneira inadequada. Há muitas autoridades que, por estarem no poder, vão evitar processos democráticos, mas são os avaliados os que sofrem as consequências, fazendo-se necessário questionar comportamentos injustos (Shohamy, 2001, p. 143). Os avaliados costumam ter plena confiança nos (e submissão aos) testes. Os instrumentos são temidos e respeitados, como se os avaliados estivessem nas mãos dos examinadores. Examinadores de línguas têm obrigações profissionais e morais. Existe uma urgência de se examinarem questões éticas na avaliação e se promoverem regras claras de conduta (ética na avaliação), devendo o examinador questionar e refletir sobre seus possíveis usos inadequados de testes e qual direção seguir, uma vez que tais usos (inadequados) sejam identificados.

Muitas vezes, os avaliados seguem regras e são cobrados por conhecimentos específicos, mas os examinadores têm autonomia para criar as suas próprias regras (Shohamy, 2001, p. 155). O examinador, muitas vezes, é quem define as informações que estarão incluídas nos testes, cabendo ao avaliado seguir e acatar, o que leva à pergunta: quais são os direitos dos avaliados? Os avaliados (ou *test takers*) são definidos em uma perspectiva mais ampla do que apenas o indivíduo submetido ao teste? (Shohamy, 2001, p. 156). Ou seja, o aluno, o professor, o diretor e os pais são direta ou indiretamente afetados.

Os avaliados devem ser encorajados a desenvolver uma visão crítica sobre a avaliação e questionar o seu uso. É provável que eles tenham sido avaliados com materiais desconhecidos, métodos desconhecidos e em situações nas quais os resultados foram usados para finalidades não planejadas/estabelecidas previamente. Muitas vezes, os avaliados não têm poder para

questionar as regras vigentes, demandando, assim, o estabelecimento de grupos de defesa e representação, visando a proteger os avaliados de abuso de poder ou de não serem avisados que serão testados ou o que será feito com os resultados. Ademais, a ética por trás do uso de distratores é questionável, além do que o aluno pode não ter sido sensibilizado com o método avaliativo em questão. Lamentavelmente, os avaliadores ainda recorrem a exercícios com o intuito de confundir os alunos acerca do conteúdo, o que mostra a urgência de que educadores e diretores sejam letrados em avaliação e que avaliem com uma postura acolhedora. Os conhecimentos recrutados ao se avaliar são complexos e diversos, e as tecnologias digitais trazem praticidade e novos desafios para a prática avaliativa.

A Avaliação Crítica revela a importância de uma abordagem ética e transparente, de modo que os envolvidos compreendam não apenas os métodos e os objetivos, mas que possam questionar e desconstruir práticas inadequadas. Essa postura crítica depende diretamente do preparo dos profissionais envolvidos no processo de avaliação. Surge, assim, a importância do letramento em avaliação, que não se limita à aplicação de instrumentos avaliativos, mas abrange o domínio de conhecimentos, habilidades e princípios que permitem aos educadores planejar, conduzir e interpretar avaliações de maneira justa, acolhedora e significativa.

3.3 Letramento em avaliação

Levando em consideração os diversos conhecimentos e habilidades recrutados ao avaliar a aprendizagem e a importância de profissionais da educação letrados em avaliação, Giraldo (2018) propõe descritores de conhecimentos, habilidades e princípios para o perfil de um professor de línguas letrado em Avaliação, como se segue.

Conhecimentos de linguística aplicada

- Compara abordagens para o ensino e a avaliação de línguas, como testes de linguagem comunicativa e avaliação baseada em tarefas;
- Explica questões relevantes em linguística aplicada, como bilinguismo, política e planejamento linguístico, pragmática e sociolinguística;
- Analisa tendências na aquisição de segunda língua e seu impacto na avaliação de línguas, como motivação, influência intercultural e estratégias de aprendizado;
- Integra teorias relacionadas à linguagem e ao uso da linguagem, como modelos de habilidade linguística, análise de discurso e ensino de gramática.

Conhecimentos de teoria e conceitos

- Descreve a história do teste e da avaliação de línguas e seu impacto nas práticas e na sociedade atuais;
- Interpreta a confiabilidade na avaliação de línguas e suas implicações: confiabilidade, a teoria clássica dos testes, análise de itens, ameaças, cálculo da confiabilidade de testes e itens, confiabilidade dos avaliadores etc.;
- Interpreta a validade na avaliação e suas implicações: validade de construto, conteúdo e critério, validade como argumento;
- Calcula procedimentos estatísticos para investigar a validade, como a correlação de produto-momento de Pearson (PPMC);
- Interpreta as qualidades principais para práticas de avaliação de línguas (além da confiabilidade e validade) e suas implicações para a avaliação de línguas: autenticidade, praticidade, interatividade, justiça, ética e impacto (incluindo o efeito retroativo);
- Compreende análises estatísticas básicas: média, intervalo, desvio padrão, distribuição de pontuação etc.;
- Diferencia os conceitos relacionados a paradigmas de avaliação: teste tradicional vs. alternativo, testes normativos e critérios;
- Diferencia os propósitos principais e as tomadas de decisão relacionadas à avaliação de línguas: nivelamento, desempenho, proficiência etc.;
- Explica importantes passos no desenvolvimento de testes: propósito do teste, definição de construto, especificações de conteúdo, especificações de teste etc.;
- Examina o significado e as implicações da avaliação crítica de línguas: poder, ética e justiça;
- Avalia as consequências (intencionais ou não) resultantes das avaliações em seu contexto;
- Avalia o tipo de efeito retroativo que as avaliações podem ter na aprendizagem, nos currículos e nas instituições;
- Contrasta métodos de avaliação com suas vantagens e desvantagens: testes, portfólios, avaliação de desempenho, autoavaliação e avaliação por pares, encenações, entre outros;
- Ressalta a natureza, o propósito e o *design* de rubricas de pontuação, como holísticas

e analíticas;

- Reconhece as implicações do *feedback* dentro de um paradigma de avaliação formativa.

Consciência do próprio contexto de avaliação de línguas

- Explica suas próprias crenças, atitudes, contexto e necessidades para avaliação;
- Avalia as políticas de testes e avaliação que influenciam seu ensino;
- Avalia as tensões existentes que influenciam a avaliação de línguas em sua escola;
- Descreve as diretrizes e políticas gerais que orientam a aprendizagem de línguas e a avaliação em seu contexto, como o tipo de currículo de línguas;
- Critica o tipo de efeito retroativo que as avaliações geralmente têm no contexto de seu ensino.

Habilidades instrucionais

- Alinha objetivos curriculares, instrução e avaliação;
- Planeja, implementa, monitora, registra e relata o desenvolvimento da linguagem dos alunos;
- Fornece *feedback* sobre o desempenho dos alunos (normativo e referenciado por critérios);
- Coleta dados formais (por exemplo, através de testes) e informais (enquanto observa em sala de aula) do desenvolvimento da linguagem dos alunos;
- Melhora a instrução com base em resultados e *feedback* de avaliações;
- Utiliza meios alternativos de avaliação, como portfólios;
- Usa métodos de avaliação de linguagem apropriadamente: para monitorar o aprendizado de línguas e nada mais;
- Fornece experiências de avaliação motivadoras, proporcionando *feedback* encorajador, ou proporcionando cenários de autoavaliação;
- Comunica resultados de testes normativos e referenciados por critérios para uma variedade de públicos: alunos, pais, diretores de escolas etc.;
- Usa múltiplos métodos de avaliação para tomar decisões com base em informações substanciais;
- Incorpora tecnologias na avaliação dos alunos.

Habilidades de design

- Identifica claramente e declara o propósito da avaliação de línguas;
- Define claramente o(s) construto(s) de linguagem sobre que um teste fornecerá informações;
- Desenha avaliações válidas não apenas em termos de conteúdo do curso, mas também sobre tarefas do curso;
- Constrói especificações de teste (ou esquemas) para desenhar formas paralelas de um teste;
- Elabora tabelas de especificações para informar os usuários dos formatos de teste disponíveis, onde aplicável;
- Desenha avaliações que sejam confiáveis, autênticas, justas, éticas, práticas e interativas;
- Escreve itens de respostas selecionadas como múltipla escolha, verdadeiro/falso e associações;
- Melhora os itens de teste após análise de itens, focalizando itens que são muito difíceis, fáceis ou pouco claros;
- Desenha questões abertas (escrita e fala) juntamente com rubricas para avaliação;
- Desenvolve rubricas para avaliações alternativas, como portfólios e autoavaliação;
- Garante segurança para que acessos indesejados aos testes sejam impedidos;
- Planeja *workshops* de treinamento para avaliadores, sempre que necessário.

Habilidades em medição educacional (habilidades avançadas nem sempre necessárias)

- Interpreta dados de testes em larga escala, ou seja, estatísticas descritivas, como médias e curvas; tem a capacidade de calcular estatísticas descritivas;
- Infere pontos fortes e fracos dos alunos com base em dados;
- Critica testes externos e suas qualidades com base em suas características psicométricas;
- Interpreta dados relacionados ao *design* de testes, como dificuldade de itens e discriminação de itens;
- Calcula índices de confiabilidade e validade usando métodos apropriados, como Kappa e PPMC;
- Investiga estatisticamente índices de facilidade e discriminação.

Habilidades tecnológicas

- Usa softwares como o Pacote Estatístico para as Ciências Sociais;
- Executa operações no Excel, como estatísticas descritivas e correlações de confiabilidade;
- Usa recursos da *internet* como tutoriais *on-line* e adapta conteúdos para suas necessidades específicas de avaliação de línguas.

Consciência e ações em relação a questões críticas na avaliação de línguas

- Informa claramente as inferências e decisões que derivam dos resultados em avaliações;
- Usa os resultados da avaliação como *feedback* para influenciar o aprendizado de línguas, não outras fontes irrelevantes para o construto (por exemplo, preconceito pessoal em relação a um estudante);
- Trata todos os alunos ou usuários da avaliação de línguas com respeito;
- Usa testes, processos de teste e resultados de testes de maneira ética;
- Fornece práticas de avaliação justas e não discriminatórias;
- Critica o impacto que testes padronizados e testes poderosos podem ter e toma uma posição sobre eles;
- Observa as diretrizes para os testes usadas na instituição em relação à avaliação de línguas;
- Critica testes externos com base em sua qualidade e impacto;
- Implementa práticas transparentes de avaliação de línguas; informa os alunos sobre o que, como e por que avaliar;
- Implementa práticas democráticas de avaliação de línguas, dando aos alunos oportunidades de compartilhar suas opiniões sobre a avaliação.

Sem o intuito ambicioso de oferecer uma lista hermética e completa de habilidades, com base no estudo de Giraldo (2018), a relação acima de conhecimentos necessários convocados pelos professores ao avaliarem seus alunos ilustra a complexidade da prática avaliativa e a responsabilidade envolvida nas diversas decisões que avaliar compreende.

O letramento em avaliação, ao propor práticas éticas, bem como transparentes, desafia as estruturas tradicionais que muitas vezes sustentam desigualdades nos contextos de ensino, aprendizagem e avaliação. Para além dos aspectos técnicos ou institucionais, as dinâmicas de

poder que influenciam a prática avaliativa também devem ser questionadas. É nesse ponto que a perspectiva decolonial se torna crucial, valorizando as vozes, experiências e práticas em contextos marginalizados e promovendo a desconstrução de padrões que perpetuam a hegemonia cultural e linguística, muitas vezes alinhada ao Norte Global.

3.4 Decolonialidade

As pesquisas em ensino de línguas, sobretudo as de ensino de língua espanhola, a meu ver, têm acrescentado mais e mais espaços para discussões sobre decolonialidade e o empoderamento das vozes até então marginalizadas, como as do Sul (Santos, 2022, p. 130-131). As pesquisas em ensino de línguas devem ir além de questões estruturalistas, mas abordar questões de poder, identidade, raça etc. A língua passou a ser entendida como comunicação em um sentido mais amplo, e entender o contexto em que é utilizada se tornou essencial.

A cultura colonizadora visa a manter no poder a homogeneidade de falas e culturas (sobretudo do Norte Global). Língua e cultura são indissociáveis e, segundo Nieto (2010 *apud* Santos, 2022, p. 131), aprendemos a língua inseridos em nossos contextos culturais, e a cultura em que estamos inseridos interfere em nossas escolhas linguísticas. A ideia fixa de identidade nacional já não é mais suficiente para definir quem alguém é.

Os professores de línguas podem abordar língua, cultura e identidade, de modo que as visões norte-americanas ou eurocêntricas não sejam o ponto de partida, já que isso acontece com certa frequência no contexto de ensino de línguas (Santos, 2022, p. 132). Ademais, cultura e identidade não devem ser consideradas algo inflexível, já que estão em constante transformação.

A hierarquização de saberes agora dá espaço a uma mistura de culturas com a globalização. Por isso é tão relevante o papel do professor de línguas, tendo em vista que discussões como essa são predominantes no contexto de sala de aula.

A apreciação pelo Norte Global – ou especificamente por países como os Estados Unidos, a Inglaterra, a Alemanha, a França e a Itália – reforça a valorização das produções acadêmicas vindas dos países mencionados, o que muitas vezes é considerado algo natural. Decolonizar é importante, pois grande parte das produções acadêmicas surgidas no Sul Global visam a atingir o Norte Global, trazendo luz para a importância de se questionar a colonialidade do saber e maior inclusão de saberes do Sul. Dessa forma, professores podem trazer discussões para seus alunos, para que questionem e compreendam seus lugares no mundo. Portanto, os materiais utilizados em sala de aula precisam ser cuidadosamente pensados, de modo a não

privilegiarem apenas o que vem da Europa ou dos Estados Unidos, mas trazerem reflexões que vão além da consolidação de estereótipos e supervalorização de sociedades ocidentais, já que a ditadura do saber ditou, e dita, gostos, preferências e pontos de discussão (Santos, 2022, p. 141).

Canagarajah (1999 *apud* Santos, 2022, p. 144) destaca a importância de se empoderarem os alunos, tendo em vista que os materiais didáticos de ensino de línguas muitas vezes reproduzem ideais de países colonizadores. Os alunos devem ser lembrados que também são produtores de conhecimento.

Kumaravadivelu (2012 *apud* Santos, 2022, p. 135) defende que os materiais didáticos devem ser produzidos por profissionais locais, que são os que mais conhecem a realidade em que estão inseridos os alunos. A propagação de ideais e de realidades dos países colonizadores nos materiais vai ao encontro da ilusão de um falante ideal, que faz com que muitos alunos busquem professores nativos, como se fossem os mais bem preparados para ensinar línguas, independentemente de sua formação. Não questionar tais crenças só aumenta o sentimento de inferioridade, que sempre foi reforçado em nós, de países colonizados, “o complexo de vira-lata”. Com isso, é importante que vozes antes marginalizadas ganhem espaço na hegemonia do saber, visando a contribuições acadêmicas de todas as partes do globo.

Segundo Souza e Hashiguti (2022, p. 150), o despertar decolonial surge com grande potencial de mudança teórica e ética, questionando a hegemonia do saber e possibilitando uma ponte entre os países do Sul Global. Os autores elucidam que a colonialidade é um resultado do colonialismo, que persiste na vida cotidiana de um sujeito, na cultura, no meio acadêmico e no senso comum, através de hierarquizações inventadas e respaldadas pela modernidade, os saberes tidos como superiores são, geralmente, aqueles eurocentrados.

O termo Virada Decolonial surgiu em 2005, segundo Maldonado-Torres (2007 *apud* Souza; Hashiguti, 2022, p. 151), nos Estados Unidos, em uma conferência na Universidade da Califórnia. Ironicamente, os Estudos Decoloniais comumente ganham espaço em universidades do Norte, enquanto o objeto de estudo se encontra no Sul, dividindo opiniões entre estudiosos da área, já que os sujeitos do Sul podem se tornar objetos de estudo e não terem necessariamente as suas vozes ouvidas. Ainda em consonância com o dito anteriormente, Cusicanqui (2012 *apud* Souza; Hashiguti, 2022, p. 153) questiona o descompasso existente entre teoria e prática decolonial, tendo em vista que as universidades do Sul muitas vezes reproduzem em seu fazer acadêmico um colonialismo interno.

O professor Lynn Mario Trindade Menezes de Souza, ao ser entrevistado por Simone Hashiguti (Souza; Hashiguti, 2022, p. 153), questiona a normatividade praticada na Linguística Aplicada e sua limitação a questões exclusivamente linguísticas. Muitas vezes, ainda segundo

Lynn Mario, substituímos termos como “língua estrangeira” por “língua adicional”, mas as práticas continuam coloniais e normativas, tendo em vista que a língua não é homogênea nem canônica. O fenômeno linguístico (Souza; Hashiguti, 2022, p. 156) ocorre em determinado contexto, com determinados interlocutores e suas histórias, inseridos sócio-historicamente e não considerados falantes ideais.

Portanto, a língua não pode ser reduzida a apenas trocas gramaticais. As trocas não são simplesmente emissão e recepção de mensagem, pois todos carregam as suas próprias interpretações baseadas em suas vivências pessoais. É preciso abrir espaço para outras formas de linguagem. O que é dado como universal é aquilo que não tem sotaque e que é válido para todos e sem prazo de validade, o que vem de uma imposição histórica, ignorando o que é local. Quando se pensa em determinada língua, naturalmente se pensa em um recorte geográfico e populacional, logo, não existe uma língua universal (Souza; Hashiguti, 2022, p. 157-158). Ademais, existem grupos de pessoas não escolarizadas e que não dominam a norma, já que foram privadas de oportunidades escolares. Portanto, vale a pena questionar a norma, considerando que a língua não é estática e existem várias realidades dentro de uma mesma língua.

Lynn Mario também reforça que questionar a norma não significa aleatoriedade, mas sim, pensar na língua em determinado contexto, com suas variações (Souza; Hashiguti, 2022, p. 160-161). O professor ainda alerta para o quão perigoso pode ser o fato de se considerar um sotaque como o “correto” e esperar que professores tenham esse determinado sotaque, de forma a negar a própria identidade em busca de outra, valorizada socialmente. Portanto, a atitude decolonial deve ultrapassar o discurso meramente teórico, o que não significa desconsiderar a bibliografia do Norte, mas questionar a hierarquização sem trocar uma norma por outra (Souza; Hashiguti, 2022, p. 172).

Outra contribuição para as discussões de Decolonialidade, é a da Decolonialidade digital e de dados. Com o crescente uso de plataformas digitais educacionais, bem como redes sociais e outras plataformas de uso profissional e pessoal, nossos dados estão cada vez mais expostos em uma grande rede de dados, controlados comumente por organizações multinacionais.

Segundo Pontes (2024), para que esse controle aconteça, as multinacionais devem gerir grande parte da arquitetura do ecossistema digital dos países em desenvolvimento. Isso significa impor o *design*, os códigos dos programas de computador e as licenças de propriedade intelectual para os *softwares* e *hardwares* que constituem a infraestrutura digital dos governos e de seus serviços públicos. Essa ação limita as possibilidades de desenvolvimento local. Como exemplo, no contexto brasileiro, sabe-se que, durante a pandemia do covid-19, a Microsoft

disponibilizou seu pacote educacional, que incluía o *software* de videoconferências (Microsoft Teams), de forma gratuita, para grande parte de universidades e institutos federais.

Ainda segundo Pontes (2024), tal movimento, de aparente caridade, foi acompanhado de outros similares, como a parceria entre a Microsoft e o Ministério da Educação (MEC) em 2022, disponibilizando o Office 365 Educacional (versão 1), incluindo o Teams para acesso por parte das secretarias municipais, estaduais, institutos e universidades federais. Tais iniciativas, que ocorreram sobretudo durante a pandemia causada pelo covid-19, facilitaram as atividades acadêmicas para os alunos beneficiados, mas também reforçam como os países do Sul Global são comumente vítimas do extrativismo de dados, já que as agências de inteligência do Norte Global fazem parcerias para realizarem vigilância em massa nos países do Sul Global.

Vale lembrar que, para que isso aconteça, não há necessariamente o uso de violência, mas o incentivo para que as pessoas compartilhem (de forma “voluntária”) mais frequentemente os seus dados, sobre diferentes esferas de suas vidas. Os mecanismos de controle são menos óbvios e mais sutis, fazendo com que as normas sejam cada vez mais estabelecidas em diversas áreas (educação, lazer, profissional) e ditando o que iremos consumir ou valorizar com cada vez menos poder de escolha. Pontes (2024) ainda questiona o que, de fato, pode estar por trás de iniciativas como as do Google Meet de acesso ilimitado, tendo em vista que a neutralidade não existe.

Ainda sobre a ditadura de consumo que é incentivada, creio comumente afetar também o meio acadêmico, através de autores e obras do Norte Global sugeridos através de divulgação massiva para consumo no Sul Global, bem como práticas educacionais trazidas de fora por diversos meios e que não necessariamente se alinham com as realidades do Sul.

Refletir sobre minha própria prática como docente pesquisador vem abrindo diversas possibilidades de desconstrução relacionadas à decolonialidade. Considero que tive acesso tardio a certas leituras sobre decolonialidade, algo que foi mais presente nas discussões com outros professores de espanhol, sobretudo.

Estudei em franquias de escolas de idiomas por muitos anos, certamente por mais de dez anos, em franquias que certamente traziam parte da cultura estadunidense ou inglesa como culturas superiores, direta ou indiretamente. Como adolescente, certamente não questionaria nada disso, fosse pela imaturidade, fosse pelo fato de a imposição cultural ter sido feita de forma sutil, quase imperceptível. Determinados sotaques eram considerados superiores, e alguns outros poderiam ser motivo de constrangimento, algo que apenas hoje entendo como “complexo de vira-lata”, que nós, de países em desenvolvimento, temos em relação ao que é nosso. É como se tudo o que fosse britânico ou estadunidense estivesse atrelado a um *status* de superioridade,

ao passo que tudo o que é da América do Sul seria considerado inferior ou pouco desenvolvido (músicas, hábitos, livros, língua, autores, filmes, obras culturais de um modo geral). Ironicamente, tal complexo de inferioridade era reforçado por professores brasileiros de língua inglesa, trabalhando em escolas de língua inglesa. O complexo era tamanho, que só o fato de falar a língua inglesa já conferia certo *status* ao professor, fazendo-o ignorar salários extremamente baixos, muitas vezes. Como estudei inglês em tais escolas de idiomas na minha infância e adolescência, período em que estava formando os meus valores e crenças, certamente absorvi imensamente as crenças que reforçavam o nosso complexo de vira-lata através de um sotaque considerado correto, de formas culturais consideradas de maior prestígio, ou me sentindo inferior por não alcançar determinadas pontuações em exames de proficiência reconhecidos internacionalmente. Vale lembrar que estudar idiomas, no Brasil, é algo de pouco incentivo em escolas públicas na educação básica e muitos dos que têm acesso aos estudos o fazem pelo privilégio de estudarem em escolas de idiomas particulares. Logo, a ambição de sucesso nos exames de proficiência é algo especialmente desafiador.

Comecei a minha carreira como professor de inglês em 2014 e, por muito tempo, novamente, trabalhei em franquias de escolas de idiomas e muito provavelmente reproduzi, por diversas vezes, falas que reforçavam a suposta superioridade dos saberes do Norte, seja “corrigindo” pronúncias, valorizando celebrações culturais específicas e incentivando meus alunos a consumirem obras específicas vindas do Norte Global. Era muito comum que nós, professores, escutássemos alunos adolescentes (ou até mesmo adultos) dizendo que a língua espanhola era feia, mas que a língua inglesa era muito importante para a vida profissional/pessoal. Pergunto-me até que ponto a língua espanhola poderia ser considerada “feia” por tais alunos, simplesmente por representar grande parte da América do Sul.

Apesar de já ter estudado espanhol na infância, retomei os meus estudos em 2014, e a minha paixão pela língua cresceu consideravelmente, abrindo espaço para diversas reflexões que iam além de questões estritamente linguísticas. Decidi estudar Letras/Espanhol e, durante o curso, participei do Congresso Brasileiro de Hispanistas em 2022 na cidade de São Paulo (SP), evento acadêmico que reuniu professores de espanhol de todo o país e serviu como um divisor de águas em minha vida. Já estava acostumado a participar de eventos de professores de inglês e percebi que o perfil dos participantes era completamente diferente. Os professores de espanhol abordaram diversas questões políticas ao longo de suas apresentações, e as discussões sobre decolonialidade e complexo de vira-lata sempre vinham à tona. Assisti a apresentações incríveis que abordavam materiais didáticos decoloniais, discussões sobre sotaque e diversas outras contribuições que me transformaram de forma definitiva. As

discussões abordavam não só a valorização de diversos sotaques, como também a representação não normativa de pessoas e celebrações em livros didáticos, favorecendo diversos tipos de beleza, casais homoafetivos, pessoas negras, pessoas com necessidades específicas e datas comemorativas da América do Sul. Eu, que, até então, estava habituado a utilizar livros didáticos estadunidenses, britânicos e espanhóis, passei a prestar mais atenção ao que era considerado padrão em tais livros.

Na minha prática como professor de espanhol, procurei de forma intencional materiais que não fossem unicamente da Espanha, além de materiais audiovisuais de diversos países hispanofalantes. O espanhol é a língua oficial de aproximadamente 21 países (Argentina, Bolívia, Chile, Colômbia, Costa Rica, Cuba, Equador, El Salvador, Espanha, Guatemala, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Porto Rico, República Dominicana, Uruguai, Venezuela, Guiné Equatorial), mas grande parte dos materiais didáticos de espanhol como língua adicional é da Espanha, que, dentre os países hispanofalantes, é o único europeu, reforçando mais a cultura europeia. As discussões sobre padrões de beleza e família sempre me acompanharam, por eu ser parte da comunidade LGBTQ+, além de sul-americano, mas nem sempre recebi autorização para certos debates em escolas de idiomas franquizadas. Até que comecei a dar aulas de forma autônoma, e diversas possibilidades se concretizaram. Minha prática como professor de espanhol, a partir de 2020, veio acompanhada de mais maturidade, munida de leituras sobre decolonialidade.

Se, em muitos casos, estudantes brasileiros de língua inglesa querem parecer nativos ou têm a ilusão de passarem despercebidos como nativos, tentando ao máximo perder seu sotaque (muitas vezes através de cursos formais voltados para tal finalidade), reflito como na língua espanhola é diferente. Em muitos casos, na minha experiência e na de colegas próximos, o “portunhol” é considerado suficiente para viajar e se comunicar, sem esforço excessivo de um estudo formal da língua espanhola. Parece existir uma cobrança interna em ser associado ao Norte Global e apagar os traços do sul.

Minha tese de doutorado, mais do que abordar questões quantificáveis, busca conhecer as referências utilizadas pelos autores dos artigos selecionados para o *corpus* e, principalmente, conhecer as contribuições de “nossos vizinhos”, valorizando suas produções acadêmicas e incorporando tais contribuições na minha prática docente, sempre que possível, e quiçá de outros colegas interessados nesse tema. Trata-se de seguir valorizando os saberes a que já temos acesso (muitos do norte global) e abrir espaço igualmente para pesquisadores da América do Sul coexistirem com os saberes até então hegemônicos. Por fim, em uma nota mais pessoal sobre a minha postura como pesquisador, os estudos decoloniais contribuíram para que eu

retomasse minha própria voz em meus textos, seja expressando a minha opinião seja usando a primeira pessoa, hábito este perdido ao longo dos anos.

A valorização de múltiplas perspectivas e a incorporação de novas práticas foram amplificadas pela necessidade de adaptação imposta por eventos globais, como a pandemia de covid-19. Nesse cenário, a avaliação *on-line* emergiu como um dos maiores desafios para profissionais da educação, exigindo tanto a reconfiguração de práticas avaliativas quanto a reflexão sobre o papel da tecnologia, a exemplo de ferramentas como o ChatGPT, que estão moldando o futuro da educação.

3.5 Mudanças no cenário da avaliação on-line e o ChatGPT

Apesar de instrumentos avaliativos virtuais não serem exclusividade da pandemia, o cenário da avaliação foi drasticamente transformado com a emergência da pandemia causada pelo vírus covid-19. Segundo Seraj *et al.* (2022) e dados da Fundação Lemann (2021), até a metade do ano de 2020, a pandemia impactou 94% dos estudantes em todo o mundo, desde o ensino básico ao superior. Embora o ensino *on-line* seja uma prática familiar, o ERT – sigla para Ensino Remoto de Emergência (*Emergency Remote Teaching*) – é relativamente novo. Enquanto o ensino *on-line* é planejado para o contexto virtual, o ERT é uma transição rápida devido a um possível desastre. Logo, um número considerável de educadores e alunos estava despreparado e desprovido de tecnologias apropriadas para as suas práticas, ainda que diversas instituições adotassem o uso obrigatório de plataformas de acesso gratuito por parte dos professores (Skype, Whatsapp, Youtube etc.).

Ainda segundo Seraj *et al.* (2022), em um estudo que analisou um número considerável de pesquisas sobre avaliação na pandemia, alguns dos principais desafios atrelados à avaliação *on-line* foram: falta de preparo, falta de interesse dos alunos, avaliação *on-line* desafiadora, facilidade para desonestidade, importação do sistema de pontuação e limitação de tempo.

Algumas práticas de avaliação impulsionadas durante a pandemia, provavelmente, permaneceram após o fim da pandemia, tendo em vista que profissionais da educação e alunos estão mais familiarizados atualmente com diversas plataformas virtuais, havendo, assim, menos ansiedade para os envolvidos. Ainda sobre as facilidades, a possível pontuação automática e nota imediata são funções que vão facilitar a vida dos profissionais da educação. A possibilidade de realização de testes em horários diversos também vai dispensar a presença de aplicadores, além de proporcionar flexibilidade de horário, tendo em vista que o aluno pode realizar o teste em um período amplo, determinado pelo professor. Novas formas de se assegurar a identidade

dos alunos também surgiram com as avaliações *on-line*, como o uso de câmera e microfone durante a sua realização, bem como gravação de tela, evitando possíveis consultas e fraudes.

O ChatGPT, assistente conversacional e um dos produtos oriundos dos estudos e das práticas da inteligência artificial, criado com base em bilhões de textos gerados por seres humanos, tem a capacidade de gerar, traduzir, resumir e parafrasear textos, explicar assuntos complexos, corrigir erros e descrever imagens (Meniado, 2023). A Unesco, em 2023, apontou diversas aplicações do ChatGPT no ensino superior, como na área de Ensino e Aprendizagem, Pesquisa, Administração e Envolvimento Comunitário (Meniado, 2023).

A ferramenta mostrou-se útil também para o desenvolvimento de materiais dedicados ao ensino de língua adicional e suas avaliações. Oferece a possibilidade de criar textos adaptados para a realidade dos alunos, assim como *feedback* personalizado e em tempo real, potencializando a experiência de aprendizagem para alunos de língua adicional.

No entanto, o ChatGPT, segundo a Unesco, apresenta também desafios e limitações (Meniado, 2023). Desde o seu lançamento em 2022, preocupações foram levantadas, sobretudo no meio acadêmico, em relação à ameaça à integridade acadêmica, à privacidade de dados, à precisão de informações, ao desenvolvimento de competências interpessoais e aos direitos de propriedade intelectual. Há o risco de que alunos possam fazer uso da ferramenta de forma equivocada ou excessiva.

Dentre os principais impactos e contribuições do ChatGPT para o ensino, aprendizagem e avaliação de línguas, podem ser mencionados:

- Identificação de erros linguísticos superficiais nas produções escritas dos alunos, propiciando melhora na escrita dos estudantes. Oferece *feedback* imediato no desempenho escrito;
- Correção de erros linguísticos, bem como explicações sobre vocabulário, gramática e sintaxe relacionadas aos erros. Oferece recomendações personalizadas;
- Proporciona prática linguística personalizada e interativa, o que pode ajudar os alunos a desenvolverem suas habilidades de produção oral, aumentando, consequentemente, a segurança ao praticar o idioma. O ChatGPT pode atuar como interlocutor em diálogos ou atividades de *role-play* e jogos interativos.

Quanto às contribuições para professores, podem ser citadas:

- Auxiliar na preparação de aulas e no desenho de atividades pensadas nas necessidades dos alunos, bem como diversos níveis e estilos de aprendizagem;
- Gerar novos conteúdos e ideias para diversos temas;

- Contribuir para que professores garantam que suas aulas sejam atrativas, apropriadas e educativas, além de estarem alinhadas aos objetivos educacionais.

Evidentemente, o ChatGPT pode auxiliar os professores a otimizarem tempo e esforço empregados no preparo de aulas, na correção e nas demais funções relacionadas à profissão, tornando-a, assim, menos cansativa.

No que concerne à avaliação, o ChatGPT pode atuar como:

- Gerador de perguntas, produz materiais para testes, gera perguntas de compreensão ou perguntas amplas que podem acompanhar atividades de leitura. Auxilia professores a gerarem tópicos de escrita feitos sob medida, apropriados aos contextos dos cursos e níveis dos alunos;
- Marcador automatizado, para fornecer pontuação em atividades de escrita. Pode atribuir notas automaticamente (baseado em critérios predefinidos), além de oferecer comentários de *feedback*. Oferece correção satisfatória de redações, potencialmente revolucionando os métodos de avaliação da habilidade escrita. Auxilia professores a destacarem pontos fortes e pontos de melhora nas correções de atividades, sobretudo quanto a critérios de adequação ao público leitor, relevância de conteúdo, clareza, propósito, profundidade e organização.

Em suma, o ChatGPT foi associado à facilitação no desenvolvimento de instrumentos avaliativos somativos e formativos, por gerar exercícios e atividades diversas. Pode auxiliar também na pontuação e na correção, atividades que, por vezes, demandam tempo excessivo. Se estabelecidos critérios bem definidos de correção, o ChatGPT pode oferecer correções consideravelmente confiáveis e precisas.

Alguns desafios são, no entanto, associados ao uso do ChatGPT:

- Respostas imprecisas: supõe o significado de expressões ambíguas usadas pelos estudantes, sem antes confirmar a real intenção. Produz respostas que parecem plausíveis, mas que são incorretas;
- Desonestidade acadêmica e plágio: pode ser usado como ferramenta que estimula a desonestidade e seus conteúdos podem não ser detectados por identificadores de plágio. Não fornece as fontes de determinadas citações e suas respostas não são totalmente originais, mas paráfrases de citações sem fonte;
- Deterioração de determinadas habilidades: o apoio excessivo nas respostas da ferramenta pode impedir o desenvolvimento de determinadas habilidades intelectuais, como pensamento criativo, solução de problemas e pensamento crítico.

O complexo processo de leitura, escrita e revisão pode ser simplificado por uma geração automática de textos, com possíveis ajustes posteriores, demandando menor esforço linguístico;

- Desigualdade educacional: alunos com acesso ao ChatGPT e domínio de seu uso podem ter resultados consideravelmente superiores àqueles que não têm acesso a ele nem domínio de seu uso, resultando em uma avaliação injusta.

A Unesco (2023) aponta que o ChatGPT desempenha diversas funções, como tutor privado, *codesigner*, companheiro de estudos, motivador e avaliador.

Considerando-se o suporte que a ferramenta oferece aos professores de línguas por suas diversas funções, é crucial que os professores recebam treinamento apropriado sobre como potencializar o uso da ferramenta para que poupem esforços e tempo dedicados às inúmeras responsabilidades associadas à profissão.

Professores e alunos devem ser treinados para usar a ferramenta de forma criteriosa e ética, visando a resultados positivos. O letramento em inteligências artificiais generativas deve ser promovido através de programas conduzidos nos ambientes escolares, orientando todos sobre o possível uso inadequado da ferramenta.

Pessoalmente, acredito que o ChatGPT tenha trazido contribuições inquestionáveis para o campo da avaliação no ensino de línguas. Do ponto de vista de um professor, é possível personalizar instrumentos avaliativos para perfis específicos de alunos, seja por nível, seja por dificuldades específicas, seja por temas específicos, sem fazer uso excessivo de tempo extraturno. Professores que comumente têm suas vidas pessoais afetadas pela quantidade de trabalho extra fora do ambiente educativo podem finalmente ter mais qualidade de vida, ao passo que consideram as diferenças dos perfis de alunos com a ajuda do ChatGPT.

O simples uso do ChatGPT, por um lado, não vai garantir instrumentos mais acolhedores e diversos, se o professor não souber buscar de forma adequada na plataforma e não tiver sido minimamente letrado em Avaliação. A falta de tempo para produzir e corrigir avaliações, que muitas vezes foi o motivo principal para o *design* de instrumentos padronizados para turmas diversas, agora não é mais uma questão limitante. Professores devem se familiarizar com o ChatGPT antes de usá-lo, bem como buscar capacitação em conhecimentos avaliativos.

Por outro lado, os alunos devem ser sensibilizados sobre o uso responsável do ChatGPT, por ser uma ferramenta que vai facilitar a busca por informações, tornando-os assim mais autônomos, mas trazendo luz para os seus desafios. Cabe aos profissionais dos ambientes educacionais (escolas e universidades) mostrar os pontos positivos e negativos do uso da ferramenta, através de exemplos práticos. Os alunos devem ser incentivados a buscar

informação através de diversas fontes, para confirmarem possíveis notícias e pesquisas (assim como já era necessário na época da mídia impressa e televisiva), bem como conhecer os desafios a curto e longo prazo para o cérebro, pelo fato de o uso excessivo de inteligências artificiais possivelmente coibir o desenvolvimento de determinadas habilidades.

A inteligência artificial (como produto geral) e o ChatGPT dificilmente poderão ser ignorados, pela praticidade que trazem para o dia a dia e, provavelmente, outras similares surgirão. Haverá, assim, uma necessidade de atualização constante por parte de profissionais da educação e dos alunos.

A presença de tecnologias como o ChatGPT no contexto educacional promove reflexões inevitáveis e urgentes sobre os conhecimentos e as habilidades que devem ser priorizadas na formação dos alunos. Tal transformação exige que os Marcos Curriculares e outros documentos oficiais de órgãos educacionais sejam revisitados para incorporar competências digitais e éticas que capacitem tanto os alunos quanto os professores a navegar por um cenário educacional em constante mudança. Os desafios impostos pela realidade tecnológica e pela prática docente, especialmente no ensino público, demandam diretrizes atualizadas que considerem as limitações estruturais e as potencialidades das novas ferramentas.

3.6 Marcos Curriculares

O Marco Curricular de um país é um dos documentos de referência para importantes tomadas de decisão no âmbito educacional. Frequentemente elaborado por autoridades, especialistas, professores e demais partes envolvidas, o Marco Curricular reúne informações como conhecimentos e habilidades essenciais para cada etapa educacional, formas de avaliar e materiais didáticos, além do intuito de promover programas educacionais consistentes em todo o país.

Apesar de o documento servir como referência, a realidade da prática docente no ensino público, com salas numerosas e diversas, professores mal remunerados e poucos recursos didáticos, faz com que a aplicação de suas diretrizes seja, muitas vezes, inviável. A precariedade do ensino também é evidenciada por pesquisas como o Programa Internacional de Avaliação de Estudantes (PISA). O programa tem como objetivo avaliar alunos em três habilidades principais: leitura, matemática e ciências. A pesquisa é feita a cada três anos – a última datada de 2022 – e avalia o desempenho de alunos de 15 anos nas habilidades mencionadas. O resultado da avaliação de 81 países revela que, na habilidade de leitura, os países latino-americanos que configuram o *ranking* estão todos abaixo da média, sendo: Chile (37), Uruguai

(43), Costa Rica (49), México (50), Brasil (53), Colômbia (54), Peru (55), Argentina (58), Panamá (59), Guatemala (67), Paraguai (68), República Dominicana (74) (Ranking..., 2023).

Documentos e pareceres oficiais do Peru, do Chile, do Uruguai e da Colômbia foram selecionados para que suas contribuições sobre avaliação da aprendizagem enriquecessem o presente estudo, sobretudo porque estão disponíveis nos *sites* dos respectivos Ministérios da Educação.

Segundo o Marco Curricular Nacional do Peru (Peru, 2014), é crucial superar o ensino enciclopédico característico do século XIX e preparar crianças, adolescentes e adultos para assumirem papéis ativos e não de meros espectadores passivos na realização de seus objetivos. Ser um protagonista ativo no desenvolvimento humano de um país como o Peru requer que as novas gerações possuam habilidades muito específicas. Elas precisam ser capazes de analisar criticamente realidades complexas e em constante mudança, além de demonstrar grande criatividade. No entanto, o sistema educacional até o presente momento não conseguiu oferecer oportunidades plenamente satisfatórias para que adquiram essas competências, apesar de essas estarem incluídas no currículo oficial desde o final do século XX.

No que concerne à avaliação, o Marco Curricular reserva poucas considerações. Em curto trecho dedicado ao assunto, o Marco Peruano define a avaliação como um processo inerente ao ensino e à aprendizagem que deve estar presente em todas as etapas. A observação e o registro contínuo do desempenho dos estudantes ao longo de suas trajetórias são essenciais, sendo, portanto, importante que o professor tenha clareza em relação ao que espera do aluno e que o aluno saiba o que é esperado dele em relação ao seu desempenho. Tais expectativas só podem ser alinhadas com um programa especialmente desenhado, visando ao desenvolvimento de competências e capacidades.

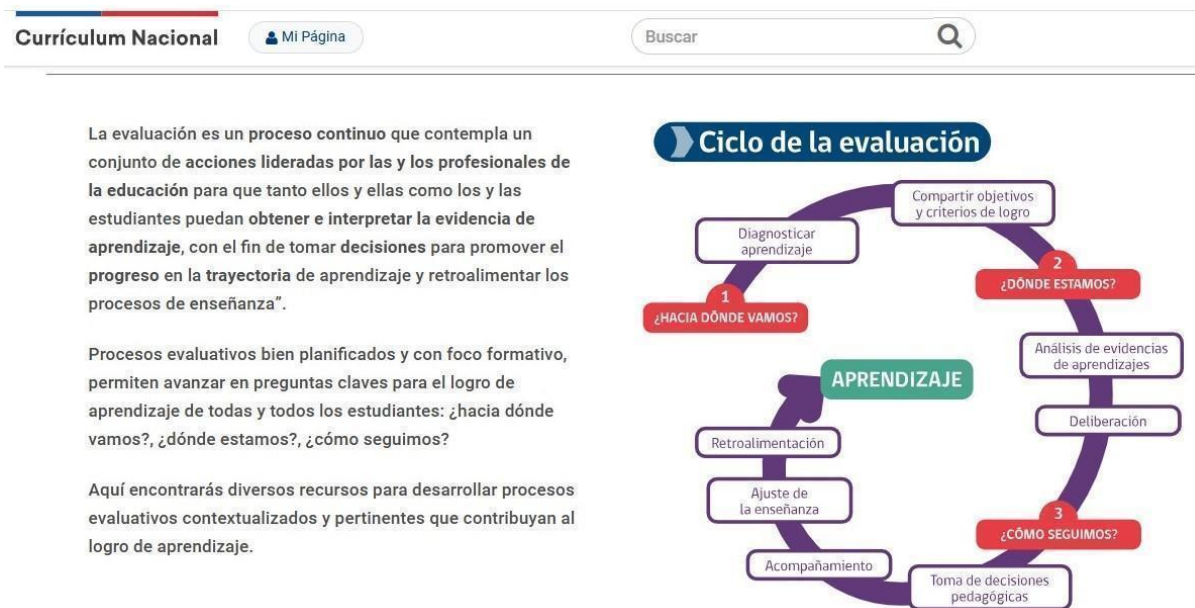
O documento distingue dois tipos de avaliação:

- **Avaliação para a aprendizagem:** serve para comprovar os avanços do aluno ao longo de todo o processo pedagógico e os seus resultados servem para orientá-lo. A reflexão entre o que foi atingido e o que era esperado é o foco das reflexões, resultando na busca por ferramentas que permitam chegar ao esperado. Portanto, a devolutiva fornecida aos alunos é crucial para promover reflexão. É essencial que seja oportuna, assertiva, descritiva, reflexiva, orientadora e ajude os alunos a se autoavaliarem. Assim, os alunos saberão avaliar a qualidade de suas produções e seus desempenhos, tendo como base critérios previamente estabelecidos;
- **Avaliação da aprendizagem ou certificadora:** serve para certificar o nível de aprendizagem alcançado pelo aluno e avaliar o nível de desempenho alcançado em

determinadas competências, desde que os mecanismos de avaliação ofereçam resultados válidos e confiáveis. É aconselhável a implementação de tarefas autênticas e complexas, que exijam o recrutamento de determinadas habilidades do aluno para resolvê-las.

A avaliação certificadora pode demandar do docente o desenho de atividades de nível considerável de exigência cognitiva, em diversos graus de complexidade, desde que com ordens claras. Os alunos devem ser submetidos a avaliações com padrões avaliativos explícitos e que estejam cientes dos seus critérios.

No contexto chileno, o *site* oficial Currículum Nacional disponibiliza, com acesso livre, as Bases Curriculares, os planos de estudo vigentes com a legislação nacional e recursos para auxílio na implementação das práticas em questão. Através de tais recursos, a prática docente é facilitada, colaborando para a formação integral dos estudantes. A disponibilidade de materiais para docentes é destacada, mas alunos e a população como um todo podem acessar os conteúdos no *site* oficial (Chile, 2024), conforme pode ser visto na Figura 1. Quanto à avaliação, especificamente, o cenário é particularmente promissor. O infográfico (Figura 1) do site chileno Currículum Nacional apresenta o ciclo da avaliação da aprendizagem. O processo é representado como um ciclo contínuo, dividido em três grandes perguntas orientadoras: 1) "Para onde vamos?" (diagnóstico da aprendizagem e compartilhamento de objetivos), 2) "Onde estamos?" (análise de evidências e deliberação), e 3) "Como seguimos?" (decisões pedagógicas, acompanhamento e ajustes), destacando a retroalimentação como elemento central.

Figura 1 – Perguntas sobre avaliação no *site* chileno

Fonte: Chile (2024).

O grande diferencial do portal chileno é que existe um apoio adicional para a criação dos instrumentos, com exemplos de avaliação formativa que podem ser adaptados a contextos específicos, bem como um banco de itens para as avaliações.

O Marco Curricular Nacional do Uruguai (Uruguai, 2022) também reserva curto trecho dedicado às reflexões e aos critérios avaliativos. No documento, o currículo é responsável por apresentar as metas de aprendizagem que o estudante deve alcançar, tendo em vista que as metas são um compromisso do sistema educativo com a sociedade. Logo, é importante que se possa identificar se os objetivos foram alcançados pelo aluno, bem como em que grau e o que fazer em caso de não se atingirem os objetivos. Este é um dos papéis da avaliação, seja em um contexto de aulas particulares ou em grupo. O aluno deve ser orientado com informações precisas do que fazer para alcançar o resultado almejado.

A avaliação é uma ferramenta no caminho para o desenvolvimento das aprendizagens e nunca um fim em si mesma. Não se aprende para a avaliação, mas sim, se avalia para que se continue aprendendo, mediante seus efeitos retroativos e ajustes de expectativas por parte do aluno e do professor.

A avaliação deve ter o intuito de desenhar os próximos passos dos alunos para que continuem aprendendo enquanto são avaliados. Dentro dessa concepção de avaliação formativa em um currículo desenvolvido para um desenvolvimento processual das competências, o corpo docente deverá avaliar o progresso do aluno e trilhar as próximas decisões com base nessa

avaliação. Logo, é importante que exista coerência entre o que se espera do aluno e a quais experiências de aprendizagem ele será exposto, já que a avaliação com foco em desenvolvimento de competências é complexa por natureza.

O Ministério da Educação da Colômbia, através do *Marco Nacional de Qualificaciones* (Colômbia, 2023), tem o objetivo de promover uma trajetória de aprendizagem flexível e diversificada. No documento em questão, no que diz respeito à avaliação da aprendizagem, declara que as instituições de ensino possuem autonomia para avaliar os seus alunos e têm como base as regras do Sistema Institucional de Avaliação dos Estudantes (SIEE). O SIEE afirma que a avaliação não é uma tarefa isolada do processo formativo, pelo que deve estar articulada e ser coerente (conceitual, pedagógica e didaticamente) com toda a proposta educativa que o Estabelecimento de Ensino definiu.

No concerne à pandemia, ao retornar presencialmente para a sala de aula, o programa *Evaluar para Avanzar* (Avaliar para Avançar) foi o cerne das práticas avaliativas. Ele consiste em três etapas: 1) a aplicação de uma avaliação formativa, para identificar os atrasos e progressos dos alunos influenciados pela pandemia; 2) suporte aos professores para que possam usar os resultados das avaliações como norteadores para intervenções futuras, 3) mentoria para professores, para apoiar os seus processos pedagógicos.

Nos documentos analisados, menções são feitas ao caráter processual da avaliação, tendo em vista que a avaliação não deve ser um fim, mas uma orientação para que se continue aprendendo. Questões como o alinhamento do que se espera do aluno e critérios claros também são apontadas, além de possíveis intervenções futuras tomadas com base nos resultados das avaliações. O impacto causado pela pandemia também foi brevemente considerado no documento colombiano, informação que deveria ser detalhada e aprofundada em todos os documentos oficiais dos mais diversos países, tendo em vista o impacto inegável causado pela pandemia do covid-19 e por tecnologias digitais como o ChatGPT nas possíveis formas de avaliação da aprendizagem. Os pareceres sobre avaliação são rígidos, superficiais e distantes da prática em diversos países. A pandemia acarretou a adoção de diversas formas de avaliar e interagir através de diferentes plataformas virtuais, promovendo novas possibilidades de pesquisa na área de Avaliação e trazendo um novo olhar para a discussão acerca da confiabilidade. Ainda assim, após anos de práticas e teorias pouco aprofundadas e distantes da realidade, é possível que alguns documentos oficiais não sejam atualizados e ignorem as mudanças causadas pela pandemia e pelas tecnologias digitais na educação e avaliação.

Finda a seção 3, que contém aportes teórico-metodológicos em torno do tema da avaliação, em contexto sobretudo de países de fala hispânica da América do Sul, passo agora a

descrever a metodologia adotada nesta tese.

4 METODOLOGIA DA PESQUISA

Os periódicos acadêmicos e suas publicações de artigos permitem uma ampla análise e problematização das principais tendências de pesquisas internacionais e nacionais em diversos campos de estudo (Vanti, 2002; Ruiz; Jorge, 2002). A bibliometria, subárea das Ciências da Informação, permite um levantamento, inicialmente quantitativo, dessas pesquisas publicadas em determinadas áreas de estudo, as principais épocas de publicação, as problematizações mais recorrentes, as regiões geográficas mais ativas, os idiomas mais utilizados, os autores mais citados, bem como outros dados igualmente relevantes. Conforme Braga (1974, p. 162 *apud* Vanti, 2002, p. 155),

A Bibliometria examina, primeiramente, as relações entre diferentes variáveis: recursos humanos-documentos, artigos-periódicos, produção-consumo etc., que apresentam diversas regularidades de distribuição. O número de artigos que originam n citações, o número de instituições produzindo anualmente n doutorados, o número de autores com n artigos, o número de revistas contendo n artigos constituem exemplos do mesmo tipo de distribuição.

Usando a bibliometria como técnica de pesquisa, esta tese busca coletar e analisar artigos sobre avaliação, sobretudo em contexto latino-americano, conforme já anunciei.

Aliada à técnica de pesquisa bibliométrica, esta tese também faz uma análise de conteúdo dos artigos encontrados com o tema avaliação no ensino de línguas. A Análise de Conteúdo, na qualidade de metodologia para a análise dos dados obtidos por levantamento bibliométrico, serve, como apontam Bardin (2011) e Mendes e Miskulin (2017), para determinar um núcleo de discussão, que, nesta tese, é a avaliação no ensino de línguas.

A Análise de Conteúdo, segundo Bardin (2011), leva em consideração as significações, isto é, os conteúdos de determinados produtos publicados e, igualmente, a sua forma e a coocorrência de distribuição desses conteúdos. Tal análise busca conhecer o que está por trás das palavras sobre as quais se debruça e visa ao conhecimento de variáveis de ordem psicológica, sociológica, histórica etc., por meio de um mecanismo de dedução com base em indicadores reconstruídos com base em uma amostra de mensagens particulares.

As análises de conteúdo ocorrem através de documentos, isto é, de uma análise primeiramente desses produtos, denominada análise documental. A análise documental ocorre, sobretudo, por classificação-indexação, seguida de uma análise categórica temática, que se configura como uma das técnicas da análise de conteúdo. O objetivo da análise documental é a

representação condensada da informação, para consulta e armazenamento; o objetivo da análise de conteúdo é a manipulação de mensagens (conteúdo e expressão do conteúdo em questão).

Para Bardin (2011), a análise de conteúdo, que engloba a documental, passa por fases. A pré-análise representa a fase de organização. Corresponde a um período de intenções do pesquisador, mas tem por objetivo sistematizar e tornar operacionais as ideias iniciais, visando a conduzir a um esquema preciso do desenvolvimento das operações sucessivas em um plano de análise. Recorrendo ou não ao computador, trata-se de estabelecer um programa que, podendo ser flexível (permitindo a introdução de novos procedimentos no decurso da análise), deve, no entanto, ser preciso.

De modo geral, esta primeira fase possui três intuitos: a escolha dos documentos a serem submetidos à análise; a formulação das hipóteses e dos objetivos; e a elaboração de indicadores que fundamentam a interpretação final. Os três fatores não se sucedem, obrigatoriamente, segundo uma ordem cronológica, embora mantenham uma forte conexão: a escolha de documentos depende dos objetivos, ou, inversamente, o objetivo só é possível em função dos documentos disponíveis; os indicadores serão construídos em função das hipóteses, ou, pelo contrário, as hipóteses serão criadas na presença de certos índices. A pré-análise, por fim, tem por objetivo a organização, embora ela própria seja composta de atividades não estruturadas, isto é, “abertas”, por oposição à exploração sistemática dos documentos.

Como parte das etapas iniciais, a leitura “flutuante” consiste em estabelecer contato com os documentos, analisar e conhecer o texto, deixando-se invadir por impressões e orientações. Gradualmente, a leitura vai se tornando mais precisa, em função de hipóteses emergentes, da projeção de teorias adaptadas sobre o material e da possível aplicação de técnicas utilizadas sobre materiais semelhantes. Posteriormente, na fase de escolha dos documentos, convém escolher o universo de documentos suscetíveis de fornecer informações sobre o problema levantado.

Com o universo demarcado (o gênero de documentos sobre os quais se pode efetuar a análise), é necessário proceder à constituição de um *corpus*. *Corpus* é o conjunto dos documentos a serem submetidos aos procedimentos analíticos. A sua constituição implica, muitas vezes, escolhas, seleções e regras, sendo as principais:

- Regra da exaustividade: uma vez definido o campo do *corpus*, é preciso levar em consideração todos os elementos desse corpus;
- Regra da representatividade: a análise pode acontecer em dada amostra desde que o material a isso se preste. A amostragem pode ser considerada rigorosa, se a amostra for uma parte representativa do universo inicial;

- Regra da homogeneidade: os documentos retidos devem ser homogêneos, isto é, devem obedecer a critérios precisos de escolha e não apresentar muitas particularidades fora desses critérios;
- Regra de pertinência: os documentos retidos devem ser adequados, como fonte de informação, de modo a corresponderem ao objetivo que suscita a análise.

No que tange à formulação das hipóteses e dos objetivos, ainda como uma das etapas explicitadas por Bardin (2011), a autora explica que a hipótese é uma afirmação provisória que nos propomos verificar (confirmar ou infirmar), recorrendo aos procedimentos de análise. Trata-se de uma suposição cuja origem é a intuição e que permanece em suspenso enquanto não for submetida à prova de dados seguros. Os objetivos, por sua vez, referem-se à finalidade geral a que nos propomos (ou que é fornecida por uma instância exterior). As hipóteses nem sempre são estabelecidas na pré-análise. Por outro lado, não é obrigatório ter como guia um *corpus* de hipóteses, para se proceder à análise. Algumas análises ocorrem sem ideias preconcebidas. Uma ou várias técnicas são consideradas adequadas a priori, para fazerem a matéria “falar”, desde que utilizadas sistematicamente. Desde a pré-análise devem ser determinadas operações de recorte do texto em unidades comparáveis de categorização para análise temática e de modalidade de codificação para o registro dos dados.

Vale ressaltar que antes das análises propriamente, tanto bibliométrica quanto de conteúdo, o material reunido deve ser preparado. Trata-se de uma preparação material e, eventualmente, de uma preparação formal (edição). Por exemplo, no caso desta tese, todos os artigos das revistas latino-americanas *on-line* podem ser descarregados em formato .pdf e armazenados em um diretório para posteriores aferições.

Com base nessa metodologia, os resultados obtidos permitem que sejam feitas inferências e interpretações correspondentes aos objetivos propostos, ou ainda, que surjam respostas não previstas.

Assim, como mencionei antes, o objetivo geral desta tese é investigar a frequência e o foco das principais pesquisas acerca da avaliação no ensino de línguas em periódicos internacionais latino-americanos, bem como compreender as principais práticas avaliativas que emergem de tais artigos. No que concerne aos objetivos específicos, esta tese pretende i. identificar possíveis tendências e diferenças de foco de pesquisas acerca do tema avaliação no ensino de línguas em periódicos internacionais; ii. problematizar possíveis diferenças de foco em pesquisas sobre avaliação; e iii. analisar certo volume de dados mais proeminentes, contrastando-os com aqueles já coletados em Borja (2020a);

Em minha visão, por meio do levantamento quantitativo, os resultados dão margem a

interpretações qualitativas, que, associadas a dados políticos e sociais, como aqueles expostos na seção 3, permitem o aprofundamento e o questionamento de tendências investigativas.

Os artigos científicos sobre avaliação foram coletados tendo a base de dados da SCImago como referência. A SCImago é uma base de dados destinada à comunidade científica que oferece uma visão mais completa e precisa sobre a qualidade e o impacto da produção acadêmica. Ao utilizar a SCImago, pesquisadores, instituições e bibliotecas podem tomar decisões mais embasadas sobre onde publicar, onde buscar informações e como avaliar o desempenho de diferentes áreas do conhecimento.

Optei por investigar e selecionar aqueles periódicos cujos *sites* fossem de domínio público e de livre acesso. Todos os artigos encontrados sobre o tema da Avaliação nesses periódicos foram lidos, em busca de contribuições que fossem além (ou não) de temas como Quadro Europeu Comum e Exames de proficiência britânicos ou norte-americanos (não desmerecendo tais contribuições, mas com anseio por ouvir outras vozes normalmente ignoradas) já encontrados em minha dissertação de mestrado (Borja, 2020a).

A busca abarcou artigos de avaliação em língua materna e língua adicional, visando a ampliar os resultados obtidos previamente. Incluir artigos em língua materna e adicional me pareceu importante para expandir significativamente o escopo da pesquisa, acessando um corpo de conhecimento que pode não estar disponível apenas em língua estrangeira. Além disso, essa expansão se mostrou particularmente importante em áreas de estudo que são fortemente influenciadas por pesquisas internacionais, como a Avaliação, nas quais as teorias e práticas muitas vezes transcendem fronteiras linguísticas.

Não foi possível detectar o mesmo recorte de tempo daquele constatado em Borja (2020a), a saber, pesquisas sobre avaliação publicadas entre os anos de 1985 e 2018. No *corpus* da SCImago, o primeiro artigo sobre Avaliação encontrado data de 1994 e o mais atual, 2023.

Ainda com base nos resultados, busquei encontrar parâmetros que explicassem até que ponto pareceres oficiais sobre avaliação influenciam na diversidade de temas pesquisados sobre avaliação no ensino de línguas, assim como a pandemia e demais questões políticas e sociais. *Sites* oficiais dos principais órgãos responsáveis pela educação foram analisados com base nos resultados quantitativos e qualitativos dos periódicos. As tecnologias digitais (ChatGPT) e os efeitos da pandemia do covid-19 na educação também são considerados como possíveis influenciadores.

5 RESULTADOS

Neste capítulo, atrelo-me aos resultados das pesquisas bibliométricas em 5.1 e na análise de conteúdo em 5.2.

5.1 Resultados bibliométricos: dados gerais

A base de dados da SCImago indexa, ao todo, 53 periódicos de diferentes nacionalidades, dentre elas: chilenos, colombianos, mexicanos, argentinos, peruanos, venezuelanos, uruguaios e brasileiros. Embora os periódicos brasileiros não sejam o foco desta tese, o Quadro 2 exhibe o número total de periódicos. Também foram descartados os periódicos cujos *sites* estavam fora do ar ou que não publicaram artigos sobre Avaliação. A busca foi realizada durante o ano de 2023, podendo, obviamente, sofrer alterações em anos subsequentes. Vale ressaltar que quando fiz as primeiras buscas, por exemplo, encontrei um artigo intitulado *Evaluación de la escritura en niños de tercer grado en Brasil*, publicado no periódico Boletín de Linguística, da Venezuela. Nos meses seguintes, esse periódico já não estava mais disponível e por isso não faz parte dos dados desta tese. Para Vanti (2002), a instabilidade das bases de dados pode se configurar como um dos problemas relacionados às análises bibliométricas.

Quadro 2 – Periódicos da base de dados da SCImago

Nº	Título do periódico	País
1	Literatura y Linguística	Chile
2	Boletim do Museu Paraense Emilio Goeldi: Ciencias Humanas	Brasil
3	Profile: Issues in Teachers' Professional Development	Colômbia
4	Nueva Revista del Pacifico	Chile
5	Boletín de la Academia Peruana de la Lengua	Peru
6	CoDAS	Brasil
7	Revista Signos	Chile
8	Boletin de Filologia	Chile
9	Logos: Revista de Linguística, Filosofía y Literatura	Chile
10	Mutatis Mutandis	Colômbia
11	Ikala	Colômbia

Nº	Título do periódico	País
12	Onomazein	Chile
13	RLA	Chile
14	Revista Brasileira de Linguística Aplicada	Brasil
15	Cuadernos de Musica, Artes Visuales y Artes Escenicas	Colômbia
16	Latin-American Journal of Discourse Studies	Brasil
17	Acta Comportamentalia	México
18	Cadernos de Tradução	Brasil
19	DELTA Documentação de Estudos em Linguística Teórica e Aplicada	Brasil
20	Estudios Filológicos	Chile
21	Texto Livre	Brasil
22	Forma y Función	Colômbia
23	Bakhtiniana	Brasil
24	Mextesol Journal	México
25	Caracol	Brasil
26	Lenguas Modernas	Chile
27	Revista Chilena de Fonoaudiología	Chile
28	Revista Argentina de Ciencias del Comportamiento	Argentina
29	Estudios de Literatura Colombiana	Colômbia
30	Cuadernos de Linguística Hispánica	Colômbia
31	Contratexto	Peru
32	Calidoscopio	Brasil
33	Comunicación y Medios	Chile
34	Metafora	Peru
35	Cogency	Chile
36	Praksis	Brasil
37	Letras (Peru)	Peru
38	Revista de Estudos da Linguagem	Brasil
39	Estudios de Asia y Africa	México
40	Nueva Revista de Filología Hispánica	México

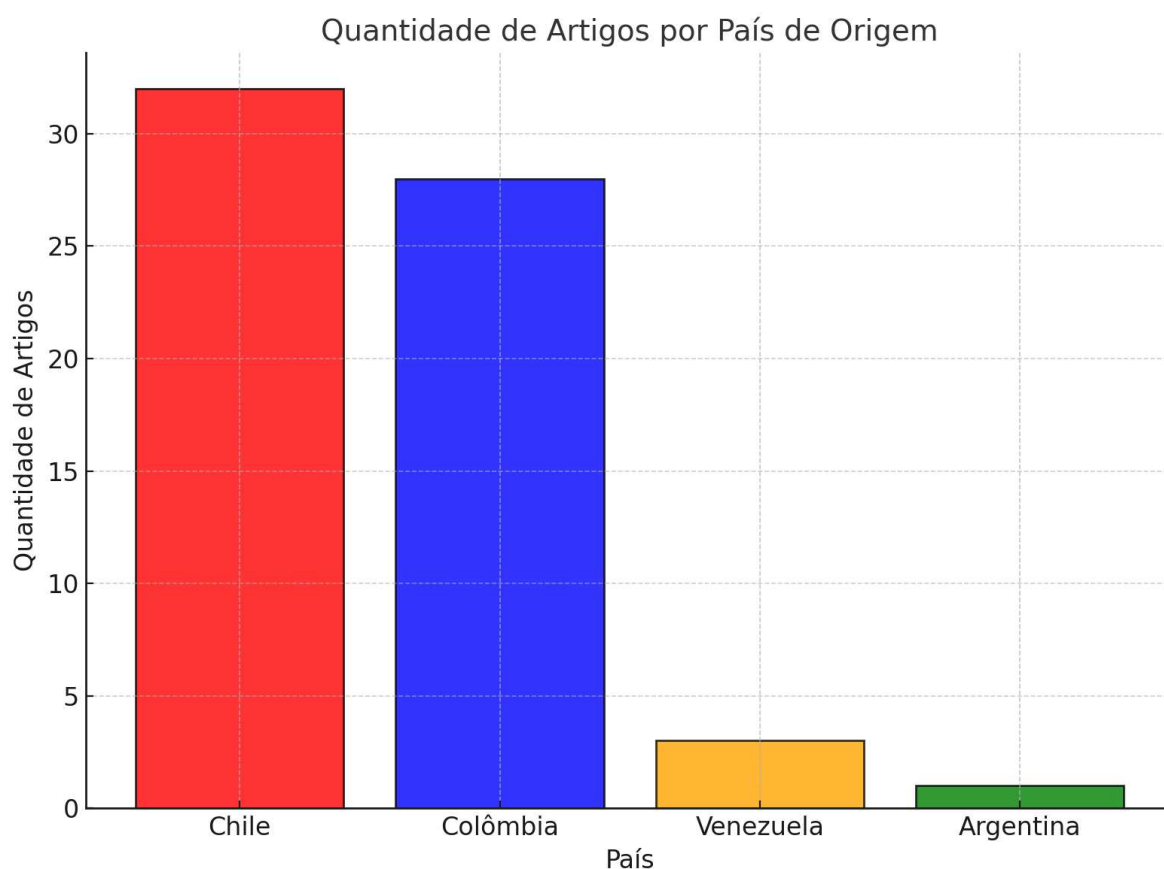
Nº	Título do periódico	País
41	Alea: estudos neolatinos	Brasil
42	Lengua y Habla	Venezuela
43	Linguística	Uruguai
44	Lexis (Peru)	Peru
45	Signo y Pensamiento	Colômbia
46	Estudios de Teoria Literaria	Argentina
47	Brazilian English Language Teaching Journal	Brasil
48	Anclajes	Argentina
49	Cadernos de Linguagem e Sociedade	Brasil
50	Acta Scientiarum Language and Culture	Brasil
51	Colloquia Germanica	Chile
52	Machado de Assis em Linha	Brasil
53	Espaço Plural	Brasil

Fonte: elaboração própria.

As palavras-chave utilizadas na busca pelos artigos nos *sites* dos periódicos acadêmicos foram em espanhol (*evaluación, prueba, examen, lenguas*) e em inglês (*test, assessment, exam, language*). Uma busca minuciosa por artigos foi feita e repetida manualmente nas abas de busca de cada *site* dos periódicos, visando a confirmar se ainda restavam artigos cujo tema era avaliação. As buscas foram feitas durante o segundo semestre de 2023.

A Figura 2 exibe o número de periódicos por país que publicaram sobre o tema avaliação no ensino de línguas.

Figura 2 – Número de periódicos por país



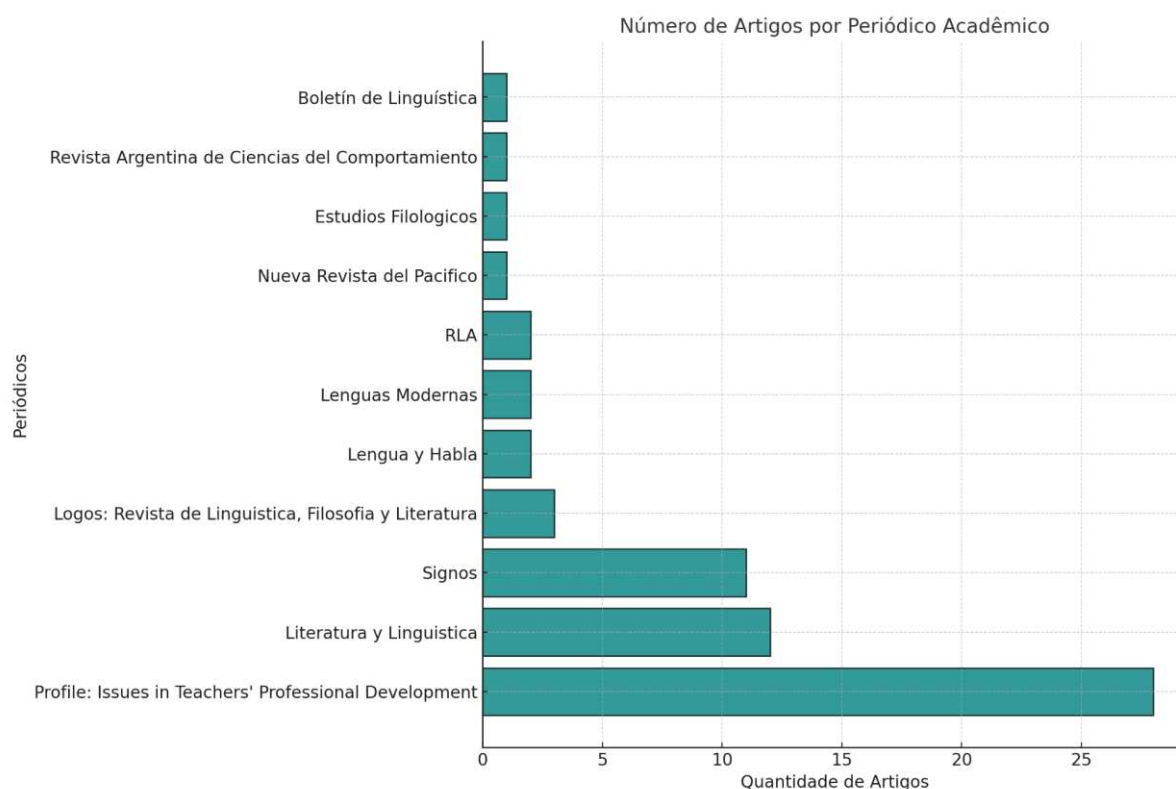
Fonte: elaboração própria.

Após seleção de artigos latino-americanos, com recorte temporal entre 1994 e 2023, descartando aqueles trabalhos que divergiam do tema, 64 artigos no total compuseram o *corpus* desta tese, compreendendo, ao todo, 11 periódicos de quatro países diferentes.

Esta tese, assim, abarcou periódicos que investigaram o ensino de línguas (materna ou adicional), ampliando meus estudos iniciais feitos unicamente com línguas adicionais.

Dentre os artigos selecionados, 28 foram publicados no periódico colombiano Profile: Issues in Teachers' Professional Development; 12, no periódico chileno Literatura y Linguística; 11, na revista chilena Signos; 3, no periódico chileno Logos: Revista de Linguística, Filosofía y Literatura; 2, no periódico venezuelano Lengua y Habla; 2, no periódico chileno Lenguas Modernas; 2, no periódico chileno RLA 1, no periódico chileno Nueva Revista del Pacífico; 1, no periódico chileno Estudios Filológicos; 1, no periódico argentino Revista Argentina de Ciencias del Comportamiento; 1, no periódico venezuelano Boletín de Lingüística, como exhibe o Gráfico 1 a seguir.

Gráfico 1 – Nomes dos periódicos e quantidade de artigos por publicação

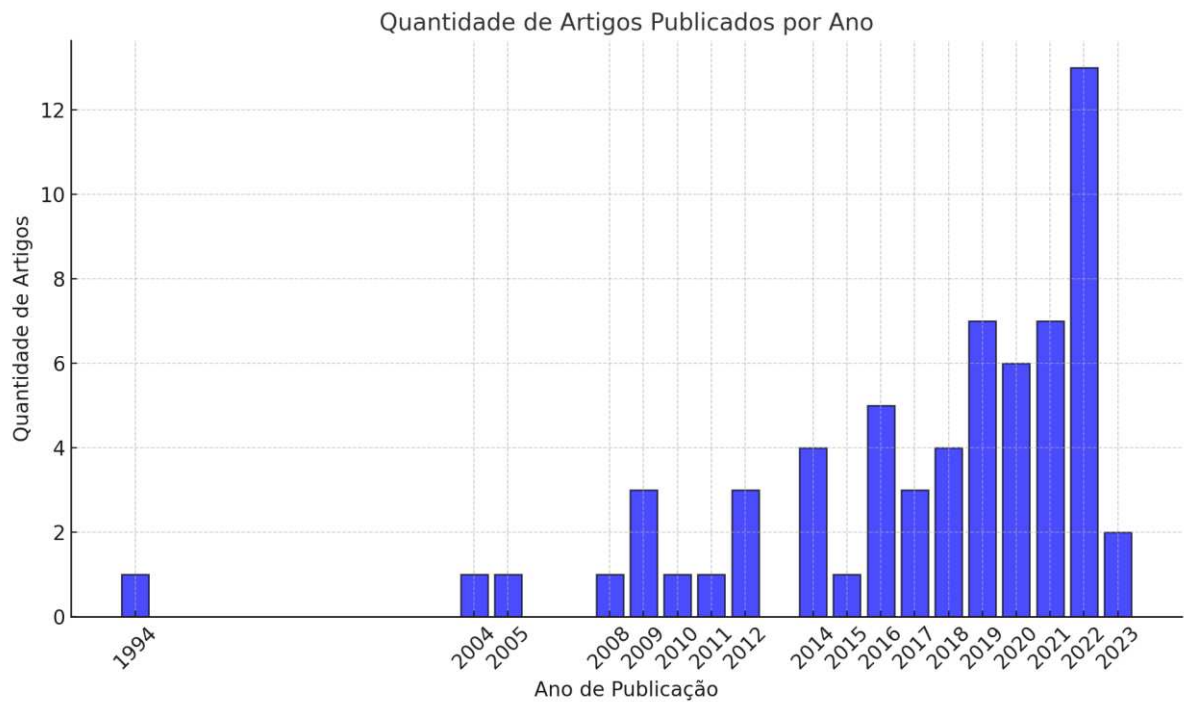


Fonte: elaboração própria.

Tais artigos foram confrontados com os 38 artigos nacionais anteriormente analisados em Borja (2020a). Se, em nível nacional, houve prevalência por artigos de um único estado (o estado de São Paulo), na pesquisa atual, em nível latino-americano amplo, o maior número de artigos advém de periódicos colombianos e chilenos.

Dentre os periódicos cujas datas de publicação constavam nos arquivos, o recorte de data de publicações variou entre 1994 e 2023, sendo o maior volume de publicações sobre Avaliação em 2014, 2016, 2018, 2019, 2020, 2021 e 2022, representando um número considerável de artigos recentes, algo extremamente positivo para a área de Avaliação. Não foram considerados nessa relação aqueles artigos que não especificavam a data de publicação, mas foram igualmente lidos e analisados.

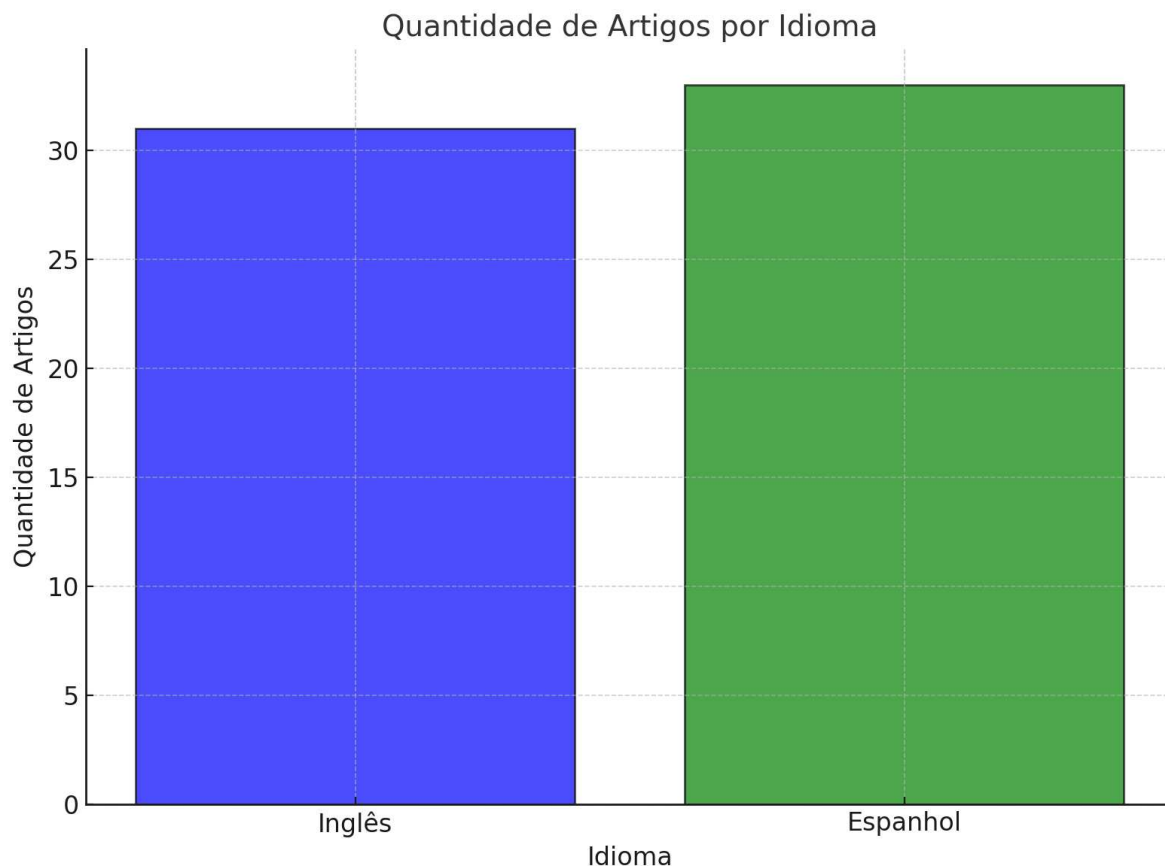
Gráfico 2 – Publicações no período de 1994 a 2023



Fonte: elaboração própria.

Dentre os artigos selecionados, 31 foram escritos em língua inglesa, e 33, em língua espanhola, o que pode representar forte motivação para atingir leitores do Norte Global, conforme exhibe o Gráfico 3 a seguir.

Gráfico 3 – Línguas nas quais os artigos foram escritos



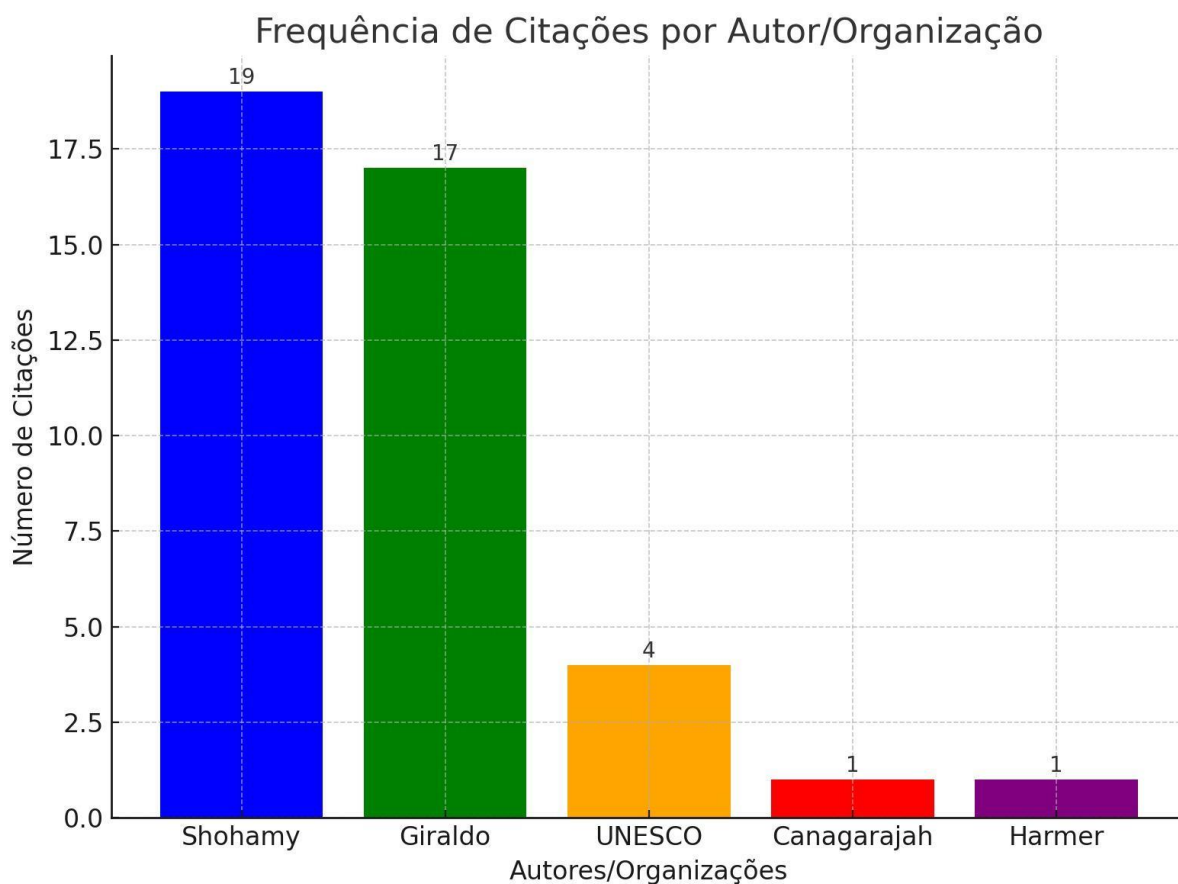
Fonte: elaboração própria.

Dentre os autores/fontes referenciados na Fundamentação Teórica de Avaliação desta tese (cf. Seção 3), poucos se repetiram nas referências bibliográficas dos 64 artigos sul-americanos.

Autores brasileiros como Méndez (2002), Scaramucci (2000), Rodrigues, Rocha e Gonçalves (2011), Luckesi (2000) e Day (2012) não foram referenciados nas pesquisas latino-americanas, o que mostra o nosso distanciamento acadêmico com os “nossos vizinhos” do Sul, já que autores brasileiros com importantes contribuições para o campo da Avaliação não foram mencionados em pesquisas desses países. Outros autores (de distintas nacionalidades) referenciados nesta tese, como Seraj *et al.* (2022), Meniado (2023) e Kumaravadivelu (2012 *apud* Santos, 2022), igualmente não se repetiram nas referências do *corpus*.

Por outro lado, alguns autores e algumas fontes foram referenciados com maior e menor frequência, conforme mostra o Gráfico 4 a seguir.

Gráfico 4 – Citações nas publicações latino-americanas



Fonte: elaboração própria.

Shohamy foi referenciada 19 vezes nos textos do *corpus*; Giraldo, 17; a Unesco, 4; Canagarajah, 1; e Harmer, 1. Vale ressaltar que as referências dos autores ou das fontes mencionados não são necessariamente do mesmo ano em que citei na Fundamentação Teórica desta tese, isto é, na Seção 3. Elana Shohamy é professora universitária em Israel; Frank Giraldo é professor universitário na Colômbia; Suresh Canagarajah é um linguista cingalês e professor universitário nos Estados Unidos; e, por fim, Jeremy Harmer é um autor/professor britânico com experiência docente no Reino Unido e no México. Nenhuma das referências mais recorrentes foi de pesquisadores brasileiros. Nas referências do *corpus*, houve predominância de referências em língua inglesa e espanhola.

Autores como Mónica Véliz, Frank Giraldo e Tania Tagle-Ochoa, que publicaram mais de um artigo do *corpus* da presente pesquisa, não tiveram, em suas referências bibliográficas mais recorrentes, obras consultadas em comum. No entanto, um número considerável de obras em inglês foi consultada como referência. No caso específico de Frank Giraldo, houve um uso comum de textos de sua própria autoria em suas referências. De todas as maneiras, ele é o autor

que mais publicou no *corpus* aqui analisado, com um total de três publicações pertencentes aos anos de 2018, 2020 e 2022, sugerindo sua preocupação um tanto persistente com o tema da avaliação.

Os demais trabalhos citados na minha Fundamentação Teórica, como Souza e Hashiguti (2022), Maldonado-Torres (2007 *apud* Souza; Hashiguti, 2022), Cusicanqui (2012 *apud* Souza; Hashiguti, 2022), Matos e Santos (2022) e Pontes (2024), não foram considerados na análise em questão, por representarem contribuições sobre Decolonialidade, sendo o meu foco aqui a comparação de recorrências de referências sobre Avaliação.

Todas as referências bibliográficas de cada um dos 64 artigos foram digitadas manualmente e salvas em um documento único do Google Docs.

5.2 Quadros com resultados da análise de conteúdo

Para apresentar, de forma clara e concisa, os resultados a seguir, exponho os dados concernentes à análise do conteúdo dos artigos analisados em quadros e, em seguida, apresento um breve resumo explicativo sobre eles (Bardin, 2011).

Vale ressaltar que as categorias presentes nos quadros – dados quanto ao tipo de avaliação, população avaliada, competências, métodos e instrumentos utilizados nas avaliações, desafios, limitações e recomendações referentes aos 64 artigos coletados, por ordem de data, isto é, de 1994 a 2023 – emergiram na fase de pré-análise dos artigos, conforme preconiza Bardin (2011).

5.2.1 *La evaluación de la comprensión de textos escritos: el problema del resumen – Revista Lenguas Modernas – 1994*

O Quadro 3 exibe partes do conteúdo do artigo de Marianne Peronard (1994). O texto explora a viabilidade de se avaliar a compreensão de leitura, focalizando o papel do leitor e considerando as limitações do texto, a língua e o conhecimento social.

Quadro 3 – Peronard (1994)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Marianne Peronard, Universidad de Chile (1994).
Tipo de Avaliação	O estudo utiliza a técnica de avaliação por meio de um teste de compreensão de textos escritos, que inclui perguntas abertas que exigem que os alunos resumam

	o conteúdo do texto lido.
População Avaliada	O público-alvo do estudo consiste em 372 estudantes, sendo 198 de colégios privados e 174 de escolas públicas, em nível de escolaridade equivalente (oitavo ano de educação básica no Chile e primeiro ano de educação secundária na Argentina).
Competências Avaliadas	As principais competências avaliadas incluem a compreensão de leitura, a capacidade de descobrir o significado de termos desconhecidos, reconhecer relações causais e estabelecer uma macroestrutura do texto.
Métodos e Instrumentos	Os testes consistem em perguntas abertas que pedem aos alunos para resumir o texto lido. O texto utilizado no teste é “El peligro de las transfusiones”, e as perguntas exigem que os alunos utilizem estratégias de compreensão para responder.
Resultados e Conclusões	Os resultados mostraram diferenças leves na extensão das respostas entre os alunos de Córdoba e Valparaíso, mas essas diferenças não foram estatisticamente significativas. O estudo conclui que a avaliação da compreensão é complexa e que a técnica do resumo pode ser uma ferramenta útil, embora não forneça uma medida completa da compreensão textual.
Desafios e Limitações	Um dos desafios mencionados é a dificuldade de explicar as diferenças observadas nos grupos, especialmente em relação ao nível de escolaridade. Além disso, o mero relato de palavras não reflete adequadamente a compreensão textual.
Recomendações e Futuras Pesquisas	O estudo não menciona explicitamente recomendações para estudos futuros, mas sugere que a diversidade de enfoques e a consideração de diferentes métodos de avaliação podem enriquecer a pesquisa sobre a compreensão de textos.

Fonte: elaboração própria.

O estudo avaliou a compreensão de leitura de 372 alunos do oitavo ano no Chile e do primeiro ano na Argentina, utilizando um teste que incluía a elaboração de resumos de texto. As competências avaliadas abrangeram a compreensão textual, identificação de significados, relações causais e estrutura textual. Os resultados identificaram pouca variação na extensão dos resumos entre os países, sem diferenças significativas. A pesquisa concluiu que a avaliação da compreensão é complexa e que o resumo é uma ferramenta útil, embora apresente limitações quanto à compreensão integral do texto e sugere a exploração de diferentes métodos de avaliação para enriquecer a pesquisa.

5.2.2 Teacher correction versus peer-marking – Revista Profile – 2004

O Quadro 4 exhibe partes do conteúdo do artigo de Mariana Correia Mourente Miguel (2004). O artigo descreve um experimento com o objetivo de avaliar os dois métodos de correção de composições mais amplamente utilizados – a correção tradicional pelo professor e a correção por pares – e seu efeito na frequência de erros.

Quadro 4 – Mourente Miguel (2004)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Mariana Correa Mourente Miguel, Universidade Federal do Rio de Janeiro (2004).
Tipo de Avaliação	Redação em língua inglesa.
População Avaliada	Estudantes de língua inglesa de uma instituição pública. Alunos do nível A2 do Quadro Europeu Comum, com idades entre 18 e 27.
Competências Avaliadas	Produção escrita.
Métodos e Instrumentos	Redações corrigida por professores x redações corrigidas por pares (analisadas empiricamente).
Resultados e Conclusões	A avaliação por pares auxilia os alunos a assumirem um papel mais ativo no processo de aprendizagem e avaliação. Os resultados entre as correções feitas por professores e por pares foram similares e sem diferenças significativas.
Desafios e Limitações	Recorte limitado, apenas com estudantes brasileiros.
Recomendações e Futuras Pesquisas	Seria apropriado utilizar ambas as formas de correção (professor e pares), atendendo a diferentes perfis de alunos. Realizar futuramente uma pesquisa em nível latino-americano.

Fonte: elaboração própria.

Com base na análise dos dados, é possível concluir que a capacidade de identificar erros não foi significativamente melhor na avaliação por pares em comparação com os alunos que foram avaliados pelos métodos tradicionais, já que ambos os grupos apresentaram uma melhoria similar. Assim, pode ser mais eficaz para os professores adotar ambos os estilos de correção em suas aulas. Dessa forma, alunos que possam se sentir desconfortáveis com um dos métodos têm a chance de se beneficiar do outro.

Uma limitação do estudo é que ele foi conduzido exclusivamente com alunos brasileiros. Portanto, o experimento não leva em conta como o contexto cultural pode influenciar a reação dos alunos a diferentes métodos de avaliação. São necessárias mais pesquisas, que incluam contextos de ensino de inglês como língua estrangeira em outros países da América Latina, para obter uma compreensão mais aprofundada de como a avaliação funciona durante o processo de aprendizagem de uma língua.

5.2.3 Using students' analysis in teaching public speaking for business – Revista Profile – 2005

O Quadro 5 exibe partes do conteúdo do artigo de M. Ashraf Rizvi (2005). O artigo investiga a participação do aluno em um programa de comunicação oral de um currículo de graduação em negócios e comércio na Universidade Sultan Qaboos, Sultanato de Omã.

Quadro 5 – Ashraf Rizvi (2005)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	M. Ashraf Rizvi, Indian School of Mines, Dhanbad, Índia (2005).
Tipo de Avaliação	Avaliação por pares da habilidade de comunicação oral (com o auxílio de folhas de avaliação e critérios estabelecidos). A habilidade de comunicação é enfatizada em seminários e grupos de discussões para promover segurança e confiança.
População Avaliada	Alunos universitários omanenses (20), com estudos focalizados em comunicação empresarial e oratória.
Competências Avaliadas	Habilidade de comunicação oral, em um curso voltado para a fala em público em inglês (com ênfase empresarial/em negócios).
Métodos e Instrumentos	Os alunos do estudo foram encorajados a opinar sobre o formato do curso, bem como avaliar a <i>performance</i> oral de seus colegas de curso (avaliação por pares) através de <i>feedback</i> oral.
Resultados e Conclusões	Concluiu-se que qualquer conhecimento técnico não poderá ser potencializado a não ser que um funcionário se comunique bem. Os alunos parecem mais interessados em aprender habilidades que sejam imediatamente ligadas aos seus objetivos e ambições. O <i>feedback</i> da avaliação por pares cria oportunidades de interação oral e um retorno imediato a respeito da <i>performance</i> oral dos colegas. Todos os alunos dão e recebem <i>feedback</i>. O <i>feedback oral</i> poupa tempo, especialmente em turmas grandes. Além de ser uma boa prática para futuros professores. Os alunos quiseram se envolver nas escolhas avaliativas e demandam o uso de diversos métodos de avaliação. Os estudantes também demonstraram precisar de mais tempo para a realização de trabalhos de qualidade, assim como expressaram preferência por <i>feedback</i> constante de seus progressos. Logo, eles precisaram manter um canal aberto com os professores para receber direcionamentos. Envolver os alunos no processo de avaliação se mostrou uma escolha frutífera, já que os alunos estavam genuinamente interessados em compartilhar/receber contribuições sobre o desempenho dos colegas/desempenho próprio. Alguns alunos quiseram maior flexibilidade de tempo para a realização de tarefas e se opuseram à ideia de penalizar estudantes que entregassem os trabalhos mais tarde.
Desafios e Limitações	Os estudantes tinham níveis distintos de proficiência na produção oral em língua inglesa, devido a bagagens sociais e educacionais distintas. Alguns alunos consideram que os critérios de pontuação eram muito rígidos.
Recomendações e Futuras Pesquisas	Algumas das sugestões dos alunos podem ser implementadas futuramente.

Fonte: elaboração própria.

As análises dos alunos podem aumentar a motivação ao envolvê-los em experiências de sala de aula interessantes e relevantes. Embora esse estudo seja limitado a um contexto específico, com participantes de um país no Oriente Médio, o enfoque e a abordagem têm implicações mais amplas para profissionais de outras regiões do mundo. De fato, os resultados parecem estar alinhados com estudos sobre aquisição de segunda língua que tratam da criação de experiências e oportunidades de aprendizagem. Com base nisso, conclui-se que estimular a participação dos alunos pode levar a resultados positivos na aprendizagem de línguas, e

professores que lecionam cursos de comunicação oral devem procurar maneiras práticas de incorporar essa abordagem em suas aulas.

Os professores têm a oportunidade de envolver os alunos no processo de aprendizagem e obter dados essenciais para revisar e avaliar o curso, com o objetivo de aprimorar e aumentar sua eficácia. Qualquer curso de comunicação oral deve não apenas atender às necessidades dos alunos, mas também ser centrado no aluno. Nesse contexto, as análises dos alunos podem ser extremamente valiosas ao oferecer uma contribuição mais abrangente para o conteúdo, a estrutura e a execução do curso. Além disso, essas análises podem criar oportunidades para experiências de sala de aula envolventes e significativas.

Primeiramente, combinar a análise de necessidades, a resposta dos colegas e o *feedback* dos alunos pode ser uma estratégia eficaz para se obter uma visão mais ampla sobre o conteúdo, o *design* e a execução de um programa, já que fornece dados cruciais para revisar e avaliar um programa existente e aumentar sua eficácia.

Envolver os alunos no processo pode promover a aprendizagem reflexiva. Ao analisar suas necessidades comunicativas, expressar suas preferências de aprendizagem e fornecer *feedback* aos colegas, os alunos tornam-se mais conscientes do que precisam como participantes do curso e desenvolvem habilidades para refletir sobre seu processo de aprendizagem.

5.2.4 Preliminary Evaluation of the Impact of a Writing Assessment System on Teaching and Learning – Revista Profile – 2008

O Quadro 6 exibe partes do conteúdo do artigo de Ana Patricia Muñoz e Martha E. Álvarez (2008). O artigo, com base na crença de que avaliações válidas e confiáveis podem motivar mudanças positivas no ensino e na aprendizagem, descreve a implementação de um sistema de avaliação de escrita no programa de inglês de um centro de idiomas na Colômbia. Para avaliar o impacto do sistema, as autoras analisaram as melhorias em 27 scripts ao longo do tempo. Foi também analisada a qualidade das práticas de ensino e de avaliação de 35 professores. Melhorias na escrita foram determinadas por meio da complexidade sintática usando o número médio de palavras por texto em unidades T como uma medida. Práticas de avaliação e ensino foram avaliadas por meio da análise de portfólios de escrita. Os resultados indicam algumas melhorias na escrita e a necessidade de mais treinamento e orientação para permitir que os professores utilizem o sistema de avaliação de forma mais adequada.

Quadro 6 – Muñoz e Álvarez (2008)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Ana Patricia Muñoz e Martha E. Álvarez, EAFIT University, Colômbia (2008).
Tipo de Avaliação	O estudo utiliza um sistema de avaliação de escrita (Writing Assessment System – WAS) que inclui pré-testes e pós-testes aplicados a tarefas de escrita específicas.
População Avaliada	O público-alvo do estudo são alunos adultos de um programa universitário de inglês, especificamente aqueles que estão em cursos intensivos de diferentes níveis de proficiência (pré-intermediário, intermediário e intermediário-avançado).
Competências Avaliadas	Habilidades de escrita, que vão desde a capacidade de preencher formulários simples até a redação de ensaios e relatórios mais complexos, dependendo do nível de proficiência dos alunos.
Métodos e Instrumentos	Os testes do estudo consistem em tarefas de escrita abertas, nas quais os alunos têm um tempo limitado para completar as atividades. Os tópicos dos exercícios incluem a redação de textos persuasivos e descritivos, com foco em diferentes gêneros e propósitos de escrita.
Resultados e Conclusões	Os resultados preliminares indicam que houve melhorias na escrita dos alunos, sugerindo que a implementação do sistema de avaliação teve um impacto positivo no aprendizado. O estudo conclui que avaliações significativas podem motivar mudanças positivas na instrução e no aprendizado da escrita. O estudo demonstra melhorias na complexidade sintática das produções escritas dos alunos. No entanto, não reflete necessariamente o progresso dos alunos na escrita, pois a produção de frases mais complexas não implica em menos erros gramaticais ou maior coerência. Para uma avaliação mais completa da melhoria na linguagem, é necessário considerar outros fatores, como coerência, coesão e revisão do texto. A relação entre o <i>feedback</i> dos professores e as revisões dos alunos também é crucial para entender como os alunos respondem às orientações e atingem os padrões de escrita. Além disso, a pesquisa identificou dificuldades na implementação de um novo sistema de avaliação, sugerindo que mudanças nas crenças e práticas dos professores podem ser necessárias para que o sistema funcione adequadamente. Finalmente, espera-se que, após mais treinamento, seja possível avaliar o impacto da instrução dos professores no progresso dos alunos na escrita, permitindo a implementação de melhorias necessárias.
Desafios e Limitações	Os principais desafios do estudo incluem a flutuação significativa da população estudantil, já que muitos alunos interrompem suas aulas em diferentes períodos do ano. Isso dificultou a continuidade do acompanhamento dos mesmos alunos a longo prazo.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas devem considerar a formação contínua dos professores em relação à aplicação de rubricas de avaliação e à consistência na avaliação, além de explorar mais a relação entre a avaliação e o currículo. Estudos futuros devem investigar como envolver os professores em processos de mudança educacional.

Fonte: elaboração própria.

5.2.5 Exploring teachers practices – Revista Profile – 2009

O Quadro 7 exibe partes do conteúdo do artigo de Jorge Hugo Muñoz Marín (2009). O artigo relata os resultados de um estudo exploratório que buscou identificar as práticas de avaliação de professores de inglês no programa de compreensão leitora da Universidade de Antioquia.

Quadro 7 – Muñoz Marín (2009)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Jorge Hugo Muñoz Marín, Universidad de Antioquia, Colômbia (2009).
Tipo de Avaliação	O estudo destaca principalmente o uso de métodos tradicionais de avaliação, particularmente testes e provas, como os principais instrumentos para avaliar a compreensão de leitura.
População Avaliada	O público-alvo avaliado no estudo consiste em alunos matriculados no programa de compreensão de leitura em língua estrangeira da Universidade de Antioquia. O programa conta com cerca de 5.000 estudantes de vários departamentos acadêmicos, e as práticas de avaliação dos professores de inglês envolvidos nesse programa são o foco da pesquisa.
Competências Avaliadas	A principal competência avaliada no estudo é a compreensão de leitura. O foco está em avaliar a capacidade dos alunos de entender e interpretar informações em textos. Isso inclui vários aspectos da leitura, como fluência, conhecimento de vocabulário, inferências, resumo e habilidades de leitura crítica. As práticas de avaliação visam a medir o quão bem os alunos conseguem interagir com os textos e demonstrar sua compreensão por meio de diferentes tarefas e atividades.
Métodos e Instrumentos	As provas e os testes utilizados no estudo consistem principalmente em perguntas de múltipla escolha e outros formatos, como questões de resposta curta ou aberta. Essas avaliações são projetadas para verificar a compreensão dos alunos sobre os materiais de leitura abordados no curso. Os principais tópicos dos exercícios nesses testes geralmente incluem a compreensão de textos específicos, o conhecimento de vocabulário, ideias principais, detalhes e habilidades de inferência.
Resultados e Conclusões	Os professores empregam uma variedade de práticas de avaliação, mas há uma predominância no uso de métodos quantitativos, como testes e provas, para avaliar a compreensão de leitura. Foi observada uma falta de familiaridade dos professores com métodos alternativos de avaliação, como autoavaliação, avaliação entre pares, diários e portfólios. Isso indica uma lacuna no uso de abordagens mais formativas e centradas nos alunos. Os professores demonstraram grande preocupação em verificar o alcance dos objetivos de aprendizagem pelos alunos. Eles acreditam que as avaliações tradicionais fornecem informações mais precisas e confiáveis sobre as habilidades de leitura dos alunos. Alguns professores relataram que suas práticas de avaliação envolvem “ensinar para o teste”, concentrando-se na preparação dos alunos para provas e questionários específicos, o que pode limitar a experiência avaliativa.
Desafios e Limitações	A pesquisa foi realizada com uma amostra limitada de 15 professores de uma única universidade, o que pode não ser representativo de práticas mais amplas em outras instituições ou contextos. Tanto alunos quanto professores exibiram uma falta de familiaridade com práticas de avaliação qualitativa. Essa lacuna pode ter influenciado os resultados.

	O uso predominante de métodos de avaliação quantitativos pode limitar a compreensão das habilidades de leitura dos alunos. A interpretação dos resultados de avaliação pode ser subjetiva, dependendo dos critérios utilizados por diferentes professores.
Recomendações e Futuras Pesquisas	Pesquisas futuras devem considerar uma amostra maior e mais diversificada de professores de várias instituições para melhorar a generalização dos achados. Isso poderia fornecer uma compreensão mais ampla das práticas de avaliação em diferentes contextos. É recomendada a implementação de programas de formação continuada focalizados em métodos alternativos de avaliação. Isso ajudaria os professores a se familiarizarem com práticas de avaliação qualitativa. Pesquisas futuras poderiam incluir as perspectivas dos alunos sobre as práticas de avaliação, particularmente suas experiências com avaliações qualitativas. Compreender as visões dos alunos poderia informar os professores sobre como alinhar melhor os métodos de avaliação às necessidades e preferências dos alunos.

Fonte: elaboração própria.

Os professores também incorporam perguntas relacionadas aos temas e conteúdos das leituras indicadas em sala de aula, garantindo que as avaliações estejam alinhadas com os materiais utilizados para desenvolver habilidades de leitura. O estudo observa que os professores relataram utilizar testes de múltipla escolha, questionários e exames para avaliar as habilidades de leitura dos alunos e que há uma falta de familiaridade dos professores com métodos alternativos de avaliação, como autoavaliação, avaliação entre pares, diários ou portfólios. É de suma importância que os professores reflitam sobre suas práticas de avaliação e considerem a integração de métodos alternativos de avaliação. Isso poderia melhorar o processo de avaliação e apoiar melhor a aprendizagem dos alunos na compreensão de leitura. Os achados indicam a necessidade de formação contínua e desenvolvimento profissional para que os professores ampliem seu repertório de avaliação.

5.2.6 Analisis descriptivo de narraciones – Revista Literatura Y Linguística – 2009

O Quadro 8 exibe partes do conteúdo do artigo de Ricardo Benítez Figari (2009). O texto explora a avaliação de ensaios escritos por alunos do ensino básico.

Quadro 8 – Benítez Figari (2009)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Ricardo Benítez Figari, Pontificia Universidad Católica de Valparaíso, Chile (2009).
Tipo de Avaliação	O estudo utiliza uma avaliação baseada em uma escala holística focalizada, que permite classificar os ensaios escritos dos alunos de acordo com critérios

	específicos. Além disso, a avaliação foi realizada por meio de pares avaliadores que analisaram as produções escritas com uma pauta de correção.
População Avaliada	O público-alvo do estudo são crianças de terceiro ano básico, com um total de 165 estudantes de escolas municipais e subvencionadas da Região de Valparaíso, no Chile.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a produção escrita, especificamente a capacidade de narrar histórias de forma clara e organizada, além de aspectos metalinguísticos e cognitivos relacionados à compreensão oral da linguagem não literal.
Métodos e Instrumentos	Os testes consistem na produção de textos narrativos, nos quais os alunos escrevem redações baseadas em tópicos sugeridos. As instruções foram dadas de forma escrita, e os alunos puderam se inspirar em imagens apresentadas em uma tela. Não se trata de múltipla escolha ou questões abertas, mas sim de uma tarefa de escrita livre.
Resultados e Conclusões	Os resultados preliminares indicam que os alunos não alcançaram os níveis mais altos de pontuação na escala de avaliação, sugerindo que: os objetivos do programa de estudos podem ser insuficientes; os professores não exigem um maior refinamento nas produções escritas dos alunos; e as habilidades de escrita narrativa dos estudantes estão aquém das expectativas. A pesquisa conclui que há uma necessidade de melhorar a formação dos alunos em aspectos narrativos e descritivos.
Desafios e Limitações	Os desafios incluem a dificuldade de obter um consenso entre os avaliadores, a falta de tempo para a correção das redações e a carga horária necessária para avaliar textos de alunos. Além disso, a escala de avaliação pode ter sido considerada muito exigente para o nível escolar dos participantes.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de revisar e possivelmente ajustar os programas de estudo para incluir mais aspectos que ajudem no desenvolvimento das habilidades de escrita. Além disso, recomenda-se a utilização de escalas de avaliação que sejam mais adequadas ao nível dos alunos e que promovam um maior consenso entre os avaliadores.

Fonte: elaboração própria.

Através de uma escala holística, o estudo classificou os ensaios escritos dos alunos através de critérios específicos, avaliação essa realizada com o auxílio de pares avaliadores responsáveis pela avaliação das produções escritas das crianças do terceiro ano básico. Ao todo, 165 estudantes chilenos compuseram a pesquisa. A produção escrita foi avaliada, especificamente, quanto à capacidade de narrar histórias de forma clara e organizada, além de aspectos metalinguísticos e cognitivos relacionados à compreensão oral da linguagem não literal. Conclui-se que há uma necessidade de se melhorarem as habilidades dos alunos nos aspectos narrativos e descritivos.

5.2.7 Language testing in Colombia: a call for more teacher education and teacher training in language assessment – Revista Profile – 2009

O Quadro 9 exibe partes do conteúdo do texto de Alexis A. López Mendoza e Ricardo Bernal Arandia (2009). O texto apresenta os resultados de um estudo que examina as percepções dos professores de línguas sobre a avaliação e como eles a utilizam em sala de aula. Os resultados sugerem que há uma diferença significativa nas percepções dos professores dependendo do seu nível de treinamento em avaliação de linguagem. Por isso, o artigo ressalta a importância de se capacitarem adequadamente em avaliação os futuros professores de línguas estrangeiras na Colômbia.

Quadro 9 – López Mendoza e Bernal Arandia (2009)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Alexis A. López Mendoza, Universidad de los Andes, Colômbia, e Ricardo Bernal Arandia, Universidad Piloto de Colombia, Colômbia (2009).
Tipo de Avaliação	A pesquisa fez um apanhado de opiniões de professores de língua inglesa sobre avaliação e suas experiências prévias com disciplinas/formação em avaliação, assim como suas práticas avaliativas.
População Avaliada	Professores colombianos de língua inglesa. 82 professores de língua inglesa, que ministram aulas em diversos contextos educacionais, com experiência entre 3 e 17 anos de sala de aula.
Competências Avaliadas	Letramento em avaliação (assim como o perfil dos professores, opiniões e experiências prévias com avaliação).
Métodos e Instrumentos	Não houve menção.
Resultados e Conclusões	A visão sobre avaliação de cada professor (se é mais positiva ou negativa/simplista), depende de formação e letramento em avaliação do profissional. Os professores que já tiveram treinamento na área de avaliação, tinham uma visão mais positiva sobre práticas avaliativas (com mais foco no aluno e no processo). Aqueles profissionais que não tiveram formação demonstraram uma visão mais simplista da avaliação e, por vezes, mais numérica/burocrática. Na Colômbia, o interesse por publicar estudos sobre avaliação é baixo. A falta de cursos de capacitação em avaliação na Colômbia é um dos motivos pelos quais os professores não conseguem monitorar as consequências do uso da avaliação. Das universidades analisadas no estudo (que ofereciam programas para formação de professores), poucas ofereciam treinamento em avaliação. Dentre as universidades analisadas em seus programas de mestrado, por outro lado, todas ofereciam ao menos uma disciplina sobre avaliação. Por outro lado, na Colômbia, o número de professores que de fato concluem um programa de mestrado é reduzido. Do total, 71 professores avaliam ao final do curso ou da unidade; apenas 11 afirmaram avaliar de forma processual; 64 afirmaram usar testes tradicionais (com exercícios de múltipla escolha/preencher lacunas); e 18 alegaram utilizar instrumentos avaliativos alternativos (redações, apresentações, entrevistas etc.). Quanto às notas e aos resultados, 59 professores alegaram utilizar números ou letras, e os outros 23 utilizavam descritores qualitativos.
Desafios e Limitações	Em algumas salas de aula, a avaliação não é processual e tende a ser mais somativa.

	Programas de formação de professores na Colômbia não costumam oferecer treinamento em avaliação. Logo, existem muitas práticas avaliativas inadequadas e antiéticas. Apenas 32 dos 82 professores já haviam recebido treinamento em avaliação.
Recomendações e Futuras Pesquisas	Conscientizar professores colombianos de língua inglesa sobre a importância do letramento em avaliação. Conduzir estudos que analisem os propósitos reais dos testes e depois comparar com o que, de fato, foi praticado, além de examinar as consequências das práticas avaliativas nas vidas dos estudantes e suas famílias. É imperativo que todo professor faça, ao menos, um curso/disciplina sobre avaliação no ensino de línguas e que continue se aperfeiçoando através de <i>workshops</i> e outros meios de capacitação. Os professores precisam conhecer diferentes tipos de avaliações de línguas e a informação que cada uma proporciona, além de garantir que os procedimentos de avaliação sejam válidos e confiáveis. Validade refere-se a procedimentos que oferecem informações precisas sobre o que está sendo medido. Confiabilidade significa que os procedimentos de avaliação devem gerar pontuações consistentes, independentemente da situação ou contexto em que são aplicados. Observou-se que há uma preferência por métodos tradicionais de avaliação em vez de abordagens alternativas. Além disso, grande parte do <i>feedback</i> é dado na forma de notas e, geralmente, é fornecido ao final do processo. Logo, dados desse estudo sugerem que a avaliação em sala de aula no ensino de inglês na Colômbia tende a ser mais somativa do que focalizada na formação contínua.

Fonte: elaboração própria.

Nesse estudo, as visões dos professores sobre a avaliação de línguas foram exploradas. Há uma possível relação entre a formação em avaliação de línguas e as percepções que os professores têm sobre o tema. Educação e treinamento adequados podem alterar a forma como os professores percebem a avaliação de línguas. Se tiverem uma visão positiva sobre a avaliação, os professores poderão escolher ou criar métodos de avaliação apropriados para seu contexto e seus alunos, possibilitando que as avaliações ofereçam informações valiosas.

Também foi abordado como os professores utilizam a avaliação de línguas em suas aulas. Contudo, a literatura também aponta uma carência de medidas de avaliação adequadas, válidas e confiáveis para aprendizes de inglês. Como os alunos são avaliados para diversos fins, é crucial rever se os instrumentos e procedimentos de avaliação usados com frequência são válidos, apropriados e se envolvem todas as partes interessadas no processo de avaliação. É essencial que os criadores de testes interajam com diferentes grupos envolvidos, para compreender melhor a cultura e o contexto da avaliação em que o teste é utilizado. Além disso, devemos investigar como essas práticas de avaliação impactam a vida dos alunos e de suas famílias. Há uma necessidade crescente de mais pesquisas sobre o efeito dos testes na aprendizagem de línguas e nos próprios aprendizes.

Acredito que os resultados de estudos como esse podem incentivar administradores, professores em formação e em exercício, além da comunidade educacional em geral, a atualizar seu desenvolvimento profissional e aprimorar suas práticas de avaliação, melhorando, assim, a qualidade do ensino de línguas e a motivação dos alunos para aprender. Atualmente, é importante reconhecer que a avaliação vai além de medir ou atribuir notas. É mais motivador e menos intimidante para os professores de línguas começarem a discutir a avaliação para a aprendizagem, em vez de apenas realizarem a avaliação da aprendizagem.

Por fim, tanto os professores quanto os futuros professores precisam de mais formação em avaliação de línguas. A responsabilidade de capacitar os professores de línguas para desenvolver, usar, pontuar e interpretar avaliações deve ser das instituições de ensino superior que oferecem programas de formação docente, das instituições que possuem programas de línguas e dos próprios professores. É fundamental que todos os futuros professores realizem pelo menos um curso sobre avaliação de línguas antes de iniciarem sua carreira, além de buscarem aprimorar suas habilidades através de formação contínua, conferências, *workshops* e outras atividades para estabelecer uma cultura de avaliação de línguas voltada para a melhoria na educação.

5.2.8 Designing proficiency tests to accredit previous knowledge in American and British literature in a bilingual education program – Revista Profile – 2010

O Quadro 10 apresenta partes do conteúdo do estudo de León Chica, D’Costa Martínez e Franco Jácome (2010). Os testes de literatura americana e britânica foram elaborados com base nas necessidades dos alunos, nos requisitos e no programa da universidade, na teoria sobre literatura e avaliação e em testes comerciais para creditar o conhecimento prévio. A proficiência e o bom domínio do inglês foram fatores essenciais para a aprovação nos testes. Os testes não foram divulgados na pesquisa para manter sua validade, uma vez que devem ser utilizados pela universidade para creditar o conhecimento prévio dos alunos.

Quadro 10 – León Chica, D’Costa Martínez e Franco Jácome (2010)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	César Julio León Chica, Catalina D’Costa Martínez e Gisela Franco Jácome, Universidad El Bosque, Colômbia (2010).
Tipo de Avaliação	Teste de proficiência em literatura americana e britânica.
População Avaliada	Alunos (10) de uma universidade privada da Colômbia. Os alunos cursam uma licenciatura em educação bilíngue, com ênfase no ensino de língua inglesa.

Competências Avaliadas	Proficiência em literatura americana e britânica (períodos históricos, origens, gêneros, obras excepcionais), além de proficiência em língua inglesa.
Métodos e Instrumentos	Questionários e entrevistas para entender as necessidades dos alunos sobre o teste que seria elaborado, além de um teste de proficiência com questões abertas, fechadas e um ensaio.
Resultados e Conclusões	Os estudantes gostariam de ter sua proficiência em literatura avaliada para que pudessem ganhar tempo e avançar em estudos novos. A voz dos alunos foi escutada em todos os seus aspectos, para que o teste de proficiência fosse o mais justo e completo possível. Eles opinaram em pontos como a quantidade de exercícios e o formato das questões. Os conhecimentos prévios dos alunos em literatura americana e britânica foram de extrema importância. Todos os estudantes obtiveram uma pontuação satisfatória para aprovação no teste. A pontuação para ser considerado proficiente não foi atingida pela maioria.
Desafios e Limitações	Os resultados para o teste de literatura americana não foram tão favoráveis quanto o teste de literatura britânica, não pelo formato do teste, mas possivelmente por falta de conhecimentos prévios dos alunos e dificuldades com a língua inglesa. O número de alunos envolvidos no teste foi significativamente pequeno e os pesquisadores tiveram desafios relacionados à disposição e à prontidão de alguns envolvidos para participação ativa e pontual na pesquisa. Existem poucos modelos de teste de proficiência em literatura para servir de exemplo para testes futuros. Focalizar perguntas como: de que modo as informações obtidas na pesquisa podem ser traduzidas em melhorias para os programas de literatura da universidade?; seria importante desenvolver um banco de questões para evitar usar o mesmo teste todo semestre?; existem outras formas de se avaliarem os conhecimentos prévios em literatura? A variação nos resultados dos testes entre os participantes que possuíam conhecimento prévio tanto em literatura quanto no domínio da segunda língua corrobora as teorias e os estudos discutidos no artigo sobre avaliação. Outra descoberta relevante foi que as habilidades avaliadas por meio da compreensão de leitura e da redação de ensaios também devem ser levadas em conta na avaliação tanto do conhecimento prévio quanto da competência em literatura. Foram elaborados e validados dois testes de proficiência em literatura, embora, segundo a literatura TESOL, haja pouca pesquisa colaborativa entre professores, e a maior parte dos estudos sobre testes de proficiência se concentra na avaliação linguística. Os testes desenvolvidos nesse estudo alcançaram uma abordagem mais comunicativa para a avaliação de literatura. Os testes avaliaram o conhecimento prévio (conteúdo teórico/literatura) através da seção de linha do tempo, das questões de múltipla escolha e das perguntas abertas, assim como as habilidades de compreensão de leitura e de redação de ensaios. Algumas limitações enfrentadas durante esse projeto de pesquisa incluíram o fato de que dez participantes representavam uma amostra pequena em relação ao total de alunos de educação bilíngue (111 na época da pesquisa). Quando as pessoas se voluntariam para participar de estudos desta esfera, frequentemente não compreendem a importância de seu compromisso para o sucesso do projeto. Cumprir suas responsabilidades não era a prioridade dos participantes, ao contrário dos pesquisadores, que dependiam da disposição e do tempo dos participantes para responder e testar os instrumentos, e tiveram que ajustar seu cronograma ao dos participantes. Além disso, algumas condições pessoais dos entrevistados causaram atrasos na pontuação dos testes e na classificação e análise dos dados.
Recomendações e Futuras Pesquisas	Após a conclusão da pesquisa, surgiram as seguintes questões para futuros estudos. Como as informações obtidas nas entrevistas com os participantes

	podem ser usadas para modificar ou aprimorar os programas de literatura atuais da universidade? Como os dois testes de proficiência resultantes dessa pesquisa poderão ser realmente utilizados pela universidade? Devem ser revisados ou modificados? Se sim, por quem? Deve-se criar um banco de questões para evitar a repetição exata dos testes a cada semestre? Como as habilidades e competências dos alunos podem ser consideradas para melhorar os programas de literatura da universidade? Existe outra forma de avaliar o conhecimento prévio?
--	---

Fonte: elaboração própria.

No geral, a aceitação dos alunos em relação ao teste foi boa, ainda que tenham sugerido modificações. Os testes de proficiência criados foram baseados nos tipos de questões que os alunos sugeriram, além de considerar os programas da universidade, outros testes existentes no mercado e diversas antologias literárias. Todos os alunos aprovados tinham um alto nível de proficiência em língua inglesa. Devido à ausência de testes-modelo de literatura que cobrissem a maioria dos aspectos em que os alunos desejavam ser avaliados, foi necessário dedicar mais tempo à sua elaboração. Devido à sua natureza, os testes desenvolvidos para esse projeto precisaram ser corrigidos pelos professores e não puderam ser pontuados por máquinas.

5.2.9 *La evaluación de un modelo metodológico mixto para mejorar las habilidades productivas y receptivas en español como segunda lengua: un estudio cuasi-experimental – Revista Lengua y Habla – 2011*

O Quadro 11 apresenta partes do conteúdo do estudo de Livingstone (2011).

Quadro 11 – Livingstone (2011)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Kerwin A. Livingstone, University of Guyana, Guiana (2011)
Tipo de Avaliação	O estudo utiliza uma avaliação composta de <i>pre-test</i> e <i>post-test</i> , que integra habilidades receptivas (compreensão auditiva e leitura) e produtivas (produção oral e escrita).
População Avaliada	O público-alvo do estudo são universitários que estão aprendendo espanhol como segunda língua.
Competências Avaliadas	As principais competências avaliadas incluem compreensão oral, compreensão escrita, produção oral e produção escrita.
Métodos e Instrumentos	Os testes consistem em uma prova escrita e uma entrevista. O <i>pre-test</i> e o <i>post-test</i> têm a mesma estrutura, mas utilizam textos diferentes para evitar influenciar os resultados. Os tópicos abordados estão relacionados aos objetivos, conteúdos e habilidades linguísticas do módulo de ensino.
Resultados e Conclusões	Os resultados indicam que o modelo metodológico misto, sustentado pelo Enfoque por Tarefas e Cooperativo, é eficaz para aumentar o aprendizado de espanhol como L2.
Desafios e	O estudo não menciona explicitamente desafios e limitações.

Limitações	
Recomendações e Futuras Pesquisas	O estudo sugere que, para o <i>design</i> de módulos de ensino de línguas, se integrem e implementem modelos metodológicos mistos, pois são considerados os mais eficazes para o aprendizado de línguas estrangeiras e segundas línguas.

Fonte: elaboração própria.

Em linhas gerais, a pesquisa conclui que a metodologia mista é propícia para o aprendizado de línguas.

5.2.10 Teachers' Beliefs about Assessment in an EFL Context in Colombia – Revista Profile – 2012

O Quadro 12 exibe partes do conteúdo do artigo de Muñoz, Palacio e Escobar (2012). O texto explora avaliações tradicionais e modernas.

Quadro 12 – Muñoz, Palacio e Escobar (2012)

Categoria	Descrição
Perfil (Ano/Autor)	Ana Patricia Muñoz, Marcela Palacio e Liliana Escobar, Universidad EAFIT, Colômbia (2012)
Tipo de Avaliação	Os testes mencionados no estudo incluem tanto avaliações tradicionais (como testes de múltipla escolha) quanto métodos mais dinâmicos e interativos, como <i>role-plays</i> e discussões.
População Avaliada	O público-alvo do estudo são os alunos do programa de inglês para adultos de um centro de idiomas de uma universidade privada na Colômbia.
Competências Avaliadas	Produção oral, produção escrita, a avaliação de habilidades comunicativas e a compreensão de conteúdo.
Métodos e Instrumentos	O estudo menciona a utilização de uma variedade de métodos de avaliação, incluindo testes centrados na comunicação, testes baseados em conteúdo, além de métodos alternativos, como discussões, dramatizações, portfólios e técnicas de autoavaliação e avaliação por pares.
Resultados e Conclusões	Os resultados indicam uma contradição entre o que os professores afirmam fazer e suas crenças sobre avaliação. A pesquisa sugere que os professores precisam de mais oportunidades para reflexão e autoavaliação, além de orientação sobre práticas de avaliação formativa. Há uma discrepância significativa entre o que os professores afirmam fazer em termos de avaliação e suas crenças subjacentes. Muitos professores continuam a seguir práticas tradicionais que enfatizam notas em vez de critérios descritivos e específicos. As crenças dos professores sobre avaliação são influenciadas por suas experiências pessoais, formação e contexto social, o que pode dificultar a adoção de novas práticas de avaliação.
Desafios e Limitações	Um dos desafios identificados foi a resistência dos professores a mudar suas práticas de avaliação, que muitas vezes ainda se baseavam em métodos tradicionais. Além disso, a falta de experiência e conhecimento sobre métodos alternativos de avaliação foi um obstáculo significativo.
Recomendações e Futuras Pesquisas	O estudo recomenda que os professores sejam mais preparados por meio de programas de formação que os ajudem a implementar uma gama mais ampla de

	instrumentos de avaliação. Também sugere a necessidade de mais pesquisa sobre a relação entre as crenças dos professores e suas práticas de avaliação.
--	--

Fonte: elaboração própria.

Os testes abordados no estudo incluem tanto avaliações tradicionais quanto métodos mais dinâmicos e interativos, como *role-plays* e discussões, com as principais habilidades avaliadas sendo a produção oral, a produção escrita, a avaliação de habilidades comunicativas e a compreensão de conteúdo. Muitos professores continuam a seguir práticas tradicionais que enfatizam notas em vez de critérios descritivos e específicos, indicando uma contradição entre o que os professores afirmam fazer e suas crenças sobre avaliação. A pesquisa sugere que os professores precisam de mais oportunidades para reflexão e autoavaliação, além de orientação sobre práticas de avaliação formativa. A falta de experiência e conhecimento sobre métodos alternativos de avaliação representou um obstáculo, considerando-se que as crenças dos professores sobre avaliação são influenciadas por suas experiências pessoais.

5.2.11 The impact of conferencing assessment on EFL students' grammar learning – Revista Profile – 2012

O Quadro 13 traz apontamentos do estudo de Baleghizadeh e Zarghami (2012), que teve como objetivo examinar o efeito da avaliação por conferência (encontros de discussão) na aprendizagem de gramática dos alunos iranianos de inglês como língua estrangeira (EFL) e suas percepções em relação ao aprendizado formal de gramática. A análise geral resultante sugere que a avaliação por encontros exerce um impacto positivo na aprendizagem de gramática desses alunos e pode alterar suas atitudes em relação ao estudo da gramática. Utilizar métodos de avaliação voltados para o processo, como a avaliação por conferência, oferece amplas oportunidades para que professores e alunos interajam efetivamente. Dessa forma, os professores podem apoiar o aprendizado dos alunos, fornecendo *feedback* descritivo apropriado e auxiliando na identificação de suas dificuldades.

Com isso, alunos e professores podem atuar como parceiros na avaliação, estabelecendo metas claras e tarefas de avaliação específicas. Esse processo pode encorajar os alunos a assumirem a responsabilidade pelo próprio sucesso e a se responsabilizar pelo próprio aprendizado. Isso tende a motivá-los a alcançar um aprendizado mais eficaz e melhores resultados. É importante ressaltar que a avaliação não deve ser vista como algo separado da instrução. Para ser mais autêntica, a avaliação deve basear-se nos comportamentos que os

alunos demonstram durante a avaliação formativa e contínua, e os alunos devem estar cientes dos resultados esperados da instrução e da avaliação, dos processos envolvidos e dos critérios de avaliação.

Quadro 13 – Baleghizadeh e Zarghami (2012)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Sasan Baleghizadeh, Shahid Beheshti University, G.C., Irã, e Zahra Zarghami, Allameh Tabataba'i University, Irã (2012).
Tipo de Avaliação	Teste de gramática com 50 itens (25 de múltipla escolha e 25 de identificação de erros).
População Avaliada	42 estudantes universitários iranianos (de áreas de estudo distintas) com nível intermediário de língua inglesa.
Competências Avaliadas	Competência gramatical em língua inglesa.
Métodos e Instrumentos	Teste de conhecimentos gramaticais em língua inglesa, assim como encontros de discussão sobre dificuldades e pontos fortes relacionados à gramática.
Resultados e Conclusões	Fornecer <i>feedback</i> envolve comentários construtivos, ser específico ao dar informações, fornecer comentários imediatamente, usar linguagem acolhedora, considerar as necessidades e vontades dos alunos, certificar-se de que os alunos entenderam o <i>feedback</i>, focalizar melhorias e um aspecto por vez. Os alunos revelaram atitudes positivas em relação ao estudo formal de gramática após o experimento de encontros de discussão gramatical. O método alternativo em questão foi mais benéfico que outros formatos somativos de avaliação, possivelmente pelo fato de os alunos se responsabilizarem mais por seus estudos. A autoavaliação e reflexão foram incentivadas, e as atitudes dos alunos mudaram significativamente em relação ao estudo formal de gramática. No passado, formas tradicionais de se ensinar gramática, sobretudo através da gramática-tradução, fizeram com que alunos criassem uma resistência em relação ao estudo dos aspectos gramaticais da língua inglesa. Nos encontros de discussão, os alunos são incentivados a falar sobre suas dificuldades e pontos fortes em relação aos conhecimentos gramaticais; logo, eles não estão pensando apenas no que aprendem, mas em como aprendem.
Desafios e Limitações	Formas tradicionais antigas de se ensinar gramática, sobretudo através da gramática-tradução, fizeram com que alunos criassem uma resistência em relação ao estudo dos aspectos gramaticais da língua inglesa.
Recomendações e Futuras Pesquisas	Estudantes e professores podem trabalhar como aliados nos processos avaliativos, incorporando objetivos claros de aprendizagem e atividades avaliativas específicas.

Fonte: elaboração própria.

5.2.12 A comparison of oral evaluation ratings by native English speaker teachers and non-native English speaker teachers – *Revista Literatura y Lingüística* – 2012

O Quadro 14 mostra partes do conteúdo do artigo de Baitman e Véliz Campos (2012), que investigam um tipo específico de avaliação: a certificadora.

Quadro 14 – Baitman e Véliz Campos (2012)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Brittany Baitman e Mauricio Véliz Campos, Universidad Católica Silva Henríquez, Santiago, Chile (2012).
Tipo de Avaliação	O estudo utiliza uma avaliação oral baseada no TOEFL iBT/Next Generation Independent Speaking Rubric, que envolve a gravação digital das respostas dos alunos e a avaliação holística por avaliadores certificados.
População Avaliada	O público-alvo do estudo são alunos de uma universidade privada, que possuem níveis de proficiência em inglês variando do nível um (mais baixo) ao nível sete (mais alto).
Competências Avaliadas	As habilidades de fala, que incluem fluência, pronúncia, precisão gramatical e uso de vocabulário.
Métodos e Instrumentos	Os testes consistem em tarefas orais em que os alunos devem responder a perguntas após ouvir uma breve apresentação. As respostas são gravadas e avaliadas com base em critérios específicos da rubrica de avaliação.
Resultados e Conclusões	O estudo concluiu que existem diferenças nas práticas de avaliação entre professores nativos e não nativos: os primeiros tendem a ser mais lenientes e focalizam mais a fluência, enquanto os segundos enfatizam a precisão gramatical e o vocabulário. Isso se reflete nas pontuações que cada grupo atribui aos alunos. As conclusões do estudo podem ter implicações significativas para a formação de professores e para o desenvolvimento de rubricas de avaliação que considerem as diferentes abordagens de nativos e não nativos.
Desafios e Limitações	Um dos desafios mencionados é a falta de pesquisa atual sobre a avaliação oral sob a perspectiva de falantes nativos de língua inglesa e falantes não-nativos, o que limita a compreensão das diferenças nas avaliações.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de mais pesquisas sobre a avaliação oral, especialmente em relação a como diferentes critérios de avaliação podem impactar os resultados dos alunos e a eficácia do ensino.

Fonte: elaboração própria.

O estudo analisa uma avaliação oral baseada no TOEFL iBT. Há uma gravação digital das respostas dos alunos de uma universidade privada, que são avaliados por profissionais certificados. O nível de proficiência dos alunos envolvidos é variável. As habilidades de fala avaliadas incluem fluência, pronúncia e precisão gramatical. Concluiu-se, por fim, que existem diferenças nas práticas de avaliação entre professores nativos e não nativos. Acredita-se que os nativos são mais lenientes e focalizam mais fluência, ao passo que os demais focalizam vocabulário e aspectos relacionados à precisão gramatical, o que pode afetar a pontuação atribuída aos alunos. A escassez de pesquisas atuais sobre avaliação oral (envolvendo professores nativos e não nativos) representa um desafio para as pesquisas na área.

5.2.13 The feasibility of assessing teenagers' oral English language performance with a rubric
 – Revista Profile – 2014

O Quadro 15 destaca aspectos do estudo de Pineda (2014). Embora alguns professores tenham relatado clareza no construto, outros enfrentaram dificuldades em definir um construto claro, evidenciando a necessidade real de capacitação em questões de avaliação. Essa capacitação deve ocorrer na fase inicial de uma inovação curricular como essa. Posteriormente, a formação pode incluir o acompanhamento dos professores durante a aplicação da rubrica, a observação das atividades de avaliação e o fornecimento de *feedback*. As atitudes dos professores e a disposição para aprender um novo procedimento de avaliação, suas crenças anteriores sobre avaliação e suas condições de trabalho são aspectos fundamentais a serem considerados por administradores que desejam implementar uma inovação curricular semelhante em qualquer programa.

Monitorar como a avaliação com rubrica facilita a avaliação dos alunos é crucial para professores que lidam com grandes quantidades de trabalhos. Com a utilização de uma rubrica, não apenas a praticidade, mas também a relevância para a aprendizagem dos alunos são favorecidas, o que é necessário nas instituições e nos programas educacionais atuais. Apesar do longo tempo necessário para se definirem as tarefas de avaliação e se criarem as rubricas inicialmente, é altamente recomendado que os professores de inglês experimentem essa abordagem e que os administradores escolares ofereçam as condições (tempo e espaço) para que os professores se reúnam e desenvolvam seus próprios critérios de avaliação. O tempo investido no início será recuperado posteriormente, quando os professores começarem a economizar tempo na avaliação, sem mencionar o impacto positivo do *feedback* gerado pelas rubricas na aprendizagem e na motivação dos alunos.

Quadro 15 – Pineda (2014)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Diana Pineda, Universidad de Antioquia, Colômbia (2014)
Tipo de Avaliação	Utilização de rubricas para a avaliação do desempenho oral.
População Avaliada	Jovens estudantes de língua inglesa em uma universidade pública colombiana.
Competências Avaliadas	Desempenho oral.
Métodos e Instrumentos	Rubricas para a avaliação, além de entrevistas com professores e alunos.
Resultados e Conclusões	Quando muitos alunos parecem se importar apenas com as notas, a rubrica é uma alternativa viável para promover reflexão nas práticas avaliativas, além de ser um instrumento formativo para professores e alunos. Os professores sinalizaram que a avaliação era a área em que eles mais precisavam de treinamento e reflexão. Os mesmos profissionais também precisam de mais formação sobre rubricas.

	Foi sugerido que os professores compartilhassem as rubricas com os seus alunos, fazendo da prática avaliativa algo mais democrático e formativo. É esperado que os alunos se sintam mais empoderados e mais responsáveis pelos seus próprios percursos educacionais. As rubricas podem fazer da avaliação um momento menos ameaçador e mais reflexivo. As rubricas também podem ser usadas em momentos de autoavaliação e avaliação por pares. Os alunos ficam mais cientes, sabendo dos critérios de avaliação, do que se fossem avaliados apenas com números. Logo, é considerada uma prática mais justa e transparente.
Desafios e Limitações	No início, criar uma rubrica e se acostumar com o seu uso pode demandar muito tempo. Mas, após se familiarizarem, os professores podem se beneficiar consideravelmente do seu uso.
Recomendações e Futuras Pesquisas	Que administradores das escolas criem oportunidades para que os professores se reúnam e criem seus próprios critérios de avaliação, visando a futuramente colher os frutos da utilização de rubricas. Para os docentes, a experiência de utilizar uma rubrica para avaliação foi, de maneira geral, positiva. Apenas um dos professores no estudo relatou dificuldades na implementação. Os professores reconheceram algumas das qualidades esperadas da rubrica, como sua utilidade e praticidade, equidade e princípios democráticos em suas orientações. Além disso, os docentes perceberam que a rubrica serve como uma ferramenta para obter evidências do desempenho dos alunos, auxiliando na identificação de suas fraquezas e pontos fortes, e responsabilização por suas necessidades de aprendizado. Os professores também destacaram a importância do planejamento, da criatividade e da elaboração de tarefas desafiadoras. Se os professores tiverem clareza sobre o que precisam avaliar e comunicarem isso de forma explícita aos alunos, isso influenciará suas práticas de ensino e o aprendizado dos alunos. Além disso, a avaliação se torna mais justa e transparente.

Fonte: elaboração própria.

5.2.14 Experiencia académica y estrategias de comprensión lectora en estudiantes universitarios de primer año – Literatura e Lingüística – 2014

O Quadro 16 mostra partes do conteúdo do artigo de Neira Martínez, Reyes Reyes, Riffo Ocares (2014). O artigo utiliza provas de leitura e protocolos verbais para tratar das experiências avaliativas de estudantes universitários.

Quadro 16 – Neira Martínez, Reyes, Riffo Ocares (2014)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Angie Carolina Neira Martínez, Fernando Teddy Reyes Reyes e Bernardo Esteban Riffo Ocares, Universidad de Concepción y Universidad del Desarrollo, Concepción, Chile (2014).
Tipo de Avaliação	Foi utilizada uma prova de compreensão de leitura (Lectum 7) e um protocolo de pensamento em voz alta. A prova consistiu em perguntas de múltipla escolha, enquanto o protocolo exigiu que os estudantes verbalizassem seus pensamentos durante a leitura.

População Avaliada	O público-alvo do estudo são estudantes universitários de primeiro ano de Pedagogia em Español da Universidade de Concepción (40 estudantes).
Competências Avaliadas	A compreensão leitora, que abrange dimensões como a textual (nível de palavra, oração, microestrutura e macroestrutura), pragmática e crítica.
Métodos e Instrumentos	Os testes consistem em perguntas de múltipla escolha, com um total de 32 perguntas que abordam diferentes dimensões da compreensão leitora. Os textos utilizados incluem dois argumentativos, um expositivo e um narrativo.
Resultados e Conclusões	Os resultados revelaram um desempenho médio geral baixo na compreensão leitora, com diferenças significativas entre leitores hábeis e menos hábeis. Além disso, estudantes com estudos superiores prévios apresentaram um desempenho melhor em comparação com aqueles sem experiência. Os leitores menos hábeis mostraram insegurança em suas respostas e frequentemente precisaram retornar ao texto para responder às perguntas, indicando uma dependência maior do material escrito. Algumas perguntas, especialmente aquelas que exigiam interpretação de gráficos ou compreensão pragmática, foram particularmente desafiadoras para todos os grupos, mas mais para os leitores menos hábeis.
Desafios e Limitações	O estudo enfrentou desafios relacionados à amostra não probabilística e ao tamanho limitado da amostra, o que pode afetar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se investigarem mais a fundo as estratégias de leitura utilizadas pelos estudantes e como essas estratégias podem ser melhoradas. Além disso, recomenda-se ampliar a amostra e incluir diferentes contextos educacionais para obter uma visão mais abrangente.

Fonte: elaboração própria.

O estudo exploratório sobre a compreensão de leitura de universitários de primeiro ano de Pedagogia em Español utilizou uma prova de múltipla escolha (Lectum 7) e protocolos de pensamento em voz alta. Os resultados indicaram um desempenho médio geral baixo, com diferenças significativas entre leitores e um melhor desempenho por parte de estudantes com estudos superiores prévios. Leitores com mais dificuldades demonstraram insegurança e maior dependência do texto, e questões que exigiam interpretação de gráficos ou compreensão pragmática foram desafiadoras para todos. O estudo, que enfrentou limitações devido ao tamanho e ao tipo da amostra, sugere a necessidade de se investigarem estratégias de leitura e se ampliar a pesquisa para diferentes contextos educacionais.

5.2.15 Competencia léxica, comprensión lectora y rendimiento académico en estudiantes de enseñanza media – Literatura y Lingüística – 2014

O Quadro 17 exhibe partes do conteúdo do artigo de Riffo Ocares *et al.* (2014). A intenção do artigo é investigar testes padronizados.

Quadro 17 – Riffo Ocares *et al.* (2014)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Bernardo Riffo Ocares, Fernando Reyes Reyes, Abraham Novoa Lagos, Mónica Véliz de Vos e Ginette Castro Yáñez, Universidad de Concepción, Universidad del Desarrollo, Concepción, Chile e Universidad Católica de Temuco, Temuco, Chile (2014).
Tipo de Avaliação	Foi utilizada uma combinação de testes padronizados, incluindo a prova de compreensão leitora (Lectum) e a tarefa de decisão lexical. Esses testes foram aplicados de forma individual e em grupo, e basearam-se em itens de múltipla escolha e decisões rápidas sobre a existência de palavras.
População Avaliada	O público-alvo do estudo são estudantes do quarto ano do ensino médio, especificamente de uma escola particular subvencionada na comuna de Concepción. A amostra incluiu 32 estudantes, dos quais 17 homens, e 15, mulheres, todos com idades entre 17 e 18 anos.
Competências Avaliadas	As competências avaliadas no estudo incluem a compreensão leitora e a competência lexical. Busca-se determinar como essas competências se relacionam com o desempenho acadêmico dos estudantes.
Métodos e Instrumentos	Os testes consistem em perguntas de múltipla escolha (no caso da prova de compreensão leitora) e decisões rápidas sobre palavras e pseudopalavras (na tarefa de decisão lexical). Os principais tópicos dos exercícios incluem aspectos textuais e pragmáticos da leitura, bem como a identificação de palavras no contexto do vocabulário.
Resultados e Conclusões	Os resultados indicam que a média geral de notas dos estudantes e a tarefa de decisão lexical são preditores significativos da compreensão leitora. Conclui-se que existe uma correlação entre a competência lexical e a compreensão leitora, sugerindo que melhorar a competência lexical poderia beneficiar a compreensão leitora e o desempenho acadêmico.
Desafios e Limitações	Um dos desafios mencionados é a falta de pontuações normativas para a prova de compreensão leitora, o que dificultou a classificação dos estudantes como bons ou maus leitores. Além disso, o tamanho da amostra (32 estudantes) pode limitar a generalização dos resultados.
Recomendações e Futuras Pesquisas	Embora o estudo não mencione explicitamente recomendações, pode-se inferir que estudos futuros poderiam se beneficiar se incluíssem um maior número de participantes e estabelecessem pontuações normativas para as provas de compreensão leitora. Também seria útil explorar mais a fundo a relação entre a competência lexical e outras variáveis demográficas ou contextuais que possam influenciar o desempenho acadêmico.

Fonte: elaboração própria.

O estudo investigou a relação entre a compreensão leitora e a competência lexical no desempenho acadêmico de 32 estudantes do ensino médio de uma escola particular em Concepción, Chile, com idades entre 17 e 18 anos. A pesquisa utilizou uma combinação de testes padronizados aplicados individualmente e em grupo, incluindo uma prova de compreensão leitora (Lectum) com itens de múltipla escolha e uma tarefa de decisão lexical com decisões rápidas sobre a existência de palavras. Os resultados revelaram que a média geral de notas dos estudantes e a tarefa de decisão lexical estão conectadas com a compreensão

leitora, indicando uma correlação entre a competência lexical e a compreensão leitora. Conclui-se que o aprimoramento da competência lexical pode beneficiar a compreensão leitora e, consequentemente, o desempenho acadêmico. No entanto, o estudo enfrentou desafios como a falta de pontuações normativas para a prova de compreensão leitora e o tamanho limitado da amostra, o que pode restringir a generalização dos achados.

5.2.16 *Ansiedad Ante Los Exámenes, Procrastinación y Síntomas – Revista Argentina de Ciencias del Comportamiento – 2014*

O Quadro 18 exibe partes do conteúdo do artigo de Furlan, Ferrero e Gallart (2014). A intenção do artigo é investigar escalas de autoavaliação.

Quadro 18 – Furlan, Ferrero e Gallart (2014)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Luis Alberto Furlan, María José Ferrero e Gabriela Gallart, Universidad Nacional de Córdoba, Argentina (2014).
Tipo de Avaliação	O estudo utiliza uma avaliação por meio de questionários e escalas de autoavaliação, como a Escala Cognitiva de Ansiedade em Exames (S-CTAS) e a Tuckman Procrastination Scale (A-TPS).
População Avaliada	O público-alvo do estudo são estudantes da Universidade Nacional de Córdoba (UNC), com uma amostra de 219 estudantes de diversas áreas.
Competências Avaliadas	As principais questões avaliadas incluem a ansiedade em situações de exame, a procrastinação acadêmica e sintomas de saúde mental, como ansiedade e depressão.
Métodos e Instrumentos	Os testes consistem em questionários de autoavaliação, que incluem itens de múltipla escolha e escalas do tipo Likert. Os principais tópicos abordados nos exercícios incluem a frequência de sintomas de ansiedade, procrastinação e outras manifestações de saúde mental.
Resultados e Conclusões	Os resultados indicam que estudantes com altos níveis de ansiedade e procrastinação apresentam uma maior frequência de sintomas mentais. O estudo conclui que há uma correlação significativa entre a ansiedade em exames, a procrastinação acadêmica e a presença de sintomas de saúde mental.
Desafios e Limitações	Um dos desafios mencionados é a necessidade de se esclarecer a estrutura interna do questionário CESIM, o que sugere que os resultados devem ser interpretados de forma preliminar.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de mais pesquisas para aprofundar a compreensão das comorbidades entre ansiedade, procrastinação e sintomas mentais. Recomenda-se também a utilização de amostras mais representativas e a aplicação de métodos de avaliação mais diversificados.

Fonte: elaboração própria.

O estudo investigou a relação entre a ansiedade em exames, a procrastinação acadêmica e sintomas de saúde mental em 219 estudantes de diversas áreas da Universidade Nacional de

Córdoba (UNC), utilizando questionários e escalas de autoavaliação como a Escala Cognitiva de Ansiedade em Exames (S-CTAS) e a Tuckman Procrastination Scale (A-TPS). As principais questões avaliadas incluíram a frequência de sintomas de ansiedade, procrastinação e questões relacionadas à saúde mental, como ansiedade e depressão, através de itens de múltipla escolha e escalas Likert. Os resultados revelaram uma correlação significativa entre altos níveis de ansiedade em exames e procrastinação com uma maior frequência de sintomas mentais. Conclui-se que existe uma ligação entre a ansiedade em exames, a procrastinação acadêmica e a presença de sintomas relacionados à saúde mental. O estudo sugere a necessidade de mais pesquisas para aprofundar a compreensão de determinadas comorbidades, recomendando o uso de amostras mais representativas e métodos de avaliação mais diversos para futuros estudos.

5.2.17 Propuesta de un modelo para una prueba con fines específicos académicos en ele – Revista Literatura Y Lingüística – 2015

O Quadro 19 exhibe partes do conteúdo do artigo de Vine Jara e Ferreira Cabrera (2015). O artigo combina avaliações que agrupam múltiplas atividades.

Quadro 19 – Vine Jara e Ferreira Cabrera (2015)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Ana Vine Jara e Anita Ferreira Cabrera, Universidad de Concepción, Chile (2015).
Tipo de Avaliação	O estudo utiliza uma avaliação que combina itens objetivos e abertos, incluindo perguntas de múltipla escolha e atividades de expressão escrita e oral. Dada a ausência de provas de Espanhol para fins acadêmicos reconhecidas nacional ou internacionalmente, o estudo destaca a necessidade de um instrumento que permita avaliar as competências acadêmicas de estudantes estrangeiros.
População Avaliada	O público-alvo da avaliação são estudantes estrangeiros de intercâmbio universitário que estão aprendendo espanhol em um nível B2.
Competências Avaliadas	As competências avaliadas incluem as quatro habilidades linguísticas: compreensão leitora, compreensão oral, produção escrita e produção oral.
Métodos e Instrumentos	Os testes do estudo consistem em itens de correção objetiva (como perguntas de múltipla escolha) e tarefas abertas. Os principais tópicos dos exercícios incluem temas acadêmicos e de atualidade.
Resultados e Conclusões	Os resultados mostraram que os itens do protótipo da prova apresentaram validade de conteúdo, pois mediram adequadamente as dimensões definidas. Os especialistas avaliaram os itens com pontuações altas na escala Likert, indicando que atendiam aos critérios estabelecidos.
Desafios e Limitações	Embora a pesquisa não detalhe explicitamente os desafios, pode-se inferir que a validação do conteúdo e a necessidade de garantir que os itens realmente meçam as dimensões desejadas são aspectos críticos que podem apresentar desafios no processo de avaliação.

Recomendações e Futuras Pesquisas	O estudo sugere que a participação de especialistas na validação do conteúdo é fundamental e que os itens e tarefas devem ser continuamente refinados para garantir sua eficácia na avaliação das competências linguísticas. No entanto, não são especificadas recomendações concretas para estudos futuros.
--	--

Fonte: elaboração própria.

Esse estudo exploratório buscou validar o conteúdo de um protótipo de avaliação para estudantes estrangeiros de intercâmbio universitário com nível de espanhol B2, dada a carência de instrumentos reconhecidos para esse público. A avaliação proposta combina itens objetivos, como múltipla escolha, e tarefas abertas de expressão escrita e oral, abrangendo as quatro habilidades linguísticas (compreensão leitora e oral, produção escrita e oral) com foco em temas acadêmicos e da atualidade. Os resultados indicaram que os itens do protótipo demonstraram validade de conteúdo, com especialistas atribuindo pontuações elevadas, sugerindo que as questões mediram adequadamente as dimensões definidas. O estudo enfatiza a importância da participação de especialistas na validação e a necessidade de refinamento contínuo dos itens para assegurar a eficácia da avaliação, apesar de não apresentar recomendações específicas para futuras pesquisas.

5.2.18 Aligning English Language – Revista Profile – 2016

O Quadro 20 exibe partes do conteúdo do artigo de Palacio, Gaviria e Brown (2016). O artigo combina avaliações que reúnem mensurações de desempenho oral e escrito.

Quadro 20 – Palacio, Gaviria e Brown (2016)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Marcela Palacio, Sandra Gaviria, Universidad EAFIT, Medellín, Colômbia, e James Dean Brown, University of Hawai'i at Mānoa, Honolulu, Estados Unidos (2016)
Tipo de Avaliação	O estudo utiliza uma combinação de avaliações em sala de aula que incluem avaliações de desempenho tanto oral quanto escrito. Especificamente, o sistema de avaliação é composto de: Provas de meio de curso e finais: essas provas são projetadas para avaliar a compreensão dos alunos sobre o material ensinado nos cursos. Cada uma dessas provas representa 30% da nota final do curso. Notas de acompanhamento: os 40% restantes da nota do curso são atribuídos através de várias atividades de avaliação que os professores podem conduzir, que podem incluir rubricas, tarefas de avaliação sugeridas e convenções de escrita. Análises de itens: O estudo também destaca o uso de técnicas de análises de itens para melhorar a qualidade das questões de prova, garantindo que as avaliações estejam alinhadas com os princípios comunicativos do ensino e aprendizagem de línguas.

	De modo geral, a abordagem de avaliação é projetada para ser mais integrada ao aprendizado e à prática em sala de aula, afastando-se dos formatos tradicionais de teste.
População Avaliada	O público-alvo avaliado no estudo são os estudantes de idiomas, especificamente aqueles nos cursos 5 a 8, que estão em processo de consolidação do nível A2 do Quadro Europeu Comum de Referência para Línguas.
Competências Avaliadas	As principais competências avaliadas no estudo incluem: Compreensão oral, leitura, uso da língua (uso de gramática e vocabulário), escrita e fala. Essas competências estão alinhadas com os indicadores de desempenho do curso e foram projetadas para refletir a competência comunicativa que a metodologia do Idiomas-EAFIT busca desenvolver nos alunos.
Métodos e Instrumentos	Os testes no estudo consistem em uma variedade de tipos de itens, incluindo questões de múltipla escolha, questões abertas, afirmativas verdadeiras/falsas e tarefas contextualizadas (evocando formas linguísticas esperadas em contextos da vida real). Os principais temas dos exercícios nesses testes estão relacionados a situações sociais, acadêmicas ou profissionais, frequentemente conectados por um tema comum. Os testes têm como objetivo medir a capacidade de descrever, comparar e contrastar pessoas, lugares e coisas, refletindo a competência comunicativa que o programa busca desenvolver.
Resultados e Conclusões	Os exames analisados foram considerados uma melhoria significativa em relação às práticas de avaliação anteriores. O estudo destacou que os novos testes estão mais bem alinhados com os princípios comunicativos de ensino e aprendizagem de línguas, tornando-os mais relevantes e eficazes para avaliar o desempenho dos alunos. As análises realizadas sugeriram que os testes são confiáveis e válidos para tomar decisões de aprovação e reprovação. Mas, apesar dos achados positivos, o estudo concluiu que sempre há espaço para mais aprimoramentos. De modo geral, o estudo defende uma avaliação contínua e adaptação dos métodos de avaliação para garantir que eles apoiem efetivamente o aprendizado dos alunos e reflitam os valores do programa educacional.
Desafios e Limitações	Muitos professores não receberam treinamento adequado sobre como elaborar atividades de acompanhamento que se assemelhem às tarefas dos testes, o que pode afetar o alinhamento entre ensino e avaliação. O estudo encontrou dificuldades em criar itens de compreensão oral que utilizassem materiais de áudio autênticos, apropriados para cada nível de proficiência, que estivessem de acordo com os temas das unidades e que avaliassem várias habilidades descritas nos padrões do curso. O suporte financeiro da administração foi, por vezes, um desafio, o que pode limitar os recursos disponíveis para desenvolver e implementar o sistema de avaliação. Os resultados podem não ser facilmente generalizáveis para outros contextos ou instituições, pois são específicos ao programa de idiomas em questão e ao seu ambiente educacional único.
Recomendações e Futuras Pesquisas	Substituição ou revisão de itens que não são úteis nas avaliações. É sugerido que haja um foco maior em fornecer treinamento para os professores sobre como usar as rubricas de forma mais confiável e como projetar atividades de acompanhamento que se alinhem com os testes. A pesquisa deve continuar a investigar a eficácia, eficiência e qualidade dos itens, utilizando análises de confiabilidade e validade para garantir que os testes estejam realmente medindo o que se propõem a medir.

	Futuras pesquisas devem explorar mais a integração entre avaliação e práticas de ensino.
--	--

Fonte: elaboração própria.

Nesse estudo, a abordagem de avaliação, de modo geral, é projetada para ser mais integrada ao aprendizado e à prática em sala de aula, afastando-se dos formatos tradicionais de teste. O público-alvo avaliado no estudo são os estudantes de idiomas em processo de consolidação do nível A2 do Quadro Europeu Comum. O estudo focaliza a avaliação das habilidades de compreensão oral, leitura, uso da língua (uso de gramática e vocabulário), escrita e fala. A pesquisa enfatizou a importância de se integrar a avaliação com o aprendizado e as práticas em sala de aula. Formatos tradicionais de teste foram considerados incongruentes com as práticas educacionais atuais, e o novo sistema de avaliação visa a fornecer medidas mais autênticas das habilidades dos alunos. Muitas vezes houve uma falta geral de concordância entre os professores em relação à correção das provas, o que pode levar a marcações subjetivas de itens abertos, devido a inconsistências entre os avaliadores. Há uma necessidade de se explorarem formatos de avaliação mais autênticos que se afastem dos testes tradicionais, que, muitas vezes, não capturam adequadamente as habilidades comunicativas dos alunos.

5.2.19 Evaluación de una intervención pedagógica de alfabetización visual en la comprensión lectora de textos multimodales en programas de formación inicial docente en Lenguaje y Comunicación – Revista Literatura y Lingüística – 2016

O Quadro 21 exibe partes do conteúdo do artigo de Farias (2016). O artigo investiga um protocolo de intervenção em alfabetização visual e compreensão leitora.

Quadro 21 – Farías (2016)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Miguel Farías, Universidad de Santiago de Chile (2016)
Tipo de Avaliação	O estudo utiliza uma avaliação semiexperimental na qual os participantes foram expostos a um protocolo de intervenção em alfabetização visual e, em seguida, realizaram testes de compreensão leitora. Os testes incluíram perguntas textuais e multimodais.
População Avaliada	O público-alvo do estudo são estudantes em formação inicial docente nos cursos de Linguagem e Comunicação de duas universidades.
Competências Avaliadas	As principais competências avaliadas no estudo são a compreensão leitora, especialmente em relação a textos multimodais, que incluem a habilidade de interpretar e relacionar texto e imagem.

Métodos e Instrumentos	Os testes consistem em perguntas textuais e multimodais. Os tópicos dos exercícios incluem a leitura de textos multimodais, identificação de propósitos de elementos textuais e visuais, e formulação da ideia principal do texto.
Resultados e Conclusões	Os resultados mostraram que não houve diferenças significativas entre os grupos expostos à intervenção e os que não foram expostos. No entanto, uma leve diferença foi observada em um dos textos, com o grupo com intervenção obtendo desempenho ligeiramente melhor.
Desafios e Limitações	Um dos principais desafios foi o baixo número de participantes, que limitou a capacidade de evidenciar tendências nos resultados. Além disso, a falta de diferenças significativas entre os grupos expostos e não expostos à intervenção foi uma limitação importante.
Recomendações e Futuras Pesquisas	O estudo recomenda que futuras aplicações incluam textos multimodais de maior complexidade e propõe que o modelo de intervenção seja adaptado para explorar mais a dimensão estética dos textos multimodais.

Fonte: elaboração própria.

Realizou-se uma avaliação semiexperimental na qual os participantes foram expostos a um protocolo de intervenção em alfabetização visual e, em seguida, realizaram testes de compreensão leitora. Os testes incluíram perguntas textuais e multimodais. O público envolvido no estudo são estudantes em formação inicial docente nos cursos de Linguagem e Comunicação de duas universidades. A principal habilidade avaliada foi a compreensão leitora, sobretudo se relacionado a textos multimodais, que envolvem o processo de interpretar e relacionar texto e imagem. O estudo, com baixo número de participantes, não apresentou diferenças significativas entre os grupos expostos à intervenção e os que não foram expostos.

5.2.20 Producción del Ensayo Argumentativo en Estudiantes de dos Establecimientos de Educación Media Técnico-Profesional de la Provincia de Ñuble. Resultados de un Diagnóstico – Revista Literatura y Lingüística – 2016

O Quadro 22 exibe partes do conteúdo do artigo de Pastene Labrín, Díaz Chavarría e Molina Castilho (2016). O artigo trata da avaliação diagnóstica.

Quadro 22 – Pastene Labrín, Díaz Chavarría e Molina Castillo (2016)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Federico Carlos Pastene Labrín, Doctor de la Universidad Austral de Chile, Académico Universidad del Bío-Bío, Chillán, Chile; Rosa Marta Díaz Chavarría, Magíster en Literatura de la Universidad de Chile, Académica Universidad del Bío-Bío, Chillán, Chile; Sandra Annette Molina Castillo, Magíster en Educación. Universidad La República, Chile, Académica Universidad del Bío-Bío, Chillán, Chile (2016)
Tipo de Avaliação	O estudo utiliza uma avaliação diagnóstica de escrita, na qual os alunos redigem um ensaio argumentativo. A avaliação é feita com base em uma rubrica que

	considera várias dimensões da escrita, como ortografia, coerência e argumentação.
População Avaliada	O público-alvo do estudo são estudantes do terceiro e quarto anos de dois estabelecimentos de Educação Média Técnico-Profissional (EMTP) da Província de Ñuble, totalizando 202 alunos.
Competências Avaliadas	As principais competências avaliadas incluem a produção escrita, especificamente a habilidade de escrever um ensaio argumentativo, que abrange dimensões como ortografia, coerência, coesão, adequação, léxico, partes do discurso e argumentação.
Métodos e Instrumentos	Os testes consistem em uma prova de escrita, por meio da qual os alunos escolhem um dos quatro temas propostos para redigir seu ensaio. Os tópicos dos exercícios incluem a estrutura do ensaio argumentativo e a capacidade de expressar e defender uma posição.
Resultados e Conclusões	Os resultados mostram que os alunos alcançaram um desempenho aceitável, com 60,5% de sucesso na produção do ensaio argumentativo, destacando-se a adequação como a dimensão mais bem-sucedida, enquanto as partes do discurso e a argumentação apresentaram maiores dificuldades.
Desafios e Limitações	Os principais desafios incluem a identificação de fatores que influenciam a variabilidade no desempenho dos alunos, além de dificuldades na estruturação e argumentação nos ensaios. As correlações entre notas e desempenho foram significativas, mas outros fatores também podem estar em jogo.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se desenvolverem melhor as habilidades de escrita e argumentação nos alunos, além de se considerarem outros fatores que podem influenciar a produção escrita. Recomenda-se também a adaptação de métodos de ensino e avaliação para atender às necessidades dos alunos.

Fonte: elaboração própria.

O presente estudo avaliou as habilidades de escrita argumentativa de 202 alunos do terceiro e quarto anos de dois estabelecimentos de Educação Média Técnico-Profissional, utilizando uma avaliação diagnóstica de ensaio argumentativo com uma rubrica que considerou ortografia, coerência e argumentação. Os resultados indicaram um desempenho aceitável dos alunos, com maiores dificuldades em partes do discurso e argumentação, sugerindo a necessidade de se aprimorarem as habilidades de escrita e argumentação e se adaptarem os métodos de ensino e avaliação para atender às necessidades dos alunos.

5.2.21 Memoria operativa, comprensión lectora y rendimiento escolar – Revista Literatura y Lingüística – 2016

O Quadro 23 exibe partes do conteúdo do artigo de Guzmán, Véliz e Reyes (2016). O estudo investigou a relação entre memória operativa e compreensão de leitura em 80 estudantes chilenos do ensino básico e médio, utilizando testes padronizados, como Lectum e TAL.

Quadro 23 – Guzmán, Véliz e Reyes (2016)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Belén Guzmán, Magíster en Lingüística Aplicada por la Universidad de Concepción, Chile; Mónica Véliz, Magíster en Lingüística por la Universidad de Concepción. Académica de la Universidad de Concepción, Chile; y Fernando Reyes, Magister en Psicología de la Salud por la Universidad de Concepción. Académico de la Universidad del Desarrollo, Concepción, Chile (2016)
Tipo de Avaliação	O estudo utiliza testes padronizados para avaliar a memória e a compreensão de leitura. Especificamente, a prova de compreensão leitora (Lectum) e o teste de amplitude de leitura (TAL) foram aplicados.
População Avaliada	O público-alvo do estudo consiste em 80 estudantes (30 mulheres e 50 homens) de um estabelecimento educacional particular-subvencionado na cidade de Concepción, Chile. Os estudantes pertencem ao sétimo ano de ensino básico, assim como ao primeiro e ao terceiro ano do ensino médio.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a memória operativa e a compreensão de leitura. O estudo também considera o desempenho dos alunos em termos de notas em diversas disciplinas, especialmente em Língua e Comunicação.
Métodos e Instrumentos	Os testes consistem em questões de múltipla escolha e tarefas que envolvem a leitura de orações e a memorização de palavras. O teste de amplitude de leitura (TAL) envolve a leitura de orações e a memorização da última palavra de cada uma, enquanto a prova de compreensão leitora (Lectum) é composta de subtestes que avaliam diferentes dimensões da leitura.
Resultados e Conclusões	Os resultados indicam que a memória operativa tem um papel significativo na compreensão de leitura e no desempenho escolar dos estudantes. O estudo conclui que há uma correlação positiva entre a memória operativa e a capacidade de compreensão leitora, o que sugere que intervenções que melhorem a memória operativa podem beneficiar o desempenho escolar.
Desafios e Limitações	Entre os desafios e as limitações do estudo, destaca-se a necessidade de garantir que todos os participantes compreendam as instruções dos testes, além de possíveis variações nas condições de aplicação dos testes. A amostra também é limitada a um único estabelecimento educacional, o que pode afetar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas poderiam incluir uma amostra mais ampla e diversificada, além de considerar outros fatores que pudessem influenciar a memória operativa e a compreensão de leitura, como o contexto socioeconômico dos alunos. Também é recomendado o desenvolvimento de intervenções específicas para melhorar a memória operativa em ambientes escolares.

Fonte: elaboração própria.

Os resultados apontaram uma correlação positiva entre memória operativa e desempenho escolar, sugerindo que intervenções para melhorar a memória operativa podem beneficiar a compreensão de leitura. O estudo destaca a importância de garantir a compreensão das instruções dos testes e a necessidade de futuras pesquisas com amostras mais amplas e diversificadas, além de considerar fatores socioeconômicos e desenvolver intervenções específicas para aprimorar a memória operativa em ambientes escolares.

5.2.22 *Strategies Instruction to Improve the Preparation for English Oral Exams – Revista Profile – 2016*

O Quadro 24 exhibe partes do conteúdo do artigo de Abad e Alzate (2016).

De modo geral, a forma como os alunos escolhem e usam estratégias para se preparar para avaliações orais é fortemente influenciada pela qualidade da instrução dos professores. Usar critérios de avaliação bem elaborados e compartilhá-los previamente com os alunos melhora a compreensão do que se espera deles e como podem atingir esses objetivos, aumentando a sensação de autonomia dos alunos. Quando os critérios de avaliação são combinados com a instrução sobre estratégias dentro de um sistema de ensino abrangente, os alunos tendem a se sentir mais no controle de sua aprendizagem e podem se preparar de maneira mais eficaz para melhorar seu desempenho em testes.

A escolha e o uso de estratégias de aprendizagem são influenciados por fatores como idade, gênero, estilo de aprendizagem e tipo de atividade. O estudo sugere que a importância atribuída por um professor a certos aspectos da avaliação, como expressão verbal, habilidades linguísticas, organização e preparação, pode influenciar as estratégias que os alunos utilizam. A prioridade que os professores dão a esses aspectos durante a instrução está diretamente relacionada ao valor que os alunos atribuem a eles.

Apesar de a instrução sobre estratégias poder parecer prescritiva, é fundamental que os professores a ajustem para atender às necessidades de aprendizagem dos alunos. No entanto, os professores devem considerar que a ênfase em determinados aspectos da língua pode influenciar a escolha das estratégias pelos alunos e, conseqüentemente, seu processo geral de aprendizagem.

Quadro 24 – Abad e Alzate (2016)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	José Vicente Abad e Paula Andrea Alzate, Fundación Universitaria Luis Amigó, Medellín, Colômbia (2016).
Tipo de Avaliação	Exames de língua inglesa, focalizados na produção oral. Análise do uso de estratégias de aprendizagem com foco na habilidade oral (rubricas e questionários). Testes formativos e somativos.
População Avaliada	Dois grupos (um com 12 e o outro com 14 participantes) de estudantes colombianos de língua inglesa com nível B1, entre 18 e 25 anos de idade, de duas universidades privadas.
Competências Avaliadas	Competência de produção oral em exames de língua inglesa.
Métodos e Instrumentos	Preparo de alunos para testes orais com base na elaboração de estratégias e rubricas.

Resultados e Conclusões	Melhoria na <i>performance</i> dos testes após instrução em estratégias de aprendizagem (ambos os grupos). Compartilhar rubricas previamente ajudou os alunos a identificarem o propósito da avaliação e promoveu sensação de empoderamento e controle no processo de preparo. Importância de rubricas claras, definindo os objetivos de aprendizagem.
Desafios e Limitações	Inconsistências entre ensino e avaliação de língua inglesa na Colômbia. Foco dos alunos em passar nos testes, em vez de aprimorar o conhecimento linguístico.
Recomendações e Futuras Pesquisas	Clareza nas instruções reduz a necessidade de esclarecimentos adicionais. Compartilhar critérios de avaliação com antecedência aumenta a compreensão do propósito da atividade e a sensação de controle sobre a preparação. Critérios devem descrever detalhadamente os objetivos de aprendizagem, requisitos e critérios de avaliação. Incorporar instrução sobre estratégias em cursos de línguas para auxiliar na preparação para exames orais, incentivando o uso de estratégias metacognitivas (planejamento, monitoramento e ajuste).

Fonte: elaboração própria.

5.2.23 *Analysing Test Takers' Views On A Computer Based Speaking Test – Revista Profile – 2017*

A pesquisa de Amengual-Pizarro e García-Laborda (2017), exibida em partes no Quadro 25, é interessante por ter escutado a voz dos estudantes a respeito do teste. Trata-se de uma revista colombiana, mas se refere a uma universidade da Espanha e de um teste britânico de proficiência.

Os resultados desse estudo mostram que, apesar das dificuldades em capturar a interação oral humana, os testes realizados por computador são considerados uma forma válida de medir a competência oral em inglês e são vistos como um método adequado para a avaliação das habilidades de fala. De maneira geral, os participantes têm uma visão positiva do teste de fala APTIS e acham a aplicação do teste muito prática, intuitiva e amigável. Além disso, os dados mostram que os candidatos têm opiniões muito favoráveis sobre o material de apoio fornecido pelos pesquisadores para a sua familiarização com o formato, conteúdo e nível de dificuldade do teste. Eles também acreditam que as sessões de prática para autorreflexão e desenvolvimento de estratégias de aprendizado foram bastante úteis para obter bons resultados no teste.

É evidente que os candidatos preferem os testes presenciais aos realizados por computador para avaliar a habilidade oral, devido à natureza social e interacional das habilidades de fala, que não parecem ser bem avaliadas em formatos computadorizados. Características pessoais dos candidatos, como a introversão, podem influenciar na escolha do

modo de avaliação mais adequado, já que candidatos introvertidos relataram estar menos ansiosos e mais confortáveis na ausência de um interlocutor.

A maioria desses participantes também acredita que as pontuações dos testes computadorizados tendem a ser mais confiáveis porque os examinadores não são influenciados pelas características pessoais dos candidatos. Um conceito fundamental em testes de linguagem é a adequação ao propósito: Os testes não são apenas válidos, mas válidos para um propósito específico. As pontuações dos testes computadorizados fornecem uma evidência sobre o desempenho de um candidato, mas não devem ser a única base para a tomada de decisões. Portanto, acredita-se que ambos os métodos devem ser vistos como complementares e não como alternativas concorrentes.

Quadro 25 – Amengual-Pizarro e García-Laborda (2017)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Marian Amengual-Pizarro, Universidad de las Islas Baleares, Espanha, e Jesús García-Laborda, Universidad de Alcalá, Espanha (2017)
Tipo de Avaliação	Teste de produção oral mediado por computador (comparado aos testes presenciais)
População Avaliada	80 estudantes universitários da Universidad de Alcalá (Espanha) que realizaram o APTIS, teste de proficiência em língua inglesa. A maioria dos estudantes avaliados era composta de mulheres (85%).
Competências Avaliadas	Produção oral em língua inglesa.
Métodos e Instrumentos	Um questionário foi aplicado a 80 estudantes que realizaram o teste APTIS, para que suas opiniões (sobre o teste de produção oral) fossem ouvidas. O APTIS é um teste desenvolvido pelo British Council, desenhado para um público acima de 15 anos. O teste engloba componentes como gramática, vocabulário, compreensão oral, compreensão escrita, produção escrita e produção oral. O APTIS pode ser administrado presencialmente, em formato papel e caneta, mas é comumente administrado <i>on-line</i> . A avaliação de produção oral leva, aproximadamente, 12 minutos. O teste de produção oral engloba questões como: perguntas pessoais, descrição de imagens e comparação, perguntas baseadas em uma imagem.
Resultados e Conclusões	Os testes <i>on-line</i> de produção oral são satisfatórios. Questões como personalidade introvertida e timidez podem influenciar na preferência por um modelo <i>on-line</i> ou presencial de teste. As vantagens dos testes <i>on-line</i> superam as suas limitações. Testes <i>on-line</i> reduzem gastos e tempo. Os testes <i>on-line</i> podem atuar como fontes complementares de testes mais convencionais. O material de apoio para a realização do teste foi de grande suporte. A familiarização com a plataforma utilizada foi traduzida em resultados mais satisfatórios. Em resposta ao questionário, 75% dos alunos disseram preferir testes presenciais de produção oral, devido ao caráter humano da interação oral (incluindo linguagem corporal, negociação de sentidos etc.).

	Um teste não é válido universalmente, mas válido para um contexto específico. Ele é apenas uma única fonte de evidência e não deve ser usado como guia único para a tomada de decisões.
Desafios e Limitações	A avaliação das habilidades de produção oral é considerada desafiadora por natureza. Parte da interação espontânea pode ser perdida devido ao uso do computador. Alguns alunos preferem a interação presencial com o tutor para o teste de produção oral, alegando que ajuda na espontaneidade da interação e que os tutores podem auxiliar quando os alunos não sabem o que falar. Quando um humano atribui a nota, ele pode estar influenciado por questões pessoais.
Recomendações e Futuras Pesquisas	A necessidade por uma proficiência oral satisfatória, requer instrumentos avaliativos que possam ser aplicados rapidamente e de forma eficiente, mantendo a qualidade do instrumento. Ambos os formatos (presencial e <i>on-line</i>) devem ser usados de modo complementar e não substitutivo.

Fonte: elaboração própria.

5.2.24 *El feedback correctivo escrito indirecto en el aprendizaje de la forma comparativa de adjetivos en inglés – Revista Logos – 2017*

O Quadro 26 exhibe partes do conteúdo do artigo de Muñoz e Ferreira (2017). Utilizando pré-teste, pós-teste imediato e diferido, além de tarefas de produção escrita como *e-mail* e carta, o estudo investigou o impacto de duas estratégias de *feedback* corretivo na habilidade de 24 alunos do oitavo ano de um colégio particular de usar a forma comparativa de adjetivos em inglês.

Quadro 26 – Muñoz e Ferreira (2017)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Belén Muñoz, Universidad de Concepción, Chile, e Anita Ferreira, Universidad de Concepción, Chile (2017)
Tipo de Avaliação	O estudo utiliza avaliações escritas, incluindo um pré-teste, um pós-teste imediato e um pós-teste diferido, além de tarefas de produção escrita.
População Avaliada	O público-alvo do estudo é composto de 24 estudantes de oitavo ano básico de um colégio particular privado, com idades entre 13 e 14 anos.
Competências Avaliadas	As principais competências avaliadas no estudo são a escrita, especificamente no uso da forma comparativa de adjetivos em inglês.
Métodos e Instrumentos	Os testes consistem em avaliações escritas em que os alunos redigem um <i>e-mail</i> simulado e uma carta. Os tópicos principais dos exercícios envolvem a utilização da forma comparativa de adjetivos em inglês.
Resultados e Conclusões	O estudo conclui que as duas estratégias de <i>feedback</i> corretivo têm efeitos positivos no uso da forma comparativa de adjetivos.
Desafios e Limitações	O estudo não menciona explicitamente os desafios e limitações.
Recomendações e Futuras Pesquisas	O estudo não fornece recomendações específicas, mas sugere que futuras pesquisas poderiam explorar diferentes métodos de <i>feedback</i> e suas aplicações.

	em contextos variados.
--	------------------------

Fonte: elaboração própria.

Os resultados indicaram que ambas as estratégias de *feedback* corretivo foram eficazes no aprimoramento dessa habilidade específica, embora o estudo não tenha detalhado desafios ou limitações e tenha sugerido futuras pesquisas para explorar diferentes métodos de *feedback* em contextos variados.

5.2.25 Habilidades metamorfológicas y su incidencia en la comprensión lectora – Revista Signos – 2017

O Quadro 27 exhibe partes do conteúdo do artigo de Vega Rodríguez, Torres Rodríguez e Del Campo Rivas (2017).

Quadro 27 – Vega Rodríguez, Torres Rodríguez e Del Campo Rivas (2017)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Yuri Vega Rodríguez, Universidad Del Bío Bío, Chile, Angélica María Torres Rodríguez, Universidad San Sebastián, Chile, Manuel Del Campo Rivas, Universidad Santo Tomás, Chile (2017).
Tipo de Avaliação	O estudo utiliza uma avaliação por meio de testes, especificamente a Prueba de Evaluación de Habilidades Metamorfológicas (PEHMM), que foi projetada para medir as habilidades metamorfológicas dos alunos.
População Avaliada	O público-alvo do estudo são estudantes de quarto ano do ensino básico, com uma amostra de 36 crianças (19 meninos e 17 meninas) de escolas regulares urbanas em Chillan Viejo, Chile.
Competências Avaliadas	As principais competências avaliadas no estudo são as habilidades metamorfológicas, que incluem a consciência morfológica, a capacidade de análise e composição de palavras, e a compreensão leitora.
Métodos e Instrumentos	Os testes consistem em provas que avaliam a habilidade metamorfológica de composição e análise. As provas incluem atividades como construir palavras usando morfemas derivativos, derivar palavras em contextos oracionais, e identificar a palavra de origem de um estímulo apresentado. Cada prova contém 15 itens.
Resultados e Conclusões	O estudo conclui que a consciência morfológica desempenha um papel importante no desenvolvimento do vocabulário e na leitura de palavras, especialmente à medida que as crianças são expostas a palavras morfológicamente complexas.
Desafios e Limitações	O estudo não menciona explicitamente os desafios e limitações do estudo.
Recomendações e Futuras Pesquisas	O estudo não fornece recomendações específicas para estudos futuros, mas há uma ênfase na importância da consciência morfológica.

Fonte: elaboração própria.

O presente estudo aborda testes como a Prueba de Evaluación de Habilidades Metamorfológicas, que visa a medir as habilidades metamorfológicas dos alunos. Estudantes de escolas regulares chilenas (do quarto ano do ensino básico) são o público-alvo (19 meninos e 17 meninas). Tais alunos são principalmente avaliados quanto à consciência morfológica, compreensão leitora, análise e composição de palavras. As provas (que incluem 15 itens) são compostas de atividades como construir, derivar e identificar palavras. Conclui-se, dentre outros, que a consciência morfológica desempenha um papel primordial no desenvolvimento do vocabulário e na leitura de palavras, sobretudo à medida que as crianças são expostas a palavras morfológicamente complexas.

5.2.26 Habilidades favorecedoras del aprendizaje de la lectura en alumnos de 5 y 6 años
Revista Signos – 2018

O Quadro 28 exibe partes do conteúdo do artigo de Gutiérrez (2018). O artigo investiga instrumentos distintos de avaliação.

Quadro 28 – Gutiérrez (2018)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Raúl Gutiérrez, Universidad de Alicante, Espanha (2018)
Tipo de Avaliação	O estudo utilizou uma avaliação pré-teste e pós-teste, aplicada de forma individual aos alunos. Os instrumentos de avaliação foram aplicados por profissionais da educação, como docentes especialistas em Audição e Linguagem e psicopedagogos, garantindo a homogeneidade na coleta de dados.
População Avaliada	O público-alvo do estudo são alunos de 5 e 6 anos, totalizando 408 participantes, que foram divididos em grupos experimentais e de controle.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a consciência fonológica, a velocidade de denominação e o conhecimento alfabético.
Métodos e Instrumentos	Os testes consistem em avaliações que medem a consciência fonológica (como a consciência silábica e fonêmica), a velocidade de denominação (RAN) de números e letras. Os exercícios incluem tarefas que avaliam a capacidade dos alunos de identificar e manipular sons em palavras, além de reconhecer rapidamente números e letras.
Resultados e Conclusões	Os resultados indicam que os alunos do grupo experimental apresentaram melhorias significativas nas habilidades de consciência fonológica, velocidade de denominação e conhecimento alfabético em comparação com o grupo de controle. As intervenções tiveram um impacto positivo no aprendizado da leitura.
Desafios e Limitações	O estudo enfrentou desafios relacionados à homogeneidade na aplicação dos testes e à necessidade de garantir que todos os participantes estivessem igualmente motivados e engajados.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de continuar explorando a eficácia de intervenções semelhantes em diferentes contextos e com populações diversas.

	Recomenda-se também a inclusão de métodos de avaliação mais variados e a realização de estudos longitudinais para acompanhar o desenvolvimento das habilidades ao longo do tempo.
--	---

Fonte: elaboração própria.

O estudo investigou o impacto de intervenções no aprendizado de leitura de 408 alunos de 5 e 6 anos, divididos em grupos experimental e de controle, utilizando avaliações pré-teste e pós-teste aplicadas individualmente por profissionais da educação para medir a consciência fonológica, velocidade de denominação e conhecimento alfabético. Os resultados demonstraram que o grupo experimental apresentou melhorias significativas nessas habilidades, indicando um impacto positivo das intervenções, apesar dos desafios relacionados à homogeneidade na aplicação dos testes e à motivação dos participantes. Sugeriu-se a continuidade da exploração da eficácia dessas intervenções em diferentes contextos e com metodologias de avaliação mais variadas, incluindo-se estudos longitudinais.

5.2.27 Validez y fiabilidad de una escala de evaluación oral en ele – 1 – Revista RLA – 2018

O Quadro 29 exibe partes do conteúdo do artigo de Martín Alonso (2018). O estudo investigou a validade e a fiabilidade de uma escala de avaliação para a prova oral de classificação em Espanhol como Língua Estrangeira (ELE) aplicada a estudantes da Pontificia Universidad Javeriana em Bogotá. A prova avalia a produção oral dos candidatos por meio de tarefas que exigem a apresentação pessoal, descrição e comparação de cidades, e discussão sobre questões ambientais, com foco em critérios como alcance, correção, fluidez e pronúncia.

Quadro 29 – Martín Alonso (2018)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Begoña Martín Alonso, Pontificia Universidad Javeriana, Colômbia (2018)
Tipo de Avaliação	O estudo utiliza uma escala de avaliação para a prova oral de classificação em ELE (Espanhol como Língua Estrangeira), que considera critérios como alcance, correção, fluidez e pronúncia.
População Avaliada	O público-alvo do estudo são estudantes de espanhol como língua estrangeira candidatos à prova oral de classificação na Pontificia Universidad Javeriana (PUJ) em Bogotá.
Competências Avaliadas	A principal competência avaliada na prova é a produção oral, especificamente em relação a critérios de correção, fluidez e pronúncia.
Métodos e Instrumentos	Os testes consistem em tarefas orais, por meio das quais os candidatos devem realizar atividades como se apresentar, descrever a cidade de Bogotá, compará-

	la com sua cidade natal e discutir questões sobre a contaminação em Bogotá. As tarefas são projetadas para avaliar a produção oral dos candidatos.
Resultados e Conclusões	Os resultados indicam que a escala de avaliação da prova oral de classificação é válida e fiável. Os dados qualitativos e quantitativos mostraram que os descritores da escala refletem adequadamente o uso da língua pelos candidatos e que as modificações feitas durante o processo de validação foram pertinentes.
Desafios e Limitações	As principais limitações incluem a amostra limitada de estudantes e avaliadores que participaram das pilotagens. Além disso, há a necessidade de mais investigações sobre como os avaliadores interpretam os descritores da escala e a possibilidade de alcançar um acordo razoável entre eles, quando não há formação prévia.
Recomendações e Futuras Pesquisas	O estudo recomenda a realização de futuras pilotagens e análises qualitativas para aprofundar a compreensão sobre a interpretação dos descritores pelos avaliadores. Também enfatiza a importância de coletar diferentes tipos de evidências para sustentar a adequação das interpretações e decisões tomadas com base nos resultados da prova.

Fonte: elaboração própria.

Os resultados, baseados em dados qualitativos e quantitativos, confirmaram a eficácia da escala, indicando que os descritores refletem adequadamente o desempenho linguístico dos candidatos. No entanto, o estudo reconhece limitações, como o tamanho reduzido da amostra e a necessidade de se investigar a interpretação dos descritores pelos avaliadores. Também recomenda futuras pesquisas com amostras maiores e análises qualitativas aprofundadas para aprimorar a escala e garantir a validade das interpretações e decisões com base nos resultados da prova.

5.2.28 Evaluación de la escritura en niños de tercer grado en Brasil – Revista Boletín De Lingüística – 2018

O Quadro 30 exibe partes do conteúdo do artigo de Scliar-Cabral (2018).

Quadro 30 – Scliar-Cabral (2018)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Leonor Scliar-Cabral, Universidade Federal de Santa Catarina (2018).
Tipo de Avaliação	O estudo utiliza a Avaliação Nacional de Alfabetização (ANA), que é uma avaliação padronizada aplicada a uma grande população de alunos. A ANA classifica a produção escrita em dois grandes blocos: o da palavra e o do texto, permitindo uma análise quantitativa e qualitativa das habilidades de escrita dos alunos.
População Avaliada	O público-alvo da avaliação são crianças de 8 anos que completam o ciclo de alfabetização, especificamente alunos do terceiro grau da educação básica

	(primária) em escolas públicas no Brasil. A avaliação abrangeu uma população de 2.160.601 crianças.
Competências Avaliadas	As principais competências avaliadas incluem a capacidade de escrever palavras com diferentes estruturas silábicas e a produção de textos com base em situações dadas. A ANA se propõe a avaliar habilidades como a ortografia e a estrutura narrativa.
Métodos e Instrumentos	Os testes da ANA consistem em questões abertas, em que os alunos devem produzir textos e escrever palavras. Os principais tópicos dos exercícios incluem a grafia de palavras com estrutura silábica canônica e não canônica, além da produção de narrativas.
Resultados e Conclusões	Os resultados indicam um perfil preocupante, considerando que 33,85% dos alunos conseguiram apenas escrever de forma imprecisa ou produzir fragmentos, e apenas 8,28% conseguiram escrever uma história completa sem erros de segmentação. A conclusão é que há uma necessidade urgente de se melhorarem as políticas de alfabetização e a formação de professores para abordar essas deficiências.
Desafios e Limitações	Os principais desafios incluem a inconsistência no <i>design</i> do instrumento de avaliação e a dificuldade em conceituar a unidade “palavra” na linguística. Além disso, a avaliação não detalha as razões por trás das dificuldades dos alunos, o que limita a compreensão das causas dos resultados insatisfatórios.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de uma reflexão crítica sobre as habilidades de escrita que devem ser desenvolvidas e recomenda que as decisões sobre metodologias e materiais pedagógicos sejam baseadas em evidências. Além disso, enfatiza a importância de se aprofundar a discussão sobre os indicadores de alfabetização e suas implicações para a prática educacional.

Fonte: elaboração própria.

A Avaliação Nacional de Alfabetização (ANA) é uma avaliação padronizada aplicada a uma grande população de alunos. A ANA classifica a produção escrita em dois grandes blocos: o da palavra e o do texto, permitindo uma análise quantitativa e qualitativa das habilidades de escrita dos alunos. Crianças de oito anos, especificamente alunos do terceiro grau da educação básica (primária) em escolas públicas no Brasil são o público-alvo do estudo. Quanto ao foco das habilidades avaliadas, a capacidade de escrever palavras com diferentes estruturas silábicas e a produção de textos com base em situações dadas representam grande peso da avaliação. A ANA se propõe a avaliar habilidades como a ortografia e a estrutura narrativa. Dentre os principais desafios apresentados como conclusão, 33,85% dos alunos conseguiram apenas escrever de forma imprecisa ou produzir fragmentos, e apenas 8,28% conseguiram escrever uma história completa sem erros de segmentação. Ademais, existe uma inconsistência no *design* do instrumento avaliativo.

5.2.29 Complejidad, precisión y fluidez en el desempeño oral de aprendices con distintos niveles de proficiencia en inglés como L2 – Revista Signos – 2018

O Quadro 31 exibe partes do conteúdo do artigo de Alfaro Chat, Gómez Álvarez e Saez Carrillo (2018). O artigo investiga testes padronizados.

Quadro 31 – Alfaro Chat, Gómez Álvarez e Saez Carrillo (2018)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Javiera Alfaro Chat, Lilian Gómez Álvarez e Katia Saez Carrillo, Universidad de Concepción, Chile (2018).
Tipo de Avaliação	O estudo utiliza uma avaliação baseada em testes padronizados, especificamente o Test of English as a Foreign Language (TOEFL iBT), que inclui tarefas de produção oral em um formato de <i>mock exam</i> .
População Avaliada	O público-alvo do estudo são 44 estudantes de Pedagogia/Letras (em inglês) de diferentes anos de formação curricular, com níveis de proficiência em inglês A2, B2 e C1, conforme o Marco Comum Europeu de Referência.
Competências Avaliadas	As principais competências avaliadas no estudo são complexidade, precisão e fluidez na produção oral em inglês.
Métodos e Instrumentos	Os testes consistem em tarefas de produção oral que fazem parte de um simulado (<i>mock exam</i>) do TOEFL iBT, que avalia a capacidade de realizar monólogos semiespontâneos. Os tópicos dos exercícios incluem a produção oral em diferentes contextos e a aplicação de habilidades linguísticas em situações práticas.
Resultados e Conclusões	O estudo conclui que os aprendizes não conseguem atender a todos os aspectos do desempenho linguístico simultaneamente, priorizando complexidade, precisão ou fluidez, dependendo da tarefa. Isso sugere que as demandas do processamento linguístico superam a capacidade de atenção dos aprendizes.
Desafios e Limitações	Um dos principais desafios mencionados é a dificuldade dos aprendizes em equilibrar os diferentes aspectos do desempenho linguístico, o que pode afetar a avaliação. Além disso, a pesquisa pode ter limitações relacionadas ao tamanho da amostra e à generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas poderiam explorar mais a fundo a interação entre complexidade, precisão e fluidez, além de considerar diferentes contextos de aprendizagem e tipos de tarefas. Recomenda-se também a inclusão de uma amostra maior e mais diversificada para validar os resultados.

Fonte: elaboração própria.

O estudo investigou a produção oral em inglês de 44 estudantes de Pedagogia/Letras, de níveis A2, B2 e C1, utilizando tarefas de um simulado do TOEFL iBT. Os resultados revelaram que os aprendizes enfrentam dificuldades em equilibrar complexidade, precisão e fluidez simultaneamente, priorizando um aspecto em detrimento dos outros, devido às demandas do processamento linguístico. A pesquisa sugere a necessidade de estudos futuros com amostras maiores e mais diversificadas, explorando a interação entre esses aspectos e considerando diferentes contextos de aprendizagem e tipos de tarefas.

5.2.30 *Language Assessment Literacy and Teachers' Professional Development: A Review of the Literature – Revista Profile – 2018*

O Quadro 32 exhibe partes do conteúdo do artigo de Giraldo (2018).

Quadro 32 – Giraldo (2018)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Frank Giraldo, Universidad de Caldas, Manizales, Colômbia (2018).
Tipo de Avaliação	O estudo menciona que os professores tendem a usar procedimentos de avaliação tradicionais, como testes finais, e evitam avaliações alternativas, como a avaliação por pares. No entanto, os programas analisados enfatizam a crítica e a criação de instrumentos de avaliação.
População Avaliada	O público-alvo do estudo inclui professores de línguas, que podem ser tanto professores em formação e professores em exercício quanto estudantes em programas de pós-graduação.
Competências Avaliadas	Compreensão oral, fala, leitura e escrita.
Métodos e Instrumentos	Os testes consistem em uma variedade de formatos, incluindo questões de múltipla escolha e questões abertas. Os principais tópicos abordados nos exercícios incluem a análise de itens, além de tópicos relacionados à validade e à confiabilidade das avaliações.
Resultados e Conclusões	O estudo conclui que os programas de formação em avaliação de linguagem priorizam a crítica e o <i>design</i> de avaliações, respondendo às necessidades expressas pelos professores. Além disso, destaca a importância de se conectarem teoria e prática no <i>design</i> de avaliações.
Desafios e Limitações	Um dos desafios mencionados é a ausência de um foco explícito em princípios éticos e de justiça nas avaliações, que são considerados fundamentais. Além disso, alguns estudos não abordaram qualidades essenciais, como validade e confiabilidade, o que limita a compreensão completa da avaliação de linguagem.
Recomendações e Futuras Pesquisas	A formação em avaliação de linguagem deve priorizar o <i>design</i> rigoroso de avaliações e a inclusão de princípios éticos e de justiça. Além disso, sugere que a teoria deve ser conectada ao <i>design</i> para que os professores possam aplicar o conhecimento de forma prática em suas avaliações.

Fonte: elaboração própria.

O estudo analisa os métodos de avaliação utilizados por professores de línguas (incluindo-se docentes em formação, em exercício e pós-graduandos), destacando a preferência por procedimentos tradicionais, como testes finais, em detrimento de abordagens alternativas, como a avaliação por pares. Apesar disso, os programas examinados enfatizam a crítica e a criação de instrumentos avaliativos, abrangendo competências linguísticas como compreensão oral, fala, leitura e escrita, por meio de formatos variados (múltipla escolha e questões abertas), com foco em análise de itens, validade e confiabilidade. Os resultados indicam que esses

programas atendem às demandas dos professores ao priorizar o *design* de avaliações, mas apontam lacunas, como a falta de discussão explícita sobre princípios éticos e de justiça, além da omissão, em alguns casos, de critérios essenciais, como validade e confiabilidade. Conclui-se que a formação em avaliação linguística deve integrar teoria e prática, assegurando rigor metodológico, inclusão de aspectos éticos e aplicabilidade do conhecimento no contexto real de ensino.

5.2.31 *La comprensión lectora de escolares de educación básica – Revista Literatura y Lingüística – 2019*

O Quadro 33 exibe partes do conteúdo do artigo de Gallego Ortega, Figueroa Sepúlveda e Rodríguez Fuentes (2019).

Quadro 33 – Gallego Ortega, Figueroa Sepúlveda e Rodríguez Fuentes (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	José Luis Gallego Ortega, Doctor en Ciencias de la Educación. Académico de la Universidad de Granada, Granada, Espanha; Susana Figueroa Sepúlveda, Doctora en Ciencias de la Educación de la Universidad de Granada. Académica de la Universidad Arturo Prat, Iquique, Chile; e Antonio Rodríguez Fuentes, Doctor en Ciencias de la Educación. Académico de la Universidad de Granada, Granada, Espanha (2019).
Tipo de Avaliação	O estudo utiliza provas de múltipla escolha para avaliar a compreensão leitora.
População Avaliada	O público-alvo do estudo são alunos da educação básica, especificamente aqueles do segundo ao oitavo ano de uma escola urbana no Chile.
Competências Avaliadas	A compreensão leitora, que é desdobrada em outros níveis (literal, inferencial e crítico), além do nível de aprendizado da leitura (insuficiente, básico e apropriado).
Métodos e Instrumentos	Os testes consistem em perguntas de múltipla escolha, em que os alunos devem escolher a resposta correta entre várias opções. Os principais tópicos dos exercícios incluem a compreensão de textos em diferentes níveis.
Resultados e Conclusões	Os resultados indicam que o nível de compreensão leitora e, portanto, o aprendizado, diminui à medida que os alunos avançam nos anos escolares. Além disso, não foram encontradas diferenças significativas na compreensão leitora de acordo com o gênero. Os resultados ainda ressaltam a importância de se desenvolverem estratégias de intervenção para melhorar a compreensão leitora, especialmente em níveis escolares nos quais o desempenho é mais fraco.
Desafios e Limitações	Os resultados restringiram-se a um único contexto (um centro urbano chileno), a avaliação da compreensão leitora foi realizada apenas com provas não padronizadas, e a amostra pertence a um único nível socioeconômico.
Recomendações e Futuras Pesquisas	Sugere-se realizarem-se novos estudos que explorem outros contextos (como áreas rurais), com provas padronizadas e diferentes tipologias textuais, além de se incluírem alunos de diversos grupos socioeconômicos para uma visão mais ampla das dificuldades na leitura.

Fonte: elaboração própria.

A compreensão leitora de alunos do segundo ao oitavo ano de uma escola urbana no Chile foi avaliada, utilizando-se provas de múltipla escolha que abrangiam os níveis literal, inferencial e crítico. Os resultados revelaram uma diminuição no nível de compreensão leitora à medida que os alunos avançam nos anos escolares, sem diferenças significativas entre gêneros. A pesquisa destaca a necessidade de intervenções para melhorar a compreensão leitora, especialmente nos anos com desempenho mais fraco, e sugere a realização de novos estudos em diferentes contextos, com provas padronizadas e amostras mais diversificadas, para uma compreensão mais ampla das dificuldades de leitura.

5.2.32 Chilean Pre Service Teachers – Revista Profile – 2019

O Quadro 34 exibe partes do conteúdo do artigo de Saavedra-Jeldres e Campos-Espinoza (2019). O estudo abordou o uso de portfólios de escrita para desenvolver a produção de texto em inglês de 51 estudantes de Pedagogia no Chile, avaliando suas percepções sobre a melhoria em gramática, vocabulário, organização de parágrafos e habilidades de *feedback* por meio de um questionário de escala Likert.

Quadro 34 – Saavedra-Jeldres e Campos-Espinoza (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Pamela Andrea Saavedra-Jeldres e Mónica Campos-Espinoza, Universidad Católica de Temuco, Temuco, Chile (2019).
Tipo de Avaliação	Portfólios de escrita, que é uma estratégia para desenvolver a habilidade de produção de textos escritos em inglês como língua estrangeira. Os portfólios incluem a compilação de trabalhos escritos, versões melhoradas e reflexões sobre o processo de escrita.
População Avaliada	O público-alvo do estudo são estudantes de Pedagogia, especificamente 51 alunos do segundo semestre de um programa de formação de professores de inglês em uma universidade do Chile.
Competências Avaliadas	Produção escrita, reflexão sobre o processo de escrita e aspectos relacionados ao conhecimento linguístico, como gramática, vocabulário e organização de parágrafos.
Métodos e Instrumentos	O estudo utilizou um questionário de escala Likert para avaliar as percepções dos professores em formação sobre sua melhoria em várias áreas, incluindo-se gramática, vocabulário, organização de parágrafos e habilidades de <i>feedback</i> e correção. O questionário continha 22 afirmações que os alunos avaliaram em cinco categorias, como “nunca” ou “sempre”.
Resultados e Conclusões	Os resultados mostraram que os professores em formação valorizam a estratégia de portfólios e acreditam ter melhorado suas habilidades de produção escrita e reflexão.
Desafios e	Os principais desafios identificados incluem aspectos organizacionais, como

Limitações	limitação de tempo e número de palavras, além da pressão para atender às demandas externas ao escrever. Os alunos relataram dificuldades em fornecer <i>feedback</i> entre pares.
Recomendações e Futuras Pesquisas	Embora o estudo não forneça recomendações explícitas, a identificação de desafios sugere que futuras pesquisas poderiam explorar estratégias para mitigar a pressão sobre os alunos e melhorar a eficácia do <i>feedback</i> entre pares. Além disso, considerar as dificuldades organizacionais e logísticas pode ser uma área de foco para aprimorar a implementação de portfólios em contextos de ensino de inglês.

Fonte: elaboração própria.

Os resultados indicaram que os alunos valorizam a estratégia e percebem melhorias na escrita e na reflexão, apesar de desafios organizacionais, como limitação de tempo e dificuldades em fornecer *feedback* entre pares. A pesquisa sugere que futuras investigações devem explorar estratégias para mitigar a pressão sobre os alunos e aprimorar o *feedback* entre pares, além de considerar as dificuldades logísticas na implementação de portfólios no ensino de inglês.

5.2.33 *La competencia intercultural en la clase de francés lengua extranjera: ¿qué y cómo evaluarla?* – Revista Lengua y Habla – 2019

O Quadro 35 exibe partes do conteúdo do artigo de Machado Key (2019). O estudo explorou a avaliação da competência intercultural de futuros professores de francês e estudantes do Instituto Pedagógico de Caracas, utilizando entrevistas e questionários para analisar o desenvolvimento de conhecimento sociocultural, destrezas interculturais e traços de atitude.

Quadro 35 – Machado Key (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Yolibeth Machado Key, Universidad Pedagógica Experimental Libertador, Venezuela (2019).
Tipo de Avaliação	O estudo focaliza a avaliação da competência intercultural através de entrevistas com docentes e questionários para estudantes. Não são mencionados portfólios, testes de múltipla escolha ou avaliação por pares, mas sim a coleta de dados qualitativos para analisar as competências culturais desenvolvidas em sala de aula.
População Avaliada	O público-alvo da avaliação são futuros professores de francês, assim como os estudantes de níveis intermediários e avançados do Programa Francês do Instituto Pedagógico de Caracas.
Competências Avaliadas	As competências avaliadas incluem conhecimento sociocultural, destrezas interculturais e traços de atitude. Isso abrange aspectos como a vida cotidiana do estrangeiro, sensibilidade cultural, capacidade de mediação cultural e abertura para outras culturas.

Métodos e Instrumentos	O estudo utiliza um roteiro de entrevista e um questionário. Os principais tópicos dos exercícios incluem conhecimento sobre a vida cotidiana, valores e crenças da cultura estrangeira, bem como a capacidade de estabelecer relações entre culturas.
Resultados e Conclusões	Embora o estudo não forneça resultados específicos, espera-se que os dados coletados ajudem a identificar as necessidades de formação em competências culturais e a fortalecer os cursos oferecidos nos planos de estudo dos professores de francês.
Desafios e Limitações	Pode-se inferir que a falta de um enfoque sistemático do ensino da competência intercultural no contexto venezuelano poderia ser um desafio.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se desenharem novos descritores que se ajustem melhor ao perfil de um futuro professor de língua estrangeira, indicando que há espaço para melhorar a avaliação da competência intercultural na formação docente.

Fonte: elaboração própria.

A pesquisa, que se concentra na coleta de dados qualitativos, identificou a necessidade de fortalecer a formação em competências culturais nos planos de estudo, destacando-se a possível falta de um enfoque sistemático do ensino da interculturalidade no contexto venezuelano. A pesquisa sugere a importância de se desenvolverem novos descritores que melhor se adaptem ao perfil do futuro professor de língua estrangeira, visando a aprimorar a avaliação da competência intercultural na formação docente.

5.2.34 Consideraciones sobre la variación fonética en español y sus implicaciones para la evaluación internacional de la lengua como LE/L2: dele, siele y celu – Revista RLA – 2019

O Quadro 36 exibe partes do conteúdo do artigo de Andión Herrero e Criado de Diego (2019). O estudo tem como foco a avaliação das competências orais em exames de certificação internacional, como o DELE, SIELE e CELU. As provas incluem tanto a expressão oral quanto a interação, o que sugere um enfoque da importância da comunicação em contextos reais (em diversas situações).

Quadro 36 – Andión Herrero e Criado de Diego (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	María Antonieta Andión Herrero e Cecilia Criado de Diego, Universidad Nacional de Educación a Distancia, Espanha (2019).
Tipo de Avaliação	O estudo focaliza a avaliação das competências orais em exames de certificação internacional, como o DELE, SIELE e CELU. Menciona-se que as provas incluem tanto a expressão oral quanto a interação, o que sugere um enfoque da avaliação de habilidades comunicativas em contextos reais.

População Avaliada	O público-alvo são os candidatos que realizam exames de certificação de espanhol como língua estrangeira ou segunda língua, podendo incluir falantes de diversas variedades do espanhol.
Competências Avaliadas	As principais competências avaliadas são as habilidades orais, incluindo a expressão e interação oral. Isso implica que os candidatos devem demonstrar sua capacidade de se comunicar efetivamente em espanhol em situações conversacionais.
Métodos e Instrumentos	Os exames analisados incluem provas orais, que podem consistir em entrevistas e tarefas de interação. Menciona-se que as provas são subjetivas e requerem a avaliação de examinadores credenciados.
Resultados e Conclusões	O estudo conclui que é fundamental considerar a variação fonética do espanhol e as atitudes dos avaliadores em relação a essa variedade para garantir uma avaliação justa e válida. Destaca-se a necessidade de formação dos avaliadores em relação à diversidade linguística do espanhol.
Desafios e Limitações	Um dos desafios mencionados é a falta de formação específica dos avaliadores quanto à variação do espanhol, o que pode introduzir vieses na avaliação. Além disso, aponta-se um possível desequilíbrio na representação das variedades do espanhol nas provas.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se desenvolverem iniciativas para valorizar a diversidade linguística na avaliação e se aprimorar a formação dos avaliadores. Recomenda-se que os exames reflitam melhor a plurinormatividade do espanhol e sejam consideradas as crenças e atitudes dos avaliadores em relação às variedades.

Fonte: elaboração própria.

O público-alvo compreende aqueles que realizam exames de certificação em língua espanhola como língua estrangeira. As provas requerem a avaliação de examinadores credenciados e conclui-se que é importante se considerarem a variação fonética do espanhol e a postura dos avaliadores quanto a isso, para que a avaliação seja justa. É considerado essencial que os examinadores estudem a respeito da variação fonética, valorizando, assim, a diversidade linguística.

5.2.35 Desempeños en la calidad de explicaciones y argumentaciones en estudiantes chilenos de 8° básico – Revista Signos – 2019

O Quadro 37 exibe partes do conteúdo do artigo de Figueroa, Meneses e Chandía (2019).

Quadro 37 – Figueroa, Meneses e Chandía (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Javiera Figueroa, Pontificia Universidad Católica de Chile, Chile; Alejandra Meneses, Pontificia Universidad Católica de Chile, Chile; e Eugenio Chandía Universidad de Chile, Chile (2019).
Tipo de Avaliação	O estudo utiliza uma rubrica para medir a qualidade da escrita dos alunos em diferentes tarefas, adaptada do instrumento da Avaliação Nacional do Progresso

	Educativo (NAEP).
População Avaliada	O público-alvo avaliado no estudo são estudantes chilenos do oitavo ano.
Competências Avaliadas	As competências avaliadas incluem a qualidade das explicações e argumentações dos alunos, focalizando dimensões como conhecimento do propósito, desenvolvimento das ideias e organização discursiva.
Métodos e Instrumentos	Os testes consistem em tarefas de escrita, que são avaliadas com base na rubrica mencionada. As questões são abertas, permitindo que os alunos desenvolvam suas respostas em forma de textos argumentativos e explicativos.
Resultados e Conclusões	Os resultados indicam que os alunos apresentam um desempenho básico nas tarefas de escrita, com variações significativas dependendo da tarefa, do sexo dos estudantes e do grupo socioeconômico a que pertencem.
Desafios e Limitações	Os principais desafios do estudo incluem a alta demanda cognitiva e linguística enfrentada pelos alunos, além de limitações relacionadas à consistência interna da avaliação.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de melhorias nas práticas de ensino e avaliação da escrita, além de recomendações para futuras pesquisas que possam explorar mais a fundo as diferenças de desempenho entre os grupos socioeconômicos e as estratégias de ensino que podem ser mais eficazes.

Fonte: elaboração própria.

Através de uma rubrica, o estudo tem como foco medir a qualidade da escrita dos alunos chilenos do oitavo ano em diferentes tarefas. A avaliação foi adaptada do instrumento da Avaliação Nacional do Progresso Educativo (NAEP).

Dentre as competências avaliadas, a qualidade das explicações e argumentações dos alunos são analisadas, focalizando dimensões como conhecimento do propósito, desenvolvimento das ideias e organização discursiva. As questões são abertas, possibilitando que os alunos desenvolvam suas respostas em forma de textos argumentativos e explicativos.

Os estudos mostram resultados que indicam desempenho básico nas tarefas escritas, com variações a depender da atividade avaliada em questão, o gênero e o grupo socioeconômico dos estudantes.

5.2.36 Complejidad textual, vocabulario pasivo y comprensión lectora en escolares chilenos de segundo ciclo básico – Revista Literatura y Lingüística – 2019

O Quadro 38 exibe partes do conteúdo do artigo de Novoa Lagos (2019).

Quadro 38 – Novoa Lagos (2019)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Abraham Novoa Lagos, Universidad de Concepción, Concepción, Chile (2019).
Tipo de Avaliação	Foi utilizada uma avaliação por meio de testes de compreensão leitora, que

	incluíam perguntas de múltipla escolha com resposta única. Foram elaboradas duas versões do mesmo teste com diferentes níveis de complexidade textual.
População Avaliada	O público-alvo do estudo são 66 alunos do sexto ano do ensino fundamental, com idades entre 11 e 12 anos, que pertencem a uma escola particular subvencionada em Chiguayante, Chile.
Competências Avaliadas	As competências avaliadas incluem habilidades de leitura, especificamente as capacidades de localizar, relacionar e interpretar informações, além de refletir sobre o texto.
Métodos e Instrumentos	Os testes consistem em perguntas de múltipla escolha. Os principais tópicos dos exercícios incluem a compreensão de textos narrativos literários, por meio dos quais os alunos devem responder a perguntas relacionadas à leitura.
Resultados e Conclusões	Os resultados indicam que a complexidade intrínseca dos textos afeta a compreensão global da leitura. No entanto, a compreensão do vocabulário passivo não parece influenciar significativamente a compreensão global, embora possa ser relevante em grupos com diferentes habilidades de leitura.
Desafios e Limitações	Um dos desafios mencionados é a variabilidade dos leitores.
Recomendações e Futuras Pesquisas	O estudo sugere que a compreensão do vocabulário passivo possa ter um papel mais significativo em grupos com diferenças nas habilidades de leitura. Recomenda-se investigar mais a fundo essa variável em estudos futuros.

Fonte: elaboração própria.

Testes de compreensão leitora foram aplicados e incluíam perguntas de múltipla escolha com resposta única. Duas versões do mesmo teste foram criadas, com diferentes níveis de complexidade textual. O público-alvo são 66 estudantes chilenos de uma escola particular, todos com idade entre 11 e 12 anos. Quanto à habilidade de leitura, as competências especificamente avaliadas foram as de localizar, relacionar e interpretar informações. Conclui-se que a complexidade do texto afeta a compreensão global da leitura. A variabilidade dos leitores é considerável.

5.2.37 A Post-Positivist and Interpretive Approach to Researching Teachers' Language Assessment Literacy – Revista Profile – 2020

O Quadro 39 exibe partes do conteúdo do artigo de Giraldo (2020).

Quadro 39 – Giraldo (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Frank Giraldo, Universidad de Caldas, Manizales, Colombia (2020).
Tipo de Avaliação	O estudo discute a utilização de diversas abordagens de avaliação, incluindo-se a avaliação alternativa, mas não se limita a um único tipo. Observações em sala de aula e questionários são mencionados como métodos para entender como os professores aplicam suas competências em avaliação.

População Avaliada	O público-alvo do estudo são os professores de línguas, especificamente aqueles que ensinam inglês como língua estrangeira.
Competências Avaliadas	O letramento em avaliação de línguas, que abrange a capacidade dos professores de projetar, implementar e interpretar avaliações de forma eficaz. Isso pode incluir habilidades em leitura, escrita e fala, dependendo do contexto de ensino.
Métodos e Instrumentos	O estudo menciona que os testes podem incluir questões de múltipla escolha e questões abertas, mas não fornece detalhes específicos sobre os tópicos dos exercícios. No entanto, discute a importância de se alinharem as avaliações com o currículo escolar e as diretrizes nacionais.
Resultados e Conclusões	O estudo conclui que a adoção de abordagens qualitativas complementares às quantitativas pode gerar dados mais abrangentes e válidos sobre o letramento em avaliação dos professores. O letramento em avaliação de línguas é crucial para o desenvolvimento profissional dos professores de línguas, pois influencia diretamente a qualidade das avaliações que eles aplicam em sala de aula.
Desafios e Limitações	Os principais desafios incluem a superficialidade das respostas em questionários e a falta de clareza em termos técnicos, que podem dificultar a compreensão dos professores sobre os conceitos de avaliação. Além disso, o letramento em avaliação dos professores é uma área subexplorada.
Recomendações e Futuras Pesquisas	É recomendável que futuras pesquisas adotem abordagens qualitativas para complementar os dados quantitativos, a fim de explorar mais profundamente as complexidades do letramento em avaliação. Além disso, é sugerido que mais estudos sejam realizados sobre o impacto das iniciativas de desenvolvimento profissional no processo de letramento em avaliação dos professores.

Fonte: elaboração própria.

O estudo investigou o letramento em avaliação de professores de inglês como língua estrangeira, explorando como eles aplicam suas competências em avaliação por meio de observações em sala de aula e questionários. Os resultados indicam que a adoção de abordagens qualitativas complementares às quantitativas pode fornecer dados mais abrangentes sobre o letramento em avaliação dos professores, que é crucial para a qualidade das avaliações em sala de aula. Os principais desafios incluem a superficialidade das respostas em questionários e a falta de clareza em termos técnicos. Recomenda-se que futuras pesquisas adotem abordagens qualitativas e explorem o impacto de iniciativas de desenvolvimento profissional no letramento em avaliação dos professores.

5.2.38 Teachers' Assessment Approaches Regarding EFL Students' Speaking Skill – Revista Profile – 2020

O Quadro 40 exibe partes do conteúdo do artigo de Duque-Aguilar (2020). A pesquisa indica que os professores preferem práticas de avaliação somativa, com entrevistas que avaliem

a habilidade de fala dos estudantes. Métodos alternativos de avaliação são mencionados, como diários, vídeos e autoavaliação.

Quadro 40 – Duque-Aguilar (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Jaime Fernando Duque-Aguilar, Universidad de la Amazonia, Florencia, Colômbia (2020)
Tipo de Avaliação	O estudo destaca a preferência dos professores por práticas de avaliação somativa, como entrevistas estruturadas ou semiestruturadas, para avaliar a habilidade de fala dos alunos. Além disso, menciona a utilização de métodos alternativos e formativos, como diários, gravações em vídeo e autoavaliação.
População Avaliada	O público-alvo do estudo são alunos de inglês como língua estrangeira de um programa universitário de formação de professores na Colômbia.
Competências Avaliadas	As principais competências avaliadas no estudo são as habilidades de produção oral, que incluem aspectos como vocabulário, pronúncia e a habilidade de transmitir ideias.
Métodos e Instrumentos	Os testes utilizados no estudo são predominantemente de avaliação somativa, que podem incluir questões abertas e interações orais. Os tópicos abordados nos exercícios geralmente envolvem a capacidade dos alunos de se expressar oralmente e demonstrar domínio do conteúdo do curso.
Resultados e Conclusões	Os resultados revelaram que os professores tendem a se concentrar em práticas de avaliação somativa e que há uma necessidade de reflexão sobre suas abordagens de ensino e avaliação. O estudo conclui que a conscientização sobre as práticas de avaliação de fala pode ser um primeiro passo importante para melhorar as práticas de avaliação na universidade e mudar as percepções de professores e alunos sobre a avaliação oral.
Desafios e Limitações	O número limitado de participantes (quatro professores) e a subjetividade das opiniões e experiências dos professores em relação à avaliação da fala podem não refletir a totalidade das práticas de avaliação em contextos mais amplos.
Recomendações e Futuras Pesquisas	O estudo sugere que os professores reflitam sobre suas práticas de avaliação e considerem a implementação de métodos alternativos e formativos. Além disso, recomenda-se a realização de estudos futuros que incluam uma amostra maior de professores e que explorem mais profundamente a relação entre as abordagens de avaliação e o desempenho dos alunos.

Fonte: elaboração própria.

O estudo tem como público-alvo estudantes de língua inglesa de um programa universitário da Colômbia. Vocabulário, pronúncia e a capacidade de transmitir ideias são os principais elementos avaliados na produção oral. Um número limitado de professores participou, e a subjetividade em suas experiências e opiniões em relação à avaliação da fala é um ponto a ser considerado. Conclui-se que métodos alternativos devem ser considerados, bem como uma amostra com maior número de professores em um possível futuro estudo.

O Quadro 41 exibe partes do conteúdo do artigo de Cárcamo (2020).

Quadro 41 – Cárcamo (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Benjamín Cárcamo, Universidad de las Américas, Viña del Mar, Chile (2020).
Tipo de Avaliação	O estudo não especifica um tipo de avaliação particular, mas discute a importância da retroalimentação corretiva escrita como um recurso fundamental na avaliação para a aprendizagem, sugerindo que a entrega oportuna de <i>feedback</i> é crucial para o processo de aprendizagem.
População Avaliada	O público-alvo avaliado no estudo são estudantes de línguas, especificamente aqueles que estão aprendendo uma segunda língua e que se beneficiam da retroalimentação corretiva escrita em suas produções textuais.
Competências Avaliadas	A principal competência avaliada no estudo é a escrita, com foco na precisão gramatical e na qualidade da escrita dos alunos.
Métodos e Instrumentos	O estudo não detalha os testes específicos utilizados, mas menciona a utilização de diferentes tipos de <i>feedback</i> corretivo, que podem incluir correções diretas e indiretas, além de <i>feedback</i> focalizado em forma e função. Os tópicos dos exercícios podem abranger aspectos gramaticais e estruturais da escrita.
Resultados e Conclusões	Os resultados e conclusões do estudo indicam que a falta de sistematicidade na correção escrita tem dificultado a comparação entre estudos.
Desafios e Limitações	Os principais desafios e limitações do estudo incluem a diversidade de definições e a falta de consenso sobre os tipos de retroalimentação corretiva escrita, o que pode levar a resultados inconsistentes em pesquisas.
Recomendações e Futuras Pesquisas	A pesquisa oferece recomendações para estudos futuros, sugerindo que a tipologia proposta pode ser utilizada para sistematizar e comparar diferentes abordagens de correção, além de incentivar os educadores a expandirem suas práticas de <i>feedback</i> .

Fonte: elaboração própria.

O estudo investiga a importância da retroalimentação corretiva escrita na avaliação da aprendizagem de línguas, focalizando estudantes de segunda língua e suas produções textuais. Embora não especifique um tipo de avaliação, o estudo destaca a relevância do *feedback* oportuno para a precisão gramatical e qualidade da escrita. A pesquisa aponta desafios como a falta de consenso na definição de tipos de *feedback* e a necessidade de sistematização para comparação entre estudos, e sugere que a tipologia proposta pode auxiliar na organização das abordagens de correção e na expansão das práticas de *feedback* pelos educadores.

5.2.40 *Level of English in Colombian Higher Education: A Decade of Stagnation – Revista Profile – 2020*

O Quadro 42 exibe partes do conteúdo do artigo de Benavides (2020).

Quadro 42 – Benavides (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Jorge Eliecer Benavides, Universidad de Nariño, Pasto, Colômbia (2020).
Tipo de Avaliação	O estudo utiliza os resultados de testes padronizados aplicados a estudantes de ensino superior na Colômbia. Esses exames avaliam diferentes competências, incluindo a proficiência em inglês.
População Avaliada	O público-alvo do estudo são os estudantes universitários colombianos que estão prestando os exames Saber Pro, especificamente aqueles que estão finalizando seus cursos de graduação.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a proficiência em inglês, que abrange os níveis de leitura, escrita, fala e compreensão oral, conforme os critérios do Quadro Europeu Comum de Referência para Línguas.
Métodos e Instrumentos	Os testes consistem em questões de múltipla escolha que avaliam a proficiência em inglês. Os tópicos dos exercícios incluem compreensão de leitura, gramática e vocabulário, mas o estudo não detalha os tópicos específicos.
Resultados e Conclusões	Os resultados indicam que, ao longo de uma década, não houve melhorias significativas nos níveis de proficiência em inglês entre os estudantes universitários colombianos. A maioria dos alunos continua a apresentar níveis baixos de proficiência (A- e A1), com apenas uma pequena porcentagem alcançando níveis mais altos (B+). O estudo conclui que é necessário revisar as políticas educacionais e implementar alternativas estruturais para melhorar a formação docente e a metodologia de ensino.
Desafios e Limitações	Os principais desafios incluem a falta de progresso significativo nos níveis de proficiência em inglês e as barreiras estruturais no sistema educacional (como o grande número de alunos por sala de aula), que dificultam a implementação eficaz das políticas de ensino de inglês.
Recomendações e Futuras Pesquisas	O estudo recomenda uma revisão das políticas educacionais por parte das universidades, com foco em alternativas estruturais que promovam a capacitação docente, a melhoria das metodologias de ensino e a organização curricular. Além disso, sugere que futuras pesquisas explorem mais a fundo as causas da estagnação nos níveis de proficiência em inglês.

Fonte: elaboração própria.

A pesquisa analisa a proficiência em inglês de estudantes universitários colombianos, utilizando os resultados dos exames Saber Pro, que avaliam diversas competências, incluindo o inglês. Os resultados revelam que, ao longo de uma década, não houve melhoria significativa nos níveis de proficiência, com a maioria dos alunos permanecendo em níveis básicos. O estudo aponta para a necessidade de revisão das políticas educacionais e implementação de alternativas estruturais para melhorar a formação docente e a metodologia de ensino, destacando desafios como o grande número de alunos por sala de aula. Recomenda-se que futuras pesquisas explorem as causas da estagnação na proficiência em inglês.

5.2.41 Evaluación de la producción escrita de estudiantes secundarios haitianos por parte de profesores de Lengua y Literatura en Chile – Revista Estudios Filológicos – 2020

O Quadro 43 exibe partes do conteúdo do artigo de Lizasoain e Toledo Veja (2020). O estudo analisa a produção escrita em espanhol de alunos haitianos aprendendo a língua e falantes nativos (principalmente chilenos e latino-americanos), avaliada tanto por professores das escolas quanto por avaliadores externos.

Quadro 43 – Lizasoain e Toledo Veja (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Andrea Lizasoain, Facultad de Filosofía y Humanidades, Universidad Austral de Chile, Chile, e Gloria Toledo Veja, Facultad de Letras, Pontificia Universidad Católica de Chile, Chile (2020).
Tipo de Avaliação	O estudo utiliza a avaliação de produções escritas, por meio da qual os textos são coletados e avaliados tanto por professores de língua que trabalham nas mesmas escolas dos alunos quanto por avaliadores externos. A avaliação se concentra em aspectos da escrita em espanhol dos alunos haitianos e falantes de espanhol.
População Avaliada	O público-alvo do estudo são estudantes haitianos que estão aprendendo espanhol como língua estrangeira, além de estudantes falantes de espanhol (principalmente chilenos e de outros países da América Latina).
Competências Avaliadas	As principais competências avaliadas incluem a produção escrita em espanhol, que é um aspecto crucial para o desenvolvimento da competência comunicativa dos alunos.
Métodos e Instrumentos	Os testes consistem em produções escritas, principalmente de respostas abertas na disciplina de Língua e Literatura. Os tópicos abordados nos exercícios não foram detalhados, mas envolvem a avaliação da escrita em espanhol dos alunos.
Resultados e Conclusões	Os resultados indicam que os alunos receberam notas semelhantes de seus professores de língua, mas os avaliadores externos aplicaram um esquema que resultou em avaliações contraditórias. Isso sugere que os princípios de avaliação não foram atendidos, destacando a necessidade de preparar melhor os professores para fornecer <i>feedback</i> útil que contribua para o desempenho escolar.
Desafios e Limitações	Um dos principais desafios mencionados é a falta de formação adequada dos professores em avaliação, o que pode levar a uma avaliação arbitrária e subjetiva. Além disso, a avaliação não é considerada uma parte integral do processo.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de preparar melhor os professores de língua para que possam oferecer <i>feedback</i> mais útil e eficaz aos alunos, especialmente no contexto de ensino de espanhol como língua estrangeira. Além disso, enfatiza a importância de refletir sobre as práticas avaliativas e como elas podem ser aprimoradas para contribuir efetivamente para o aprendizado dos alunos, especialmente em situações multiculturais.

Fonte: elaboração própria.

As competências focalizadas incluem a escrita como elemento central do desenvolvimento comunicativo, com testes baseados em respostas abertas em Língua e Literatura. Os resultados revelam discrepâncias nas avaliações: enquanto os professores atribuíram notas semelhantes, os avaliadores externos aplicaram critérios que geraram contradições, indicando falhas nos princípios avaliativos. O estudo identifica como desafio a falta de preparo dos professores em avaliação, levando a critérios subjetivos, e a desconexão

entre avaliação e processo de aprendizagem. Conclui-se que é essencial capacitar os docentes para fornecer *feedback* mais eficaz, além de repensar as práticas avaliativas, especialmente em contextos multiculturais, a fim de melhorar o desempenho dos alunos.

5.2.42 *La influencia del componente no verbal asociado al feedback correctivo en la producción oral en ELE – Revista Signos – 2020*

O Quadro 44 exhibe partes do conteúdo do artigo de Muñoz Moreno e Mavrou (2020).

Quadro 44 – Muñoz Moreno e Mavrou (2020)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Elvira Muñoz Moreno, Leiden University Centre For Linguistics (Lucl), Países Baixos, e Irini Mavrou, Universidad Antonio de Nebrija, Espanha (2020).
Tipo de Avaliação	O estudo não menciona especificamente um tipo de avaliação formal. Em vez disso, a pesquisa se concentra na análise da interação oral e na retroalimentação corretiva durante as aulas, utilizando gravações audiovisuais para observar a dinâmica entre o professor e os alunos.
População Avaliada	O público-alvo do estudo é composto de 15 estudantes de diferentes nacionalidades, com idades entre 18 e 22 anos, que estão matriculados em um curso de espanhol como língua estrangeira na University College de Utrecht, Holanda. A maioria dos alunos tem o inglês como segunda língua.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a produção oral dos alunos, especificamente em relação à interação durante as atividades de classe. O foco está na correção linguística e na fluência na língua espanhola.
Métodos e Instrumentos	O estudo não utiliza testes formais, mas analisa atividades de interação oral que incluem conversas, descrições de fotos e discussões sobre temas culturais. As atividades foram projetadas para promover a interação e a prática da língua, com foco na correção de erros linguísticos.
Resultados e Conclusões	O estudo conclui que a combinação de técnicas de foco na forma ajuda a direcionar a atenção dos alunos para a forma linguística sem desviar o foco do significado. A técnica mais utilizada foi a elicitación, com ênfase na correção linguística, o que pode estar relacionado ao nível de proficiência dos alunos.
Desafios e Limitações	Um dos desafios mencionados é que os resultados se baseiam exclusivamente na análise da atuação do professor, sem considerar a reflexão docente ou a interpretação dos alunos. Isso limita a compreensão completa da dinâmica de aprendizagem e interação.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas devem incluir a reflexão dos professores e a interpretação dos alunos para obter uma visão mais abrangente da interação em sala de aula. Além disso, recomenda-se explorar mais a fundo o impacto do componente não verbal na aprendizagem de línguas.

Fonte: elaboração própria.

O estudo analisa a interação oral e a retroalimentação corretiva em aulas de espanhol como língua estrangeira, com foco em 15 estudantes universitários de diversas nacionalidades. Através de gravações audiovisuais, a pesquisa avalia a produção oral dos alunos, especialmente

em atividades de interação, como conversas e discussões culturais, buscando entender como as técnicas de foco na forma auxiliam na correção linguística e fluência. O estudo conclui que a elicitación é uma técnica eficaz para direcionar a atenção dos alunos para a forma linguística sem desviar do significado, mas aponta a limitação de se basear exclusivamente na atuação do professor. Recomenda-se que pesquisas futuras incluam a reflexão de professores e alunos, além de explorar o impacto da comunicação não verbal na aprendizagem de línguas.

5.2.43 A Reflection on Initiatives for Teachers' Professional Development Through Language Assessment Literacy – Revista Profile – 2021

O Quadro 45 exibe partes do conteúdo do artigo de Giraldo (2021).

Quadro 45 – Giraldo (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Frank Giraldo, Universidad de Caldas, Manizales, Colômbia (2021).
Tipo de Avaliação	O estudo discute várias formas de avaliação, incluindo-se as avaliações formativa e somativa, além de práticas como a pesquisa-ação colaborativa e <i>workshops</i> que abordam o desenvolvimento de instrumentos de avaliação, como testes e rubricas.
População Avaliada	O público-alvo avaliado no estudo são professores de línguas que estão buscando desenvolver suas competências em avaliação de linguagem e melhorar sua prática de ensino.
Competências Avaliadas	As principais competências avaliadas incluem a habilidade comunicativa, que abrange leitura, escrita, fala e compreensão oral, além de conceitos-chave como validade e confiabilidade na avaliação.
Métodos e Instrumentos	Os testes discutidos no estudo incluem tanto questões de múltipla escolha quanto questões abertas. Os principais tópicos dos exercícios desses testes envolvem a criação de itens de teste para leitura, gramática, vocabulário e escrita, além da análise de itens de teste existentes.
Resultados e Conclusões	Os resultados do estudo indicam que os participantes desenvolveram uma maior consciência crítica em relação aos itens de teste de larga escala, aprenderam a criar questões para compreensão de leitura e se tornaram mais conscientes da importância de tarefas de vocabulário para o ensino e avaliação. As conclusões sugerem que a formação em avaliação de linguagem é crucial para a prática docente. As iniciativas de formação em avaliação de linguagem não apenas ensinaram aspectos específicos, como validade e confiabilidade, mas também ajudaram os professores a se tornarem mais críticos em relação às suas práticas e sistemas de avaliação. Os professores expressam uma necessidade maior de treinamento prático em avaliação de linguagem, focalizando aspectos práticos em vez de reflexivos.
Desafios e Limitações	Os principais desafios e limitações do estudo incluem a necessidade de mais formação prática em avaliação de linguagem e a falta de foco em tópicos éticos e de justiça na avaliação, que são frequentemente negligenciados pelos

	professores.
Recomendações e Futuras Pesquisas	A pesquisa oferece recomendações para que futuros estudos incluam uma maior ênfase em ética e justiça na avaliação, além de sugerir que programas de formação continuada sejam desenvolvidos para atender às necessidades práticas dos professores em relação à avaliação de linguagem.

Fonte: elaboração própria.

Diversas formas de avaliação são discutidas, como a formativa e somativa, além de práticas, como *workshops*, que trabalham o desenvolvimento de instrumentos avaliativos. Professores de línguas que buscam aperfeiçoar seus conhecimentos e sua prática de ensino e de avaliação são o público-alvo. A habilidade comunicativa é avaliada, bem como conceitos como confiabilidade e validade na avaliação. Dentre os testes estudados e abordados na pesquisa, estão incluídas questões de múltipla escolha e questões abertas. Os principais tópicos dos exercícios dos testes mencionados envolvem a criação de itens de teste para leitura, gramática, vocabulário e escrita, além da análise de itens de teste existentes. Os participantes do estudo desenvolveram maior consciência crítica em relação aos instrumentos avaliativos, mas há necessidade de mais exemplos práticos.

5.2.44 Assessment of Students' Oral Communicative Competence in English through a Web Conferencing Platform – Revista Profile – 2021

Avaliar a competência comunicativa oral dos alunos no ambiente *on-line* não é uma prática comum na instituição onde o estudo de Cruz-Ramos e Herrera Díaz (2021), em partes exibido no Quadro 46, foi realizado. No entanto, em relação ao primeiro objetivo da pesquisa, a avaliação *on-line* se mostrou viável, principalmente devido ao uso de uma plataforma de videoconferência que os alunos consideraram fácil de utilizar, além do suporte de um instrutor experiente. Assim, foi demonstrado que a avaliação *on-line* do desempenho oral dos alunos não só é válida, como também eficaz, desde que seja planejada com cuidado.

Os alunos que participaram da avaliação oral *on-line* usando a ferramenta Zoom não encontraram grandes problemas e conseguiram se desempenhar como fariam em um ambiente presencial. Uma das principais vantagens de usar essa plataforma de videoconferência, em vez de fazer com que os alunos se encontrassem pessoalmente com o instrutor dentro de um tempo limitado, foi tornar o processo de avaliação mais compatível com a modalidade *on-line* do curso. A presença de um segundo instrutor ou assistente pode ajudar nos aspectos técnicos e na transmissão das instruções, melhorando a qualidade da experiência para os alunos. O assistente

estava sempre disponível para responder a perguntas adicionais que poderiam não ser abordadas pelo instrutor principal durante a avaliação oral na plataforma de videoconferência.

Sobre o segundo objetivo da pesquisa, o estudo também demonstrou que a competência comunicativa dos alunos melhorou de forma consistente desde o início até o fim da implementação, e essa melhoria foi resultado das mudanças feitas no *design* do curso de Inglês II. Isso ressalta a importância de se alinharem recursos, atividades, suporte e avaliação ao se trabalhar com um curso *on-line*.

Vale destacar que o estudo apresentou algumas limitações, como amostra pequena e ausência de análise estatística independente para cada aspecto da competência comunicativa.

Quadro 46 – Cruz-Ramos e Herrera-Díaz (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	María de los Milagros Cruz-Ramos, Escuela Normal “Juan Enriquez”, Veracruz, México, e Luz Edith Herrera-Díaz, Universidad Veracruzana, Veracruz, México (2021).
Tipo de Avaliação	Avaliação <i>on-line</i> de língua inglesa com foco na competência comunicativa. Os testes comunicativos diferem de outros testes de língua por pretenderem prever como os estudantes se saem em uma situação comunicativa real.
População Avaliada	Alunos universitários de uma universidade mexicana.
Competências Avaliadas	Competência comunicativa.
Métodos e Instrumentos	O teste é dividido em dois momentos: no primeiro, o professor interage com o aluno; no segundo, são formados duplas/trios para que haja interação (mesmo para aqueles que não tiveram momentos de prática), baseado na sugestão de tema de uma carta.
Resultados e Conclusões	As avaliações praticadas devem ser diversas e similares às atividades às quais os alunos foram expostos ao longo do curso. Logo, interações síncronas <i>on-line</i> no curso podem ser aliadas a interações <i>on-line</i> na avaliação, para que alunos desenvolvam as habilidades e competências necessárias para a comunicação e a negociação de sentidos em língua inglesa. A avaliação da competência comunicativa deve considerar a forma da língua e seu uso. A avaliação <i>on-line</i> é viável, desde que com instrutores experientes e em uma plataforma com a qual os alunos estejam familiarizados. Com as adaptações feitas no curso, os alunos que utilizaram o Zoom para a avaliação oral não enfrentaram desafios, tendo em vista que a prática do curso era semelhante ao momento da avaliação. A participação de um segundo instrutor pode ser benéfica para o momento da avaliação.
Desafios e Limitações	Os cursos <i>on-line</i> são criticados na instituição em questão. Poucas oportunidades de interação entre alunos e professores e poucas oportunidades de exposição ao idioma, para que, de fato, aprendam. A avaliação segue os mesmos moldes, quer o aluno estude na modalidade <i>on-line</i> , quer na presencial, ainda que algumas modalidades ofereçam mais oportunidades de prática e interação.

	Os testes orais são os mais desafiadores para os alunos, inclusive para os matriculados na modalidade presencial. Espera-se que os alunos que fizeram o curso <i>on-line</i> também realizem uma avaliação da habilidade de produção oral, mesmo que não tenham recebido qualquer tipo de prática ao longo do curso. Ademais, os testes de avaliação oral acontecem presencialmente inclusive para os alunos da modalidade <i>on-line</i> (desconsiderando a natureza virtual do curso em questão). A faculdade parece não considerar a situação. Avaliar os alunos em critérios de (apenas) forma e uso seria simplificar a natureza de interações comunicativas. Os alunos demonstraram dificuldades com o Simple Past. A quantidade reduzida de amostras para o estudo pode configurar uma possível limitação.
Recomendações e Futuras Pesquisas	Considerar, em um novo formato, não só forma e uso da língua, mas uma avaliação mais ampla. Os aspectos tecnológicos e pedagógicos do curso podem ser melhorados, visando a mais oportunidades para que os alunos desenvolvam suas competências comunicativas. Os avaliadores apenas veem um fragmento do que os alunos são capazes de produzir em um recorte limitado de tempo. As avaliações devem ter relação com situações de uso real do aluno. Expandir o estudo com grupos maiores de estudantes <i>on-line</i>. Um momento extra de avaliação foi adicionado aos dois antes comumente praticados, totalizando em três. O Zoom, utilizado na avaliação <i>on-line</i>, agora inclui também um professor assistente.

Fonte: elaboração própria.

5.2.45 Self and Peer Assessment of Student Generated Podcasts – Revista Profile – 2021

O Quadro 47 exibe partes do conteúdo do artigo de Cadena-Aguilar e Álvarez-Ayure (2021).

Quadro 47 – Cadena-Aguilar e Álvarez-Ayure (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Albedro Cadena-Aguilar e Claudia Patricia Álvarez-Ayure, Universidad de La Sabana, Chía, Colômbia (2021).
Tipo de Avaliação	O estudo aborda a avaliação formativa, especificamente através de autoavaliação e avaliação por pares, por meio das quais os alunos refletem sobre seu próprio desempenho e o de seus colegas em relação à produção de <i>podcasts</i> .
População Avaliada	O público-alvo do estudo são estudantes de graduação em um contexto de ensino de inglês como língua estrangeira.
Competências Avaliadas	Produção oral, com foco em algumas questões, como ênfase de frases e entonação.
Métodos e Instrumentos	O estudo não menciona testes tradicionais, como múltipla escolha ou questões abertas, mas um questionário que coleta percepções sobre a autoavaliação e a avaliação por pares, focalizando a habilidade oral.
Resultados e Conclusões	Os resultados indicam que a prática de autoavaliação e avaliação por pares ajudou os alunos a melhorarem sua compreensibilidade ao gravar <i>podcasts</i> . Os

	participantes mostraram evolução, de uma participação limitada para um envolvimento ativo nas práticas de avaliação, o que sugere uma mudança positiva nas atitudes em relação à autoavaliação. A pesquisa enfatizou a necessidade de um treinamento adequado para os alunos em relação às práticas de autoavaliação, sugerindo que a familiarização com os instrumentos de avaliação e os critérios é crucial para o sucesso dessas práticas. O uso de avaliação formativa, em vez de apenas avaliações somativas, foi considerado benéfico para o desenvolvimento das habilidades de fala dos alunos, permitindo um aprendizado mais significativo e incentivando-os a refletir criticamente sobre sua própria produção oral.
Desafios e Limitações	O estudo menciona que alguns participantes estavam inicialmente apreensivos em relação à autoavaliação e esperavam principalmente a avaliação do professor, o que pode ter limitado a eficácia das práticas de avaliação por pares no início.
Recomendações e Futuras Pesquisas	Embora o estudo não forneça recomendações específicas, a ênfase na necessidade de treinamento adequado para os alunos em relação às práticas de autoavaliação sugere que futuras pesquisas podem explorar métodos para melhorar a formação dos alunos nesse aspecto, além de investigar mais a fundo a eficácia de diferentes abordagens de avaliação formativa.

Fonte: elaboração própria.

O estudo investiga o impacto da autoavaliação e avaliação por pares na produção oral de *podcasts* por estudantes de inglês como língua estrangeira, utilizando um questionário para coletar percepções sobre essas práticas. Os resultados indicam que essas abordagens formativas melhoraram a compreensibilidade dos alunos e promoveram uma mudança positiva em suas atitudes em relação à autoavaliação. A pesquisa destaca a necessidade de treinamento adequado para os alunos em relação aos critérios de avaliação e sugere que o uso de avaliação formativa, em vez de apenas somativa, beneficia o desenvolvimento das habilidades de fala. Embora alguns alunos estivessem inicialmente apreensivos, a prática mostrou-se eficaz, indicando a importância de se explorarem métodos para aprimorar a formação em autoavaliação e se investigarem diferentes abordagens de avaliação formativa em futuras pesquisas.

5.2.46 Estructura hipertextual y habilidad lectora: Análisis de la producción digital de textos argumentativos en estudiantes de 4º Año Medio – Revista Signos – 2021

O Quadro 48 exibe partes do conteúdo do artigo de Vallejos Carrasco, Véliz de Vos e Sáez Carrillo (2021). A pesquisa investiga a relação entre o uso de hipertextos como recurso de leitura e a qualidade da produção de textos argumentativos em estudantes do quarto ano do ensino médio no Chile.

Quadro 48 – Vallejos Carrasco, Véliz de Vos e Sáez Carrillo (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Gabriela Vallejos Carrasco, Mónica Véliz de Vos e Katia Sáez Carrillo, Universidad de Concepción, Chile (2021).
Tipo de Avaliação	O estudo utiliza uma avaliação que inclui a aplicação de uma prova de compreensão leitora (Lectum 7) e um teste de memória operativa (Test de Dígitos Inversos). Além disso, foi aplicada uma rubrica de avaliação adaptada para a produção de textos argumentativos, que foi validada estatisticamente.
População Avaliada	O público-alvo do estudo são estudantes do quarto ano médio (décimo segundo ano) de um estabelecimento educacional municipal na Região do Biobío, no Chile. A amostra inicial consistiu em 131 estudantes, com idades entre 17 e 19 anos.
Competências Avaliadas	As principais competências avaliadas no estudo incluem a habilidade de leitura e a produção escrita, especificamente a capacidade de escrever textos argumentativos.
Métodos e Instrumentos	Os testes aplicados no estudo incluem uma versão resumida da prova de compreensão leitora (Lectum 7), que consiste em dois textos e um teste de memória operativa. A prova de compreensão leitora é utilizada para selecionar os sujeitos com pontuações extremas, enquanto o teste de memória operativa envolve a repetição de séries de números.
Resultados e Conclusões	O estudo conclui que há uma relação entre a estrutura hipertextual utilizada como recurso de leitura e a qualidade da produção escrita argumentativa dos estudantes. Os resultados indicam que a habilidade leitora e a estrutura do hipertexto influenciam positivamente a produção textual.
Desafios e Limitações	Os principais desafios e limitações do estudo incluem a seleção da amostra e a dependência de autoavaliações dos estudantes sobre suas habilidades de leitura e escrita. Além disso, a pesquisa pode ter limitações relacionadas à generalização dos resultados para outras populações.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas podem explorar mais a fundo a relação entre diferentes tipos de hipertexto e a produção textual, além de considerar a inclusão de variáveis adicionais que possam influenciar a habilidade leitora e a escrita dos estudantes.

Fonte: elaboração própria.

Através de uma avaliação que combina provas de compreensão leitora, testes de memória operativa e uma rubrica validada para textos argumentativos, a pesquisa conclui que a habilidade de leitura e a estrutura do hipertexto influenciam positivamente a produção textual dos alunos. O estudo destaca a importância da estrutura hipertextual no processo de leitura e escrita, mas reconhece limitações na seleção da amostra e na dependência de autoavaliações. Recomenda-se que pesquisas futuras explorem diferentes tipos de hipertexto e incluam variáveis adicionais para uma compreensão mais abrangente do tema.

5.2.47 Disciplinary Power Lying Behind the Requisite of English Language Mastery in International Scholarships – Revista Profile – 2021

O Quadro 49 exibe partes do conteúdo do artigo de Forero-Mondragón e Quintero-Polo (2021).

Quadro 49 – Forero-Mondragón e Quintero-Polo (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Alber Josué Forero-Mondragón e Álvaro Hernán Quintero-Polo, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia (2021).
Tipo de Avaliação	Testes de proficiência em inglês.
População Avaliada	Candidatos a bolsas de estudo internacionais, que precisam demonstrar proficiência em inglês para atender aos requisitos das chamadas de bolsas.
Competências Avaliadas	Habilidades de leitura, escrita, fala e compreensão oral, conforme exigido pelos testes de certificação.
Métodos e Instrumentos	Os testes consistem em avaliações padronizadas que podem incluir questões de múltipla escolha e questões abertas, focalizando tópicos como gramática, vocabulário, compreensão de texto e produção escrita.
Resultados e Conclusões	O estudo conclui que o discurso do inglês padrão está intimamente entrelaçado com os discursos de globalização, qualidade na educação e competitividade. Essa conexão posiciona a escola como um espaço que utiliza técnicas disciplinares para preparar os indivíduos para atender às exigências do mercado de trabalho global, especialmente as demandas de corporações multinacionais.
Desafios e Limitações	O estudo enfrenta desafios relacionados à complexidade das interações sociais e políticas que influenciam a avaliação da proficiência em inglês, além de limitações na generalização dos resultados para contextos fora das bolsas de estudo analisadas.
Recomendações e Futuras Pesquisas	Embora o estudo não forneça recomendações explícitas, a análise sugere a necessidade de uma reflexão crítica sobre as práticas de avaliação e de inclusão de vozes diversas nos processos de certificação de proficiência em inglês, o que pode ser uma área para investigações futuras.

Fonte: elaboração própria.

O foco do estudo recai na análise de testes de proficiência em inglês aplicados a candidatos a bolsas de estudo internacionais, avaliando habilidades de leitura, escrita, fala e compreensão oral. A pesquisa conclui que o discurso do inglês padrão está intrinsecamente ligado a conceitos de globalização, qualidade educacional e competitividade, posicionando a escola como um espaço de preparação para o mercado de trabalho global. O estudo destaca a complexidade das interações sociais e políticas na avaliação da proficiência e sugere a necessidade de uma reflexão crítica sobre as práticas de avaliação, incentivando a inclusão de diversas perspectivas nos processos de certificação.

5.2.48 *The Impact of Assessment Training – Revista Profile – 2021*

O Quadro 50 exibe partes do conteúdo do artigo de González (2021). O estudo examina o impacto da formação em avaliação nas práticas de professores de inglês como língua

estrangeira em universidades mexicanas que utilizam portfólios e provas mensais com componentes de escrita, além de rubricas analíticas.

Quadro 50 – González (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Elsa Fernanda González, Universidad Autónoma de Tamaulipas, Ciudad Victoria, Mexico (2021).
Tipo de Avaliação	O estudo menciona a utilização de portfólios e exames mensais que incluem componentes de escrita. Os professores também refletem sobre a utilização de rubricas analíticas para a avaliação da escrita.
População Avaliada	O público-alvo do estudo são os alunos de instituições de ensino superior no México que estão aprendendo inglês como língua estrangeira.
Competências Avaliadas	Os professores são incentivados a avaliar as quatro habilidades linguísticas: leitura, escrita, fala e compreensão oral, mas a ênfase do estudo está na escrita.
Métodos e Instrumentos	Os testes consistem em avaliações que incluem componentes de escrita, como exames mensais e trabalhos de portfólio. Os tópicos dos exercícios não são detalhados especificamente, mas os professores mencionam a importância de refletir sobre a avaliação e a utilização de portfólios para estimular o desenvolvimento e a reflexão dos alunos.
Resultados e Conclusões	Os resultados indicam que a formação em avaliação teve um impacto positivo nas práticas de avaliação dos professores, levando a uma maior reflexão sobre suas metodologias de ensino e avaliação. Os professores relataram sentir mais confiança em suas avaliações e mais capacidade de explicar as notas aos alunos. A formação ajudou a melhorar a clareza e a eficácia das práticas de avaliação. Os professores demonstraram um aumento na reflexão sobre suas práticas de ensino e avaliação. Esse autoconhecimento levou a uma maior consciência sobre a necessidade de se fornecer <i>feedback</i> mais construtivo. Os professores começaram a implementar mais atividades de escrita em suas aulas e a utilizar rubricas de forma mais sistemática. Isso resultou em um foco maior nas necessidades dos alunos e na qualidade do <i>feedback</i> fornecido. O estudo sugere que a formação em avaliação deve ser vista como um processo contínuo, em vez de uma oportunidade única.
Desafios e Limitações	O estudo identificou que alguns participantes não tinham experiência anterior em formação de avaliação, o que pode ter limitado a profundidade da mudança em suas práticas.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas podem explorar mais a fundo a eficácia de diferentes métodos de formação em avaliação e como esses métodos podem ser adaptados para atender às necessidades específicas dos professores. Além disso, recomenda-se a inclusão de uma amostra mais diversificada de participantes para se obter uma visão mais abrangente das práticas de avaliação em diferentes contextos.

Fonte: elaboração própria.

Os resultados indicam que a formação impulsionou a reflexão dos educadores sobre suas metodologias, aumentando a confiança e a clareza nas avaliações, com maior foco no *feedback* construtivo e nas necessidades dos alunos. A pesquisa destaca a necessidade de formação contínua em avaliação e sugere que futuras investigações explorem diferentes métodos de formação e incluam amostras diversificadas para uma análise mais abrangente.

5.2.49 *Assessment practices in some language centers in public Universities of Ecuador and student proficiency outcomes – Revista Literatura y Linguística – 2021*

O Quadro 51 exibe partes do conteúdo do artigo de Almeida, Castilho e Rosero (2021).

Quadro 51 – Almeida, Castillo e Rosero (2021)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Evelyn Almeida, Universidad Central del Ecuador, Universidad de las Fuerzas Armadas ESPE. Quito, Equador; Diana Castillo, Universidad Central del Ecuador. Quito, Equador; e Andrea Rosero, Universidad Central de Ecuador. Quito, Equador (2021).
Tipo de Avaliação	O estudo utiliza uma combinação de avaliações, incluindo testes padronizados, avaliações escritas e apresentações orais. As avaliações são tanto formativas quanto somativas, com ênfase em testes de múltipla escolha e questões abertas.
População Avaliada	Alunos que estão cursando o sexto nível de inglês como língua estrangeira nos centros de idiomas das universidades analisadas, que é equivalente ao nível B1 do Quadro Europeu Comum.
Competências Avaliadas	Leitura, escrita, compreensão oral e fala.
Métodos e Instrumentos	Os testes consistem em uma combinação de questões de múltipla escolha e questões abertas. Os tópicos dos exercícios incluem gramática, vocabulário, compreensão de leitura e produção escrita, com base nos conteúdos dos livros utilizados nos cursos.
Resultados e Conclusões	O estudo conclui que existe uma conexão entre o desempenho dos alunos e os sistemas de avaliação aplicados. Alunos que estudam em centros que utilizam um sistema de avaliação mais estruturado tendem a obter pontuações mais altas. Além disso, a padronização das avaliações contribui para uma melhor compreensão do progresso dos alunos. Conclui-se também que o uso de diferentes tipos de avaliação proporciona uma visão mais abrangente das competências dos alunos, permitindo que os professores identifiquem áreas que necessitam de mais atenção.
Desafios e Limitações	Os principais desafios incluem a variabilidade nas práticas de avaliação entre diferentes professores e a falta de parâmetros claros para as avaliações orais.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de uma maior padronização nas práticas de avaliação e a implementação de um sistema de avaliação mais estruturado. Recomenda-se também a formação contínua dos professores em métodos de avaliação e a criação de um banco de testes para garantir a consistência e a objetividade nas avaliações.

Fonte: elaboração própria.

O presente estudo une e analisa diferentes tipos de avaliação, como testes padronizados, avaliações escritas e apresentações orais (tanto formativas como somativas). O público-alvo da pesquisa são alunos de cursos língua inglesa, equivalente ao nível B1. Tais cursos são ministrados em centros de idiomas de universidades. As quatro habilidades são trabalhadas nos instrumentos avaliativos (produção oral e escrita, compreensão oral e escrita) e envolvem

questões abertas e de múltipla escolha. Conclui-se que centros de idiomas com um sistema de avaliação mais estruturado contribuem para o desempenho dos alunos e que a diversidade de instrumentos também pode oferecer um panorama mais amplo do desenvolvimento dos estudantes. No entanto, dentre os desafios identificados, notou-se que distintos professores adotaram avaliações diferentes e que os critérios para a avaliação oral não foram bem estruturados. Por fim, o estudo ressalta a importância de o sistema avaliativo dos centros de idiomas seguir uma prática padronizada e os professores terem formação contínua nas práticas avaliativas.

5.2.50 The Factors Contributing to Language Teachers' Effectiveness in an EFL Learning Context: A Questionnaire Validation Study – Revista Profile – 2022

O Quadro 52 exibe partes do conteúdo do artigo de Nayernia, Nosrati e Mohebbi (2022).

Quadro 52 – Nayernia, Nosrati e Mohebbi (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Akram Nayernia e Rana Nosrati, Iran University of Science and Technology, Tehran, Irã, e Hassan Mohebbi, European Knowledge Development Institute (EUROKD), Ankara, Turquia (2022).
Tipo de Avaliação	O estudo utiliza um questionário baseado em uma escala Likert para avaliar a eficiência dos professores de idiomas. O questionário foi desenvolvido com base em uma revisão da literatura e em entrevistas com especialistas e inclui itens que abordam diferentes aspectos da eficácia docente, como conhecimento de conteúdo, pedagogia e experiência.
População Avaliada	O público-alvo do estudo são professores de inglês como língua estrangeira (EFL) que atuam em instituições de ensino, tanto em universidades quanto em institutos de idiomas privados. A amostra incluiu 93 professores de EFL com diferentes níveis de experiência.
Competências Avaliadas	As principais competências avaliadas no estudo incluem: - conhecimento de conteúdo e pedagogia; - proficiência oral; - autoeficácia; - criatividade na apresentação de aulas; e - uso apropriado da língua nativa (L1), quando necessário.
Métodos e Instrumentos	O teste consiste em um questionário com itens de múltipla escolha, possibilitando aos participantes avaliar a sua própria eficácia em várias áreas, como comunicação em inglês, avaliação justa dos alunos, preparação para perguntas e uso de técnicas de ensino variadas. Os tópicos principais incluem a comunicação, a criatividade, a autoeficácia e a adaptação ao contexto de ensino.
Resultados e Conclusões	Os resultados indicaram que os fatores como conhecimento de conteúdo, experiência e autoeficácia são cruciais para a eficácia dos professores de EFL. O estudo concluiu que a eficácia docente pode ser melhorada através do desenvolvimento contínuo dessas competências e da reflexão sobre a prática pedagógica. A autoeficácia dos professores foi destacada como um fator significativo que influencia não apenas a percepção que os professores têm de sua própria eficácia,

	mas também o impacto que têm sobre o aprendizado dos alunos. Professores que se sentem mais confiantes tendem a ser mais eficazes em suas práticas de ensino.
Desafios e Limitações	Entre os desafios e as limitações do estudo, destacam-se a amostra relativamente pequena de professores e a possibilidade de viés nas autoavaliações. Além disso, a pesquisa foi realizada em um contexto específico (Irã), o que pode limitar a generalização dos resultados para outros contextos educacionais.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas podem incluir uma amostra mais diversificada de professores de diferentes contextos e níveis de ensino. Além disso, recomenda-se a exploração de métodos de avaliação alternativos, como observações em sala de aula e <i>feedback</i> de alunos, para complementar as autoavaliações e obter uma visão mais abrangente da eficácia docente.

Fonte: elaboração própria.

A eficácia dos professores de idiomas foi avaliada através de um questionário baseado em escala Likert. O questionário foi construído com base na revisão de entrevistas com especialistas e revisão da literatura e focaliza tópicos como pedagogia, conhecimento de conteúdo, proficiência oral, criatividade nas aulas e experiência. Os professores avaliados são docentes de língua inglesa (inglês como língua estrangeira) que atuam em universidades e em institutos de idiomas. São 93 profissionais com distintos níveis de experiência.

O teste consiste em questões de múltipla escolha com que os participantes avaliam a própria eficácia e conhecimento. O conhecimento que os professores têm sobre a própria prática traz resultados positivos para a carreira e para os estudantes. O estudo representa, no entanto, uma amostra relativamente pequena, de 93 alunos, em um contexto único, o Irã.

5.2.51 Mapudungun: cuestionario para evaluar competencia oral en mapudungun en niños mapuche – Revista Literatura y Lingüística – 2022

O Quadro 53 exibe partes do conteúdo do artigo de Boudon *et al.* (2022). A pesquisa em questão utiliza um instrumento avaliativo focalizado na competência oral em mapudungun.

Quadro 53 – Boudon *et al.* (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Paula Alonqueo Boudon, Doctora en Psicología, Universidad de La Frontera, Temuco, Chile; Rodrigo Becerra Parra, Magíster en Lingüística. Doctorando Universidad de Alberta, Alberta, Canadá; Jacqueline Caniguán Caniguán, Magíster en Lingüística. Universidad de La Frontera, Temuco, Chile; Nataly Huenchunao Huenchunao, Magíster en Psicología, Colegio América, Vilcún, Chile; e Eloy Oliva Vásquez, Magíster en Psicología, Universidad Santo Tomás, sede Puerto Montt, Puerto Montt, Chile (2022).
Tipo de Avaliação	O estudo utiliza um questionário como instrumento de avaliação, que é projetado para medir a competência oral em mapudungun. A avaliação é baseada em

	respostas dos participantes, evitando critérios de desempenho padrão e normativo da língua.
População Avaliada	O público-alvo do estudo são crianças mapuches, especificamente aquelas escolarizadas em estabelecimentos rurais de Padre Las Casas, com idades entre 6 e 12 anos.
Competências Avaliadas	As principais competências avaliadas incluem a produção e compreensão oral em mapudungun, abrangendo seções como léxico, saudações, termos espaciais, temporais e numéricos, além de narrativas.
Métodos e Instrumentos	Os testes consistem em uma combinação de questões de múltipla escolha e questões abertas e abordam tópicos como léxico, saudações e termos espaciais, temporais e numéricos, com diferentes níveis de dificuldade para cada item.
Resultados e Conclusões	Os resultados indicaram que os níveis de desempenho das crianças em mapudungun foram baixos. O questionário demonstrou ser uma ferramenta confiável para avaliar a competência linguística, mas também revelou a necessidade de ajustes para melhorar a validade e a clareza dos itens.
Desafios e Limitações	Os desafios incluíram a complexidade de alguns itens, que foram considerados descontextualizados e difíceis para as crianças. Além disso, a carga cognitiva excessiva ao avaliar compreensão e produção de forma independente foi uma limitação identificada. Outra limitação foi o uso de amostra pequena, o que pode comprometer a generalização dos resultados. A administração do questionário a uma amostra maior foi considerada inviável, devido ao tempo necessário para sua aplicação.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de revisar e modificar itens que apresentaram complexidade excessiva. Também recomenda que futuras pesquisas considerem a validade e a sensibilidade cultural, ao desenvolver instrumentos de avaliação para populações indígenas.

Fonte: elaboração própria.

As respostas dos participantes guiam o estudo e não padrões normativos da língua. As crianças mapuches são o público-alvo, entre 6 e 12 anos de idade, especificamente as que frequentam escolas rurais. Foram avaliadas através de questões abertas e de múltipla escolha, principalmente quanto à compreensão e à produção em mapudungun, incluindo seções como léxico, saudações, termos numéricos e narrativas. O teste revelou o baixo desempenho dos alunos em mapudungun. Apesar de os itens do instrumento avaliativo terem-se mostrado adequados para avaliar as questões linguísticas, o teste ainda precisa de ajustes quanto à clareza e ao nível de complexidade das questões.

5.2.52 English Language Teachers Perceived Classroom – Revista Profile – 2022

O Quadro 54 exhibe partes do conteúdo do artigo de Tajeddin, Saeedi e Panahzadeh (2022).

Quadro 54 – Tajeddin, Saeedi e Panahzadeh (2022)

Categoria	Descrição
------------------	------------------

Perfil (Autor, Instituição e Ano)	Zia Tajeddin, Tarbiat Modares University, Teerã, Irã; e Zari Saeedi e Vahid Panahzadeh, Allameh Tabataba'i University, Teerã, Irã (2022).
Tipo de Avaliação	Testes (como <i>quizzes</i>) para identificar pontos fortes e de melhoria dos alunos e planejar intervenções futuras. Além disso, o estudo aborda a prática de <i>feedback</i> construtivo e a autoavaliação dos alunos.
População Avaliada	O público-alvo do estudo são professores de inglês como língua estrangeira no Irã (especificamente 348 professores).
Competências Avaliadas	Leitura, escrita e fala.
Métodos e Instrumentos	Questões de múltipla escolha e questões abertas. Os tópicos principais incluem a formatação dos itens de teste (por exemplo, múltipla escolha ou dissertativa) e a participação dos alunos na discussão sobre mudanças curriculares com base nos resultados dos testes.
Resultados e Conclusões	Os resultados indicam que a maioria dos professores se considera moderadamente letrada em avaliação e expressa a necessidade de cursos específicos sobre avaliação em suas formações. O estudo conclui que há uma lacuna na formação de professores em relação à avaliação em sala de aula, sugerindo a inclusão de um componente prático de avaliação nas formações.
Desafios e Limitações	Um dos principais desafios identificados foi a falta de cursos focalizados em avaliação que preparem os professores com as habilidades necessárias. Além disso, a pesquisa revelou que muitos professores não se sentem adequadamente preparados para realizar avaliações eficazes, o que limita a interpretação dos resultados.
Recomendações e Futuras Pesquisas	O estudo recomenda que mais tempo e foco sejam dedicados à avaliação em cursos de formação de professores, tanto na formação inicial quanto na formação continuada. Além disso, sugere a necessidade de mais pesquisas em contextos locais para entender melhor as dinâmicas do letramento em avaliação entre professores de línguas.

Fonte: elaboração própria.

O estudo investigou o letramento em avaliação de 348 professores de inglês como língua estrangeira no Irã, utilizando testes (*quizzes*) para identificar pontos fortes e de melhoria dos alunos em leitura, escrita e fala, além de abordar *feedback* construtivo e autoavaliação. Os resultados indicam que a maioria dos professores se considera moderadamente letrada em avaliação, mas expressa a necessidade de formação específica na área, revelando uma lacuna na sua preparação e sugerindo a inclusão de um componente prático de avaliação nos cursos de formação inicial e continuada, além de mais pesquisas em contextos locais.

5.2.53 El desarrollo del conocimiento de los componentes y funciones del lenguaje escrito en estudiantes prelectores – Revista Signos – 2022

O Quadro 55 exhibe partes do conteúdo do artigo de Martínez Giménez e Sellés Nohales

(2022).

Quadro 55 – Martínez Giménez e Sellés Nohales (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Tomás Martínez Giménez, Universidad de Valencia, Espanha, e Pilar Sellés Nohales, Universidad Católica de Valencia, Espanha (2022).
Tipo de Avaliação	O estudo utiliza uma avaliação observacional-descritiva com testes padronizados, especificamente a prova BIL 3-6, que avalia as habilidades pré-leitoras e o conhecimento metalinguístico de crianças de 3 a 6 anos.
População Avaliada	O público-alvo do estudo são 344 estudantes de 3 a 6 anos com desenvolvimento psicomotor típico (exceto na área de linguagem), que têm como língua veicular o espanhol.
Competências Avaliadas	A habilidade de leitura, o reconhecimento de palavras e frases e as funções da leitura.
Métodos e Instrumentos	Os testes consistem em avaliações individuais em que os estudantes devem identificar frases e palavras, além de reconhecer funções da leitura. Os exercícios incluem itens como identificar se uma sequência de palavras forma uma frase ou não, com uma pontuação máxima atribuída a cada item.
Resultados e Conclusões	Os resultados indicam que existem diferenças significativas nas habilidades avaliadas entre os grupos de idade, permitindo determinar uma sequência de aquisição das habilidades. As crianças de 3 anos já conhecem algumas funções da leitura; aos 6 anos, reconhecem praticamente todas.
Desafios e Limitações	Um dos desafios mencionados é a variabilidade no desenvolvimento das crianças, que pode afetar a aquisição das habilidades. Além disso, a pesquisa não incluiu centros privados, o que pode limitar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de mais pesquisas que estabeleçam um padrão de desenvolvimento normativo nas habilidades pré-leitoras, especialmente para identificar e estimular fatores críticos em crianças que apresentam déficits.

Fonte: elaboração própria.

A pesquisa observacional-descritiva avaliou as habilidades pré-leitoras e o conhecimento metalinguístico de 344 crianças de 3 a 6 anos com desenvolvimento psicomotor típico que falam espanhol. Foi utilizada a prova padronizada BIL 3-6 contendo testes individuais de reconhecimento de palavras, frases e funções da leitura. Os resultados revelaram diferenças significativas nas habilidades entre as faixas etárias, estabelecendo uma sequência de aquisição, embora a variabilidade no desenvolvimento infantil e a exclusão de centros privados sejam desafios. O estudo sugere mais pesquisas para normatizar o desenvolvimento das habilidades pré-leitoras e identificar fatores críticos para crianças com déficits.

5.2.54 Evaluación del lenguaje oral en niños y niñas con hipoacusia: los tests estandarizados y la edad auditiva – Revista Signos – 2022

O Quadro 56 exibe partes do conteúdo do artigo de Taboh *et al.* (2022).

Quadro 56 – Taboh *et al.* (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Analí R. Taboh e Diego E. Shalom, Universidad de Buenos Aires, IFIBA, CONICET, Argentina; Valeria B. Bosco e Percival J. Denham, Instituto Oral Modelo, Argentina; e Carolina A. Gattei, Universidad Torcuato Di Tella, Argentina (2022).
Tipo de Avaliação	O estudo utiliza testes padronizados para avaliar o desenvolvimento linguístico das crianças, como o Teste de Vocabulário em Imagens Peabody (TVIP), e outros testes, como CELF, TACL-R e RDLS, que seguem pautas preestabelecidas para administração, pontuação e interpretação.
População Avaliada	O público-alvo do estudo são crianças e adolescentes com hipoacusia (NHA), que são comparados com crianças ouvintes (NDT) em termos de desenvolvimento linguístico.
Competências Avaliadas	As principais competências avaliadas incluem vocabulário, compreensão e produção oral. O estudo mede o desenvolvimento linguístico em comparação com o esperado para a idade cronológica.
Métodos e Instrumentos	Os testes consistem em avaliações que podem incluir questões de múltipla escolha, como no caso do TVIP, no qual a criança deve indicar qual imagem corresponde à palavra apresentada oralmente. Os tópicos dos exercícios envolvem vocabulário e compreensão oral.
Resultados e Conclusões	O estudo conclui que é importante complementar o uso de testes padronizados com a análise de pontuações baseadas na idade auditiva, pois isso pode fornecer uma visão mais precisa das competências linguísticas das crianças com hipoacusia. Isso é crucial para o desenvolvimento de programas educacionais que ajudem a reduzir a lacuna entre o desenvolvimento linguístico dessas crianças e o esperado para sua idade.
Desafios e Limitações	O estudo menciona que um dos desafios é a defasagem entre a idade cronológica das crianças com hipoacusia e sua experiência auditiva, o que pode dificultar a avaliação precisa de suas habilidades linguísticas. Além disso, a dependência de testes padronizados pode não capturar completamente as nuances do desenvolvimento linguístico em NHA.
Recomendações e Futuras Pesquisas	O estudo recomenda que futuras pesquisas considerem a utilização de medidas independentes da idade cronológica, como a idade auditiva, para uma avaliação mais precisa das competências linguísticas. Além disso, sugere que a combinação de diferentes métodos de avaliação pode enriquecer a compreensão do desenvolvimento linguístico em crianças com hipoacusia.

Fonte: elaboração própria.

A pesquisa em questão faz uso de testes padronizados, através dos quais avalia crianças quanto ao desenvolvimento linguístico. Alguns dos testes utilizados são o Teste de Vocabulário em Imagens Peabody e o CELF, seguindo pautas preestabelecidas de pontuação, administração e interpretação. Crianças e adolescentes com hipoacusia, bem como crianças ouvintes, são o público-alvo do estudo. O desenvolvimento linguístico delas foi foco do estudo, sobretudo em termos de vocabulário, compreensão e produção oral. Interessou saber como está o desenvolvimento quanto ao desenvolvimento cronológico esperado. Dentre as questões nos

instrumentos avaliativos, uma requeria que o aluno escolhesse, dentre as opções, a imagem que correspondesse à palavra dita oralmente. Dentre os desafios apresentados pelo estudo está a defasagem entre a idade cronológica das crianças com hipoacusia e sua experiência auditiva. Ademais, a dependência de testes padronizados pode não capturar completamente as nuances do desenvolvimento linguístico em crianças com hipoacusia.

5.2.55 Estrategias y desafíos en la creación de un instrumento de evaluación lingüística infantil
– Revista Signos – 2022

O Quadro 57 exhibe partes do conteúdo do artigo de Iacobuzio *et al.* (2022).

Quadro 57 – Iacobuzio *et al.* (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Sofía Victoria Iacobuzio, María Luisa Silva, María Victoria Gasparini e Yamila Rubbo, Centro de Investigación Interdisciplinario en Psicología y Matemática Experimental (CIIPME-CONICET), Universidad de Buenos Aires, Argentina (2022).
Tipo de Avaliação	O estudo utiliza uma avaliação que envolve a produção discursiva infantil, focalizando narrativas e recontos com base em estímulos visuais e linguísticos. Não se trata de um teste tradicional, mas de uma atividade projetada para avaliar a produção oral das crianças em um formato interativo.
População Avaliada	O público-alvo do estudo são crianças de 5 a 14 anos que apresentam dificuldades na atribuição de estados mentais (AMS).
Competências Avaliadas	As principais competências avaliadas incluem a produção de narrativas, a capacidade de recontar histórias e a utilização de verbos psicológicos, que são essenciais para a compreensão e expressão de eventos psicológicos.
Métodos e Instrumentos	Os testes consistem em atividades de narração com base em imagens e recontos de histórias. As crianças são convidadas a produzir relatos orais que são então analisados qualitativamente. Os tópicos principais incluem a estrutura narrativa e os eventos psicológicos nas narrativas.
Resultados e Conclusões	Os resultados indicam que o processo de <i>design</i> do estímulo foi otimizado através de revisões e <i>feedback</i> dos participantes, resultando em uma ferramenta de avaliação que é dinâmica e capaz de captar aspectos complexos da produção discursiva infantil. As conclusões destacam a importância de um <i>design</i> flexível que se adapte às necessidades dos avaliados.
Desafios e Limitações	Os principais desafios incluíram a necessidade de se criar um material que fosse atraente e acessível para as crianças, além de garantir que as condições de elicitación fossem adequadas. As limitações podem estar relacionadas à amostra pequena e à dependência de <i>feedback</i> qualitativo, que pode ser subjetivo.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas devam continuar a refinar os instrumentos de avaliação, incorporando <i>feedback</i> contínuo dos participantes e considerando a diversidade das experiências das crianças. A flexibilidade no <i>design</i> e a adaptação às necessidades dos avaliados são enfatizadas como essenciais para a eficácia das avaliações.

Fonte: elaboração própria.

A pesquisa avalia a produção oral de crianças de forma não tradicional, através de interações e recontos em que são utilizados estímulos visuais e linguísticos. Crianças de 5 a 14 anos são o público-alvo, com foco em crianças que apresentam dificuldades na atribuição de estados mentais. Dentre os principais focos de avaliação podemos citar a capacidade de recontar histórias, a produção de narrativas e a capacidade de utilização de verbos psicológicos. As crianças foram estimuladas a contar histórias oralmente, e as narrativas envolveram eventos psicológicos. O instrumento avaliativo – dinâmico – foi adaptado às necessidades dos avaliados. O *feedback* dos participantes foi essencial para que o instrumento avaliativo fosse o mais adequado e atraente possível para o público em questão.

5.2.56 Competencia comunicativa en educación superior: hacia una evaluación diagnóstica integral – Revista Logos – 2022

O Quadro 58 exibe partes do conteúdo do artigo de Valdés-Léon, González Riffo e Molina Olivares (2022). O presente estudo avalia a produção escrita com base em critérios do Ministério da Educação chileno.

Quadro 58 – Valdés-Léon, González Riffo e Molina Olivares (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Gabriel Valdés-Léon, Universidad de Las Palmas de Gran Canaria, Universidad Bernardo O'Higgins, Chile; Javier González Riffo, Universidad Católica Silva Henríquez, Chile; e Mario Molina Olivares, Universidad Católica Silva Henríquez, Chile (2022).
Tipo de Avaliação	O estudo utiliza uma avaliação diagnóstica que inclui um teste de disponibilidade léxica, uma adaptação da prova de compreensão e uma rubrica para avaliar a produção escrita. A avaliação é realizada com base em critérios do Ministério da Educação chileno.
População Avaliada	O público-alvo do estudo são 39 estudantes do primeiro ano do curso de Pedagogia de uma universidade privada chilena.
Competências Avaliadas	As principais competências avaliadas no estudo são: - competência léxica; - compreensão leitora; e - produção escrita.
Métodos e Instrumentos	As avaliações consistem em: - um teste de disponibilidade léxica, que os estudantes têm dois minutos para concluir; - uma adaptação da prova de compreensão leitora; e - uma avaliação da produção escrita utilizando uma rubrica.
Resultados e Conclusões	Os resultados indicam que os estudantes apresentaram um melhor desempenho na compreensão leitora, seguida pela produção escrita e, por último, na competência léxica. As conclusões sugerem a necessidade de intervenções

	pedagógicas para melhorar a competência léxica dos alunos.
Desafios e Limitações	O estudo enfrentou desafios relacionados à amostra, já que dois estudantes não puderam participar por não atenderem a todos os critérios de inclusão. Além disso, a pesquisa se limita a uma única instituição, o que pode afetar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se desenvolverem estratégias pedagógicas que abordem especificamente a competência léxica e recomenda a realização de pesquisas em diferentes contextos educacionais para validar os achados.

Fonte: elaboração própria.

Trinta e nove estudantes do primeiro ano de um curso de Pedagogia de uma universidade privada chilena foram avaliados sobretudo quanto à competência léxica, compreensão leitora e produção escrita. Concluiu-se que os estudantes tiveram maior desempenho na compreensão leitora em detrimento das outras. A amostra de estudantes foi pequena, pelo fato de representar uma única instituição e nem todos os estudantes terem tido a chance de participar.

5.2.57 Desempeño escrito de los estudiantes de secundaria en Chile: una evaluación directa e indirecta de la escritura – Revista Logos – 2022

O Quadro 59 exibe partes do conteúdo do artigo de Valenzuela Muñoz (2022).

Quadro 59 – Valenzuela Muñoz (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Ángel Valenzuela Muñoz, Universidad de Talca, Chile (2022).
Tipo de Avaliação	O estudo utiliza uma avaliação direta da escrita, em que os alunos redigem um ensaio argumentativo, e uma avaliação indireta, que consiste em exercícios de produção de textos, como o uso de conectores e a ordem das frases.
População Avaliada	O público-alvo avaliado no estudo são estudantes de quarto ano de ensino médio (último ano da educação secundária) no Chile, com uma amostra de 143 alunos.
Competências Avaliadas	As principais competências avaliadas no estudo são relacionadas à escrita, especificamente a capacidade de redigir um ensaio argumentativo, que inclui dimensões como ortografia, coerência, coesão e uso de argumentos.
Métodos e Instrumentos	Os testes consistem em uma avaliação direta, na qual os alunos escrevem um ensaio argumentativo, e uma avaliação indireta, que envolve exercícios de produção textual. Os principais tópicos dos exercícios incluem o uso de conectores e a organização das frases.
Resultados e Conclusões	Os resultados mostraram que os alunos apresentaram um desempenho médio tanto na avaliação direta quanto na indireta. As dimensões de ortografia, coerência, coesão e uso de argumentos foram as de menor desempenho. Além disso, não foi encontrada correlação entre o desempenho dos estudantes nas duas avaliações.
Desafios e Limitações	O estudo reconhece que a pesquisa sobre produção escrita não abrangeu de maneira uniforme os diferentes contextos de aprendizagem, o que pode ser uma

	limitação. Além disso, a amostra foi homogênea em termos socioeconômicos, o que pode limitar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo sugere que a avaliação direta da escrita pode servir como fator motivacional para o desenvolvimento de atividades que melhorem a produção escrita dos alunos. Recomenda-se também a realização de mais pesquisas que considerem diferentes contextos de aprendizagem e que explorem mais a relação entre as avaliações diretas e indiretas.

Fonte: elaboração própria.

O estudo em questão avalia a escrita de alunos através de um ensaio argumentativo, além de exercícios de produção de textos, podendo ter como foco o uso de conectores e a ordem das frases, assim como coerência e coesão do texto. Estudantes chilenos do quarto ano são o público-alvo da pesquisa, que tem 143 alunos compondo o público total. Por fim, recomenda-se que sejam realizadas pesquisas abrangendo diversos contextos de aprendizagem.

5.2.58 EFL In-Service Teachers' Self-Reported Perspectives on English Language Assessment – Nueva Revista Del Pacífico – 2022

O Quadro 60 exibe partes do conteúdo do artigo de Díaz Larenas *et al.* (2022).

Quadro 60 – Díaz Larenas *et al.* (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Claudio Díaz Larenas, Universidad de Concepción, Chile; Tania Tagle Ochoa Universidad Católica de Temuco, Chile; Mabel Ortiz Navarrete, Universidad Católica de Concepción, Chile; e Juan Fernando Gómez Paniagua, Tecnológico de Antioquia, Institución Universitaria, Colômbia (2022).
Tipo de Avaliação	O estudo menciona que os professores utilizam principalmente testes escritos e orais, além de trabalhos individuais e em grupo. Os exames e questionários são usados como estratégias somativas, e os professores também consideram o uso de rubricas como mecanismo de justiça e objetividade na avaliação.
População Avaliada	O público-alvo do estudo são os alunos de inglês como língua estrangeira (EFL), discentes dos professores em serviço participantes da pesquisa.
Competências Avaliadas	As principais competências avaliadas incluem leitura, escrita e fala, uma vez que os professores utilizam diferentes tipos de testes para avaliar o aprendizado dos alunos nessas áreas.
Métodos e Instrumentos	Os testes consistem em avaliações escritas e orais, com ênfase em questões de múltipla escolha e questões abertas. Os tópicos dos exercícios geralmente abrangem o conteúdo ensinado em sala de aula, embora o estudo indique que os contextos dos alunos não são sempre considerados durante a avaliação.
Resultados e Conclusões	Os resultados indicam que os professores percebem o monitoramento do aprendizado como o principal propósito da avaliação e valorizam a avaliação diagnóstica. As práticas mais frequentes estão relacionadas ao <i>design</i> e à retroalimentação dos instrumentos de avaliação, e os professores se sentem mais competentes nessas áreas.

Desafios e Limitações	O estudo é de natureza quantitativa e não experimental, o que pode limitar a profundidade das análises qualitativas. Além disso, a amostra é composta principalmente de professores do sexo feminino, o que pode não representar a diversidade total da população de professores de inglês no Chile.
Recomendações e Futuras Pesquisas	O estudo discute implicações para a formação de professores em serviço, sugerindo que é necessário um foco maior no desenvolvimento de habilidades de análise e comunicação dos resultados de avaliação, que foram identificadas como áreas de menor competência entre os participantes. Além disso, recomenda-se a consideração de contextos e trajetórias dos alunos nas práticas de avaliação.

Fonte: elaboração própria.

No estudo em questão, os professores fazem uso de testes escritos e orais, bem como trabalhos. As rubricas também são consideradas, para garantir a justiça e objetividade da experiência avaliativa. Os estudantes de inglês são avaliados sobretudo quanto às habilidades de leitura, escrita e fala, através de questões de múltipla escolha, bem como questões abertas. Conteúdos trabalhados em sala são utilizados em instrumentos avaliativos, mas a realidade dos alunos não é necessariamente considerada na experiência avaliativa. O estudo é composto, em sua maioria, de professoras do sexo feminino, o que pode não representar a diversidade de professores de língua inglesa no Chile.

5.2.59 La ansiedad hacia la evaluación en el aprendizaje del inglés como lengua extranjera en estudiantes de enseñanza media – Revista Lenguas Modernas – 2022

O Quadro 61 exibe partes do conteúdo do artigo de Leiva *et al.* (2022). O estudo investigou a ansiedade em relação a avaliações na disciplina de inglês em 1.093 estudantes do quarto ano do ensino médio de diversas regiões do Brasil, utilizando um inventário específico para medir essa ansiedade em situações de teste.

Quadro 61 – Ramos-Leiva *et al.* (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Lucía Ramos-Leiva, Universidad Católica del Norte, Chile; Marcela Quintana-Lara, Universidad Arturo Prat, Chile; Claudio Díaz-Larenas, Universidad de Concepción, Chile; Tania Tagle-Ochoa, Universidad Católica de Temuco, Chile; Michel Riquelme-Sanderson, Universidad Arturo Prat, Chile; e Paola Alarcón-Hernández, Universidad de Concepción, Chile (2022).
Tipo de Avaliação	O estudo utiliza um inventário de ansiedade específico que mede as diferenças individuais de ansiedade em situações de avaliação, como testes e exames na disciplina de inglês.
População Avaliada	O público-alvo do estudo são estudantes de quarto ano de ensino médio, com uma amostra de 1.093 alunos de diferentes regiões do país, sendo 504 mulheres e 589 homens, com idades entre 16 e 20 anos.

Competências Avaliadas	O foco do estudo é a ansiedade em relação a avaliações na disciplina de inglês, mas não especifica diretamente as competências como leitura, escrita ou fala. A ênfase está na percepção dos alunos sobre suas experiências durante as avaliações.
Métodos e Instrumentos	O estudo utiliza um inventário de ansiedade que pode incluir questões de múltipla escolha e afirmações que os alunos devem classificar em uma escala de concordância. Os tópicos abordam sentimentos de confiança, ansiedade e interferência nas notas durante as avaliações de inglês.
Resultados e Conclusões	Os resultados indicam que a ansiedade é um fator significativo que influencia o desempenho dos alunos. O estudo sugere que a criação de um ambiente de avaliação mais confortável e a congruência entre a avaliação e os conteúdos ensinados podem ajudar a reduzir a ansiedade dos alunos.
Desafios e Limitações	O estudo menciona que a amostra foi coletada de forma não probabilística, o que pode limitar a generalização dos resultados. Além disso, a coleta de dados foi feita em um único momento, o que pode não capturar a variabilidade da ansiedade ao longo do tempo.
Recomendações e Futuras Pesquisas	O estudo recomenda a implementação de ambientes de avaliação que promovam conforto e segurança para os alunos, além de considerar a retroalimentação como uma forma de apoiar o aprendizado e reduzir a ansiedade. Sugere-se também a necessidade de mais pesquisas sobre a percepção dos alunos em relação aos professores e às aulas de inglês.

Fonte: elaboração própria.

Os resultados indicam que a ansiedade influencia significativamente o desempenho dos alunos, sugerindo a importância de ambientes de avaliação mais confortáveis e alinhados com o conteúdo ensinado, para reduzir essa ansiedade. A natureza não probabilística da amostra e a coleta de dados em um único momento, no entanto, limitam a generalização dos achados, reforçando a necessidade de mais pesquisas sobre a percepção dos alunos e a implementação de práticas avaliativas que promovam conforto, segurança e retroalimentação.

5.2.60 Escribir en L2 en la escuela chilena: Una caracterización de la escritura en español de estudiantes de origen haitiano en 5º básico – Revista Signos – 2022

O Quadro 62 exibe partes do conteúdo do artigo de Espinosa Aguirre, Figueroa Miralles e Ávila Reyes (2022). O estudo analisou a produção escrita de 42 alunos do quinto ano de escolas municipais de Santiago do Chile, incluindo estudantes migrantes haitianos, por meio de uma avaliação qualitativa com foco no cumprimento do propósito comunicativo de um díptico sobre a prevenção da poluição marinha causada por plásticos.

Quadro 62 – Espinosa Aguirre, Figueroa Miralles e Ávila Reyes (2022)

Categoria	Descrição
Perfil (Autor,	María Jesús Espinosa Aguirre, Universidad Diego Portales, Chile; Javiera

Instituição e Ano)	Figuerola Miralles, Universidad Alberto Hurtado, Chile; e Natalia Ávila Reyes, Pontificia Universidad Católica de Chile, Chile (2022).
Tipo de Avaliação	O estudo utiliza uma avaliação de desempenho escrita. Os textos dos alunos são analisados e pontuados com base em uma rubrica específica. A avaliação é feita por meio de codificação dos textos, considerando-se aspectos como sintaxe e léxico.
População Avaliada	O público-alvo avaliado no estudo são estudantes de quinto ano de escolas municipais em Santiago do Chile, com uma alta população migrante, incluindo especificamente alunos de origem haitiana.
Competências Avaliadas	As principais competências avaliadas no estudo são as habilidades de escrita, com foco na capacidade de cumprir o propósito comunicativo dos textos elaborados pelos alunos.
Métodos e Instrumentos	Os testes consistem na produção de textos escritos e os alunos devem criar um dítico sobre a prevenção da poluição marinha causada por plásticos. A avaliação é qualitativa, com pontuações que variam de 1 a 6, dependendo do cumprimento do propósito comunicativo.
Resultados e Conclusões	Os resultados mostram que, apesar das barreiras linguísticas, os alunos migrantes haitianos apresentam condições favoráveis para desenvolver habilidades complexas de escrita.
Desafios e Limitações	Um dos principais desafios do estudo é a barreira linguística que os alunos enfrentam ao escrever em uma segunda língua (L2). Além disso, os resultados não são representativos de toda a população, pois a amostra é limitada a 42 textos.
Recomendações e Futuras Pesquisas	O estudo sugere a necessidade de se promover o uso da língua materna (L1) para apoiar a aprendizagem da L2, além de enfatizar a importância de estratégias de ensino que considerem as particularidades dos alunos migrantes. Recomenda-se também a realização de estudos futuros que ampliem a amostra e explorem mais a fundo as habilidades de escrita em contextos diversos.

Fonte: elaboração própria.

Revelou-se que metade dos textos não atingiu o objetivo, apesar de um potencial identificado nos alunos migrantes para desenvolver habilidades complexas de escrita. A barreira linguística é um desafio significativo, e a amostra, limitada, o que aponta para a necessidade de se promover o uso da língua materna, adaptarem-se estratégias de ensino e realizarem-se pesquisas futuras com amostras maiores para melhor compreender o desenvolvimento da escrita em contextos multiculturais.

5.2.61 Colombian' Scholars Discussion About Language Assessment: A Review of Five Journals – Revista Profile – 2022

A pesquisa de Hernández-Ocampo (2022), em partes exibida no Quadro 63, demonstra que os interesses dos acadêmicos colombianos em avaliação no ensino de línguas têm sido diversos na última década, refletindo a variedade nas práticas de avaliação que alguns deles descobriram em suas pesquisas. Essa diversidade não só enriquece o debate na área, mas

também é crucial para uma compreensão mais completa da avaliação, incluindo suas possibilidades e implicações. Nenhum dos estudos tratou do tema da justiça na avaliação de línguas. No entanto, houve interesse em discutir a equidade, como quando alguns pesquisadores sugeriram práticas avaliativas mais democráticas ou abordaram os princípios de avaliação. O letramento em avaliação, por sua vez, surgiu como um novo foco de interesse na comunidade de ensino de línguas da Colômbia.

Os pesquisadores que exploraram as práticas de avaliação encontraram uma grande variação, atribuída à falta de formação adequada, às experiências pessoais e à ausência de uma abordagem padronizada nos programas de formação de professores. Há uma necessidade de intervenção para auxiliar os professores na realização de práticas de avaliação mais robustas. A recomendação é utilizar uma variedade de avaliações, desde que sejam práticas eficazes, adaptadas ao contexto em que são aplicadas e que contribuam para a melhoria da aprendizagem.

Além disso, as crenças desempenham um papel significativo nas práticas avaliativas dos professores. Por isso, alguns estudiosos sugeriram que considerar as crenças dos professores é fundamental para que as mudanças sejam efetivas. A relação entre crenças, práticas e formação é um ciclo que pode se tornar positivo, se os professores estiverem cientes de suas práticas, refletirem sobre elas, identificarem suas necessidades e ajustarem suas crenças sobre avaliação.

Do ponto de vista dos alunos, foram estudadas as crenças, para entender como eles vivenciam diferentes formas de avaliação. Isso também é importante se para refletir sobre as práticas e se identificarem necessidades no ciclo avaliativo.

No que diz respeito ao aprimoramento das habilidades linguísticas, citam-se o uso de autoavaliação e avaliação entre pares foi comum. Sempre que se mencionava a avaliação por pares, a autoavaliação também estava envolvida. Isso mostra um forte interesse acadêmico em se promoverem métodos alternativos de avaliação que incentivem a autonomia dos alunos e o aprendizado de línguas. Se os objetivos de aprendizagem forem claros e conhecidos pelos alunos, eles podem monitorar seu próprio progresso e o de seus colegas.

Finalmente, houve um apelo generalizado por formação em avaliação no ensino de línguas nos programas de formação e desenvolvimento de professores, particularmente em relação à discussão sobre letramento em avaliação. Também foram feitas recomendações para usar instrumentos que ajudem os professores a interpretar e aproveitarem os resultados dos testes, além de sugestões sobre metodologias de pesquisa e conceitos para expandir o conhecimento de letramento entre os professores. No entanto, a maioria das propostas foi apresentada em artigos de reflexão, baseados na experiência pessoal ou em revisões teóricas. É essencial haver mais pesquisas empíricas que investiguem as particularidades dos contextos.

Isso é um convite para que os professores de línguas estejam mais atentos às suas práticas de avaliação e registrem de maneira mais sistemática o que ocorre em suas salas de aula, para que mais pesquisas-ação possam ajudar a construir um conhecimento mais sólido sobre a avaliação em sala de aula na comunidade de professores de línguas.

Quadro 63 – Hernández-Ocampo (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Sonia Patricia Hernández-Ocampo, Pontificia Universidad Javeriana & Universidad de los Andes, Bogotá, Colômbia (2022).
Tipo de Avaliação	A pesquisa faz um apanhado de diversos artigos sobre avaliação publicados em cinco periódicos colombianos. Logo, aborda vários tipos de avaliação (somativas, formativas, qualitativas, quantitativas, testes padronizados, testes alternativos).
População Avaliada	Considerando-se a diversidade de resultados encontrados nas pesquisas, as populações avaliadas incluem: professores universitários de inglês, estudantes de doutorado, professores colombianos em formação e professores atuantes.
Competências Avaliadas	Compreensão oral, letramento em avaliação, compreensão escrita, produção oral, habilidades de argumentação.
Métodos e Instrumentos	Gravações de vídeo, textos escritos, questionários, portfólios, diários, entrevistas, discussões, testes, conversas gravadas, autoavaliação, avaliação de pares.
Resultados e Conclusões	Poucos artigos sobre crenças dos professores em avaliação. Poucas disciplinas sobre avaliação nos cursos em universidades. Professores alegam valorizar a avaliação formativa, mas praticam a somativa. Alunos geralmente avaliam seus pares com notas abaixo do esperado (na avaliação de pares). Professores precisam refletir sobre os instrumentos avaliativos que criam. A formação em avaliação é importante. Os pesquisadores colombianos se interessam por vários aspectos diferentes no âmbito da avaliação.
Desafios e Limitações	Poucos artigos na área. Professores pouco letrados em avaliação. Discrepância entre teoria e prática de avaliação. Alunos não acostumados com avaliações qualitativas. Necessidade de familiarização com avaliações alternativas. Criar testes mais desafiadores. Pouca padronização nos instrumentos avaliativos por falta de preparo dos profissionais. Nenhum dos artigos abordou a questão da justiça na avaliação.
Recomendações e Futuras Pesquisas	Estudos mais empíricos do que teóricos. Mais capacitação e letramento em avaliação para os professores e profissionais da educação.

Fonte: elaboração própria.

5.2.62 In-Service University Level Efl Instructors Language Assessment Literacy and Training Needs Ali Sayyadi – Revista Profile – 2022

Apesar das limitações, o estudo de Sayyadi (2022), em partes exibido no Quadro 64, oferece contribuições práticas para melhorar a formação em avaliação para professores de inglês universitários no Irã. Os resultados podem ajudar a expandir o conhecimento na área e aumentar a conscientização dos formadores de professores sobre o letramento em avaliação dos professores universitários iranianos. A maior conscientização pode incentivá-los a criar e implementar programas de treinamento em avaliação mais eficazes que atendam melhor às necessidades reais e às lacunas dos instrutores. Além disso, os resultados podem motivar administradores universitários e chefes de departamento no Irã a valorizarem mais as práticas de avaliação de seus instrutores, considerarem programas de treinamento em avaliação mais práticos e eficazes para instrutores em formação e em serviço e oferecerem apoio financeiro adequado para aprimorar o letramento em avaliação.

Quadro 64 – Sayyadi (2022)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Ali Sayyadi, University of Tehran, Teerã, Irã (2022).
Tipo de Avaliação	Questionário aplicado a professores universitários iranianos para identificar o nível de letramento em avaliação de cada um.
População Avaliada	Sessenta e oito professores universitários iranianos (28 homens e 40 mulheres) docentes de língua inglesa, todos mestres e doutores, entre 30 e 58 anos de idade.
Competências Avaliadas	Letramento em avaliação.
Métodos e Instrumentos	Questionários <i>on-line</i> e entrevistas semiestruturadas para coletar dados qualitativos e quantitativos sobre os conhecimentos sobre avaliação de cada professor (o estudo foi realizado durante a pandemia).
Resultados e Conclusões	Professores com pouca capacitação em avaliação, que, no entanto, preferem não receber treinamentos longos em avaliação. O letramento em avaliação é pouco explorado no contexto de ensino superior. Há uma escassez de estudos sobre avaliação e letramento no contexto de ensino superior iraniano. A cultura de avaliação de cada país pode variar; logo, as prioridades relacionadas à avaliação também. Para professores, é desafiador se especializar em estudos sobre avaliação, considerando-se o volume de trabalho rotineiro. Aproximadamente 90% dos respondentes expressaram vontade de fazer cursos e treinamentos sobre avaliação (sobretudo conhecimentos práticos). Professores iranianos são comumente apresentados à teoria sobre avaliação, mas têm dificuldade de colocá-la em prática.
Desafios e Limitações	Alguns professores não souberam identificar suas necessidades em conhecimentos de avaliação. Muitos dos entrevistados afirmaram não ter buscado fontes adicionais sobre avaliação após a graduação, por não ser uma área de interesse pessoal e por precisarem seguir as normas da universidade em que trabalham. Um deles reconhece que sua prática é pouco científica. Muitos são avaliados pelas publicações, não pelas práticas docentes.

Recomendações e Futuras Pesquisas	O estudo procurou investigar as experiências de formação em avaliação de professores de inglês universitários, as práticas de avaliação em sala de aula e as necessidades de conhecimento em avaliação no Irã. De forma geral, os resultados mostraram que os instrutores não tinham recebido treinamento adequado para melhorar suas práticas de avaliação, tanto no que diz respeito ao letramento em avaliação quanto à avaliação em sala de aula. Havia sido expostos apenas aos princípios teóricos de avaliação de línguas em cursos limitados oferecidos a futuros professores nas universidades, com foco mais em conceitos e teorias do que em aspectos práticos. Apesar dessa falta de treinamento prático, os instrutores mostraram uma preferência maior por programas de treinamento de um a dois dias (ou seja, programas básicos) em vez de programas mais avançados com duração superior a três dias. Essa preferência pode ser atribuída a restrições pessoais e contextuais diversas, como o desinteresse dos instrutores pela área de testes e avaliação de linguagem e a falta de apoio das universidades. O estudo, entretanto, apresenta algumas limitações que podem ser abordadas em pesquisas futuras. Em primeiro lugar, trata-se de um estudo com uma amostra pequena, composta de 68 participantes que responderam ao questionário e oito entrevistados. Além disso, alguns termos no questionário (como “básico” e “avançado”) podem ainda ser vagos, embora sejam definidos em termos de dias. Essa falta de clareza pode levar a interpretações variadas e, portanto, a respostas distintas dos participantes.
--	---

Fonte: elaboração própria.

5.2.63 Intervención en Lectura Temprana en Alumnado con Trastorno del Desarrollo del Lenguaje – Revista Signos – 2023

O Quadro 65 exhibe partes do conteúdo do artigo de Ramírez Santana *et al.* (2023).

Quadro 65 – Ramírez Santana *et al.* (2023)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Gustavo M. Ramírez Santana, Víctor M. Acosta Rodríguez, Sergio Hernández Expósito, Atteneri Delgado Cruz, Universidad de la Laguna, Espanha (2023).
Tipo de Avaliação	O estudo utiliza uma avaliação pré-teste e pós-teste, por meio das quais as habilidades dos alunos são medidas antes e depois da implementação do programa de intervenção. A avaliação inclui tarefas específicas para medir habilidades alfabéticas e processamento fonológico.
População Avaliada	O público-alvo do estudo são crianças de 5 anos com Transtorno do Desenvolvimento da Linguagem (TDL) e um grupo de controle de crianças com desenvolvimento típico.
Competências Avaliadas	As principais competências avaliadas incluem habilidades alfabéticas iniciais, como reconhecimento de letras, associação grafema-fonema, leitura de palavras e pseudopalavras, além de consciência fonológica.
Métodos e Instrumentos	Os testes consistem em uma bateria de avaliações que inclui tarefas de múltipla escolha e exercícios práticos. Os principais tópicos dos exercícios incluem reconhecimento de letras, leitura de palavras e pseudopalavras, e diferentes aspectos da consciência fonológica, como identificação e manipulação de sílabas e fonemas.
Resultados e Conclusões	Os resultados indicam que os alunos com TDL apresentaram resultados significativamente mais baixos em comparação com os alunos com

	desenvolvimento típico. A intervenção mostrou ser eficaz em melhorar as habilidades de leitura e processamento fonológico, sugerindo a necessidade de programas de intervenção desde idades muito precoces.
Desafios e Limitações	O estudo menciona a complexidade de garantir que os grupos de alunos sejam comparáveis em outras variáveis que possam influenciar os resultados. Além disso, a amostra foi limitada a um número específico de alunos, o que pode afetar a generalização dos resultados.
Recomendações e Futuras Pesquisas	O estudo recomenda a implementação de programas de intervenção desde idades muito precoces e sugere a necessidade de mais pesquisas para explorar diferentes abordagens e metodologias que possam ser eficazes para alunos com TDL.

Fonte: elaboração própria.

O estudo analisou a produção escrita de 42 alunos do quinto ano de escolas municipais de Santiago (Chile), incluindo estudantes migrantes haitianos, por meio de uma avaliação qualitativa com foco no cumprimento do propósito comunicativo de um díptico sobre a prevenção da poluição por plásticos, revelando que metade dos textos não atingiu o objetivo, apesar de um potencial identificado nos alunos migrantes para desenvolver habilidades complexas de escrita, sendo a barreira linguística um desafio significativo e a amostra limitada, o que aponta para a necessidade de se promover o uso da língua materna, adaptarem-se as estratégias de ensino e realizarem-se pesquisas futuras com amostras maiores, para melhor compreender o desenvolvimento da escrita em contextos multiculturais.

5.2.64 Challenges and Complexities in Teacher Conceptualization of Dynamic Assessment in the L2 Classroom: A Case Study in Türkiye – Revista Profile – 2023

O Quadro 66 exibe partes do conteúdo do artigo de Özdemir-Yılmaz e Özkan (2023).

Quadro 66 – Özdemir-Yılmaz e Özkan (2023)

Categoria	Descrição
Perfil (Autor, Instituição e Ano)	Meryem Özdemir-Yılmaz, Cukurova University, Adana, Türkiye, and Eötvös Loránd University, Budapest, Hungria, e Yonca Özkan, Cukurova University, Adana, Turquia (2023).
Tipo de Avaliação	O estudo explora a avaliação dinâmica, que é um procedimento dialético que requer que os professores avaliem o progresso dos alunos prestando atenção aos erros e fornecendo indicações graduais para ajudá-los a se corrigirem.
População Avaliada	O público-alvo do estudo são alunos de inglês como segunda língua, especificamente aqueles que participam de cursos de conversação em inglês.
Competências Avaliadas	A principal competência avaliada no estudo é a produção oral durante as aulas de conversação.
Métodos e Instrumentos	O estudo não menciona testes formais, como múltipla escolha ou questões abertas, mas discute a avaliação da proficiência oral e a necessidade de se desenvolverem critérios claros para avaliar o desempenho dos alunos em atividades de fala.

Resultados e Conclusões	Os resultados mostraram que uma das professoras redefiniu seu papel como fornecedora de indicações graduais, enquanto a outra resistiu a adotar as funções requeridas pela avaliação dinâmica. O estudo conclui que é importante se considerarem fatores pessoais e contextuais que moldam as suposições dos professores, para promover mudanças na prática docente.
Desafios e Limitações	Os principais desafios incluem a resistência de um dos professores em adotar a avaliação dinâmica e a inconsistência nas práticas de avaliação em relação ao que era ensinado. Além disso, a falta de clareza sobre como implementar a avaliação dinâmica nas aulas foi um obstáculo significativo.
Recomendações e Futuras Pesquisas	O estudo sugere que futuras pesquisas devem considerar mais profundamente os fatores pessoais e contextuais que influenciam as crenças dos professores sobre avaliação. Além disso, recomenda-se o desenvolvimento profissional contínuo para ajudar os professores a integrarem a avaliação dinâmica em suas práticas de ensino.

Fonte: elaboração própria.

A pesquisa investigou a implementação da avaliação dinâmica na produção oral de alunos de inglês como segunda língua em cursos de conversação, revelando que, enquanto uma professora adaptou seu papel para fornecer indicações graduais aos alunos, a outra resistiu, demonstrando a influência de fatores pessoais e contextuais nas suposições dos professores sobre avaliação. Os desafios incluíram a resistência à mudança, a inconsistência entre ensino e avaliação e a falta de clareza na implementação da avaliação dinâmica, sugerindo a necessidade de se considerarem as crenças dos professores e se oferecer desenvolvimento profissional contínuo para facilitar a adoção dessa abordagem.

6 DISCUSSÃO

Os estudos analisados refletem uma rica diversidade de métodos e abordagens de avaliação no ensino de línguas, abrangendo práticas tradicionais, alternativas e inovadoras. Uma semelhança significativa é o foco em equilibrar avaliações formativas e somativas, com o uso frequente de instrumentos como testes padronizados (*e.g.* TOEFL, Lectum) e rubricas para produção escrita e oral. Além disso, o *feedback* corretivo emerge como um componente essencial para orientar o aprendizado, sendo aplicado tanto em avaliações formais quanto em atividades dinâmicas, como dramatizações (*role-plays*).

A integração de tecnologia, como plataformas digitais, *podcasts* e rubricas *on-line*, demonstra uma tendência crescente de modernização das práticas avaliativas. Contudo, os estudos também mostram uma convivência com métodos tradicionais, como provas de múltipla escolha e avaliações estruturadas. As diferenças aparecem, sobretudo, nos formatos e objetivos das avaliações: enquanto alguns focalizam habilidades específicas, como a leitura e a produção oral, outros buscam abordar competências mais amplas, incluindo aspectos emocionais e culturais.

O uso de autoavaliação, portfólios e avaliação por pares reflete uma preocupação crescente em promover a autonomia do aluno. Por outro lado, desafios culturais e regionais também influenciam as práticas avaliativas, como evidenciado nos estudos sobre falantes de línguas indígenas e crianças com necessidades específicas. A inovação aparece em métodos que utilizam gravações digitais, instrumentos interativos e textos multimodais, ampliando o escopo tradicional das avaliações.

Destaca-se a importância da capacitação docente para aprimorar o letramento em avaliação, com *workshops* e questionários para preparar professores na criação de instrumentos eficazes. Os estudos analisados sugerem que a avaliação no ensino de línguas está em constante evolução, adaptando-se às demandas pedagógicas e tecnológicas, ao mesmo tempo em que valoriza práticas inclusivas e culturalmente sensíveis.

Em resumo, os estudos analisados têm:

1. foco em competências comunicativas: a avaliação oral e escrita destaca-se, especialmente em estudos voltados para o ensino de línguas;
2. adoção de métodos alternativos: portfólios, rubricas e avaliação por pares estão ganhando espaço como opções mais integradoras;
3. persistência de testes tradicionais: apesar da inovação, os testes padronizados e de múltipla escolha permanecem amplamente usados;

4. valorização do *feedback*: a retroalimentação contínua aparece como um elemento-chave para a aprendizagem; e
5. adoção de ferramentas tecnológicas: avaliações mediadas por computador, como no TOEFL e em *podcasts*, começam a ser mais comuns.

6.1 Tipos de Avaliação

Os tipos mais frequentes refletem métodos tradicionais, avaliações padronizadas e abordagens centradas em habilidades linguísticas:

6.1.1 Testes padronizados

- Exames como TOEFL, Lectum e outros testes de proficiência foram citados repetidamente;
- Esses instrumentos medem habilidades específicas, como compreensão leitora, produção oral ou conhecimentos linguísticos.

6.1.2 Rubricas e critérios avaliativos

- O uso de rubricas para avaliar produções orais ou escritas aparece frequentemente;
- As rubricas são usadas para garantir maior objetividade e clareza nos processos avaliativos.

6.1.3 Provas de compreensão leitora e escrita

- Testes que incluem questões de múltipla escolha e produção textual são amplamente utilizados.

6.1.4 Questionários e escalas

- Ferramentas como questionários baseados na escala Likert e inventários de ansiedade em exames foram destacadas;
- Estudos exploratórios e diagnósticos foram bastante adotados.

Com relação aos tipos menos comuns de avaliação, podemos destacar a avaliação por portfólios, autoavaliações e métodos qualitativos para avaliar competências culturais.

6.2 População avaliada

A análise dos 64 artigos revelou uma grande diversidade nas populações avaliadas. As pesquisas abrangem uma vasta gama de faixas etárias, níveis de proficiência, contextos geográficos, culturais e metodologias, com foco particular em estudantes de línguas estrangeiras, tanto universitários quanto adolescentes e crianças. As abordagens variam de avaliações gerais de proficiência até estudos de necessidades específicas, como dificuldades cognitivas e distúrbios de linguagem.

Os artigos analisados oferecem uma visão abrangente e detalhada das populações avaliadas nos contextos de ensino de línguas, com uma diversidade notável em termos de idade, nível de proficiência, formação educacional e contextos culturais. A maioria dos estudos se concentra em estudantes universitários, cujas idades variam principalmente entre 18 e 30 anos, com alguns artigos envolvendo adultos em programas de educação para línguas ou cursos intensivos.

A proficiência dos alunos, quando mencionada, abrange os níveis de A2 a B1 do Quadro Europeu Comum de Referência para Línguas (CEFR), o que reflete a presença de alunos em diferentes estágios do aprendizado de línguas, desde os níveis iniciais até os intermediários.

Esses artigos também revelam uma grande diversidade de contextos educacionais, muitos com foco em programas específicos, como educação bilíngue, pedagogia, comunicação empresarial e formação de professores.

A formação acadêmica dos professores analisados é predominantemente avançada. A maioria possui títulos de mestrado ou doutorado e uma experiência de ensino variada, com muitos professores atuando por períodos que variam de 3 a 17 anos.

Além disso, os estudos analisam populações específicas, como estudantes de Pedagogia ou grupos menores – de 10 a 20 alunos –, permitindo um olhar mais aprofundado sobre práticas de avaliação em contextos mais restritos e especializados.

A análise do público-alvo dos artigos também destaca uma ampla gama de contextos geográficos e culturais. Diversos estudos incluem populações de diferentes países, como Espanha, Colômbia, Irã, México, Omã e Chile, refletindo práticas de ensino e avaliação que

podem ser modeladas por realidades locais, ao mesmo tempo que se alinham a padrões internacionais de proficiência.

Além disso, alguns artigos põem em foco estudantes internacionais ou de programas de intercâmbio universitário, o que indica uma tendência crescente para a inclusão de populações multiculturais e a adaptação de práticas de avaliação para diferentes culturas e contextos educacionais.

Outro ponto importante é a diversidade de faixas etárias e contextos educacionais nas populações analisadas. Enquanto muitos artigos se concentram em estudantes universitários, outros investigam crianças e adolescentes com idades entre 6 e 18 anos, com um foco especial em escolas urbanas e rurais na América Latina, como no Chile. Nesse contexto, alguns estudos lidam com grupos de alunos com necessidades específicas, como crianças com deficiências auditivas, transtornos de desenvolvimento da linguagem (TDL) ou dificuldades na atribuição de estados mentais. Essa variedade demonstra como os métodos de avaliação precisam ser adaptados para atender a diferentes perfis de alunos, considerando suas necessidades cognitivas, sociais e culturais.

O uso de grupos experimentais e de controle também é uma característica comum, principalmente em estudos que envolvem crianças com diferentes tipos de desenvolvimento cognitivo ou linguístico, como no caso de crianças com desenvolvimento típico vs. atípico. Esse enfoque permite aos pesquisadores testarem diferentes métodos de avaliação e sua eficiência em diferentes contextos, especialmente em grupos com características específicas.

Em termos de proficiência, muitos estudos lidam com alunos do nível básico ao intermediário; alguns também exploram a avaliação de habilidades em níveis mais avançados (como no caso de alunos de níveis de A2 a C1), além de examinar o uso de certificações linguísticas, como exames de espanhol e inglês como língua estrangeira.

Por fim, a análise das populações nos artigos confirma a diversidade de contextos e de necessidades que os métodos de avaliação devem abranger. A presença significativa de crianças e adolescentes em estudos, especialmente em contextos de ensino básico e secundário, destaca a importância de se adaptarem as práticas de avaliação para diferentes faixas etárias e níveis de escolaridade. Além disso, a inclusão de estudantes de diferentes origens culturais e contextos de aprendizagem, como haitianos aprendendo espanhol no Chile, revela a necessidade de uma avaliação culturalmente sensível e capaz de atender às complexas realidades dos alunos.

Em resumo, os 64 artigos analisados oferecem uma visão ampla e rica das populações avaliadas nos contextos de ensino de línguas, evidenciando a diversidade de perfis de alunos, professores e contextos educacionais. Esses estudos ressaltam a importância de se adaptarem

as práticas de avaliação às necessidades específicas de diferentes grupos, considerando suas idades, níveis de proficiência, contextos culturais e necessidades cognitivas, ao mesmo tempo que se busca alinhá-las a padrões internacionais de ensino e avaliação linguística. Para resumir, temos:

- **Diversidade de contextos geográficos e culturais:**
 - as populações avaliadas abrangem diferentes países: Brasil, Chile, Colômbia, Espanha, Haiti, Irã, México e Omã;
 - a diversidade geográfica indica uma variedade de contextos educacionais e culturais, refletindo práticas de ensino e avaliação distintas.
- **Público universitário predominante:**
 - em muitos artigos, o público-alvo é composto de estudantes universitários de instituições públicas e privadas;
 - a faixa etária desses alunos varia, com maior foco em estudantes entre 18 e 30 anos;
 - há estudos com foco em estudantes de Pedagogia, formação de professores e programas de línguas.
- **Perfil dos professores:**
 - vários artigos analisam professores de inglês como língua estrangeira, especialmente em contextos universitários (Irã, Colômbia, México etc.);
 - muitos desses professores têm formação de mestres e doutores e experiência variada – de 3 a 17 anos de prática.
- **Alunos com diversos níveis de proficiência:**
 - a maioria dos artigos tem foco em alunos com níveis de proficiência de A2 a B1 no CEFR;
 - alguns artigos detalham estudantes que estão no nível intermediário de inglês (A2 e B1) ou aqueles em processo de consolidação do idioma.
- **Alunos de programas específicos:**
 - alguns estudos têm foco em alunos de programas específicos, como educação bilíngue, Pedagogia, comunicação empresarial e formação de professores.
 - há também menção de programas intensivos e cursos voltados para adultos, principalmente em universidades privadas e centros de idiomas.
- **Populações menores ou específicas:**

- alguns artigos avaliam pequenos grupos de estudantes (*e.g.* de 10 a 20 alunos) em contextos mais restritos, como licenciaturas específicas ou programas focalizados em disciplinas específicas.
- Diversidade de idade e faixa etária:
 - estudantes adultos (*e.g.* 18 a 27 anos) são mencionados em alguns estudos, com maior foco em estudantes universitários de 18 a 25 anos;
 - há também um foco considerável em crianças e adolescentes (*e.g.* idades variando 5 e 18 anos), com muitos estudos abordando crianças de 5 a 14 anos e adolescentes em ensino básico e ensino médio.
- Profissão e formação dos participantes:
 - alguns estudos têm foco em professores de inglês (*e.g.* professores em formação ou em exercício) e alunos de pós-graduação, especialmente na formação de professores.
 - a maioria dos artigos trata de estudantes universitários (*e.g.* 18-25 anos), refletindo a formação acadêmica desses alunos nas áreas de linguística, ensino de línguas e pedagogia.
- Nível de proficiência:
 - muitos estudos têm foco em alunos com níveis de proficiência A2, B1, B2 (*e.g.* artigos sobre o CEFR);
 - alguns artigos também investigam diferentes níveis de proficiência dentro de um mesmo grupo de estudantes (*e.g.* A2, B2, C1 no mesmo estudo).
- Contextos específicos de ensino:
 - diversos estudos abordam o inglês como língua estrangeira (EFL) e inglês como segunda língua (ESL);
 - alguns estudos investigam estudantes de intercâmbio universitário e aqueles em programas de bolsas de estudo internacionais.
- Escolas e localidades:
 - muitos artigos tratam de escolas urbanas e universidades. Há destaque para escolas públicas e privadas na América Latina (*e.g.* Chile e Brasil);
 - o contexto rural também é abordado em alguns casos, como em escolas de zonas rurais no Chile.
- Diversidade cultural e nacionalidade:
 - alguns estudos mencionam estudantes internacionais (*e.g.* artigos sobre multiculturalismo, comparando crianças com deficiência auditiva e ouvintes);

- muitos estudos também se concentram em estudantes migrantes (*e.g.* estudantes de origem haitiana ou migrantes em países da América Latina).
- Distribuição por gênero:
 - a maior parte dos estudos não especifica o gênero dos participantes. Porém, em alguns artigos é mencionada a distribuição de gênero, com informações sobre o número de meninos e meninas, especialmente em estudos sobre escolas.
- Populações com necessidades específicas:
 - alguns artigos investigam populações com dificuldades cognitivas ou de linguagem, como crianças com TDL;
 - grupos experimentais e de controle são utilizados para testar diferentes condições, como o desenvolvimento de crianças com desenvolvimento típico *vs.* atípico.
- Diversidade de linguagens e línguas estrangeiras:
 - vários artigos abordam estudantes aprendendo espanhol como língua estrangeira;
 - outros têm foco em candidatos a certificações de espanhol ou estudantes de francês;
 - a língua-alvo também pode ser inglês, com ênfase em avaliações orais e de fluência.
- Métodos de avaliação:
 - muitos artigos abordam avaliações de proficiência, principalmente nas línguas estrangeiras, como espanhol e inglês;
 - avaliações orais e escritas são frequentemente mencionadas, especialmente em contextos de exames de certificação de línguas.
- Amostras de estudo:
 - a maioria dos estudos trabalha com amostras médias a grandes (*e.g.* entre 100 e 2 mil participantes), embora alguns artigos usem grupos menores, de 10 a 50 alunos, dependendo do contexto de estudo;
 - em alguns casos, a amostra é bem equilibrada entre escolas públicas e privadas.

As populações mais comumente investigadas foram:

1. Público universitário como alvo principal:

- presente na maioria dos estudos, foco em estudantes entre 18 e 30 anos, com destaque para áreas como pedagogia, línguas estrangeiras e programas de intercâmbio;
2. Nível de proficiência em idiomas:
 - grande número de estudos analisa estudantes de nível A2 a B2, de acordo com o CEFR;
 - esses níveis refletem aprendizes intermediários e avançados de línguas como inglês, espanhol e francês;
 3. Diversidade geográfica e cultural:
 - estudos abrangem principalmente países da América Latina (Chile, Colômbia, México), do Oriente Médio (Irã, Omã) e a Espanha;
 - contextos educacionais variados (urbanos, rurais e multiculturais) são recorrentes;
 4. Participação de professores:
 - professores de inglês como língua estrangeira são foco em muitos artigos, com destaque para sua formação (mestre/doutor) e experiência (3 a 17 anos);
 5. Crianças e adolescentes no ensino básico:
 - representados em estudos sobre ensino fundamental e médio, com idades variando de 6 a 18 anos;
 - foco em alunos do quarto ao oitavo anos é relevante.

Em contrapartida, as populações menos comumente investigadas foram:

1. Populações com necessidades específicas:
 - apenas uma fração dos estudos aborda crianças com dificuldades cognitivas ou transtornos (como TDL);
 - grupos experimentais e de controle para desenvolvimento típico *vs.* atípico aparecem esporadicamente;
2. Estudantes em contextos rurais:
 - a maioria dos estudos prioriza escolas urbanas, com exceções em zonas rurais, como no Chile;

3. Diversidade de linguagens estrangeiras (além do inglês):

- espanhol e francês como línguas-alvo são abordados, mas com menor frequência;
- certificação em espanhol e estudos de francês são menos recorrentes;

4. Distribuição por gênero:

- a distribuição por gênero é pouco detalhada em grande parte dos estudos. Quando mencionada, ocorre em grupos escolares específicos;

5. Programas de bolsas e intercâmbio internacional:

- menos frequente, embora presente em alguns artigos com foco em estudantes de bolsas internacionais e intercâmbios.

Esse *ranking* evidencia uma preferência por estudos envolvendo públicos universitários e intermediários de idiomas, com menor atenção para populações específicas e contextos menos abrangentes.

6.3 Competências avaliadas

Os estudos analisados exploram uma ampla gama de competências relacionadas ao ensino de línguas. Entre as mais frequentes estão as quatro macro-habilidades: produção oral, produção escrita, compreensão oral e compreensão leitora. A produção oral destaca-se por ser avaliada em diversos estudos, com ênfase em fluência, pronúncia e interação. A produção escrita também recebe atenção significativa, incluindo textos argumentativos e narrativos, com foco em coesão, coerência e correção gramatical.

A compreensão leitora é frequentemente dividida em níveis, como inferencial, literal e crítica, além de integrações multimodais envolvendo texto e imagem. A compreensão oral, embora menos abordada, é explorada em contextos variados, incluindo crianças mapuches e alunos hipoacúsicos.

Competências específicas, como conhecimentos em literatura e fala pública, são menos frequentes, mas representam contribuições importantes para contextos especializados. Os estudos também exploram habilidades fonológicas, cognição e aspectos emocionais, como a ansiedade em avaliações. A competência intercultural e o letramento em avaliação destacam-se em suas respectivas áreas, refletindo um esforço para alinhar as práticas avaliativas às demandas globais e culturais.

6.3.1 Ranking de competências mais e menos comuns

- Mais comuns
 1. Produção escrita.
 2. Produção oral.
 3. Compreensão leitora.
 4. Compreensão oral.
 5. Letramento em avaliação.
- Menos comuns
 1. Literatura.
 2. Fala pública em negócios.
 3. Competências interculturais.
 4. Habilidades fonológicas e alfabéticas.
 5. Aspectos emocionais.

A análise dos 64 artigos revela um foco claro nas quatro macro-habilidades (produção oral, escrita, compreensão oral e leitora), além de esforços para detalhar as competências em níveis específicos, como leitura crítica e escrita argumentativa. As diferenças entre os estudos refletem a diversidade metodológica e cultural: enquanto alguns abordam contextos específicos (*e.g.* crianças mapuches), outros exploram aspectos inovadores, como avaliações dinâmicas e ansiedade.

A presença de competências menos comuns, como habilidades interculturais e fonológicas, sugere que as pesquisas futuras podem se beneficiar de uma maior exploração dessas áreas para atender a uma gama ainda mais ampla de necessidades educacionais.

6.3.2 Análise comparativa das competências nos 64 artigos sobre avaliação no ensino de línguas

Ao compararmos as competências avaliadas nos artigos, temos:

- Produção oral
 - Foco em habilidades como vocabulário, pronúncia e capacidade de transmitir ideias.

- Ênfase em frases e entonação.
 - Fluência, pronúncia, precisão gramatical e uso de vocabulário.
 - Correção linguística e fluência.
 - Complexidade, precisão e fluência em inglês como L2.
 - Foco em habilidades como interação, fluência e correção.
 - Uso específico de comparativos e estruturas gramaticais.
- Compreensão oral
 - Parte de avaliação das quatro habilidades.
 - Específica para crianças mapuches.
 - Associada ao desenvolvimento linguístico em hipoacúsicos.
- Produção escrita
 - Correção de erros em redações.
 - Habilidades narrativas.
 - Escrita argumentativa: coesão, coerência, ortografia.
 - Escrita narrativa e cumprimento de propósito comunicativo.
 - Escrita em L2 com foco em morfossintaxe e ortografia.
- Compreensão leitora
 - Dividida em inferência, literal e crítica.
 - Análise de estrutura textual e relações com outros elementos.
 - Conexão com desempenho acadêmico.
 - Avaliação para pré-leitores em relação à função da leitura.
 - Foco em macroestrutura textual e identificação de relações causais.
 - Leitura multimodal: interpretação de texto e imagem.
 - Compreensão crítica e inferencial.
- Competência comunicativa
- Competências interculturais
 - Conhecimento sociocultural e atitudes interculturais.
- Letramento em avaliação
- Habilidades gramaticais
 - Foco em gramática e vocabulário.
 - Uso apropriado de L1.
- Habilidades fonológicas e alfabéticas

- Consciência fonológica, velocidade de denominação, associação grafema-fonema.
- Habilidades cognitivas e emocionais
 - Ansiedade em avaliações.
 - Metacognição: análise e composição morfológica.
- Pensamento crítico
 - Avaliação associada às habilidades de leitura e escrita.
- Habilidades específicas
 - Literatura.
 - Fala pública: foco em inglês voltado para negócios.
 - Criatividade.

6.4 Métodos de instrumentos utilizados

Os estudos analisados apresentam uma ampla diversidade de métodos e instrumentos avaliativos no ensino de línguas, com tendências que vão desde o uso de instrumentos tradicionais até abordagens mais inovadoras e personalizadas. As principais tendências identificadas foram:

- Testes de múltipla escolha e perguntas abertas
 - Esses formatos são amplamente utilizados pela sua objetividade e facilidade de implementação.
 - Os estudos combinam esses formatos em avaliações que medem tanto conhecimentos específicos quanto habilidades analíticas.
- Uso de questionários
 - Captam percepções de estudantes e professores ou medem fatores emocionais como ansiedade.
- Rubricas
 - Ferramenta essencial em estudos voltados para a produção oral e escrita. Garantem maior clareza e consistência na correção.
- Portfólios
 - Instrumento destacado em estudos como promotor de reflexão e acompanhamento do progresso dos alunos.
- Produção oral e interação social

- Estudos destacam tarefas orais que promovem a comunicação em contextos reais.
- *Feedback* corretivo
 - Explorados nos estudos tanto em formas diretas quanto indiretas, para auxiliar na melhoria contínua dos alunos.

Ainda em relação às principais tendências identificadas nos artigos, detectamos algumas abordagens inovadoras, tais como:

- Autoavaliação e avaliação por pares
 - Promovem reflexão crítica e maior engajamento.
- Dinâmicas interativas e dramatização (*role-play*)
 - Incentivam habilidades colaborativas e criativas.
- Exploração de habilidades metacognitivas
 - Abordam consciência fonológica e manipulação de morfemas.
- Integração de temas globais e multimodais
 - Conectam habilidades linguísticas a temas como poluição e cultura, além de explorar elementos textuais e visuais.

Nesse quesito das abordagens inovadoras, podemos igualmente verificar algumas diferenças relevantes entre os estudos no que concerne a:

- Formato da interação avaliativa
 - Interações em duplas/trios aparecem em poucos estudos, enquanto outros mantêm abordagens individuais ou padronizadas.
- Uso de instrumentos digitais vs. tradicionais
 - Testes *on-line* (como APTIS) contrastam com ferramentas presenciais, como papel e caneta.
- Dimensões emocionais e comportamentais
 - Consideram a ansiedade e o impacto emocional durante as avaliações.

Os artigos ainda revelam alguns pontos extremamente interessantes, quais sejam:

- Interdisciplinaridade
 - Alguns estudos conectam habilidades linguísticas a outros campos, como cultura e ciência ambiental.

- Avaliações metacognitivas e criativas
 - A exploração de habilidades reflexivas ou tarefas baseadas em temas reais enriquecem o processo avaliativo.

Para resumir, temos o seguinte *ranking* de tendências:

- Mais comuns
 - Testes de múltipla escolha e perguntas abertas.
 - Uso de rubricas em avaliações escritas.
 - Avaliações de compreensão leitora e produção oral.
- Moderadamente comuns
 - Questionários para capturar percepções ou medir fatores emocionais.
 - Combinação de formatos avaliativos (questões abertas + múltipla escolha).
- Menos comuns
 - Avaliações inovadoras, como dramatização, autoavaliação e *feedback* corretivo em tempo real.
 - Exploração de habilidades metacognitivas e multimodais.

Os estudos revisados mostram que, embora as práticas tradicionais sejam preponderantes, algumas abordagens inovadoras estão emergindo como uma alternativa para tornar as avaliações mais dinâmicas e centradas no aluno. Essa diversidade de enfoques é essencial para atender às diferentes necessidades do aprendizado de línguas, promovendo uma abordagem mais holística e contextualizada.

6.5 Resultados e conclusões presentes nos 64 artigos sobre avaliação

Ao examinar detidamente os 64 artigos, foi possível perceber alguns pontos convergentes e divergentes entre eles, conforme elencados a seguir.

6.5.1 Pontos em comum

- Necessidade de formação e capacitação docente
 - Lacunas em treinamento formal para práticas avaliativas eficazes.
 - Formação contínua melhora confiança e eficácia dos professores.

- Impacto do *feedback*
 - *Feedback* detalhado e imediato é essencial para o aprendizado, promovendo reflexão crítica, autonomia e motivação.
 - *Feedback* corretivo ajuda no desenvolvimento de habilidades específicas.
- Integração entre ensino e avaliação
 - Avaliações alinhadas aos objetivos de ensino apresentam melhores resultados.
 - A familiarização dos alunos com o formato avaliativo melhora o desempenho.
- Efetividade de métodos alternativos
 - Rubricas, autoavaliação, avaliação por pares e portfólios promovem maior autonomia e transparência.
 - Métodos dinâmicos e formativos superam práticas tradicionais em muitos contextos.
- Fatores contextuais e psicológicos
 - Diversidade cultural e socioeconômica impacta práticas avaliativas.
 - Ansiedade, motivação e outros fatores emocionais influenciam o desempenho.

6.5.2 Diferenças

- Tipos de avaliação
 - Preferência por avaliações presenciais em contextos tradicionais vs. maior aceitação de avaliações *on-line* em ambientes inovadores.
 - Divergências entre personalização (foco no aluno) e padronização (equidade e consistência).
- Abordagem por população
 - Estudos variam desde minorias (indígenas, alunos com deficiência) até contextos multilíngues e diversificados.
- Resultados de intervenções
 - Alguns métodos mostram impacto significativo, enquanto outros têm resultados limitados.
- Objetivos avaliativos
 - Foco em competências específicas, como leitura (estratégias e compreensão lexical) e produção textual (estrutura hipertextual).

6.5.3 Tendências mais comuns

- Formação docente contínua e suporte pedagógico;
- Uso de *feedback* formativo e corretivo para melhora de habilidades;
- Impacto positivo de práticas alternativas e dinâmicas na avaliação;
- Inclusão de fatores culturais e emocionais no *design* avaliativo.

6.5.4 Tendências menos comuns

- Estudos detalhados sobre populações vulneráveis (minorias culturais e socioeconômicas);
- Exploração de inovações digitais no ensino e na avaliação;
- Análise de fatores estruturais, como políticas educacionais.

A análise dos 64 artigos revela um cenário multifacetado sobre práticas de avaliação e ensino de línguas, destacando avanços, desafios e lacunas. Um dos pontos centrais é a necessidade de formação docente contínua, visto que muitos professores carecem de preparo adequado para implementar práticas avaliativas reflexivas e eficazes. A integração entre ensino e avaliação aparece como uma prática fundamental, com ênfase na importância de alinhar atividades e avaliações aos objetivos educacionais.

Os métodos alternativos, como rubricas, autoavaliação, portfólios e avaliação por pares, emergem como práticas eficazes para promover autonomia e pensamento crítico. Esses métodos contrastam com avaliações tradicionais, que ainda predominam, mas muitas vezes carecem de equidade e flexibilidade. O *feedback*, quando construtivo e imediato, é destacado como uma ferramenta poderosa para o aprendizado, impactando positivamente o desenvolvimento de competências específicas, como leitura, escrita e habilidades orais.

Os aspectos contextuais, como fatores culturais, socioeconômicos e emocionais, também são recorrentes. A globalização e a diversidade linguística exigem instrumentos avaliativos adaptados, enquanto questões como ansiedade e motivação dos alunos afetam diretamente o desempenho. Em contextos multilíngues e de populações vulneráveis, os estudos apontam para a necessidade de maior inclusão e personalização das avaliações.

Os resultados também mostram a eficácia de intervenções pedagógicas bem planejadas. Modelos mistos e inovações tecnológicas, embora menos comuns, têm demonstrado potencial

para enriquecer práticas avaliativas. Por outro lado, ainda há desafios na implementação de abordagens dinâmicas e no alcance de grupos minoritários.

Em síntese, o campo de avaliação em línguas aponta para uma transição gradual de métodos tradicionais para práticas mais formativas, inclusivas e centradas no aluno. A superação das lacunas identificadas exige maior investimento em formação docente, políticas educacionais integradas e inovações que considerem as diversidades culturais e emocionais dos alunos.

6.6 Análise quantitativa das tendências (1 a 64)

Nas subseções que se seguem, analiso, quantitativamente, as tendências avaliativas presentes nos 64 artigos. São elas:

6.6.1 Tendências gerais

- Formação docente e capacitação
 - 42 artigos destacam a importância da formação contínua e da melhoria das práticas avaliativas.
- Impacto do *feedback* no aprendizado
 - 38 artigos abordam o papel crucial do *feedback* formativo e corretivo no desenvolvimento de habilidades.
- Práticas alternativas de avaliação
 - 35 artigos mencionam métodos como autoavaliação, rubricas e avaliação por pares como mais eficazes do que os métodos tradicionais.
- Integração entre ensino e avaliação
 - 30 artigos enfatizam a necessidade de se alinharem avaliações às práticas e objetivos educacionais.
- Influência de fatores culturais e contextuais
 - 28 artigos discutem a forma como aspectos culturais, socioeconômicos e emocionais impactam as avaliações.
- Desempenho linguístico preocupante
 - 25 artigos identificam lacunas significativas nas habilidades de leitura, escrita e compreensão.

- Uso de tecnologias em avaliação
 - 12 artigos destacam a implementação de ferramentas digitais, como plataformas *on-line* e métodos automatizados.

6.6.2 Fatores específicos

- Autoavaliação e avaliação por pares:
 - 20 artigos relatam benefícios evidentes dessas práticas para promover autonomia e reflexão crítica.
- Avaliação formativa:
 - 18 artigos reforçam a eficácia de práticas formativas na melhoria do aprendizado.
- Desafios em contextos vulneráveis:
 - 15 artigos focalizam populações específicas, como indígenas, alunos com deficiências ou de baixa renda.
- Impacto de fatores emocionais (ansiedade, motivação):
 - 14 artigos analisam como fatores psicológicos influenciam o desempenho.

6.6.3 Métodos de avaliação

- Predominância de métodos tradicionais
 - 22 artigos criticam o uso excessivo de testes tradicionais, como múltipla escolha.
- Valorização de métodos personalizados
 - 16 artigos exploram a personalização e a adaptação das práticas avaliativas.
- Instrumentos de avaliação inovadores
 - 10 artigos sugerem abordagens híbridas ou tecnológicas, como o uso de rubricas digitais e estruturação hipertextual.

6.7 Desafios, limitações e recomendações

Nesta subseção, analiso os principais desafios e as limitações e recomendações provenientes dos artigos.

6.7.1 Semelhanças identificadas nos 64 artigos

- Falta de treinamento e letramento em avaliação
 - Apontada como uma barreira central na maioria dos estudos. Professores frequentemente carecem de formação específica para aplicar práticas avaliativas adequadas.
 - Resistência a métodos inovadores e preferência por modelos tradicionais (*e.g.* gramática-tradução) são comuns.
- Tamanho e representatividade das amostras
 - Estudos relatam dificuldade em generalizar resultados devido ao número reduzido de participantes e à homogeneidade amostral (cultural, socioeconômica ou de habilidades).
- Discrepância entre teoria e prática
 - A implementação prática muitas vezes não reflete os objetivos teóricos ou metodológicos declarados.
 - Uso de abordagens avaliativas que privilegiam quantitativos, mesmo em contextos qualitativos.
- Subjetividade e confiabilidade nos resultados
 - Avaliações influenciadas por vieses pessoais ou pela falta de padronização em rubricas e critérios.
 - Problemas de validade e confiabilidade dos instrumentos são frequentes.
- Desafios logísticos e institucionais
 - Recursos financeiros e limitações de tempo comprometem o uso de ferramentas e práticas avançadas.
 - Contextos institucionais específicos dificultam extrapolações para outros cenários.

6.7.2 Diferenças destacadas

- Desafios específicos ao contexto cultural:
 - Estudos realizados em países em desenvolvimento destacam problemas estruturais, como falta de padronização e preparo docente (*e.g.* Colômbia, mapuches no Chile).

- A influência da cultura dos participantes na recepção de práticas avaliativas também é mencionada.
- Dificuldades específicas de habilidades
 - Avaliação oral enfrenta desafios de espontaneidade, especialmente em contextos *on-line*.
 - Avaliação escrita lida com debates sobre o tipo e a eficácia do *feedback* correto.
- Questões éticas e inclusão
 - Menção rara, mas significativa, em relação a populações específicas, como crianças com hipoacusia e minorias linguísticas (*e.g.* falantes de mapudungun).
- Dinamismo em instrumentos inovadores
 - Estudos sobre rubricas mostram seu potencial, mas enfrentam resistência inicial e altos custos de implementação.
 - Dimensões linguísticas como fluência, precisão e complexidade são abordadas, mas ainda são difíceis de mensurar com precisão.

Os 64 artigos analisados trazem desafios recorrentes no campo da avaliação em ensino de línguas, mas também apresentam uma rica diversidade de abordagens contextuais. A falta de treinamento adequado para professores aparece como o desafio mais constante e transversal, impactando a qualidade e a eficácia das práticas avaliativas. A discrepância entre o que é proposto em teoria e o que é implementado na prática reflete dificuldades em alinhar metodologias inovadoras aos recursos e às realidades educacionais disponíveis.

Outro padrão significativo é a limitação metodológica relacionada ao tamanho das amostras e à falta de representatividade. Os estudos frequentemente se restringem a populações homogêneas ou a contextos locais específicos, dificultando a generalização dos resultados. Além disso, problemas na validade e confiabilidade de instrumentos avaliativos são recorrentes, destacando a necessidade de maior rigor no *design* e na aplicação de ferramentas.

As diferenças regionais e culturais também trazem nuances importantes. Nos países em desenvolvimento, desafios estruturais e econômicos aparecem com mais frequência, enquanto em contextos tecnologicamente avançados, o foco está na integração de ferramentas digitais e em dinâmicas como a perda de espontaneidade em avaliações orais. Questões éticas e inclusivas, embora menos mencionadas, enriquecem a discussão ao abordar populações marginalizadas ou com necessidades específicas.

Por fim, iniciativas que se utilizam de rubricas e avaliação de múltiplas dimensões do desempenho linguístico são promissoras, mas enfrentam barreiras iniciais de implementação. O equilíbrio entre inovação e aplicabilidade prática permanece um ponto de reflexão central.

6.7.3 *Ranking de tendências identificadas*

- Mais comuns
 1. Falta de treinamento ou formação em avaliação.
 2. Limitações em tamanho e representatividade da amostra.
 3. Discrepância entre teoria e prática.
 4. Problemas de confiabilidade e validade dos instrumentos.
 5. Desafios logísticos e institucionais.
- Menos comuns
 1. Questões éticas e justiça na avaliação.
 2. Populações específicas e minoritárias.
 3. Alta demanda cognitiva ou técnica.
 4. Influências culturais específicas.

As análises revelam que desafios como a falta de formação docente e limitações metodológicas ainda predominam nas discussões sobre avaliação em ensino de línguas. No entanto, aspectos menos explorados, como questões éticas e inclusivas, trazem perspectivas valiosas e mostram a complexidade desse campo de pesquisa. O futuro das práticas avaliativas requer um equilíbrio entre inovação, contexto e acessibilidade para atender às demandas educacionais globais.

6.7.4 *Semelhanças nas recomendações*

1. Capacitação e letramento em avaliação para docentes
 - Propostas incluem *workshops*, formação continuada e disciplinas específicas em currículos de graduação e pós-graduação.
2. Integração entre ensino e avaliação
 - Sugere que as práticas avaliativas sejam alinhadas aos objetivos pedagógicos e às necessidades dos alunos.

3. Exploração de contextos e populações diversas

- Ampliação para diferentes contextos multiculturais, populações minoritárias ou com necessidades específicas (e.g., crianças com hipoacusia ou comunidades indígenas).

4. Uso de ferramentas tecnológicas

- Inclui plataformas digitais, avaliações *on-line*, *podcasts*, hipertextos e alfabetização visual.

5. Abordagens qualitativas e diversificadas

- Destaque para avaliações autênticas, portfólios, rubricas e *feedback* colaborativo.

6. Foco na retroalimentação (*feedback*)

- Incentivo a um *feedback* mais eficaz, formativo e individualizado.

7. Preocupação com fatores emocionais

- Inclui redução de ansiedade, atenção ao bem-estar e consideração de fatores socioemocionais.

A análise das recomendações e sugestões para futuras pesquisas nos 64 artigos revela um panorama abrangente e diversificado sobre a avaliação no ensino de línguas. De maneira geral, há consenso sobre a necessidade de capacitação docente, abordagens avaliativas diversificadas e integração mais estreita entre avaliação e ensino.

A capacitação docente é um dos temas mais recorrentes, sendo mencionada frequentemente em estudos que reforçam a necessidade de formação continuada para melhorar a aplicação de instrumentos avaliativos. Além disso, o alinhamento entre objetivos pedagógicos e práticas avaliativas é amplamente recomendado como condição essencial para promover maior coesão no processo de ensino-aprendizagem.

Os estudos também sugerem uma transição gradual para abordagens mais qualitativas e autênticas, buscando-se reduzir a dependência de testes padronizados. Essa tendência propõe o uso de rubricas, *feedback* formativo e observações em sala como alternativas mais holísticas e significativas. Em paralelo, tecnologias emergem como ferramentas promissoras para avaliação, ainda que com desafios relacionados à infraestrutura e à desigualdade de acesso.

Por outro lado, diferenças importantes surgem no enfoque dos estudos. Enquanto alguns priorizam a validação técnica de instrumentos avaliativos, outros exploram aplicações práticas

em sala de aula. Populações específicas também são tema de investigações direcionadas, com alguns estudos analisando o impacto da avaliação em comunidades marginalizadas, bem como em crianças com necessidades especiais. Abordagens inovadoras, como o uso de *podcasts*, hipertextos digitais e plurinormatividade linguística, enriquecem ainda mais o debate, sugerindo novos caminhos para pesquisa.

Aspectos éticos e emocionais, embora menos frequentes, destacam-se como áreas de interesse emergente. A necessidade de se incorporarem questões éticas e se considerar o impacto emocional dos processos avaliativos é apontada em poucos estudos, mas representa um apelo para uma abordagem mais humana e inclusiva.

Em resumo, o campo da avaliação no ensino de línguas encontra-se em constante evolução, equilibrando avanços tecnológicos e metodológicos com preocupações sociais e éticas. As recomendações consolidadas nos 64 artigos fornecem uma base robusta para futuras pesquisas e práticas mais eficazes.

6.8 Pontos semelhantes e divergentes entre os artigos

6.8.1 Pontos semelhantes

1. Diversidade de métodos de avaliação

- Os estudos destacam o uso de diversas estratégias de avaliação, como testes tradicionais, avaliações alternativas (portfólios, autoavaliação e avaliação por pares) e instrumentos multimodais.
- Testes de produção oral, compreensão leitora e produção escrita são frequentemente priorizados para medir habilidades linguísticas específicas.

2. Uso de rubricas e *feedback*

- A aplicação de rubricas detalhadas para análises escritas e orais, como o WAS, promove avaliações objetivas e consistentes.
- O *feedback* corretivo e construtivo é amplamente valorizado como ferramenta para orientar o aprendizado e planejar intervenções pedagógicas.

3. Foco em avaliações formativas e somativas

- Combinações de práticas formativas (*e.g.* autoavaliação, dramatizações e *feedback*) e somativas (*e.g.* provas finais e testes padronizados) aparecem frequentemente nos estudos.

4. Capacitação de professores

- *Workshops*, questionários e atividades práticas visam a aprimorar as competências dos professores em letramento avaliativo e na criação de instrumentos de avaliação eficazes.

5. Integração tecnológica

- Plataformas digitais e rubricas *on-line* estão em expansão, comumente utilizadas em avaliações orais, análises de gravações e *feedback* mediado por tecnologia.

6.8.2 Pontos diferentes

1. Formato de avaliação

- Estudos com foco em gravações digitais e rubricas holísticas contrastam com abordagens baseadas em narrativas e interação oral.
- Alguns artigos evitam testes formais, priorizando a interação em sala de aula e o *feedback* corretivo.

2. Objetivo da avaliação

- Diferença entre medir proficiência geral (*e.g.* TOEFL) e habilidades específicas, como leitura ou aspectos morfológicos.

3. Foco em populações específicas

- Avaliações voltadas para crianças em idade pré-escolar, falantes de línguas indígenas, como o mapudungun, e estudantes com hipoacusia exigem adaptações culturais e pedagógicas.

4. Modalidade de avaliação

- O uso de textos multimodais, integrando elementos visuais e audiovisuais, difere das práticas centradas exclusivamente em textos impressos.

5. Atenção a aspectos não-linguísticos

- Estudos sobre ansiedade exploram o impacto emocional do processo avaliativo no desempenho acadêmico.

6.8.3 Pontos interessantes

1. Avaliações alternativas

- Autoavaliação, avaliação por pares e portfólios são práticas crescentes que promovem autonomia e reflexão crítica.

2. Inovação tecnológica

- A inclusão de ferramentas como *podcasts*, plataformas de conferência e instrumentos interativos reflete a busca por métodos mais dinâmicos.

3. Competência intercultural

- Avaliar habilidades interculturais é uma abordagem inovadora que considera aspectos emocionais e sociais no aprendizado.

4. Ênfase em avaliações multimodais

- A combinação de leitura com análise de vídeos, imagens e outros elementos amplia o escopo tradicional das avaliações.

5. Adaptações regionais

- Estudos sobre populações específicas, como falantes de línguas indígenas, destacam a necessidade de metodologias culturalmente sensíveis.

6.9 Comparações entre os resultados desta tese e Borja (2020a)

No recorte de artigos nacionais sobre avaliação (de Qualis A1) na Plataforma Sucupira de Borja (2020a), a região geográfico-política do sudeste do Brasil predominou nas pesquisas, representando mais da metade de todos os artigos publicados na área (sendo São Paulo o estado responsável por 50% das publicações) no período (1985 a 2018) e no contexto em questão. Na presente pesquisa, em contexto latino-americano amplo, a grande maioria das pesquisas é chilena (32 artigos) e colombiana (28). Em ambos os recortes (nacional e internacional), a quantidade de artigos sobre avaliação foi reduzida, sendo 38 em âmbito nacional (ao longo de 33 anos) e 64 em âmbito latino-americano amplo (não incluindo artigos brasileiros), ao longo de 29 anos. Vale ressaltar que a pesquisa de mestrado abarcou artigos que versavam sobre avaliação no ensino de línguas adicionais/estrangeiras, ao passo que a presente pesquisa,

visando a ampliar os resultados, abordou vários países e o ensino de língua (materna ou adicional) como um todo.

Em Borja (2020a), 38 artigos foram analisados no total, advindos de sete revistas; 34 escritos em língua portuguesa, e quatro, em língua inglesa. As revistas selecionadas e o respectivo número de artigos encontrados foram: *Trabalhos em Linguística Aplicada* (22 artigos), *Revista Brasileira de Linguística Aplicada* (8 artigos), *Linguagem & Ensino* (3 artigos), *D.E.L.T.A.* (2 artigos), *Ilha do Desterro* (1 artigo), *Alfa* (1 artigo), *Linguagem em (Dis)curso* (1 artigo).

A quantidade de publicações ao longo dos anos não apresentou crescimentos significativos no recorte nacional. Os anos de produção mais volumosa foram: 1996, 2004, 2006, 2014, 2015. No entanto, os anos de publicação mais robusta no corpus da presente pesquisa foram 2019, 2021 e 2022. Dentre os artigos selecionados no presente estudo, 28 advêm do periódico colombiano *Profile: Issues in Teachers' Professional Development*; 12, do periódico chileno *Literatura y Linguística*; 11, da revista chilena *Signos*; 3, do periódico chileno *Logos: Revista de Linguística, Filosofía y Literatura*; 2, do periódico venezuelano *Lengua y Habla*; 2, do periódico chileno *Lenguas Modernas*; 2, do periódico chileno *RLA*; 1, do periódico chileno *Nueva Revista del Pacífico*; 1, do periódico chileno *Estudios Filológicos*; 1, do periódico argentino *Revista Argentina de Ciencias del Comportamiento*; e 1, do periódico venezuelano *Boletín de Linguística*.

A grande incidência de artigos acerca de proficiência em língua adicional identificada em Borja (2020a) (totalizando 15 artigos) também foi identificada no presente trabalho, com testes como o TOEFL e Lectum e outros testes de proficiência citados em diversas pesquisas, medindo habilidades específicas como as quatro macro-habilidades ou até mesmo outras nuances do campo linguístico. Grande número de estudos (cerca de 50%) analisa estudantes do nível A2 ao B2, de acordo com o CEFR, e em línguas como inglês, espanhol e francês.

De um modo geral, em todos os instrumentos analisados no *corpus* do presente trabalho (e não só os que versam sobre proficiência), as principais competências avaliadas (por ordem de recorrência) foram a produção escrita, produção oral, compreensão escrita, compreensão oral e letramento em avaliação. Vale destacar que avaliar o letramento em avaliação representou um avanço e trouxe um novo panorama em relação ao encontrado em Borja (2020a). No entanto, a produção oral seguiu representando grande parte do foco das pesquisas também no presente trabalho.

Em Borja (2020a), dentre os temas mais recorrentes nos artigos analisados, podem ser citados: proficiência em língua adicional, o efeito retroativo (associado ao vestibular ou não) e

questões da prova de língua estrangeira do vestibular da Unicamp. No entanto, exames de admissão em universidades não representaram o foco do *corpus* latino-americano. Por outro lado, o público universitário foi majoritário em ambos os *corpora*, em detrimento do contexto de ensino básico, que representa a base da educação. O público universitário foi abordado em 60% das pesquisas do presente trabalho, com foco em alunos entre 18 e 30 anos e em áreas como pedagogia, línguas adicionais e até programas de intercâmbio.

Com relação a tendências mais comumente abordadas no *corpus* latino-americano em comparação ao *corpus* nacional, a adoção de métodos alternativos ganhou mais espaço, como portfólios, rubricas e avaliação por pares, trazendo maior clareza nos critérios avaliativos. No entanto, os testes tradicionais padronizados e de múltipla escolha seguem como fortes tendências. Nota-se um movimento de mudanças graduais rumo a avaliações mais centradas no aluno, com valorização do *feedback* formativo. Outro fator crucial também comumente abordado nos artigos latino-americanos foi a conscientização sobre a falta de capacitação docente. Alguns docentes mostraram-se mais cientes da falta de capacitação em avaliação ao longo de suas trajetórias, o que já demonstra um fator propulsor de futuras mudanças.

As pesquisas analisadas abarcam uma rica gama de tipos de avaliação no ensino de línguas, seja através de práticas tradicionais ou alternativas. Vale ressaltar que houve por parte de determinados estudos, a tentativa de equilibrar avaliações formativas e somativas, com o uso de instrumentos como testes padronizados, bem como a implementação de rubricas para a produção escrita e oral ou, até mesmo, dramatizações.

A integração da tecnologia pode apontar uma tentativa de modernização das práticas avaliativas, como plataformas digitais, *podcasts*, rubricas *on-line*, gravações digitais, instrumentos interativos e textos multimodais, ampliando o escopo tradicional das avaliações. No entanto, métodos tradicionais seguem fortemente presentes, como provas de múltipla escolha e avaliações estruturadas.

A autonomia do aluno foi promovida através da adoção crescente da autoavaliação, portfólios e avaliação por pares, demonstrando uma diluição do papel normalmente autoritário das avaliações e dando voz aos avaliados. Desafios culturais também foram tema de estudos sobre práticas avaliativas, como evidenciado nos estudos sobre falantes de línguas indígenas e crianças com necessidades específicas. O histórico de cada aluno e suas bagagens devem receber maior atenção; no entanto, apesar de serem mencionados, não houve um movimento considerável.

Finalmente, a capacitação docente foi comumente abordada com preocupação crescente em aprimorar o letramento em avaliação, por meio de *workshops* e questionários, visando a

preparar professores para o desenho de instrumentos eficazes. Apesar de nem todos os professores estarem abertos a capacitações longas e teóricas sobre avaliação, muitos reconhecem a importância e estão mais cientes da sua importância, com preferência para cursos práticos. Acredito ser um movimento relativamente recente, tendo em vista que a avaliação nem sempre foi questionada por parte de professores, que estavam estagnados com suas práticas, ainda que não tivessem recebido qualquer tipo de treinamento avaliativo.

Quanto à população avaliada, assim como a pesquisa realizada em Borja (2020a), o público universitário (com idades entre 18 e 30) segue recebendo grande atenção dos estudos na área.

A diversidade de níveis de proficiência também pôde ser notada, já que, quando o Quadro Europeu Comum foi mencionado, algumas pesquisas abordaram alunos principalmente entre os níveis A2 e B1. A diversidade também se estendeu para os contextos educacionais, como a educação bilíngue, o ensino superior, o contexto empresarial e a formação de professores.

Os professores em questão, majoritariamente experientes, com vasta experiência em seus currículos (entre 3 e 17 anos de atividade docente), são majoritariamente mestres ou doutores.

A diversidade regional também enriqueceu o *corpus* com populações analisadas advindas da Colômbia, do Chile, do Irã, de Omã, do México e da Espanha, bem como estudantes de intercâmbio em contexto universitário. Contextos urbanos e rurais também foram foco das pesquisas, englobando diferentes grupos etários (crianças, adolescentes ou adultos), além de alunos com necessidades específicas.

Quanto à gama de competências avaliadas, relacionadas ao ensino de línguas, as mais frequentes foram as quatro macro-habilidades: produção oral, produção escrita, compreensão oral e compreensão escrita. Assim como observado em Borja (2020), a avaliação da produção oral e de suas particularidades/desafios recebeu destaque, por ser abordada em diversos estudos, com ênfase em fluência, pronúncia e interação. A produção escrita também recebeu atenção significativa, incluindo textos argumentativos e narrativos, com foco em coesão, coerência e correção gramatical. A compreensão leitora foi comumente dividida em níveis, como inferência, literal e crítica, além de integrações multimodais envolvendo texto e imagem. A compreensão oral, com frequência menos robusta, foi inserida em contextos como a avaliação de crianças mapuches e alunos hipoacústicos.

Competências diversas, como o letramento em literatura e a habilidade de falar em público, são menos frequentes, mas representam contribuições importantes que fogem de temas

comumente abordados em diversos estudos. Outros estudos também exploram habilidades fonológicas, cognição e aspectos emocionais, como a ansiedade em avaliações. O letramento em avaliação também recebeu destaque e promoveu a “conscientização da falta” em diversos docentes, o que já é um ponto de partida considerável: reconhecer a falta de formação. Os professores, com frequência considerável, reconhecem a importância da constante formação docente em práticas avaliativas, para que possam oferecer experiências avaliativas mais inclusivas e diversas. Outro ponto a se considerar é o alinhamento entre as práticas em sala de aula com os instrumentos avaliativos, para que as avaliações sejam condizentes com o que, de fato, é praticado em sala de aula.

A autonomia dos alunos pôde ser considerada em diversos estudos que abordaram autoavaliação, avaliação por pares, portfólios e rubricas. Acarretando maior participação dos alunos no processo avaliativo, diluindo a ansiedade por estarem cientes dos critérios e trazendo luz para o papel ativo do aluno no processo avaliativo como um todo. Dessa forma, a equidade é priorizada, bem como a formação de alunos mais críticos e participativos.

O campo de avaliação em ensino de línguas, no recorte em questão, apontou para uma transição gradual de métodos tradicionais para práticas mais formativas, inclusivas e centradas no aluno. A superação das lacunas identificadas exige maior investimento em formação docente, bem como políticas educacionais e a conscientização de fatores como a diversidade cultural e as necessidades específicas dos alunos.

No que concerne aos desafios enfrentados nos estudos em contexto de avaliação em ensino de línguas, a falta de treinamento e capacitação adequados para professores representou o desafio mais constante e transversal, impactando consideravelmente a experiência avaliativa. Há também um distanciamento entre práticas avaliativas e práticas concretas em sala de aula, assim como práticas avaliativas inovadoras que se deparam com desafios de escassez de recursos no contexto educacional.

Outro padrão significativo é a limitação relacionada ao tamanho das amostras. Os estudos frequentemente se restringem a populações homogêneas ou a contextos locais específicos, dificultando a generalização dos resultados.

As diferenças regionais, econômicas e culturais também apontam para nuances relevantes. Em lugares que enfrentam alguma situação de vulnerabilidade, desafios estruturais e econômicos aparecem com mais frequência, ao passo que, em contextos tecnologicamente avançados, o foco está na integração de ferramentas digitais e em dinâmicas como a perda de espontaneidade em avaliações orais. Questões éticas e inclusivas, apesar de pouco

mencionadas, enriquecem o *corpus*, ao abordarem populações marginalizadas ou com necessidades específicas.

Em última análise, iniciativas como a adoção de rubricas e demais práticas menos recorrentes/disseminadas enfrentam resistência inicial em seus contextos, seja por representarem práticas avaliativas desconhecidas, seja pelo tempo dedicado ao início até que os docentes se familiarizem, seja pelos seus custos.

Por fim, no que diz respeito às recomendações para futuras pesquisas, há um consenso sobre a necessidade de capacitação docente quanto a práticas avaliativas e diversidade nas avaliações e nos instrumentos, além de integração mais estreita entre avaliação e práticas docentes em sala de aula.

A capacitação docente representou um dos temas mais frequentemente abordados, que enfatizando a necessidade de formação continuada para promover melhora nas práticas avaliativas. Além disso, o alinhamento entre os objetivos pedagógicos e as práticas avaliativas, outro tema recorrente, foi abordado como uma condição essencial para promover maior coesão no processo de ensino-aprendizagem-avaliação.

Os estudos também sugerem uma transição gradual para abordagens mais qualitativas e autênticas. Através de implementações como a adoção de rubricas, *feedback* formativo e observações em sala, busca-se reduzir a dependência de testes padronizados. Tais exemplos de possibilidades de práticas avaliativas, aliadas também a tecnologias emergentes, proporcionam experiências avaliativas mais significativas e promissoras, apesar da possível resistência inicial na implementação e da desigualdade de acesso.

7 CONSIDERAÇÕES FINAIS

Ler sobre Avaliação atrelada a países diversos ampliou os meus horizontes para práticas avaliativas distintas. Professores, coordenadores e diretores podem se beneficiar grandemente dos resultados dessa leitura acerca dos artigos acadêmicos latino-americanos e suas contribuições sobre avaliação.

Através do *corpus* copilado na presente tese, assim como em nível de mestrado em Borja (2020), novas reflexões podem ser desencadeadas, seja através da reflexão sobre a importância da capacitação docente constante, seja através de avaliações mais acolhedoras ou das contribuições de “nossos vizinhos” e suas práticas avaliativas diversas.

Nas minhas hipóteses, levantadas nas páginas iniciais desta tese, questionei o fato de a avaliação possivelmente não ser um tema recorrente nos periódicos latino-americanos também, bem como não o foi no recorte nacional de Borja (2020); o fato de as discussões encontradas nos artigos latino-americanos darem ênfase aos testes externos (do Norte Global); e até que ponto os principais pareceres oficiais sobre avaliação no ensino de línguas podem influenciar as principais pesquisas e práticas do país em questão. Por último, também me interessou saber quais as principais tendências de pesquisa em avaliação e se houve alguma mudança significativa no direcionamento das principais pesquisas da área, comparando-se o *corpus* latino-americano ao *corpus* nacional de Borja (2020).

A primeira hipótese levantada, de que o volume de pesquisas sobre avaliação seria reduzido, foi confirmada. O *corpus* teve volume de 64 artigos latino-americanos publicados em quatro países e no período de 29 anos (entre 1994 e 2023). Considerando-se a quantidade de países originalmente abarcados no *ranking* da SCImago, o volume de pesquisas sobre avaliação no ensino de línguas foi considerado reduzido. As pesquisas brasileiras não foram o foco do presente trabalho, o que poderia ter alterado significativamente o volume de artigos encontrados, caso tivessem sido incluídas. Ao todo, 11 periódicos foram considerados, advindos de quatro países (Colômbia, Chile, Venezuela e Argentina). Os anos de 2014, 2016, 2018, 2019, 2020, 2021 e 2022 foram os anos de maior volume de publicações na área para o recorte em questão, o que revela grande volume de publicações atuais. A introdução de tecnologias digitais e a ocorrência da pandemia podem ter trazido novos interesses e inquietações para pesquisas na área de avaliação da aprendizagem no ensino de línguas. O ano de 2022 foi o de maior volume de publicações, e acredito que o período pós-pandêmico trouxe um novo panorama de pesquisas, devido à urgência de se adaptarem as formas de avaliação até então pouco questionadas. Possivelmente, no futuro, novas pesquisas sobre avaliação poderão

ter foco em temas como a inteligência artificial. Vale ressaltar que pode levar um tempo até que fenômenos mundiais passem a influenciar as tendências investigativas.

Quanto à segunda hipótese, de que o Norte Global e suas formas de avaliar seriam o foco, a hipótese foi parcialmente confirmada. Questões como o Quadro Europeu Comum e testes de proficiência também foram comuns no *corpus* latino-americano, mas acompanhados também de outras pesquisas e inovações. Portanto, novas formas de avaliar já estão recebendo mais visibilidade, o que pode mostrar uma possível ampliação de possibilidades de pesquisa. No entanto, 31 artigos foram escritos em língua inglesa, e 33, em língua espanhola. Tal fato nos mostra um provável foco em atingir o Norte Global com nossas contribuições acadêmicas. As referências bibliográficas utilizadas pelos artigos do *corpus* também se dividiram em língua inglesa e espanhola, o que mostra um distanciamento de nossos teóricos brasileiros da área de avaliação, como Scaramucci (2000), Luckesi (2000) e Méndez (2002). O mesmo ocorreu com os resultados analisados no *corpus* brasileiro (Borja, 2020), sendo poucas as referências utilizadas em língua espanhola. Vale ressaltar que grande parte da nossa produção acadêmica é de acesso livre e, em determinadas universidades americanas ou europeias, o acesso a artigos é pago, bem como se paga para publicar, o que poderia nos levar ainda mais para um caminho de valorização das contribuições de “nossos vizinhos”. Dificilmente resumos de pesquisas brasileiras são escritos em espanhol, da mesma forma que pesquisas de países hispanofalantes não colocam resumos em português. A proximidade geográfica parece não contribuir e seguimos com complexo de extrema valorização do Norte Global em detrimento de nós mesmos e de nossos colegas/compatriotas latino-americanos. A valorização exclusiva das contribuições do Norte Global é um processo que pode levar anos para ser desconstruído. Defendo que a valorização equitativa de todos os saberes e contribuições, independentemente de sua origem geográfica, constitui um cenário profícuo para todos os envolvidos.

Compreendo a avaliação da aprendizagem no ensino de línguas materna e adicional, bem como em nível mais abrangente, como uma prática sistemática para a coleta de informações sobre o progresso dos alunos, visando a intervenções posteriores de forma acolhedora e qualitativa. Nessa perspectiva, os instrumentos aplicados podem e idealmente serão diversos, considerando distintos instrumentos distribuídos ao longo da jornada educacional. Testes, portfólios, debates, entrevistas e demais formas são consideradas válidas e, se aplicadas ao longo de todo o processo, podem obter dados concretos sobre o processo de aprendizagem do aluno, com vistas a poder guiar melhor o seu trajeto, graças à riqueza de instrumentos.

Os resultados dos instrumentos avaliativos devem ser acompanhados de *feedback* e intervenção acolhedoras para a superação de obstáculos anteriores. Mudanças na ementa devem ser incentivadas, dado que os programas não são fixos e as avaliações trazem recursos para possíveis mudanças. Todo o contexto educacional deve ser considerado em suas especificidades, e todo o corpo de profissionais da escola deve ser sensibilizado sobre questões avaliativas. Tendo em vista que comumente professores não foram incentivados a estudar e questionar as práticas avaliativas aplicadas, é de suma importância que programas de capacitação sejam conduzidos, com vistas a oferecer cursos teóricos e práticos sobre avaliação, facilitando, assim, a rotina do professor, e promovendo um ambiente mais acolhedor aos alunos. Caso a capacitação contínua em avaliação não ocorra, é possível que práticas avaliativas a que os professores foram submetidos quando eram alunos, sejam ainda praticadas sem questionar o contexto atual em que trabalham.

Frequentemente, os professores são considerados competentes com base na quantidade de alunos reprovados e na complexidade de seus instrumentos avaliativos, o que reflete uma visão limitadora e pouco inclusiva. Para promover um perfil questionador nos seus alunos, o docente pode acolher e escutar a voz dos alunos submetidos aos instrumentos avaliativos, ouvindo suas angústias e dificuldades e promovendo um canal aberto para questionamento de conteúdos e práticas avaliativas. A validade, ética e justiça devem perpassar toda a prática avaliativa docente, levando em consideração o efeito retroativo causado por aquele instrumento em curto e em longo prazo.

Acredito ser importante reconhecer a realidade regional do contexto educacional no qual alunos serão avaliados, bem como estar aberto a conhecer práticas avaliativas praticadas em outros contextos, que possivelmente contribuirão para novas formas de se avaliar. A leitura de artigos, dissertações, teses e pesquisas acadêmicas trazem um panorama amplo de como as avaliações são compreendidas ao redor do mundo. Sendo importante considerar cada realidade, bem como países diversos. Ler sobre avaliações atreladas ao uso de tecnologias em países em guerra abriu os meus horizontes para práticas avaliativas que só teríamos conhecimento prático no Brasil com a pandemia causada pelo covid-19 (Abduh; Borja; Esqueda, 2023). Professores, coordenadores, diretores se beneficiariam grandemente da leitura dos principais periódicos acadêmicos ao redor do mundo e suas contribuições sobre avaliação. Acredito que através do *corpus* copilado na presente pesquisa (assim como em nível de mestrado em Borja 2020), novas reflexões poderão ser desencadeadas, seja através da reflexão sobre a importância da capacitação docente constante, sobre avaliações mais acolhedoras, as contribuições de “nossos vizinhos” e da possibilidade de práticas avaliativas diversas.

Vale ressaltar que, dentre os artigos selecionados para o *corpus*, advindos de periódicos latino-americanos, pesquisadores que estudaram e trabalharam em diversos países publicaram seus artigos. Países como Chile, Brasil, Índia, Colômbia, Guiana, Irã, Argentina, Estados Unidos, Espanha, Venezuela, Holanda, México, Equador, Hungria e Turquia fazem parte da lista. Os pesquisadores do Sul representam a parte majoritária dos pesquisadores; no entanto, alguns dos artigos publicados são de coautoria entre pesquisadores que estudaram no Sul e outros do Norte, não deixando clara a nacionalidade exata de cada um, apenas onde cada um estudou ou trabalhou. Logo, vozes da Índia e do Irã também enriqueceram as contribuições do *corpus*. Embora esta pesquisa não tenha analisado o perfil étnico-racial, de gênero ou a localização epistêmica dos autores, reconheço que a ausência de vozes de grupos minoritários é um dado que comunica e evidencia desafios para uma ciência plural e situada. Investigações futuras podem aprofundar essa temática para ampliar a crítica à colonialidade do saber.

No cenário latino-americano, observa-se uma diversidade de práticas e iniciativas. Na Colômbia, por exemplo, o Marco Nacional de Qualificações (2023) promove uma trajetória de aprendizagem flexível e diversificada. O Sistema Institucional de Avaliação dos Estudantes (SIEE) enfatiza que a avaliação deve ser um processo articulado e coerente com a proposta educativa da instituição. Durante a pandemia, o programa Avaliar para Avançar foi implementado com foco em três etapas: avaliação formativa, suporte aos professores e mentoria pedagógica. Apesar disso, muitos documentos oficiais ainda apresentam pareceres rígidos e superficiais, ignorando os impactos da pandemia e das tecnologias digitais, como o ChatGPT, nas formas de avaliação.

No Chile, o *site* oficial Currículum Nacional disponibiliza recursos abrangentes para a implementação de práticas avaliativas, incluindo exemplos de avaliação formativa e um banco de itens adaptáveis. O portal também destaca o caráter processual da avaliação, orientando as práticas docentes com perguntas-chave, como: “Para onde vamos?”, “Onde estamos?” e “Como seguimos?”. Essa abordagem colabora para a formação integral dos estudantes e facilita o trabalho docente.

A maioria dos artigos foi escrita em espanhol ou em inglês, indicando um esforço para alcançar leitores do Norte Global. Contudo, autores brasileiros, como Méndez (2002) e Scaramucci (2000), foram pouco referenciados, evidenciando o distanciamento acadêmico entre o Brasil e seus vizinhos sul-americanos. É essencial encurtar essas distâncias, promovendo um maior intercâmbio de ideias e referências entre pesquisadores da região.

A avaliação da aprendizagem no ensino de línguas, portanto, precisa ser constantemente repensada, com práticas que desafiem estruturas tradicionais e promovam justiça, transparência

e inclusão. Reconhecer as contribuições regionais e dialogar com os desafios globais é um passo fundamental para avançar na área.

REFERÊNCIAS

- ABDUH, M. Y. M.; BORJA, B. C.; ESQUEDA, M. D. Avaliação on-line em tempo integral durante a pandemia da covid-19: percepções dos professores de inglês como língua estrangeira. *Letras & Letras*, Uberlândia, v. 39, p. 1–27, 2023. DOI: 10.14393/ll63-v39-2023-01. Disponível em: <https://seer.ufu.br/index.php/letraseletras/article/view/66837/35883>. Acesso em: 15 abr. 2024.
- ASSESSMENT. In: RICHARDS, J. C.; SCHMIDT, R. *Longman: dictionary of language teaching and applied linguistics*. 3. ed. Abingdon: Routledge, 2002. p. 35.
- BARDIN, L. *Análise de conteúdo*. Tradução de Luís Antero Reto e Augusto Pinheiro. São Paulo: Casa de Ideias, 2011.
- BAKHTIN, M. *Marxismo e filosofia da linguagem*. 6. ed. São Paulo: Hucitec, 1992. 196 p.
- BORJA, B. C. *Avaliação no ensino de línguas estrangeiras: análise bibliométrica de artigos publicados em periódicos nacionais (1985 a 2018)*. 2020. 64 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Uberlândia, Uberlândia, 2020a. Disponível em: <http://doi.org/10.14393/ufu.di.2020.3015>. Acesso em: 11 jun. 2022.
- BORJA, B. C. Avaliação no ensino de línguas estrangeiras: análise bibliométrica de artigos publicados em periódicos nacionais (de 1985 a 2018). In: ESQUEDA, M. D. (Org.). *Estudos bibliométricos e cienciométricos em tradução: tendências, métodos e aplicações*. Curitiba, PR: Editora CRV, 2020b. p. 161–177.
- CHILE. Ministerio de Educación. *Evaluación*. Currículum Nacional, Chile. Disponível em: <https://www.curriculumnacional.cl/portal/Evaluacion/>. Acesso em: 10 abril. 2024.
- COLÔMBIA. Ministerio de Educación Nacional. Sistema Institucional de Evaluación de los Estudiantes -SIEE-. In: MINISTERIO DE EDUCACIÓN NACIONAL (Colômbia). *Evaluación*. Bogotá: Ministerio de Educación Nacional, 23 maio 2023. Disponível em: <https://www.mineduacion.gov.co/portal/micrositios-preescolar-basica-y-media/Evaluacion/Evaluacion-de-estudiantes/397381:Sistema-Institucional-de-Evaluacion-de-los-Estudiantes-SIEE>. Acesso em: 18 jul. 2024.
- DAY, K. Ensino de língua estrangeira no Brasil: entre a escolha obrigatória e a obrigatoriedade voluntária. *Escrita*, Rio de Janeiro, v. 15, n. 15, p. 1-13, 2012, . DOI: <https://doi.org/10.17771/PUCRio.escrita.20850>. Disponível em: <https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=20850&idi=1>. Acesso em: 15 abr. 2024.
- FREIRE, P. *Pedagogia do oprimido*. 17. ed. Rio de Janeiro: Paz e Terra, 1994.
- FUNDAÇÃO LEMANN. 94% dos estudantes mudaram o comportamento na pandemia. In: FUNDAÇÃO LEMANN. *Institucional*. São Paulo: Fundação Lemann, 2 ago. 2021. Disponível em: <https://fundacaolemann.org.br/noticias/94-dos-estudantes-mudaram-o-comportamento-na-pandemia/>. Acesso em: 20 maio 2025.
- HARMER, J. *The practice of English language teaching*. 3. ed. Londres: Pearson Education ESL, 2001.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Exame Nacional do Ensino Médio (Enem). In: IBGE. *Comitê de Estatísticas Sociais*. Brasília: IBGE, 2011. Disponível em: <https://ces.ibge.gov.br/base-de-dados/metadados/inep/exame-nacional-do-ensino-medio-enem.html>. Acesso em: 12 out. 2020.

LUCKESI, C. C. O que é mesmo o ato de avaliar a aprendizagem?. *Pátio*, Porto Alegre, ano 3, n. 12, fev.-abr. 2000.

MATOS, D. C. V. da S.; SANTOS, J. A. Olhares suleados à educação linguística do espanhol. *Revista Linguagem & Ensino*, Pelotas, v. 25, n. 2, p. 109–129, 2022.

MENDES, R. M.; MISKULIN, R. G. S. A análise de conteúdo como uma metodologia. *Cadernos de pesquisa*, São Paulo, v. 47, n. 165, 2017.

MÉNDEZ, J. M. A. *Avaliar para conhecer, examinar para excluir*. Porto Alegre: Artmed, 2002.

MENIADO, J. The Impact of ChatGPT on English Language Teaching, Learning, and Assessment: A Rapid Review of Literature. *Arab World English Journal*, [s. l.], v. 14, n. 4, p. 3–18, dec. 2023. DOI: 10.24093/awej/vol14no4.1. Disponível em: <https://awej.org/the-impact-of-chatgpt-on-english-language-teaching-learning-and-assessment-a-rapid-review-of-literature/>. Acesso em: 14 jul. 2024.

PERU. Ministerio de Educación. *Marco Curricular Nacional*: propuesta para el diálogo: segunda versión. Lima: Ministerio de Educación, 2014. 89 p.

PONTES, Marco Aurélio Costa. *Subversão e/ou alinhamento ao colonialismo digital*: uma pesquisa sobre o Grupo de Educadores Google (GEG) em sua complexidade. 2024. Tese (Doutorado em Estudos Linguísticos) – Instituto de Letras e Linguística, Universidade Federal de Uberlândia, Uberlândia, 2024.

RANKING da educação: Brasil está nas últimas posições no Pisa 2022; veja notas de 81 países em matemática, ciências e leitura. In: G1. *Educação*. [s. l.]: Globo Comunicação e Participações S.A., 5 dez. 2023. Disponível em: <https://g1.globo.com/educacao/noticia/2023/12/05/ranking-da-educacao-brasil-esta-nas-ultimas-posicoes-no-pisa-2022-veja-notas-de-81-paises-em-matematica-ciencias-e-leitura.ghtml>. Acesso em: 17 jul. 2024.

RODRIGUES, B. C. *et al.* Um olhar sobre a prática avaliativa de língua inglesa. *Via Litterae*, Anápolis, v. 3, n. 1, p. 5-21, jan./jun. 2011. Disponível em:

<https://www.revista.ueg.br/index.php/vialitterae/article/view/5337>. Acesso em: 11 out. 2020.

RUIZ, Jorge; JORGE, Ricardo Arencibia. Infometría, bibliometría y cienciometría: aspectos teórico-prácticos. *ACIMED*, Ciudad de la Habana, v. 10, n. 4, jul.-ago. 2002.

SANTOS, E. M. O ensino do PLE/PLA no Projeto Destino Brasil: preparação de materiais a partir de uma atitude decolonial. *Linguagem & Ensino*, Pelotas, v. 25, n. 2, p. 130-149, maio-ago. 2022.

SCARAMUCCI, M. V. R. Proficiência em LE: Considerações terminológicas e conceituais. *Trab. Ling. Apl.*, Campinas, n. 36, p. 22-22, jul.-dez. 2000.

SCARAMUCCI, M. V. R. Validade e consequências sociais das avaliações em contextos de ensino de línguas. *Lingvarvm Arena*, Porto, v. 2, p. 103-120, 2011.

SCIMAGOJR. Página inicial. In: SCIMAGO LAB. *ScimagoJR*. Scimago Lab: Madrid, 2023. Disponível em: www.scimagojr.com. Acesso em: 16 fev. 2023.

SERAJ, P. M. I. *et al.* A Systematic Review on Pedagogical Trends and Assessment Practices during the COVID-19 Pandemic: Teachers' and Students' Perspectives. *Education Research International*, [s. l.], v. 2022, p. 1–10, abr. 2022. DOI: <https://doi.org/10.1155/2022/1534018>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1155/2022/1534018>. Acesso em: 14 jul. 2024.

SHOHAMY, E. *The Power of Tests: a critical perspective on the uses of language tests*. Harlow: Person Education, 2001.

SOUZA, L. M. T. M. de; HASHIGUTI, S. Decolonialidade e(m) Linguística Aplicada: uma entrevista com Lynn Mario Trindade Menezes de Souza. *Polifonia*, Cuiabá, v. 29, n. 53, p. 1-262, jan./mar. 2022.

URUGUAI. Administración Nacional de Educación Pública. *Marco Curricular Nacional: documento preliminar en proceso de elaboración y consulta 2022*. Montevideo: ANEP, 2022. 58 p.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ci. Inf.*, Brasília, v. 31, n. 2, p. 152-162, maio-ago. 2002.

REFERÊNCIAS DO CORPUS

ABAD, J. V.; ALZATE, P. A. Strategies Instruction to Improve the Preparation for English Oral Exams. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 18, n. 1, p. 129–147, jan.-jul. 2016. DOI: <http://dx.doi.org/10.15446/profile.v18n1.49592>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902016000100008. Acesso em: 27 dez. 2024.

ALFARO CHAT, J.; GÓMEZ ÁLVAREZ, L.; SAEZ CARRILLO, K. Complejidad, precisión y fluidez en el desempeño oral de aprendices con distintos niveles de proficiencia en inglés como L2. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 51, n. 96, p. 3–23, 2018. DOI: 10.4067/S0718-09342018000100003. Disponível em: https://www.scielo.cl/scielo.php?pid=S0718-09342018000100003&script=sci_abstract. Acesso em: 27 dez. 2024.

ALMEIDA, E.; CASTILLO, D.; ROSERO, A. Assessment practices in some language centers in public Universities of Ecuador and student proficiency outcomes. *Literatura y Lingüística*, Santiago, n. 44, p. 241–262, 2021. DOI: 10.29344/0717621X.44.3017. Disponível em: https://www.scielo.cl/scielo.php?pid=S0716-58112021000200241&script=sci_abstract&tlng=en. Acesso em: 27 dez. 2024.

ALONQUEO BOUDON, P. et al. Mapudunguken: cuestionario para evaluar la competencia oral en mapudungun en niños mapuche. *Literatura y Lingüística*, Santiago, n. 46, p. 281–304, 2022. DOI: 10.29344/0717621X.46.2532. Disponível em: https://www.scielo.cl/scielo.php?pid=S0716-58112022000200281&script=sci_abstract. Acesso em: 27 dez. 2024.

AMENGUAL-PIZARRO, M.; GARCÍA-LABORDA, J. Analysing Test-Takers' Views on a Computer-Based Speaking Test. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 19, p. 23–38, 2017. Supl. 1. DOI: http://dx.doi.org/10.15446/profile.v19n_sup1.68447. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/68447>. Acesso em: 27 dez. 2024.

ANDIÓN HERRERO, M. A. et al. Consideraciones sobre la variación fonética en español y sus implicaciones para la evaluación internacional de la lengua como LE/L 2: DELE, SIELE Y CELU. *Revista de Lingüística Teórica y Aplicada*, Concepción, v. 57, n. 1, p. 13–32, jan.-jul. 2019. DOI: 10.4067/S0718-48832019000100013. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-48832019000100013. Acesso em: 27 dez. 2024.

ASHRAF RIZVI, M. Using Students' Analysis in Teaching Public Speaking for Business. *Profile: Issues in Teachers' Professional Development*, Bogotá, n. 6, p. 107–118, 2005. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902005000100010&lng=en&nrm=iso. Acesso em: 27 dez. 2024.

BAITMAN, B.; VÉLIZ CAMPOS, M. A comparison of oral evaluation ratings by native English speaker teachers and non-native English speaker teachers. *Literatura y lingüística*, Santiago, n. 27, p. 171–200, 2013. DOI: 10.4067/s0716-58112013000100010. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-58112013000100010&lng=en&nrm=iso&tlng=en. Acesso em: 27 dez. 2024.

- BALEGHIZADEH, S.; ZARGHAMI, Z. The Impact of Conferencing Assessment on EFL Students' Grammar Learning. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 14, n. 2, p. 131-144, 2012. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902012000200009&lng=en&nrm=iso. Acesso em: 27 dez. 2024.
- BENAVIDES, J. E. Level of English in Colombian Higher Education: A Decade of Stagnation. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 23, n. 1, p. 57–73, jan.-jun. 2021. DOI: <https://doi.org/10.15446/profile.v23n1.83135>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/83135>. Acesso em: 27 dez. 2024.
- BÉNITEZ FIGARI, R. Análisis descriptivo de narraciones escritas por niños y niñas de tercer año básico. *Literatura y Lingüística*, Santiago, n. 20, p. 103–123, 2018. DOI: 10.29344/0717621x.20.1565. Disponível em: <https://ediciones.ucsh.cl/index.php/lyl/article/view/1565>. Acesso em: 27 dez. 2024.
- CADENA-AGUILAR, A.; ÁLVAREZ-AYURE, C. P. Self- and Peer-Assessment of Student-Generated Podcasts to Improve Comprehensibility in Undergraduate EFL Students. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 23, n. 2, p. 67–85, jul. dez. 2021. DOI: <https://doi.org/10.15446/profile.v23n2.88928>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902021000200067. Acesso em: 27 dez. 2024.
- CÁRCAMO, B. Classifying Written Corrective Feedback for Research and Educational Purposes: A Typology Proposal. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 22, n. 2, p. 211–222, jul.-dez. 2020. DOI: <https://doi.org/10.15446/profile.v22n2.79924>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/79924>. Acesso em: 27 dez. 2024.
- CRUZ-RAMOS, M. DE LOS M.; HERRERA-DÍAZ, L. E. Assessment of Students' Oral Communicative Competence in English through a Web Conferencing Platform. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 1, p. 143–156, jan.- jun. 2022. DOI: <https://doi.org/10.15446/profile.v24n1.91282>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/91282>. Acesso em: 22 dez. 2024.
- DUQUE-AGUILAR, J. F. Teachers' Assessment Approaches Regarding EFL Students' Speaking Skill. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 23, n. 1, p. 161–177, jan.-jun. 2021. DOI: <https://doi.org/10.15446/profile.v23n1.85964>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902021000100161. Acesso em: 27 dez. 2024.
- ESPINOSA AGUIRRE, M. J.; FIGUEROA MIRALLES, J.; ÁVILA REYES, N. Escribir en L2 en la escuela chilena: Una caracterización de la escritura en español de estudiantes de origen haitiano en 5o básico. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 55, n. 108, p. 37–60, 2022. DOI: 10.4067/S0718-09342022000100037. Disponível em: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342022000100037&lng=es&nrm=iso. Acesso em: 27 dez. 2024.
- FARÍAS, M. F. Evaluación de una intervención pedagógica de alfabetización visual en la

comprensión lectora de textos multimodales en programas de formación inicial docente en Lenguaje y Comunicación. *Literatura y Lingüística*, Santiago, n. 35, p. 405–420, 2017. DOI: <http://dx.doi.org/10.4067/S0716-58112017000100403>. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-58112017000100403. Acesso em: 27 dez. 2024.

FIGUEROA, J.; MENESES, A.; CHANDÍA, E. Desempeños en la calidad de explicaciones y argumentaciones en estudiantes chilenos de 8o básico. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 52, n. 99, p. 31–54, 2019. DOI: 10.4067/S0718-09342019000100031. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342019000100031. Acesso em: 27 dez. 2024.

FORERO-MONDRAGÓN, A. J.; QUINTERO-POLO, Á. H. Disciplinary Power Lying Behind the Requisite of English Language Mastery in International Scholarships. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 1, p. 175–191, jan.-jul. 2022. DOI: <https://doi.org/10.15446/profile.v24n1.91112>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/91112>. Acesso em: 27 dez. 2024.

FURLAN, L. A.; FERRERO, M. J.; GALLART, G. Ansiedad ante los exámenes, procrastinación y síntomas mentales en estudiantes universitarios. *Revista Argentina de Ciencias del Comportamiento*, Córdoba, v. 6, n. 3, p. 31–39, 2014. DOI: 10.32348/1852.4206.v6.n3.8726. Disponível em: <https://revistas.unc.edu.ar/index.php/racc/article/view/8726>. Acesso em: 27 dez. 2024.

GALLEGO ORTEGA, J. L.; FIGUEROA SEPÚLVEDA, S; RODRÍGUEZ FUENTES, A. La comprensión lectora de escolares de educación básica. *Literatura y Lingüística*, Santiago, n. 40, p. 187–208, 2019. DOI: 10.29344/0717621X.40.2066. Disponível em: https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0716-58112019000200187. Acesso em: 27 dez. 2024.

GIRALDO, F. A Post-Positivist and Interpretive Approach to Researching Teachers' Language Assessment Literacy. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 22, n. 1, p. 189–200, 2020. DOI: 10.15446/profile.v22n1.78188. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/78188>. Acesso em: 27 dez. 2024.

GIRALDO, F. A Reflection on Initiatives for Teachers' Professional Development through Language Assessment Literacy. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 23, n. 1, p. 197–213, 2021. DOI: <https://doi.org/10.15446/profile.v23n1.83094>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/83094>. Acesso em: 27 dez. 2024.

GIRALDO, F. Language Assessment Literacy: Implications for Language Teachers. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 20, n. 1, p. 179–195, 2018. DOI: <https://doi.org/10.15446/profile.v20n1.62089>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/62089>. Acesso em: 27 dez. 2024.

GONZÁLEZ, E. F. The Impact of Assessment Training on EFL Writing Classroom Assessment: Voices of Mexican University Teachers. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 23, n. 1, p. 107–124, jan.-jun. 2021. DOI:

<https://doi.org/10.15446/profile.v23n1.85019>. Disponível em:
http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902021000100107.
 Acesso em: 27 dez. 2024.

GUTIÉRREZ, R. Habilidades favorecedoras del aprendizaje de la lectura en alumnos de 5 y 6 años. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 51, n. 96, p. 45–60, 2018. DOI: 10.4067/S0718-09342018000100045. Disponível em:
https://rua.ua.es/dspace/bitstream/10045/70215/1/2018_Gutierrez_Signos.pdf. Acesso em: 27 dez. 2024.

GUZMÁN, B.; VÉLIZ, M.; REYES, F. Memoria operativa, comprensión lectora y rendimiento escolar. *Literatura y Lingüística*, Santiago, n. 35, p. 379, 2018. DOI: 10.29344/0717621x.35.1432. Disponível em:
<https://ediciones.ucsh.cl/index.php/lyl/article/view/1432>. Acesso em: 27 dez. 2024.

HERNANDEZ-OCAMPO, S. P. Colombian Scholars' Discussion about Language Assessment: A Review of Five Journals. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 2, p. 231–245, jul.-dez. 2022. DOI: <https://doi.org/10.15446/profile.v24n2.94468>. Disponível em:
<https://revistas.unal.edu.co/index.php/profile/article/view/94468>. Acesso em: 22 dez. 2024.

IACOBUZIO, S. V. *et al.* Estrategias y desafíos en la creación de un instrumento de evaluación lingüística infantil. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 55, n. 110, p. 948–975, 2022. DOI: 10.4067/S0718-09342022000300948. Disponível em:
https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342022000300948. Acesso em: 27 dez. 2024.

LAGOS, A. N. Complejidad textual, vocabulario pasivo y comprensión lectora en escolares chilenos de segundo ciclo básico. *Literatura y lingüística*, Santiago, n. 40, p. 251–272, 2019. DOI: 10.4067/S0716-58112014000200009. Disponível em:
https://scielo.conicyt.cl/scielo.php?pid=S0716-58112014000200009&script=sci_arttext&tlng=en. Acesso em: 27 dez. 2024.

LARENAS, C. D. *et al.* EFL in-service teachers' self-reported perspectives on English language assessment. *Nueva Revista del Pacífico*, Valparaíso, n. 76, p. 40–61, 2022. DOI: 10.4067/s0719-51762022000100040. Disponível em:
https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0719-51762022000100040. Acesso: 27 dez. 2024.

LEÓN CHICA, J.; D'COSTA MARTÍNEZ, C.; FRANCO JÁCOME, G. Designing Proficiency Tests to Accredite Previous Knowledge in American and British Literature in a Bilingual Education Program. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 12, n. 1, p. 159-179, jan.-jun. 2010. Disponível em:
http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902010000100011. Acesso em: 27 dez. 2024.

LIVINGSTONE, K. A. La evaluación de un modelo metodológico mixto para mejorar las habilidades productivas y receptivas en español como segunda lengua: un estudio cuasi-experimental. *Lengua y Habla*, Mérida, n. 15, p. 33-58, 2011. Disponível em:
<https://www.redalyc.org/pdf/5119/511951371002.pdf>. Acesso em: 27 dez. 2024.

LIZASOAIN, A.; TOLEDO VEGA, G. Evaluación de la producción escrita de estudiantes secundarios haitianos por parte de profesores de Lengua y Literatura en Chile. *Estudios Filológicos*, Valdivia, n. 66, p. 185–205, 2020. DOI: 10.4067/s0071-17132020000200185. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0071-17132020000200185&lng=es&nrm=iso&tlng=es. Acesso em: 27 dez. 2024.

LÓPEZ MENDOZA, A. A.; BERNAL ARANDIA, R. Language Testing in Colombia: A Call for More Teacher Education and Teacher Training in Language Assessment. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 11, n. 2, p. 55–70, 2009. DOI: <https://doi.org/10.15446/profile.v20n1.62089>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902009000200005&lng=en&nrm=iso. Acesso em: 27 dez. 2024.

MACHADO, Y. La competencia intercultural en la clase de francés lengua extranjera: ¿qué y cómo evaluarla?. *Lengua y Habla*, Mérida, v. 23, n. 23, p. 550–556, 2019. Disponível em: <https://www.redalyc.org/journal/5119/511966657035/html/>. Acesso em: 27 dez. 2024.

MARTIN ALONSO, B. Validez y habilidad de una escala de evaluación oral en ELE. *Revista de Lingüística Teórica y Aplicada*, Concepción, v. 56, n. 2, p. 65–89, 2018. DOI: <http://dx.doi.org/10.4067/S0718-48832018000200065>. Disponível em: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-48832018000200065&lng=es&nrm=iso. Acesso em: 27 dez. 2024.

MARTÍNEZ GIMÉNEZ, T.; SELLÉS NOHALES, P. El desarrollo del conocimiento de los componentes y funciones del lenguaje escrito en estudiantes prelectores. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 55, n. 108, p. 117–135, 2022. DOI: 10.4067/S0718-09342022000100117. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342022000100117. Acesso em: 27 dez. 2024.

MOURENTE MIGUEL, M. C. Teacher Correction versus Peer-Marking. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 5, n. 1, p. 23–28, 2004. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/11211>. Acesso em: 27 dez. 2024.

MUÑOZ MARÍN, J. H. Exploring Teachers' Practices for Assessing Reading Comprehension Abilities in English as a Foreign Language. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 11, n. 2, p. 71–84, 2009. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902009000200006&lng=en&nrm=iso&tlng=en. Acesso em: 27 dez. 2024.

MUÑOZ MORENO, E.; MAVROU, I. La influencia del componente no verbal asociado al feedback correctivo en la producción oral en ELE. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 53, n. 103, p. 468–495, 2020. DOI: 10.4067/S0718-09342020000200468. Disponível em: https://www.scielo.cl/scielo.php?pid=S0718-09342020000200468&script=sci_abstract. Acesso em: 27 dez. 2024.

MUÑOZ, A. P.; ALVAREZ, M. E. Preliminary Evaluation of the Impact of a Writing Assessment System on Teaching and Learning. *Profile: Issues in Teachers' Professional Development*, Bogotá, n. 9, p. 89–110, 2008. Disponível em:

http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902008000100006&lng=en&nrm=iso. Acesso em: 27 dez. 2024.

MUÑOZ, A. P.; PALACIO, M.; ESCOBAR, L. Teachers' Beliefs About Assessment in an EFL Context in Colombia. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 14, n. 1, p. 143-158, 2012. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902012000100010&lng=en&nrm=iso. Acesso em: 27 dez. 2024.

MUÑOZ, A. V. Desempeño escrito de los estudiantes de secundaria en Chile: una evaluación directa e indirecta de la escritura. *Logos: Revista de Lingüística Filosofía y Literatura*, Coquimbo, v. 32, n. 1, p. 91-105, 2022. DOI: doi.org/10.15443/RL3206. Disponível em: <https://revistas.userena.cl/index.php/logos/article/view/1643>. Acesso em: 27 dez. 2024.

MUÑOZ, B.; FERREIRA, A. El feedback correctivo escrito indirecto en el aprendizaje de la forma comparativa de adjetivos en inglés. *Logos: Revista de Lingüística Filosofía y Literatura*, Concepción, v. 27, n. 1, p. 73-89, 2017. DOI: [dx.doi.org/10.15443/RL2706](https://doi.org/10.15443/RL2706). Disponível em: https://www.scielo.cl/scielo.php?pid=S0719-32622017000100073&script=sci_abstract&tlng=en. Acesso em: 27 dez. 2024.

NAYERNIA, A.; NOSRATI, R.; MOHEBBI, H. The Factors Contributing to Language Teachers' Effectiveness in an EFL Learning Context: A Questionnaire Validation Study. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 2, p. 63-79, jul.-dez. 2022. DOI: <https://doi.org/10.15446/profile.v24n2.93571>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902022000200063. Acesso em: 27 dez. 2024.

NEIRA MARTÍNEZ, A. C.; REYES REYES, F. T.; RIFFO OCARES, B. E. Experiencia académica y estrategias de comprensión lectora en estudiantes universitarios de primer año. *Literatura y Lingüística*, Santiago, n. 31, p. 221-244, 2015. DOI: [10.4067/S0716-58112015000100012](https://doi.org/10.4067/S0716-58112015000100012). Disponível em: https://www.scielo.cl/scielo.php?pid=S0716-58112015000100012&script=sci_arttext. Acesso em: 27 dez. 2024.

ÖZDEMİR-YILMAZER, M.; ÖZKAN, Y. Challenges and Complexities in Teacher Conceptualization of Dynamic Assessment in the L2 Classroom: A Case Study in Türkiye. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 25, n. 1, p. 83-96, jan.-jun. 2023. DOI: <https://doi.org/10.15446/profile.v25n1.101065>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/101065>. Acesso em: 27 dez. 2024.

PALACIO, M.; GAVIRIA, S.; BROWN, J. D. Aligning English Language Testing With Curriculum. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 18, n. 2, p. 63, 2016. DOI: [http://dx.doi.org/10.15446/profile.v18n2.53302](https://doi.org/10.15446/profile.v18n2.53302). Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/53302>. Acesso em: 27 dez. 2024.

PASTENE LABRÍ-N, F. C.; DÍ-AZ CHAVARRÍ-A, R. M.; MOLINA CASTILLO, S. A. Producción del Ensayo Argumentativo en Estudiantes de dos Establecimientos de Educación Media Técnico-Profesional de la Provincia de Ñuble: Resultados de un Diagnóstico. *Literatura y Lingüística*, Santiago, n. 34, p. 395-418, 2018. DOI: [10.29344/0717621x.34.1455](https://doi.org/10.29344/0717621x.34.1455). Disponível em: <https://ediciones.ucsh.cl/index.php/lyl/article/view/1455>. Acesso em: 27 dez. 2024.

PERONARD, M. La Evaluación y comprensión de textos escritos: el problema del resumen. *Lenguas Modernas*, Santiago, n. 21, p. 81–93, 2017. Disponível em: <https://lenguasmodernas.uchile.cl/index.php/LM/article/view/45591>. Acesso em: 27 dez. 2024.

PINEDA, D. The Feasibility of Assessing Teenagers' Oral English Language Performance with a Rubric. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 16, n. 1, p. 181–198, abr. 2014. DOI: <http://dx.doi.org/10.15446/profile.v16n1.43203>. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1657-07902014000100013. Acesso em: 27 dez. 2024.

RAMÍREZ SANTANA, G. M. *et al.* Intervención en Lectura Temprana en Alumnado con Trastorno del Desarrollo del Lenguaje. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 56, n. 112, p. 306–324, 2023. DOI: 10.4067/S0718-09342023000200306. Disponível em: https://www.scielo.cl/scielo.php?pid=S0718-09342023000200306&script=sci_abstract. Acesso em: 27 dez. 2024.

RAMOS-LEIVA, L. *et al.* La ansiedad hacia la evaluación en el aprendizaje del inglés como lengua extranjera en estudiantes de enseñanza media. *Lenguas Modernas*, Santiago, n. 59, p. 9–29, 2022. Disponível em: <https://lenguasmodernas.uchile.cl/index.php/LM/article/view/67931>. Acesso em: 27 dez. 2024.

RIFFO OCARES, B. *et al.* Competencia léxica, comprensión lectora y rendimiento académico en estudiantes de enseñanza media. *Literatura y lingüística*, Santiago, n. 30, p. 136–165, 2014. DOI: 10.4067/S0716-58112014000200009. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0716-58112014000200009. Acesso em: 27 dez. 2024.

SAAVEDRA-JELDRES, P. A.; CAMPOS-ESPINOZA, M. Chilean Pre-Service Teachers' Perceptions Towards Benefits and Challenges of EFL Writing Portfolios. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 21, n. 2, p. 79–96, jul.-dez. 2019. DOI: <https://doi.org/10.15446/profile.v21n2.73116>. Disponível em: <http://revistas.unal.edu.co/index.php/profile/article/view/73116>. Acesso em: 27 dez. 2024.

SAYYADI, A. In-Service University-Level EFL Instructors' Language Assessment Literacy and Training Needs. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 1, p. 77–95, 2022. DOI: <https://doi.org/10.15446/profile.v24n1.93676>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/93676>. Acesso em: 27 dez. 2024.

SCLIAR-CABRAL, L. Evaluación de la escritura en niños de tercer grado en Brasil. *Boletín de Lingüística*, Caracas, v. 30, n. 49-50, p. 158–173, 2018. Disponível em: <http://bdigital2.ula.ve:8080/xmlui/654321/9466>. Acesso em: 27 dez. 2024.

TABOH, A. R. *et al.* Evaluación del lenguaje oral en niños y niñas con hipoacusia: Los tests estandarizados y la edad auditiva. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 55, n. 110, p. 928–947, 2022. DOI: 10.4067/S0718-09342022000300928. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342022000300928. Acesso em: 27 dez. 2024.

TAJEDDIN, Z.; SAEEDI, Z.; PANAHAZADEH, V. English Language Teachers' Perceived Classroom Assessment Knowledge and Practice: Developing and Validating a Scale. *Profile: Issues in Teachers' Professional Development*, Bogotá, v. 24, n. 2, p. 247–264, jul.-dez. 2022. DOI: <https://doi.org/10.15446/profile.v24n2.90518>. Disponível em: <https://revistas.unal.edu.co/index.php/profile/article/view/90518>. Acesso em: 27 dez. 2024.

VALDÉS-LÉON, G.; RIFFO, J. G.; OLIVARES, M. M. Competencia comunicativa en educación superior: hacia una evaluación diagnóstica integral. *Logos: Revista de Lingüística Filosofía y Literatura*, Coquimbo, v. 32, n. 1, p. 106–119, 2022. DOI: doi.org/10.15443/RL3207. Disponível em: <https://revistas.userena.cl/index.php/logos/article/view/1644>. Acesso em: 27 dez. 2024.

VALLEJOS CARRASCO, G.; VÉLIZ DE VOS, M.; SÁEZ CARRILLO, K. Estructura hipertextual y habilidad lectora: Análisis de la producción digital de textos argumentativos en estudiantes de 4o Año Medio. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 54, n. 105, p. 97–119, 2021. DOI: [10.4067/S0718-09342021000100097](https://doi.org/10.4067/S0718-09342021000100097). Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09342021000100097. Acesso em: 27 dez. 2024.

VEGA RODRÍGUEZ, Y.; TORRES RODRÍGUEZ, A. M.; DEL CAMPO RIVAS, M. Habilidades metamorfológicas y su incidencia en la comprensión lectora. *Revista Signos: Estudios de Lingüística*, Viña del Mar, v. 50, n. 95, p. 453–471, 2017. DOI: [10.4067/S0718-09342017000300453](https://doi.org/10.4067/S0718-09342017000300453). Disponível em: <https://scielo.conicyt.cl/pdf/signos/v50n95/0718-0934-signos-50-95-00453.pdf>. Acesso em: 27 dez. 2024.

VINE JARA, A.; FERREIRA CABRERA, A. Propuesta de un modelo para una prueba con fines específicos académicos en ELE. *Literatura y Lingüística*, Santiago, n. 33, p. 369–390, 2018. DOI: [10.29344/0717621x.33.1490](https://doi.org/10.29344/0717621x.33.1490). Disponível em: <https://ediciones.ucsh.cl/index.php/lyl/article/view/1490>. Acesso em: 27 dez. 2024.