

Tipo: Artículo original - Dossier: Temática variada

Explorando uma oficina de tarefas orais integradas: desafios e oportunidades de letramento em avaliação para professores de Espanhol em Formação Inicial

Exploring a workshop on integrated oral tasks: challenges and opportunities for assessment literacy of preservice Spanish teachers

Aline Netto Brum-Barreto

Universidade de Brasília, Instituto de Letras, Brasília - Brasil.
<https://orcid.org/0000-0003-3438-0374>

e-mail: alineanb@hotmail.com

Vanessa Borges-Almeida

Universidade de Brasília, Instituto de Letras, Brasília - Brasil.
<https://orcid.org/0000-0003-0502-5162>

e-mail: borgesalmeida@unb.br

Recibido: 3/6/2024
Aprobado: 21/12/2024

RESUMO

Letramento em avaliação de professores de línguas pode ser entendido como a capacidade de discernir criticamente o quê, como, quando e por que avaliar (Inbar-Lourie, 2008). No contexto brasileiro, no entanto, estudos mostram que os professores de línguas em formação inicial, em geral, não estão sendo adequadamente preparados para planejar, desenvolver e executar processos de avaliação de qualidade (Quevedo-Camargo, 2020). Portanto, em consonância com diversos autores, defendemos que o desenvolvimento do letramento em avaliação deve estar presente desde a formação inicial do professor (Brum-Barreto, 2023; Borges-Almeida, 2023; Giraldo & Murcia, 2019; Harding & Kremmel, 2016; Jaramillo-Delgado & Bedoya, 2019; Restrepo-Bolívar, 2020). Neste sentido, este trabalho tem por objetivo relatar uma experiência com uma oficina de elaboração de tarefas para avaliar habilidades orais integradas para professores de espanhol em formação inicial, em uma universidade pública brasileira, com foco na promoção do letramento em avaliação. Trata-se de um estudo exploratório (Gil, 2002) que analisa o conteúdo (Bardin, 2006) das tarefas integradas produzidas pelos participantes da oficina, investigando suas principais características, bem como as visões de língua e avaliação subjacentes aos instrumentos. À luz dessas análises, buscamos identificar os indícios de impactos da oficina no letramento em avaliação dos participantes, bem como os desafios e potencialidades dessa proposta. Nossos resultados indicam que oficinas curtas podem sensibilizar os participantes para a operacionalização de um método de avaliação mais alinhado a uma visão de língua em uso, embora não sejam suficientes para consolidar habilidades voltadas para a elaboração de instrumentos de avaliação realmente autênticos e confiáveis, ou para modificar concepções de língua. Nesse sentido, reforçamos a inclusão de componentes curriculares voltados para a temática da avaliação na formação inicial de professores de línguas.

Palavras-chave: letramento em avaliação em contexto de línguas; tarefas integradas; habilidades orais; espanhol como língua estrangeira; professores em formação inicial.

Conflictos de Interés: ninguno que declarar

Rol autoral: los autores han participado en todo el proceso de elaboración del artículo.

Fuente de financiamiento: sin fuente de financiamiento.

DOI: <https://doi.org/10.47133/NEMITYRA20250702b-A14>

BIBLID: 2707-1642, 7, 2, pp. 217-236

Editores responsables: Valentina Canese. Universidad Nacional de Asunción, Facultad de Filosofía, Instituto Superior de Lenguas, Paraguay.

ABSTRACT

Language teacher assessment literacy can be understood as the ability to critically discern what, how, when and why to assess (Inbar-Lourie, 2008). However, in the Brazilian context studies show that preservice language teachers, in general, have not being adequately prepared to plan, develop, and execute quality assessment processes (Quevedo-Camargo, 2020). Therefore, in line with several authors, we argue that the development of assessment literacy should be present from initial education of the teacher (Brum-Barreto, 2023; Borges-Almeida, 2023; Giraldo & Murcia, 2019; Harding & Kremmel, 2016; Jaramillo-Delgado & Bedoya, 2019; Restrepo-Bolívar, 2020). In this sense, this paper aims to report our experience with a workshop on task design to assess integrated oral skills conducted with Spanish teachers in initial training at a Brazilian public university, with a focus on the promotion of their assessment literacy. This is an exploratory study (Gil, 2002) that analyses the content (Bardin, 2006) of the integrated tasks produced by the participants at the workshop, by investigating their main characteristics, as well as the views on language and assessment underlying those instruments. In the light of these analyses, we seek to identify indications of impact of the workshop on the participants' assessment literacy, as well as the challenges and potentialities of this proposal. Results indicate that short workshops can sensitize participants to the operationalization of an assessment method which is more aligned with a view of the language in use, although they are not enough to consolidate the skills required to design really authentic and reliable assessment instruments, nor to change one's conceptions of language. In this sense, we reinforce the inclusion of curricular components focused on the theme of assessment in language teacher initial education.

Keywords: language assessment literacy; integrated tasks; oral skills; Spanish as a foreign language; preservice teachers.

Introdução

Letramento em avaliação nos contextos de ensino-aprendizagem de línguas (LAL) pode ser compreendido como o conjunto de competências avaliativas fundamentadas em construtos relacionados à língua (como a visão de língua e de proficiência) e adequadas ao contexto de atuação do professor (Inbar-Lourie, 2008; Scaramucci, 2016a). Estudos mostram que propostas visando ao desenvolvimento do LAL podem sensibilizar os professores em formação inicial para práticas de avaliação mais alinhadas à sua dimensão formativa (Brum-Barreto, 2023; Borges-Almeida, 2023a; Giraldo & Murcia, 2019).

No Brasil, conforme aponta Quevedo-Camargo (2020), os cursos de formação de professores de línguas não têm dedicado espaço suficiente para o LAL. Não é por acaso que, de forma consistente, alguns autores demonstram que professores de línguas em serviço se sentem despreparados para avaliar (Fulcher, 2012; Vogt & Tsagari, 2014). Diante disso, em consonância com diversos autores, defendemos a necessidade de iniciativas que promovam o LAL desde a formação inicial (Brum-Barreto, 2023; Borges-Almeida, 2023; Giraldo & Murcia, 2019; Harding & Kremmel, 2016; Jaramillo-Delgado & Bedoya, 2019; Restrepo-Bolívar, 2020). No Brasil e no mundo, ainda há poucas experiências empíricas exclusivamente voltadas para o LAL de professores de línguas em formação inicial (Giraldo & Murcia, 2019). Além disso, de forma geral, o volume de publicações que descrevem experiências voltadas para o desenvolvimento do LAL é limitado na América Latina (Giraldo, 2021).

Visando preencher essa lacuna, este artigo relata a experiência em uma oficina de elaboração de tarefas para avaliar habilidades orais integradas, voltada para professores de espanhol em formação inicial, em uma universidade pública brasileira, com foco na promoção do LAL. A escolha desse tema foi baseada em estudos exploratórios e diagnósticos que

identificaram as necessidades de LAL no contexto da investigação (Borges-Almeida, 2023b; Brum-Barreto, 2023).

Com este trabalho, buscamos refletir sobre o impacto de uma iniciativa como essa no LAL dos participantes, suas potencialidades e desafios, orientadas pela seguinte pergunta de pesquisa: *o que revelam as características das tarefas produzidas pelos participantes quanto aos impactos de uma oficina sobre (1) sua concepção de língua (como uso situado) e (2) seus conhecimentos e habilidades para elaborar instrumentos de avaliação?* Acreditamos que este trabalho possa colaborar com as discussões acerca das metodologias aplicadas em ações pedagógicas voltadas para o LAL em contextos de ensino-aprendizagem de línguas.

Além desta introdução, o artigo está dividido em quatro partes. Na próxima seção, apresentamos o conceito de tarefas integradas e as especificidades da avaliação de habilidades orais. Em seguida, descrevemos a metodologia aplicada na oficina de elaboração de tarefas e detalhamos a geração e análise dos dados. Após isso, apresentamos e analisamos os resultados obtidos. Finalmente, encerramos o trabalho com recomendações e considerações finais.

Referencial teórico

Tarefas de habilidades integradas

O conceito de avaliações integradoras (*integrated assessment*) ou conduzidas com tarefas integradas (*integrated tasks*)¹ não é recente (Scaramucci, 2016b). Trata-se de um método de avaliação de desempenho que pressupõe a demonstração direta da proficiência por meio de métodos que simulam situações de comunicação realistas. Este método surgiu na década de 1960 como uma alternativa aos testes baseados em itens isolados (*lápiz e papel*), que avaliam as habilidades de forma indireta e focam nos conhecimentos sobre o código linguístico, típicos do movimento psicométrico-estruturalista. Desde então, diferentes métodos de avaliação de habilidades integradas foram propostos (Scaramucci, 2016b). A literatura apresenta várias possibilidades de operacionalização da integração de habilidades em uma tarefa, como: compreensão e produção escrita (*reading-to-write* ou *reading-based writing*); compreensão e produção oral (*listening-to-speak* ou *listening speaking*); compreensão oral e produção escrita (*listening-to-write* ou *listening-based writing*); compreensão escrita e produção oral (*reading-to-speak*); e produção oral ou escrita baseada em gráficos/imagens (*graph/image-based writing or speaking*).

Partindo de uma visão holística e fundamentada no uso da língua em situações comunicativas similares às do mundo real, nas quais as habilidades linguísticas raramente ocorrem de forma isolada, o conceito de tarefas integradas, embora não tenha sido sistematicamente definido (Scaramucci, 2016b), geralmente, refere-se a avaliações de desempenho em escrita ou fala que incluem a compreensão e o processamento de um insumo, como um texto escrito ou oral (Plakans, 2015), preferencialmente autêntico (ou seja, produzido

¹ Cabe-nos aqui fazer uma breve diferenciação entre testes integrativos (*integrative test*) e tarefas integradas (*integrated tasks*). Testes integrativos fundamentam-se no conceito de “competência unitária” proposto por Oller (1979) e estão associados a métodos psicométricos de testagem, como o teste *cloze*.

com fins não didáticos). Para Cumming (2013), essas tarefas exigem que os avaliandos processem (leiam, ouçam, interpretem, resumam etc.) o insumo fornecido e utilizem as informações obtidas por meio desse processamento para gerar uma resposta². Assim, a integração de habilidades, sob nossa ótica, permite avaliar habilidades de compreensão (oral e escrita) que não podem ser observadas diretamente, mas podem ser inferidas por meio das habilidades de produção (Brown, 1996). Ao mobilizar informações do insumo para a produção de sua resposta, o avaliando demonstra sua capacidade de compreender e interpretar — seja lendo ou ouvindo o texto. Nesse sentido, os propósitos das atividades de produção determinam os propósitos das atividades de compreensão.

Um dos principais argumentos favoráveis ao uso desse tipo de tarefa é sua autenticidade, ou seja, seu grau de similaridade com tarefas de linguagem realizadas no mundo real. Plakans (2013) salienta que, ao projetar uma tarefa de habilidades integradas, é necessário, primeiramente, estabelecer o construto da avaliação com base no uso da língua pelo avaliando. Em outras palavras, o construto e o objeto da avaliação são determinados com base no "critério da situação comunicativa alvo" (Scaramucci, 2016b, p. 406). Em seguida, a partir da definição do construto, uma série de ações deve ser desenvolvida, incluindo: definir os objetivos de avaliação; escolher o insumo e identificar as informações principais que serão mobilizadas pela tarefa; elaborar instruções (ou o comando) da tarefa; estabelecer critérios válidos e consistentes de avaliação com base em expectativas de respostas; e, se possível, realizar testes piloto para avaliar o insumo, as instruções, o tempo de execução e a validade, praticidade e confiabilidade da rubrica de avaliação.

Scaramucci (2016b) e Plakans e Gebril (2017) destacam que, ao fornecer o insumo, as tarefas integradas aumentam a equidade de conhecimento sobre a temática entre os avaliandos e, consequentemente, minimizam o impacto do conhecimento prévio, da experiência de vida ou da criatividade dos avaliandos, tornando a avaliação mais justa. Assim, o insumo desempenha várias funções, como moldar a opinião dos avaliandos, fornecer ideias sobre a temática, oferecer recursos linguísticos e possibilitar a avaliação da habilidade de compreensão.

Além disso, as tarefas integradas têm o potencial de impactar positivamente os processos de ensino e aprendizagem, gerando um efeito retroativo positivo³ (Scaramucci, 2016b) e estimulando práticas mais alinhadas com uma visão de língua em uso. Cumming (2013) afirma que as tarefas integradas podem contrabalançar os efeitos negativos dos métodos

² Contudo, Plakans (2013), como outros autores, considera que nem sempre o uso do insumo na resposta é obrigatório. Para a autora, há casos em que o insumo pode servir apenas de estímulo ou ilustração para o desenvolvimento da tarefa. Nesta perspectiva, considera-se que embora “a avaliação resultante desse tipo de tarefa não esteja necessariamente relacionada à integração de habilidades, os participantes da prova utilizarão mais de uma habilidade em seu processo de desenvolvimento” (Plakans, 2013, p. 256, tradução nossa).

³ Neste trabalho o termo “efeito retroativo” é utilizado, em consonância com Quevedo-Camargo (2014), para fazer referência ao impacto, à influência ou às consequências sociais advindas da presença de algum tipo avaliação.

de avaliação baseados em correção objetiva (itens) e avaliar as habilidades linguísticas de forma consistente com as teorias de multiletramentos.

A literatura, no entanto, aponta algumas limitações associadas ao uso de tarefas integradas, tais como: a falta de uma definição mais clara do construto e dos critérios de avaliação; a possibilidade de confusão entre as habilidades, o que pode dificultar a inferência precisa do desempenho dos avaliandos, especialmente para fins de diagnóstico; a dificuldade em identificar empréstimos dos insumos nas respostas (como cópias e paráfrases); e o alto custo de aplicação (Scaramucci, 2016b).

Nesta investigação, alinhados à perspectiva acional proposta pelo Quadro Europeu Comum de Referência (QCER) (Conselho da Europa, 2001; 2020), adotamos uma visão de língua em uso situado, na qual ser proficiente em uma língua adicional equivale a ser capaz de usá-la para ser, estar e agir no mundo (Bakhtin, 1997, 2016). Sob nossa ótica, uma tarefa integrada é uma atividade de ensino, aprendizagem ou avaliação que busca aproximar-se o máximo possível das situações comunicativas do cotidiano dos avaliandos.

Operacionalmente, as tarefas integradas consistem em um conjunto de elementos que incluem as instruções ou comando da tarefa, o insumo ou texto-base e a rubrica de avaliação. Juntos, eles definem aspectos importantes da tarefa, como os interlocutores envolvidos na situação de comunicação e seus papéis sociais, os propósitos comunicativos da atividade (como narrar, argumentar, elogiar, enumerar, resumir, entre outros), o gênero discursivo que será utilizado (como uma entrevista de emprego, um seminário acadêmico ou um *e-mail* profissional) e o tipo de insumo a ser usado (um texto autêntico, seja ele oral, escrito ou multimodal) (Díaz, Quevedo-Camargo, & Borges-Almeida, 2020).

As instruções de uma tarefa integrada devem esclarecer os resultados esperados pelo avaliador para garantir a validade, tanto de construto quanto de conteúdo, da avaliação. São essenciais, também, para manter a confiabilidade, assegurando que os resultados sejam estáveis e consistentes (Scaramucci, 1995; 2016b). Além disso, as instruções devem especificar claramente o uso do insumo: qual tipo de informação deve ser “aproveitada” do insumo na produção do texto oral e os mecanismos para tal, bem como o nível de precisão necessário para representar essas informações nas respostas da tarefa (Frost, Elder & Wigglesworth, 2011). Além disso, o comando da tarefa deve se concentrar na descrição da situação comunicativa e evitar referências explícitas à situação avaliativa. Isso porque, conforme aponta Pileggi (2015), referências diretas ao insumo e instruções formuladas com imperativos, como “escreva”, “relate” ou “anote”, podem reduzir a autenticidade da tarefa.

No que diz respeito aos critérios de avaliação, o foco deve estar na capacidade do avaliando de construir e negociar significados dentro do discurso para alcançar o propósito comunicativo da tarefa. Como as tarefas são métodos de avaliação de respostas abertas, as rubricas são essenciais para operacionalizar o construto da avaliação (Luoma, 2004). Elas ajudam a aumentar a confiabilidade ao reduzir a subjetividade do processo de correção. Em contextos de ensino e aprendizagem, as rubricas comunicam aos avaliandos os resultados esperados e possibilitam o *feedback* do professor, contribuindo para a dimensão formativa da avaliação (Brookhart, 2018).

Rubricas são escalas que descrevem os critérios de avaliação em diferentes níveis de desempenho e qualidade.⁴ De acordo com Plakans (2013), existem várias abordagens para a elaboração de critérios de avaliação de tarefas integradas: (a) com base na integração de habilidades e conhecimentos estabelecidos pelo construto da avaliação; (b) a partir da análise das necessidades dos avaliandos; ou (c) através da pilotagem da tarefa, na qual amostras de desempenho são coletadas, analisadas e transformadas em critérios de pontuação.

Avaliação das habilidades orais

Em nosso contexto de investigação, estudos indicam que os professores em formação inicial frequentemente relatam que avaliar as habilidades orais é uma prática desafiadora (Brum-Barreto, 2023; Borges-Almeida, 2023b; Damacena-Dutra, 2023). A habilidade oral em línguas adicionais é complexa, pois envolve aspectos intrapessoais — psicolinguísticos, relacionados ao conhecimento e à capacidade de processamento linguístico — e interpessoais, que dizem respeito ao contexto comunicativo. Dessa forma, o falante atua simultaneamente como processador e interlocutor de um discurso (Ortiz-Preuss, 2014). A avaliação das habilidades orais, portanto, exige a adequada compreensão de sua natureza e das condições em que as interações orais ocorrem.

No que diz respeito aos aspectos psicolinguísticos, a produção da fala envolve uma sequência de processos cognitivos interligados. Nessa sequência, o falante define o que será dito (o conteúdo da fala), formula, articula e monitora os aspectos formais do código linguístico (Levelt, 1995). Para que esses processos ocorram de maneira eficaz, é necessário algum nível de automação, devido à capacidade limitada de atenção do cérebro humano para controlar conscientemente todos esses processos simultaneamente. Shekan (1996) sugere que o desempenho de um aprendiz em uma tarefa oral pode ser dividido entre três dimensões principais que competem pela atenção do falante: precisão (ou acurácia), complexidade e fluência. A precisão e a complexidade estão relacionadas ao domínio do código linguístico (gramática da língua), enquanto a fluência está vinculada à construção de significado. Assim, quando o aprendiz aumenta o foco na precisão e na complexidade, a tendência é que a fluência seja comprometida, e vice-versa. Aprendizes iniciantes de línguas adicionais frequentemente enfrentam dificuldades em gerenciar um discurso oral de maneira fluente e gramaticalmente precisa, devido à falta de automação e/ou conhecimento do código linguístico (Bygate, 2001). Portanto, Coombe, Folse e Hubley (2007) aconselham que, ao planejar uma avaliação de habilidades orais, o avaliador deve decidir se o foco estará na fluência ou na precisão. A utilização de uma variedade de métodos de avaliação com focos distintos pode contribuir para um processo de avaliação mais confiável e justo.

A habilidade oral, outrossim, é fortemente influenciada pelo contexto comunicativo. A fala cotidiana tende a ser predominantemente espontânea, com pouco planejamento, exceto

⁴ Rubricas são ferramentas de correção de tarefas ou questões abertas que apresentam os critérios avaliativos escalonados em níveis. Para cada critério em cada nível, há um descritor. Nesta oficina, devido ao limite de tempo, as ferramentas propostas apresentam critérios e níveis, porém, não descritores. Por essa razão, referimos a essas especificamente como “grades” ou “escalas” de correção.

em gêneros discursivos orais mais formais, como um “seminário acadêmico”, por exemplo. Muitas informações podem ser omitidas no discurso oral, seja devido às condições psicolinguísticas já mencionadas, seja porque já estão implícitas no contexto comunicativo ou porque são expressas por outros recursos verbais e não verbais, como expressões faciais, gestos, pausas, entonação, repetições, entre outros. Essas condições de produção impactam diretamente o produto da fala, que costuma ser menos estruturado e complexo, mais fragmentado, e com características como repetições, hesitações, anacolutos, autocorreções entre outras (Baralo, 2000; Bygate, 2001; Cantero Serena, 2020; Coombe, Folse & Hubley, 2007; Ortiz-Preuss, 2014). Por sua vez, a fala em interação é um ato recíproco no qual dois ou mais falantes (interlocutores) colaboram simultaneamente para construir um discurso.

De acordo com o Quadro Europeu Comum de Referência (Conselho da Europa, 2020), ao qual nos alinhamos, a fluência está situada no âmbito da competência pragmática, isto é, de uso da língua. Ela pode ser compreendida sob duas perspectivas: de forma holística, referindo-se à capacidade de articular uma mensagem e de manter uma fala sustentada por um certo período de tempo, adaptando-se o falante a diversos contextos comunicativos; ou de forma mais restrita e técnica, equivalendo à capacidade de falar de maneira contínua, sem interrupções ou pausas excessivamente longas.

Portanto, o objetivo de uma tarefa de avaliação de habilidades orais integradas, no contexto desta investigação, é estimar a capacidade de comunicação, de construção de significados no discurso oral e de alcance do propósito comunicativo – o que requer habilidades de compreensão oral, mediação, interação e produção oral –, respeitando as especificidades da língua oral previamente mencionadas. Nesse sentido, o foco não deve recair sobre a precisão e a complexidade gramatical, que podem ser avaliadas por outros instrumentos, como testes objetivos. Reiteramos, assim, nosso posicionamento favorável ao emprego de uma diversidade de instrumentos de avaliação, de modo a contribuir para um processo avaliativo mais confiável e justo.

Metodologia

Esta investigação, de natureza qualitativa e interpretativista (Sampieri, Collado & Lucio, 2013), configura-se como um estudo exploratório (Gil, 2002), com foco na interpretação de uma ação pedagógica em um contexto específico.

A oficina intitulada “Avalia-ção: oficina de elaboração de tarefas para avaliar habilidades orais integradas em língua espanhola como língua adicional (E/LA)” foi realizada em agosto de 2022, como parte das atividades da Semana Universitária (SU22) da Universidade de Brasília (UnB), Brasil. A Semana Universitária (SU22) é um evento anual organizado pelo Decanato de Extensão (DEX), que visa integrar atividades de ensino, pesquisa e extensão, reunindo tanto a comunidade acadêmica interna quanto a externa à instituição.

Idealizada por ambas as autoras deste artigo e ministrada pela primeira, a oficina contou com a participação de 16 discentes do curso de Letras-Língua Espanhola e Literaturas Espanhola e Hispano-Americana da Universidade de Brasília, todos com pouca ou nenhuma experiência docente. Com carga horária total de três horas e baseada em uma metodologia de

ensino socioconstrutivista (Inbar-Lourie, 2008), teve foco no desenvolvimento de habilidades para a elaboração de tarefas integradas e foi dividida em dois momentos.

No primeiro momento, durante aproximadamente uma hora, foram apresentados modelos de tarefas integradas e realizadas reflexões teóricas sobre conhecimentos, habilidades e princípios (Davies, 2008) relacionados a essas avaliações. Os temas discutidos incluíram a visão de língua subjacente, como socialmente situada, as características da língua oral, a operacionalização da integração de habilidades e os critérios de avaliação. Em outras palavras, foram abordadas questões referentes a validade, autenticidade e confiabilidade em avaliações.

No segundo momento da oficina, com carga horária de aproximadamente duas horas, foi conduzida uma atividade prática na qual, em grupos, os participantes elaboraram propostas de tarefas integradas para avaliar as habilidades de compreensão e produção oral em E/LA, num contexto de sala de aula hipotético. Essa avaliação de rendimento estava vinculada aos objetivos de aprendizagem e ensino de um processo pedagógico. A atividade prática consistiu numa “Tarefa de Formação” (quadro 1), que incluía um passo a passo detalhado para o desenvolvimento da tarefa integrada. Os participantes receberam informações sobre o perfil fictício dos estudantes para os quais a tarefa seria elaborada, além de sugestões de insumos, como textos orais do gênero *podcast*.

Quadro 1. Tarefa de Formação da oficina

Tarefa de Formação	
Objetivos de aprendizagem	Aprender a elaborar tarefas para avaliar habilidades orais integradas em língua espanhola como língua adicional (E/LA)
Descrição da Tarefa de Formação	Elaborar uma tarefa para avaliar habilidades orais integradas (compreensão oral e produção/interação oral) e a respectiva ferramenta de correção/pontuação (grade), conforme as instruções a seguir.
Produto (O que você deve entregar no final da oficina?)	A instrução (ou comando) de uma tarefa que avalia habilidades orais integradas (compreensão oral e produção/interação oral) em língua espanhola e sua respectiva ferramenta de correção/pontuação (grade) devem explicitar claramente os critérios de avaliação e as expectativas de resposta.
Elaborando tarefas passo-a-passo	
Passo 1: A validade	<p>Considerar a temática (conteúdos funcionais e linguísticos) e o gênero discursivo da produção como elementos de engajamento para as práticas de compreensão oral e produção oral, com base nos objetivos de ensino, aprendizagem e avaliação.</p> <p><i>Para efeitos práticos, nesta oficina sugerimos a elaboração de uma tarefa integrada a partir dos seguintes elementos:</i></p> <p>Perfil dos(as) aprendizes: <i>Aprendizes brasileiros(as) de língua espanhola como língua adicional (E/LA), jovens universitários(as), de nível de proficiência B1 (conforme o Quadro Comum Europeu de Referência);</i></p> <p>Integração de habilidades: <i>compreensão e produção oral</i></p> <p>Temática(s): <i>Livre [conferir os insumos disponíveis no Passo 2]</i></p> <p>Conteúdos funcionais e os recursos linguísticos: <i>Livre</i></p> <p>Gênero discursivo: <i>Livre</i></p>
Passo 2: A autenticidade	<p>Escolher um texto autêntico (insumo) e refletir sobre a integração de habilidades com base no uso do insumo.</p> <p><i>Para a elaboração da tarefa integrada proposta nesta oficina, sugerimos os seguintes insumos orais:</i></p> <p>Insumo 1: <i>Ventajas de hablar más de un idioma</i> https://open.spotify.com/episode/6bzeC7zV4Ampu8YOeuniS4?si=8a0QMrDEQx-rT_XbUg97sw)</p>

	<p>Insumo 2: Autonomía para aprender https://open.spotify.com/episode/5yggFzEp6Xbmr0O4ArmhkA?si=u0CPqsg7Rb6mqO7mul4Ybg Insumo 3: Vivir sin plástico (https://youtu.be/hpmLHorl4NE) [Até 6'45 min.] Insumo 4: ¿Qué carrera estudiar? https://open.spotify.com/episode/32OxOlqSyOaghuS6ZYhhtM</p>
Passo 3: Operacionalização do construto	<p>Elaborar a instrução (ou comando) da tarefa para avaliar habilidades orais integradas. Delimitar a situação comunicativa necessária para a produção oral do(a)s aprendiz(es). A partir disso, explicitar: (1) o enunciador (quem produz o texto/discurso); (2) o interlocutor (para quem o texto/discurso se dirige); (3) o propósito enunciativo (a finalidade do texto); (4) o uso do insumo exigido para a conclusão da tarefa; e (5) as ações e resultados esperados para a conclusão da tarefa.</p>
Passo 4: A confiabilidade	<p>Estabelecer critérios de avaliação da produção oral do(a)s aprendiz(es) com base nos objetivos de ensino, aprendizagem e avaliação; na perspectiva de uso da língua, construção de significados e alcance dos propósitos comunicativos. Em seguida, elaborar uma ferramenta de correção/pontuação (grade). Consulte o quadro 3.</p>

Fonte: Elaborado pelas autoras.

Com base nas discussões teóricas e reflexivas realizadas durante a oficina, os participantes foram orientados a desenvolver suas propostas de tarefas integradas seguindo uma série de etapas. Primeiramente, eles deveriam estabelecer os objetivos de ensino, aprendizagem e avaliação da tarefa. Em seguida, era necessário escolher uma temática e um insumo apropriado⁵, decidindo quais informações seriam extraídas do insumo para a elaboração da resposta da tarefa. Também era importante pensar em uma situação comunicativa autêntica na qual essas informações poderiam ser inseridas.

A próxima etapa envolvia a elaboração das instruções da tarefa, que deveriam especificar os interlocutores da situação comunicativa, os propósitos comunicativos, o gênero discursivo e o uso do insumo. Finalmente, os participantes deveriam sistematizar os critérios de avaliação em uma grade de correção, conforme o modelo do quadro 2. Essas atividades foram desenhadas para ajudar os participantes a desenvolver habilidades na criação de tarefas integradas que fossem autênticas e eficazes para avaliar as habilidades de compreensão e produção oral em contextos de sala de aula.

Quadro 2. Modelo de grade de correção apresentado na Tarefa de Formação

<i>Criterios de evaluación</i>	<i>Cumple la tarea</i>	<i>Cumple la tarea parcialmente</i>	<i>No cumple</i>	<i>Sugerencia del profesor</i>

Fonte: Elaborado pelas autoras.

Por fim, os participantes foram orientados a analisar as tarefas integradas que produziram com base na lista de verificação proposta no quadro 3 (autoavaliação).

⁵ Considerando a limitação do tempo, preferimos oferecer algumas opções de temáticas e de insumos para os participantes.

Quadro 3. Lista de verificação para autoavaliação da tarefa de habilidades integradas

Lista de Verificação para Tarefas Integradas					
As instruções (ou comando) da tarefa					
	Si m	Parcial mente	Não	Como aperfeiçoar ?	Obs .
1. Estão claras o suficiente? Despertam o interesse dos(as) aprendizes?					
2. Especificam o contexto em que a ação ocorre? Esse contexto é autêntico, ou seja, reflete situações de uso cotidiano da língua?					
3. Especificam os participantes e/ou interlocutores, assim como seus papéis sociais?					
4. Deixam claro qual é o gênero discursivo esperado na resposta?					
5. Tem insumo(s)? Especificam qual é o uso esperado do insumo?					
6. Integram habilidades?					
7. Especificam quais são as ações, isto é, as funções ou propósitos comunicativos envolvidos (como elogiar, argumentar, recomendar, aconselhar, desculpar, explicar, descrever, narrar, convencer etc.)?					
As características da resposta					
8. Especificam qual é resultado final, o produto esperado? O(s) objetivo(s) a ser (em) alcançado(s) está claro?					
9. Especificam qual é a relação entre o insumo e a resposta? A resposta deve conter parte da informação do insumo? Em caso afirmativo, de que forma?					
Critérios de correção					
10. Quais são os critérios, ou seja, os elementos relevantes na resposta? Esses critérios têm relação com os objetivos de ensino e aprendizagem do curso ou da disciplina?					
11. Enfatizam a construção de significados?					
12. O êxito é julgado com base no resultado alcançado pela tarefa?					

Fonte: Elaborado pelas autoras com base em Bachman e Palmer (1996).

Após a oficina, com autorização prévia⁶, as cinco tarefas integradas produzidas pelos participantes foram sistematicamente analisadas com base na técnica de análise de conteúdo simbólica (Bardin, 2006). As categorias de análise foram: (a) características gerais das instruções; (b) gênero discursivo; (c) uso do insumo nas respostas (avaliação da habilidade de compreensão oral); e (d) critérios de avaliação. Na próxima seção, apresentamos e descrevemos cada uma das tarefas integradas elaboradas pelos participantes da oficina e discutimos os aspectos mais relevantes de cada uma delas.

Análise dos resultados

Os grupos 1 e 2 elaboraram as tarefas que são apresentadas nos quadros 4 a 7. Como insumo, os participantes escolheram o *podcast* intitulado “Ventajas de hablar más de un idioma”, extraído do programa “*TED en Español*” (publicado originalmente em janeiro de

⁶ Os participantes assinaram um Termo de Consentimento Livre e Esclarecido autorizando a utilização de suas produções para fins de investigação. Os nomes e dados pessoais dos participantes serão omitidos.

2022). O *podcast* apresenta as vantagens associadas ao bilinguismo, na visão do neurocientista Esti Blanco-Elorrieta. Durante sua fala, o especialista detalha a natureza da linguagem e sua associação com a cultura e identidade dos falantes. Ele também discorre sobre a relação entre língua e políticas, narra a trajetória do bilinguismo no âmbito das ciências e enumera algumas vantagens do cérebro bilingue que, segundo sua ótica, são mais socioculturais do que cognitivas.

Quadro 4 .Tarefa Grupo 1⁷

Tarea Grupo 1: Vas a escuchar el audio “Ventajas de hablar más de un idioma”. El podcast habla de pesquisas desarrolladas a lo largo de los años acerca de las ventajas y desventajas de ser bilingüe. Anota algunos de los puntos negativos y positivos mencionados en el audio. En parejas, discuta acerca de los puntos negativos y positivos. Un alumno va a ser la defensa y el otro será la acusación. Cinco minutos por pareja [tiempo].

Insumo 1: Ventajas de hablar más de un idioma

Fonte: Participantes do grupo 1

Quadro 5. Grade de Correção Grupo 1

Criterios de evaluación	Cumple la tarea	Cumple la tarea parcialmente	No cumple	Sugerencia del profesor
Pronunciación, los estudiantes pronunciaron de manera correcta				
Argumentación, los estudiantes argumentaron de manera clara				
Los objetivos fueron alcanzados				

Fonte: Participantes do grupo 1

Quadro 6. Tarefa Grupo 2

Tarea Grupo 2: Usted es un agente del programa de podcast “TED en Español”. Escuche el audio, luego haga un comentario de opinión defendiendo u oponiéndose al tema.

Insumo 1: Ventajas de hablar más de un idioma

Fonte: Participantes do grupo 2

Quadro 7. Grade de Correção Grupo 2

Criterios de evaluación	Cumple la tarea	Cumple la tarea parcialmente	No cumple	Sugerencia del profesor
Uso correcto de la estructura de opinión				
Coherencia con el tema citado				
Adecuación del lenguaje				

Fonte: Participantes do grupo 2

⁷ As tarefas integradas são apresentadas no idioma em que foram produzidas.

No que tange à elaboração das instruções para as tarefas integradas dos grupos 1 e 2 (quadros 4 e 6), observamos que ambas não apresentam uma situação comunicativa autêntica, semelhante ao mundo real, com interlocutores, papéis sociais e propósitos comunicativos claros. Além disso, as instruções fazem referência à situação avaliativa utilizando, por exemplo, o modo imperativo (“*Anota algunos ejemplos*”; “*Escucha el audio*”) e menções explícitas e descontextualizadas ao áudio (“*Vas a escuchar el audio ‘Ventajas de hablar más de un idioma’*”). Ou seja, as instruções não são elaboradas considerando os papéis sociais assumidos pelos interlocutores na situação comunicativa proposta, mas focam no avaliando em si. Ademais, na tarefa integrada 1 (quadro 4), nota-se que o comando solicita que os avaliandos “*anotem*” algumas informações do áudio. Nesse caso, a produção oral proposta pela tarefa seria apoiada pela produção escrita, o que poderia resultar, inclusive, em uma leitura em voz alta e não em uma produção oral espontânea e pouco monitorada.

O gênero discursivo também não aparece de forma clara em ambas as tarefas integradas. Na tarefa integrada do grupo 1 (quadro 4), é possível inferir que se trata de uma espécie de debate embora não haja especificações sobre sua operacionalização. Na tarefa integrada do grupo 2 (quadro 6), o gênero identificado é “*comentario de opinión*”, contudo, como não há o estabelecimento de interlocutores e da situação de comunicação propriamente dita, é difícil prever quais seriam as características dessa produção.

De acordo com Scaramucci (2016b) e Vicentini (2019), o comando da tarefa deve especificar o gênero discursivo da produção, preferencialmente um gênero diferente daquele do insumo, para evitar cópias. Pileggi (2015) verificou, com base na análise de tarefas integradas de compreensão e produção escrita propostas pelo exame Certificado de Proficiência em Língua Portuguesa (Celpe-Bras), que a falta de definição do gênero na situação comunicativa proposta pode diminuir a confiabilidade dos resultados. A autora, com base em Yu (2013), argumenta que uma definição mais precisa do gênero beneficiaria a correção da tarefa integrada, uma vez que os critérios de qualidade estariam especificados no próprio enunciado, em conformidade com as características da produção e o uso social do gênero.

Uma outra questão relevante é o uso do insumo estabelecido em ambas as tarefas integradas. Na tarefa do grupo 1 (quadro 5), mencionam-se “*puntos positivos y negativos mencionados en el audio*”; no entanto, o áudio apenas apresenta “vantagens” (ou seja, “pontos positivos”) de se falar mais de um idioma, e não “pontos negativos”. Da mesma forma, a tarefa do grupo 2 (quadro 7) solicita um “*comentario de opinión defendiendo u oponiéndose al tema*”, mas o áudio não apresenta informações que se opõem à tese defendida pelo neurocientista. O fato de a tarefa exigir informações que não estão disponíveis no insumo pode favorecer os avaliandos que têm mais conhecimento sobre a temática. Além disso, prejudica a avaliação da habilidade de compreensão oral e a integração de habilidades na tarefa (sub-representação do construto), uma vez que a resposta pode ser produzida sem as informações do insumo.

Na visão de Scaramucci (2016b), é fundamental definir os propósitos da compreensão oral e quais informações devem ser recuperadas do insumo para a determinação dos objetivos da produção oral (Scaramucci, 2016b; Vicentini, 2019). Esses resultados podem revelar que os professores em formação inicial não tiveram a preocupação de analisar o insumo para refletir sobre como ele poderia ser explorado em atividades alinhadas com o construto de avaliação

proposto na oficina. Não nos pareceu o caso de que eles não haviam compreendido o insumo. Vale considerar, porém, que essa dificuldade poderia ser justificada pela falta de tempo, dado que o desenvolvimento da atividade na oficina foi bastante curto.

Por outro lado, esse tipo de tarefa, que mobiliza informações não explícitas no insumo, pode assumir um caráter mais desafiador, pois exige que o avaliando faça inferências e associações a partir de seu repertório e conhecimento de mundo. Nesse caso, a avaliação estaria embasada em uma visão mais ampla de compreensão oral, que não se restringe apenas à recuperação de informações explícitas no insumo (Scaramucci, 2016b). Além disso, tende a evitar cópias e/ou paráfrases do insumo. Teríamos, assim, um construto de compreensão oral que considera que o significado do texto vai além do que está explicitamente apresentado e é construído pelo sujeito que interage com ele.

Com relação às grades de correção das tarefas integradas dos grupos 1 e 2 (quadros 6 e 8), observamos que foram escolhidos critérios bastante vagos, que não operacionalizam o construto da avaliação, pouco dialogam com as instruções da tarefa e, conseqüentemente, não comunicam claramente aos avaliandos o que se espera em termos de alcance dos resultados.

O grupo 3 (quadro 8) escolheu como insumo para a tarefa integrada o podcast “Autonomia para aprender”, também do programa “TED en Español” (publicado originalmente em setembro de 2021). Neste áudio, a educadora Melina Furman discute a importância de ensinar aos estudantes a serem autônomos e a “aprender a aprender”, com base em reflexões sobre o ensino remoto emergencial durante a pandemia de COVID-19.

Quadro 8. Tarefa Grupo 3

<i>Tarea Grupo 3: La pandemia de COVID 19 ha cambiado nuestras vidas, principalmente la de los estudiantes, que tuvieron que estudiar desde sus casas. Para eso fue necesario el estímulo de la autonomía de los alumnos. A partir del contexto en que vivimos y del podcast que van a escuchar, dividiéndose en dos equipos, un para las ventajas y otro para los inconvenientes de la autonomía para los aprendientes.</i>
<i>Insumo 2: Autonomía para aprender</i>

Fonte: Participantes do grupo 3

Quadro 9. Grade de Correção Grupo 3

<i>Criterios de evaluación</i>	<i>Cumple la tarea</i>	<i>Cumple la tarea parcialmente</i>	<i>No cumple</i>	<i>Sugerencia del profesor</i>
<i>Contexto auténtico</i>				
<i>Argumentación</i>				
<i>Objetivo alcanzado</i>				
<i>Relevancia</i>				

Fonte: Participantes do grupo 3

Na tarefa elaborada pelo grupo 3 (quadro 8), observam-se os mesmos problemas identificados nas tarefas dos grupos 1 e 2, a saber, a falta de clareza na contextualização da situação comunicativa. Além disso, no áudio não há menção às “vantagens” e aos “inconvenientes” da autonomia para os estudantes, conforme indicado nas instruções da tarefa.

Isso reforça nosso argumento anterior sobre as dificuldades dos participantes em relação aos processos envolvidos na elaboração da tarefa, especialmente no que diz respeito ao uso do insumo: a localização das informações mais relevantes do texto, o reconhecimento da situação comunicativa (quem fala, para quem fala, com que objetivo, em que registro etc.), entre outros aspectos.

Conforme já mencionamos, de acordo com Scaramucci (1995; 2006), as instruções das tarefas integradas devem esclarecer os usos esperados para os insumos, com o objetivo de aumentar a validade e a confiabilidade. No entanto, nem no comando nem na grade de correção da tarefa do grupo 3 (quadros 8 e 9) observamos esse elemento. Especialmente na grade da tarefa integrada do grupo 3 (quadro 9), notamos uma confusão entre as características das tarefas e os critérios de avaliação.

O grupo 4 escolheu trabalhar com o insumo “*Vivir sin plástico*”, que é um fragmento de uma palestra de “TED en Español”. Nesse trecho, os autores do blog “*Vivir sin plástico*” relatam sua jornada para reduzir o consumo de plástico. O casal menciona dados estatísticos que demonstram que o uso excessivo de plástico pode levar o planeta a um colapso ambiental e enumera algumas estratégias que utilizaram para diminuir seu consumo de plástico.

Quadro 10. Tarefa Grupo 4

Tarea Grupo 4: *En parejas, imagine una situación en que una de las personas no consume plásticos y otro va a pedir un consejo para sacar el plástico de su vida después de haber visto el video: “Vivir sin plástico”.*

Insumo 3: *Vivir sin plástico*

Fonte: Participantes do grupo 4

Quadro 11. Grade de Correção Grupo 4

<i>Criterios de evaluación</i>	<i>Cumple la tarea</i>	<i>Cumple la tarea parcialmente</i>	<i>No cumple</i>	<i>Sugerencia del profesor</i>
<i>Si el interlocutor consiguió aconsejar el otro</i>				
<i>Si consiguió internalizar la idea del video para pedir consejo al otro</i>				
<i>Si hay un diálogo fluido entre las personas</i>				
<i>Si hay errores gramaticales de acuerdo con el nivel de los estudiantes</i>				

Fonte: Participantes do grupo 4

Observando o comando da tarefa do grupo 4 (quadro 10), identificamos que os participantes conseguiram realizar uma boa análise do insumo e apresentar um propósito comunicativo (“*dar/pedir consejos*”) que exige uma interpretação mínima do texto e a capacidade de selecionar informações relevantes para a construção da resposta. No entanto, a situação comunicativa da tarefa não está adequadamente construída e revela claramente a situação avaliativa, o que, como já mencionamos, diminui a validade e a autenticidade da tarefa. Veja-se, inclusive, que só temos a certeza daquilo que os elaboradores esperavam como produto (um diálogo) quando identificamos o terceiro critério de correção (quadro 11).

Com relação à grade de correção elaborada pelo grupo 4 (quadro 11), observamos dois critérios bastante adequados: o primeiro contempla o propósito da tarefa (“*Si el interlocutor consiguió aconsejar al otro*”), e o segundo refere-se ao uso do insumo (“*Si consiguió internalizar la idea del video para pedir consejo al otro*”). O critério de fluidez, embora bastante comum em avaliações de habilidades orais, tem uma definição imprecisa. Ele pode ser entendido de maneira mais generalista, como a capacidade de articular uma mensagem de forma eficaz em um contexto comunicativo específico, neste caso, referindo-se à colaboração dialógica entre os avaliandos-interlocutores, ou de maneira mais técnica, como a habilidade de falar sem interrupções ou pausas excessivas que possam trazer algum prejuízo ao cumprimento dos objetivos comunicativos do falante, ou seja, o conceito de fluência. Como os participantes não forneceram descritores para esse critério, não é possível identificar qual conceito de fluidez foi adotado.

Por outro lado, a noção de “erro gramatical” mencionada na grade (“*Si hay errores gramaticales de acuerdo con el nivel de los estudiantes*”) é bastante vaga, mas evidencia o impacto da tradição normativa no ensino: o de buscar – e penalizar – os erros. Ainda que a precisão gramatical seja um dos importantes indicadores de desempenho em uma tarefa, as grades de correção devem orientar o olhar do professor às características de *qualidade* do produto, e não da falta dela.

Quadro 12. Tarefa Grupo 5

Tarea Grupo 5: *Dos alumnos se encuentran en una cola para el autobús, uno de ellos reconoce al otro y empiezan a charlar. Después de oír el audio [audio gravado por los propios participantes], imagínate en una situación similar en que encuentras un compañero de clase que solo conocía virtualmente. Haz un diálogo como si estuvieras en la misma situación.*

Insumo: *producido por los participantes*

Fonte: Participantes do grupo 5

Quadro 13. Grade de Correção Grupo 5

<i>Criterios de evaluación</i>	<i>Cumple la tarea</i>	<i>Cumple la tarea parcialmente</i>	<i>No cumple</i>	<i>Sugerencia del profesor</i>
<i>Comprensión auditiva: si el alumno comprende el contexto del audio</i>				
<i>Interlocutor: es posible identificar los interlocutores de la tarea</i>				
<i>Propósito: si el alumno consigue hacer un dialogo conforme se pide en el enunciado</i>				

Recursos lingüísticos: entonación, pronunciación, además de la estructura del diálogo				
---	--	--	--	--

Fonte: Participantes do grupo 5

A tarefa do grupo 5 (quadro 12) apresenta uma particularidade. Os participantes optaram por criar seu próprio insumo gravando um áudio que simula o (re)encontro presencial de dois colegas que se conheciam apenas virtualmente, devido ao ensino remoto emergencial adotado durante a pandemia de COVID-19. O uso desse insumo limita-se à compreensão da situação comunicativa simulada, que deveria ser reproduzida na resposta da tarefa. Essa abordagem pode levar à produção de cópias e/ou paráfrases, conforme apontado por Scaramucci (2016b). Além disso, a criação do insumo próprio especificamente para a atividade resultou em um texto pouco autêntico, embora a situação comunicativa apresentada no insumo e no comando da tarefa seja realista.

Assim, a proposta da tarefa evidencia uma concepção de língua arraigada nas práticas estruturalistas, e teria, como resultado, não uma produção, mas uma *reprodução*. Esses resultados corroboram Brum-Barreto (2023), que, nesse mesmo contexto, observou a predominância de uma visão de língua reducionista e estruturalista entre os professores de espanhol em formação inicial que participaram de seu estudo. Tal concepção de linguagem é diametralmente oposta àquela proposta na oficina e dificulta o desenvolvimento de ações de formação fundamentadas em teorias sociointeracionistas ou acionais.

Em contrapartida, observamos, na construção da grade de correção da tarefa do grupo 5 (quadro 13), um movimento em direção a critérios relevantes, como o uso do insumo (“*si el alumno comprende el contexto del audio*”) e a referência a aspectos linguísticos condizentes com a modalidade oral da língua (“*recursos lingüísticos: entonación, pronúncia, além da estrutura do diálogo*”).

À luz dessas análises, que evidenciaram que os comandos apresentaram situações comunicativas realísticas podemos considerar que a atividade de elaboração de tarefas integradas proporcionou a sensibilização dos participantes quanto ao entendimento de língua em uso e quanto ao fato de que é preciso haver instrumentos, como grades, para orientar a correção e para que haja transparência quanto aos critérios junto aos alunos.

Porém, pelas características das grades de correção, cujos critérios não dialogam com as instruções das tarefas, percebemos que a oficina não foi suficiente para que a maioria compreendesse língua como ação situada. Destarte, podemos concluir que houve impactos positivos dessa ação de formação, porém limitados.

Evidentemente, deve-se considerar o tempo limitado da oficina e o contexto em que foi desenvolvida, ou seja, com alunos de um curso de formação inicial de professores que não tem um componente curricular específico sobre avaliação. Uma oficina de três horas não é suficiente para produzir mudanças em concepções de linguagem e construir conhecimentos, habilidades e princípios – ainda que básicos – em avaliação, embora possa despertar o interesse dos participantes e motivá-los a buscar novas oportunidades para desenvolver seu LAL.

Diante disso, recomendamos fortemente a inclusão de um componente curricular obrigatório que aborde os conhecimentos, habilidades e princípios subjacentes às práticas avaliativas considerando o uso situado da língua.

Considerações finais

Esta investigação teve como objetivo relatar a experiência de uma oficina de elaboração de tarefas para avaliar habilidades orais integradas, com foco na promoção do LAL de professores de espanhol em formação inicial, explorando suas potencialidades e limitações. Sob nossa perspectiva, embora tenha proporcionado o contato com uma visão de língua em uso e sua operacionalização em um instrumento de avaliação para contextos de ensino-aprendizagem de E/LA (avaliação de rendimento), a oficina teve um impacto limitado na concepção de língua e nos conhecimentos e habilidades práticas dos participantes. Para irmos além da sensibilização sobre a importância e a complexidade técnica de elaborar instrumentos avaliativos de qualidade, em direção a um nível de letramento adequado para professores de línguas, precisamos de ações de formação mais extensas. Nesse sentido, dada a complexidade do tema, reiteramos que a temática da avaliação deve ser incluída nos currículos dos cursos de formação inicial de professores de línguas como componente obrigatório.

Referências

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1). Oxford University Press.
- Bakhtin, M. (1997). *Estética da criação verbal* (2a ed.). Martins Fontes.
- Bakhtin, M. (2016). *Os gêneros do discurso* (5a ed.). Editora 34.
- Baralo, M. (2000). El desarrollo de la expresión oral en el aula de E/LE. *Carabela*, 47, 5-36. https://cvc.cervantes.es/ensenanza/biblioteca_ele/carabela/pdf/47/47_005.pdf
- Bardin, L. (2006). *Análise de conteúdo*. Edições 70.
- Borges-Almeida, V. (2023a, septiembre). *Insights and challenges for preservice language teacher assessment literacy in Brazil* [Presentación en mesa redonda]. V Seminário de Avaliação de Proficiência em Língua Estrangeira – SEMAPLE 2023, São José do Rio Preto, Brasil.
- Borges-Almeida, V. (2023b, mayo). *Preservice language teacher needs and interests in assessment: towards an emergency agenda* [Presentación oral]. New Directions LATAM, São Paulo, Brasil.
- Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, 3, 1-12. <https://doi.org/10.3389/educ.2018.00022>
- Brown, J. D. (1996). *Testing in language programs*. Prentice Hall Regents.
- Brum-Barreto, A. N. (2023). *Letramento em avaliação na formação do professor de espanhol: visão de língua e cultura de avaliar no estágio obrigatório* [Tesis de maestría, Universidade de Brasília].

- Bygate, M. (2010). Speaking. En R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 14-20). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667206>
- Cantero Serena, F. J. (2020). Didáctica de la pronunciación: de la corrección fonética al Enfoque Oral. En F. J. Cantero Serena & M. Giralt (Coords.), *Pronunciación y enfoque oral en lenguas extranjeras* (pp. 1-42). Octaedro.
- Consejo de Europa. (2020). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación. Volumen complementario*. Servicio de publicaciones del Consejo de Europa. www.coe.int/lang-cefr
- Conselho da Europa. (2001). *Quadro europeu comum de referência para as línguas: Aprendizagem, ensino, avaliação. Edição portuguesa*. Edições Asa. https://area.dge.mec.pt/gramatica/quadro_europeu_total.pdf
- Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*. Michigan Teacher Training.
- Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, 10(1), 1-8. <https://doi.org/10.1080/15434303.2011.622016>
- Damacena-Dutra, M. M. (2023). *A construção do letramento em avaliação na formação inicial de professores de inglês: avaliação de produção oral em foco* [Tesis de maestría, Universidade de Brasília].
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <https://doi.org/10.1177/0265532208090156>
- Díaz, D., Quevedo-Camargo, G., & Borges-Almeida, V. (2019). As situações de comunicação do gênero e seu papel na preparação do Celpe-Bras. *Letra Magna*, 16, 823-845.
- Frost, K., Elder, C., & Wigglesworth, G. (2012). Investigating the validity of an integrated listening-speaking task: A discourse-based analysis of test takers' oral performances. *Language Testing*, 29(3), 345-369. <https://doi.org/10.1177/0265532211424479>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. <https://doi.org/10.1080/15434303.2011.642041>
- Gil, A. C. (2002). Como classificar as pesquisas. *Como elaborar projetos de pesquisa*, 4(1), 44-45.
- Giraldo, F. (2021). A reflection on initiatives for teachers' professional development through language assessment literacy. *Profile Issues in Teachers Professional Development*, 23(1), 197-213. <https://doi.org/10.15446/profile.v23n1.83094>
- Giraldo, F., & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal*, 21(2), 243-259. <https://doi.org/10.14483/22487085.14514>
- Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. En D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 413-428). De Gruyter. <https://doi.org/10.1111/inc3.12124>

- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385-402. <https://doi.org/10.1177/0265532208090158>
- Jaramillo Delgado, J. J., & Bedoya, A. M. G. (2019). Pre-service English language teachers' use of reflective journals in an assessment and testing course. *JSR Funlam Journal of Students' Research*, 4, 210-218. <https://doi.org/10.21501/25007858.3010>
- Levelt, W. J. (1995). The ability to speak: From intentions to spoken words. *European Review*, 3(1), 13-23. <https://doi.org/10.1017/S1062798700001290>
- Luoma, S. (2004). Speaking scales. En S. Luoma, *Assessing speaking* (pp. 59-95). Cambridge University Press.
- Ortiz-Preuss, E. (2014). Habilidade oral em L2: da cognição à interação. *Revista Horizontes de Linguística Aplicada*, 13(2), 167-186. <https://doi.org/10.26512/rhla.v13i2.1370>
- Pileggi, M. G. S. (2015). *Tarefas integradas nos exames de proficiência CELPE-BRAS e TOEFL iBT* [Tesis de maestría, Universidade Estadual de Campinas]. Repositório da produção científica e intelectual da Unicamp. <https://hdl.handle.net/20.500.12733/1626538>
- Plakans, L. (2013). Writing integrated items. En G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 263-275). Routledge. <https://doi.org/10.4324/9780203181287>
- Plakans, L. (2015). Integrated second language writing assessment: Why? What? How? *Language and Linguistics Compass*, 9(4), 159-167.
- Plakans, L., & Gebril, A. (2017). Exploring the relationship of organization and connection with scores in integrated writing assessment. *Assessing Writing*, 31, 98-112. <https://doi.org/10.1016/j.asw.2016.08.005>
- Quevedo-Camargo, G. (2014). Efeito retroativo da avaliação na aprendizagem de línguas estrangeiras: que fenômeno é esse? En K. B. Mulik & M. E. Retorta (Eds.), *Avaliação no ensino-aprendizagem de línguas estrangeiras: diálogos pesquisas e reflexões* (pp. 77-93). Pontes Editores.
- Quevedo-Camargo, G. (2020). Formação de professores de línguas adicionais e letramento em avaliação: breve panorama e desafios para os cursos de licenciatura em LEM no Brasil. *Calidoscópio*, 18(2), 435-459. <https://doi.org/10.4013/cld.2020.182.10>
- Restrepo Bolívar, E. M. (2020). Monitoring preservice teachers' language assessment literacy development through journal writing. *Malaysian Journal of ELT Research*, 17(1), 38-52. <https://meltajournals.com/index.php/majer/article/view/558/541>
- Sampieri, R. H., Collado, C. F., & Lucio, M. P. B. (2013). *Metodologia de pesquisa* (5a ed.). Penso.
- Scaramucci, M. V. R. (1995). O projeto Celpe-Bras no âmbito do Mercosul: contribuições para uma definição de proficiência comunicativa. En J. C. P. Almeida Filho (Ed.), *Português para Estrangeiros: Interface com o Espanhol* (pp. 77-90). Pontes.

- Scaramucci, M. V. R. (2006). O professor avaliador: sobre a importância da avaliação na formação do professor de língua estrangeira. En L. Rottava (Ed.), *Ensino-aprendizagem de línguas: língua estrangeira* (pp. 49-64). Editora da UNIJUÍ.
- Scaramucci, M. V. R. (2016a). Letramento em avaliação (em contexto de línguas): contribuições para a Linguística Aplicada, educação e sociedade. En C. M. Jordão (Ed.), *A Linguística Aplicada no Brasil: Rumos e passagens* (pp. 141-166). Pontes.
- Scaramucci, M. V. R. (2016b). A avaliação de habilidades integradas na parte escrita do Exame Celpe-Bras. En M. L. O. Alvarez & L. Gonçalves (Eds.), *O mundo do português e o português no mundo afora: Especificidades, implicações e ações* (pp. 391-425). Pontes Editores.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1), 38-62. <https://doi.org/10.1093/applin/17.1.38>
- Vicentini, M. P. (2019). As tarefas de compreensão oral para produção escrita no Celpe-Bras: construto e operacionalização. *Estudos Linguísticos*, 48(1), 561-580. <https://doi.org/10.21165/el.v48i1.2141>
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402. <https://doi.org/10.1080/15434303.2014.960046>