

# Avaliação em contexto de línguas: conceitos teóricos e exemplos práticos em língua portuguesa

Juliana Roquele Schoffen

RETORTA, Miriam Sester; MAROCHI, Thaís Barbosa.  
*Avaliação em línguas estrangeiras: da teoria à prática.*  
Curitiba: CRV, 2018. 162 p.

As possibilidades contemporâneas de mobilidade acadêmica e profissional, bem como as amplas facilidades de comunicação criadas pelas novas tecnologias, que facilitam a relação entre pessoas de diferentes lugares, têm contribuído para a discussão sobre o ensino de línguas, visto que a proficiência em mais de uma língua possibilita maior gama de oportunidades. Também na área da pesquisa acadêmica, o ensino e a avaliação de línguas, tanto materna como adicionais, têm ensejado um crescente número de estudos no cenário internacional, pois, além de ser necessário preparar os alunos para essa nova realidade, as próprias pesquisas acadêmicas, para terem maior alcance e, portanto, maior impacto, devem preferencialmente ser publicadas em mais de uma língua. Esse cenário e suas particularidades fizeram com que a área de pesquisa em avaliação em contexto de línguas se solidificasse internacionalmente nos últimos anos, contando com um grande número de novas publicações no âmbito da avaliação escolar e da avaliação em larga escala.

No Brasil, especialmente no contexto de línguas, a área de avaliação é ainda bastante tímida, contando com poucos pesquisadores e um número relativamente pequeno de trabalhos publicados até o momento. Nesse sentido, iniciativas como a publicação do livro *Avaliação em línguas estrangeiras: da teoria à prática*, de Miriam Sester Retorta e Thaís Barbosa Marochi, são muito bem-vindas, visto que vêm preencher uma lacuna na produção nacional. A obra trata, com muita objetividade, das questões mais centrais para essa área, apresentando conceitos-chave para se compreender os principais desafios na elaboração, aplicação e análise de um bom instrumento de avaliação.

Ao optarem por iniciar o livro com um capítulo sobre ética, as autoras, experientes professoras universitárias, doutoras em Letras e Linguística Aplicada na área de concentração Avaliação no Ensino de Línguas, sinalizam um entendimento relevante: um bom sistema de avaliação não pode prescindir de um bom código de ética. A avaliação, especialmente em larga escala, tem um impacto grande na vida das pessoas e em toda a sociedade. Mesmo a avaliação feita em sala de aula é de extrema relevância para os envolvidos e precisa ser eticamente conduzida. É imprescindível, então, que esses processos sejam pautados pela ética, tanto de quem elabora e aplica os exames quanto de quem faz uso deles. Ao mencionarem essa questão, as autoras deixam claro que a principal qualidade de uma boa avaliação, mais importante que a confiabilidade ou a validade, é ser eticamente conduzida.

A avaliação sempre abarca relações de poder. Scaramucci (2018, p. 11), no prefácio do livro em questão, afirma que se trata de “um processo que envolve a coleta sistemática de evidências sobre o que as pessoas sabem ou conseguem fazer e é sempre orientada para a tomada de decisões”, que “podem ser positivas, mas também nefastas e prejudiciais, se a coleta e a interpretação dessas evidências não forem informadas pela ética e por conhecimentos retóricos adequados”. Uma avaliação eticamente conduzida e adequadamente elaborada, portanto, é fundamental para se pensar em como resolver questões educacionais e sociais do Brasil, pois, se é na avaliação que se encontram os maiores problemas, é também nela que estão as soluções (Scaramucci, 2006).

O segundo capítulo é dedicado à história da avaliação de línguas. Seguindo o entendimento de que ela é inerente à organização social, as autoras apresentam suas tendências desde a antiguidade chinesa, passando pela Grécia antiga e chegando à Inglaterra do século 19, época em que os primeiros exames de proficiência em larga escala na área de línguas foram desenvolvidos na Universidade de Cambridge.

O percurso histórico da avaliação sempre esteve relacionado às concepções de língua e linguagem e às abordagens de ensino em voga nos diferentes momentos. Nesse sentido, para compreender a história dos instrumentos de avaliação, é preciso compreender as diferentes concepções de linguagem e abordagens de ensino e aprendizagem adotadas ao longo do tempo.

Para contextualizar adequadamente o leitor que talvez esteja começando suas leituras nessa área, é apresentada, inicialmente, a divisão clássica e muito difundida dos períodos da história da avaliação feita por Spolsky (1977). Segundo essa divisão, o período anterior aos anos 1940 era chamado de período pré-científico ou Jardim do Éden. Nele, as avaliações decorriam das observações empíricas de seus criadores, já que os testes realizados antes da Segunda Guerra Mundial não possuíam embasamento científico. O período entre 1940 e 1960 era chamado de psicométrico-estruturalista ou Vale de Lágrimas, uma vez que os sistemas de avaliação se baseavam nas teorias estruturalistas de ensino e aprendizagem. Nessa época, a linguagem era vista como um sistema de comunicação, e a avaliação, embasada em uma visão psicométrica, privilegiava as provas de itens isolados. Já o período posterior aos anos 1970 foi nomeado de psicolinguístico-sociolinguístico ou Terra Prometida, quando as culturas de ensino e avaliação passaram a considerar a dimensão social da linguagem. De acordo com essa perspectiva, além de regras gramaticais, os

enunciados produzidos em determinada língua também deveriam obedecer a regras de uso, que seriam tão importantes quanto as de gramática e, portanto, também deveriam ser ensinadas e avaliadas.

Depois da divisão clássica da história da avaliação em períodos, as autoras trazem a atualização nos estudos da área, apresentando o novo entendimento de Spolsky (2017) sobre esses períodos. O autor, contemporaneamente, esboça uma série de eventos importantes na história da avaliação e da avaliação de línguas, nomeando-os de acordo com as abordagens recorrentes em diferentes momentos históricos para avaliar as habilidades humanas.

No capítulo três, são apresentados diferentes tipos de testes, propósitos e critérios de avaliação. Os tipos de testes são: a) de proficiência, b) de nivelamento, c) de rendimento, d) diagnósticos, e) de aptidão, f) de suficiência, e g) de entrada. Os propósitos de uma avaliação podem ser selecionar, classificar e diagnosticar, mas as autoras afirmam que os dois primeiros costumam ser utilizados concomitantemente. Em relação aos critérios, quando os resultados dos testes são referenciados em norma significa que o escore individual é interpretado em relação ao desempenho de um grupo; quando são referenciados em critério, o escore individual é mensurado em relação a objetivos pré-estabelecidos. Ao mencionar essa diferença as autoras afirmam que, enquanto o propósito principal dos testes referenciados em critério é medir a capacidade de um determinado indivíduo, o principal propósito dos referenciados em norma é classificar, discriminando os examinandos em bons, regulares e fracos. Os tipos de testes e os critérios apresentados são ilustrados com exemplos, e há menção explícita a quais tipos podem funcionar melhor em qual contexto, a depender dos propósitos para os quais a avaliação está sendo feita.

No capítulo quatro, intitulado “Tendências atuais nos testes de língua”, são expostas as diferenças entre os diretos, que exigem o desempenho do examinando na habilidade que se quer avaliar – como os de escrita, por exemplo, que solicitam ao examinando a produção de um texto –, e os indiretos, que verificam as habilidades subjacentes aos conhecimentos linguísticos que estão sendo avaliados. O livro traz também a diferenciação entre testes de itens isolados (que avaliam apenas um conhecimento linguístico de cada vez e, em geral, uma questão não está relacionada à outra) e testes integradores (que procuram combinar vários elementos linguísticos de uma só vez, como em uma situação real de uso da língua), bem como entre testes de conhecimento (que tendem a enfatizar o conhecimento do sistema da língua e, por isso, costumam ser indiretos e de itens isolados) e testes de desempenho (que exigem o uso linguístico em tarefas que simulam uma ação possível de ocorrer na vida real). Por fim, são retratados os testes comunicativos, os quais, segundo as autoras, são a tendência contemporânea em avaliação e buscam verificar a capacidade do aluno em usar a língua-alvo em situações reais.

Ao apresentar cada tipo de teste e suas características, as autoras mencionam vantagens e desvantagens de cada um, apontando para quais propósitos e para quais grupos um e outro tipo seriam mais indicados.

O capítulo cinco trata da validação, apontando as principais características de que um teste precisa para ser considerado um bom instrumento de avaliação:

validade, efeito retroativo, confiabilidade e praticidade. Na validade estão incluídos: 1) o conceito tradicional, que mencionava validade de construto, de conteúdo, de face e validade referenciada em critério; e 2) o conceito contemporâneo, que se unifica em torno da validade de construto, mas no qual a dimensão consequencial é tão importante quanto a evidencial, visto que contemporaneamente se entende o teste como uma prática social, considerando-se assim as dimensões sociais, históricas e políticas nas quais a avaliação está inserida.

Alinhadas com o entendimento contemporâneo da área de avaliação, as autoras expressam que, em si mesmo, um teste não é válido ou inválido; o que valida um teste é seu uso em um determinado contexto. Portanto, para validar uma determinada avaliação, é imprescindível considerar os resultados de um teste no contexto em que é utilizado.

No capítulo seis, que é o mais extenso, são apresentados os estágios de elaboração de um teste, que passam por determinar o seu propósito, planejar, selecionar o método mais adequado de avaliação, administrar o teste, corrigir e analisar os resultados. São exibidos diversos métodos de avaliação, com explicações sobre como devem ser elaboradas as questões em cada um deles, menção às suas vantagens e desvantagens, considerações sobre sua autenticidade, explicação de como é feita a correção, indicação de seus melhores e piores usos e relatos da visão do aluno/candidato. As autoras acrescentam, ainda, uma crítica a cada um dos métodos, problematizando-os e enfatizando suas limitações.

Há ainda uma seção específica sobre a correção dos testes, mencionando as diferenças entre as escalas holísticas, que costumam ser mais válidas, já que fornecem uma visão mais geral do teste e nas quais o avaliador não se concentra em aspectos específicos da língua, e as analíticas, que são mais detalhadas e costumam avaliar separadamente distintos aspectos linguísticos, como vocabulário, gramática e fluência. Há também uma seção informativa sobre a análise dos resultados de um teste, incluindo exemplos para determinar o grau de dificuldade e o índice de discriminação de um teste ou de uma questão. A última seção do capítulo trata das questões que precisam ser consideradas na apresentação dos resultados de um teste, uma das etapas mais importantes de um processo de avaliação.

Apesar de não trazer grandes novidades à área de avaliação, visto que não relata resultados de pesquisas, o livro *Avaliação em línguas estrangeiras: da teoria à prática* é uma das poucas publicações em língua portuguesa que apresentam os conceitos fundantes da área, o que, por si só, é uma contribuição bastante importante. Por mesclar os conceitos teóricos com indicações práticas, percebe-se que pode ser útil para a formação inicial de professores de línguas, dada a dificuldade de encontrar materiais em português para disciplinas de cursos de graduação, e mesmo para a formação continuada, visto que pode ensejar um debate interessante se utilizado em escolas ou cursos de atualização para professores.

## Referências bibliográficas

---

SCARAMUCCI, M. V. R. Prefácio. In: RETORTA, M. S.; MAROCHI, T. B. *Avaliação em línguas estrangeiras: da teoria à prática*. Curitiba: CRV, 2018. p. 11-14.

SCARAMUCCI, M. V. R. O professor avaliador: sobre a importância da avaliação na formação do professor de língua estrangeira. In: ROTTAVA, L.; SANTOS, S. R. (Org.). *Ensino-aprendizagem de línguas: língua estrangeira*. Ijuí: Ed. da Unijuí, 2006. p. 49-64.

SPOLSKY, B. Language testing: art or Science. In: NICKEL, G. (Ed.). *Proceedings of the Fourth International Congress of Applied Linguistics*. Stuttgart: Hochschulverlag, 1977. v. 3, p.7-28.

SPOLSKY, B. History of language testing. In: SHOHAMY, E.; OR, I. G.; MAY, S. (Ed.). *Language testing and assessment*. Auckland: Springer Reference, 2017. p. 375-384.

---

Juliana Roquele Schoffen é professora associada do Instituto de Letras da Universidade Federal do Rio Grande do Sul (UFRGS). Realiza pesquisas na área de avaliação em contexto de línguas e lidera o Grupo de Pesquisa Avaliação de Uso da Linguagem (Avalia).

[julianaschoffen@gmail.com](mailto:julianaschoffen@gmail.com)

Recebido em 29 de janeiro de 2018

Aprovado em 4 de fevereiro de 2019