

AVALIAÇÕES EM LARGA ESCALA DE LÍNGUA PORTUGUESA: UMA PESQUISA SOBRE A COMPLEXIDADE DOS TEXTOS QUE DÃO SUPORTE A ITENS QUE AVALIAM A LEITURA

Begma Tavares Barbosa¹

Hilda Linhares Micarello²

Rosângela Veiga Ferreira³

¹ Professora titular aposentada da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: begmatb@gmail.com

² Professora associada da Faculdade de Educação da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: hilda.micarello@uab.ufjf.br

³ Professora do Colégio de Aplicação João XXIII, da Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail para contato: rosangelaveiga.ferreira@ufjf.edu.br

Resumo

Este artigo discute o tema da complexidade dos textos que dão suporte a itens de avaliações em larga escala de Língua Portuguesa, com base em resultados de uma pesquisa no âmbito da qual foram analisados 763 textos utilizados em itens de avaliações em larga escala de Língua Portuguesa. A metodologia da pesquisa consistiu em análise dos itens sustentada pelas teorias da cognição, da linguística textual, e nos parâmetros estatísticos dos itens com base na Teoria da Resposta ao Item, modelo de três parâmetros. Com base na pesquisa realizada, foram definidos quatro níveis de complexidade dos textos. Os resultados apontam a relevância de critérios claros para definir a complexidade dos textos na elaboração de novos itens e para uma descrição mais precisa de padrões de desempenho com base na proficiência em leitura demonstrada pelos estudantes avaliados.

Palavras-chave: Avaliação da leitura. Avaliação em larga escala de Língua Portuguesa. Complexidade dos Textos.

Abstract

This article discusses the complexity of texts that support large scale Portuguese Language assessment items, based on the results of a research in which 763 texts used in large scale Portuguese Language assessment items were analyzed. The research methodology consisted of item analysis supported by theories of cognition, textual linguistics, and the statistical parameters of the items based on the Item Response Theory, a three parameter model. Based on the research, four levels of text complexity were defined. The results point to the relevance of clear criteria for defining the complexity of texts in the elaboration of new items and for a more accurate description of performance standards based on the reading proficiency demonstrated by the students evaluated.

Key words: Reading assessment. Large-scale Portuguese Language Assessment. Text Complexity.

INTRODUÇÃO

A pesquisa, cujos resultados são apresentados neste artigo, foi apoiada e financiada pela Fundação CAEd, da Universidade Federal de Juiz de Fora. O objetivo da pesquisa foi produzir conhecimentos sobre a validade e confiabilidade de itens aplicados em avaliações sistêmicas da leitura aplicadas em municípios e estados da federação e em avaliações nacionais. O recorte da pesquisa tratado no presente artigo se refere ao tema da complexidade dos textos que dão suporte a esses itens.

A avaliação da leitura é um tema desafiador, por ser este um construto⁴ complexo e multidimensional. Kleiman (2019, s/p) afirma que

na compreensão leitora, estão envolvidos: um texto – objeto linguístico e cultural portador de um significado -; um leitor - com saberes, experiências, capacidades e habilidades; e uma situação comunicativa de interação entre leitor e autor via texto escrito, que determina em grande parte o *que* e *como* se compreende.

No caso específico da presente pesquisa, o foco se coloca no texto como objeto linguístico portador de um significado e nos desafios que ele coloca ao leitor para que este seja capaz de compreender o que lê.

O tema da complexidade dos textos que são utilizados como suporte de itens que avaliam a leitura em avaliações de larga escala de Língua Portuguesa vem se apresentando como relevante, na medida em que a consideração dessa complexidade é fundamental para a compreensão do nível de dificuldade das tarefas de leitura propostas pelos itens aplicados com o intuito de aferir a proficiência em leitura dos estudantes, nas diversas etapas da educação básica. Além disso, critérios mais claros para a definição da complexidade dos textos contribuem para maior confiabilidade dos instrumentos de avaliação e, conseqüentemente, para a validade dos testes. De acordo com Primi (2012, p.300), em psicometria o conceito de validade “relaciona-se à questão que investiga se o teste está medindo o construto que se propõe medir”. Nesse sentido, sedo a leitura uma atividade cognitiva multidimensional, uma maior clareza sobre a natureza da atividade de leitura que está sendo requerida pelos itens que compõem os testes terá como conseqüência um melhor ajuste do teste ao que se pretende avaliar em cada etapa da escolarização.

⁴ Em psicometria, o constructo é um atributo intangível, com manifestação variável entre indivíduos, que só pode ser avaliado indiretamente, quando expresso sob a forma de performances físicas, emocionais, atitudinais ou cognitivas. Disponível em <http://portal.inep.gov.br/documents/186968/484421/Relat%C3%B3rio+Saeb+2017/e683ba93-d9ac-4c2c-8f36-10493e99f9b7?version=1.0>. Acesso em: 10 dez. 2019.

Sousa e Hübner (2014) afirmam que a opção por um tipo de tarefa para avaliar a compreensão leitora requer a explicitação da concepção de leitura que fundamenta essa opção. No caso específico de que trata o presente artigo, a concepção de leitura que tem fundamentado não apenas os processos de avaliação desse construto, mas, principalmente, as práticas de ensino é a de que a leitura é uma construção subjetiva de sentidos com base nos significados do texto, ou seja, os leitores atuam sobre o texto a partir de suas vivências em uma determinada cultura. Diante de um texto, os leitores acionam conhecimentos mais ou menos formalizados, como aqueles sobre os textos e sobre a própria língua, adquiridos na escola ou fora dela. Desse modo, o texto não porta um sentido pré-existente à leitura, ou seja, o significado não está no texto, mas se constitui na interação entre o leitor e o texto. Este, por sua vez, oferece um conjunto de pistas que guiam o leitor na tarefa de construção de sentido que é a leitura.

Os Parâmetros Curriculares Nacionais de Língua Portuguesa (BRASIL, 1998, p.36) consagraram uma virada pragmática no ensino da língua materna ao definirem o texto como “unidade básica de ensino”. Essa definição implica em assumir que ensinar a língua é, principalmente, ensinar como identificar as pistas que o texto oferece e a se guiar por essas pistas para produzir sentido para o que se lê.

Mais recentemente, a Base Nacional Comum Curricular (BRASIL, 2017) para o componente Língua Portuguesa manteve a ênfase no texto como unidade de ensino da língua, ao organizar as competências e habilidades do currículo em campos de atuação dos sujeitos na vida social. À noção de campos de atuação está subjacente a assunção de que os sujeitos atuam na sociedade por meio dos textos que nela circulam e que se materializam em gêneros que apresentam características comuns, face ao campo da vida social em que circulam. Nesse sentido, cabe pensar que nas práticas de ensino e, conseqüentemente, de avaliação, uma questão fulcral é definir de que modo progride a abordagem dos textos e quais os percursos cognitivos dos estudantes na interação que estabelecem com eles. Essa interação vai depender da estrutura do texto e do tipo de sinalização que ele oferece ao leitor para construir sentido para o lido. Assim, textos de um mesmo gênero podem apresentar estruturas distintas, o que demandará ao leitor a mobilização de habilidades também distintas.

Em geral, a complexidade dos textos é tratada, nas práticas de ensino e nas avaliações em larga escala, com base na noção de gênero textual. Essa noção tem sua origem nos estudos de Mikhail Bakhtin, filósofo da linguagem,

que afirma: “Evidentemente, cada enunciado particular é individual, mas cada campo de utilização da língua elabora seus *tipos relativamente estáveis de enunciados*, os quais denominamos *gêneros do discurso*.” (BAKHTIN, 2003, p. 262 – grifo do autor). Os textos que têm origem nas diferentes esferas da vida social apresentam, portanto, características comuns, decorrentes da relativa estabilidade dos enunciados que se originam nessas esferas. Graças a tais características comuns, pode-se identificar o gênero textual. Uma carta pessoal, por exemplo, tem características diversas de um conto.

Sem dúvida, a noção de gênero textual é relevante para orientar a progressão das práticas de ensino da leitura, visto que é possível pensar que o ensino deva partir de gêneros que circulem em esferas mais próximas à vida cotidiana dos estudantes, portanto menos complexos, e avançar para gêneros que circulem em esferas mais especializadas, portanto, mais complexos. Também as avaliações em larga escala têm se valido dessa noção, tanto para a construção dos testes quanto para a divulgação de seus resultados. Entretanto, a categoria gênero não é suficiente para revelar, em maior detalhe, como os estudantes interagem com os textos, uma vez que tal processo de interação envolve diferentes habilidades que são mobilizadas pelo leitor não apenas em função do gênero textual com o qual se põe em interação, mas também em função da própria estrutura dos textos e do tipo de tarefa de leitura que é demandada do leitor.

Neste artigo se pretende, com base nos resultados obtidos em uma pesquisa desenvolvida pelo Centro de Avaliação e Políticas Públicas para a Educação (CAEd) da Universidade Federal de Juiz de Fora, apresentar uma classificação da complexidade dos textos em níveis, definidos a partir de algumas categorias, com base numa incursão a um conjunto amplo de textos utilizados como suportes de itens que avaliam a leitura, aplicados a estudantes do 5º ao 12º ano da educação básica, em diferentes estados e municípios da federação e também em avaliações nacionais, no Brasil.

Inicialmente, serão apresentadas algumas considerações sobre como o tema da complexidade dos textos tem sido tratado na divulgação dos resultados do *Programme for International Student Assessment (PISA)*, relatório de 2018, e no relatório do SAEB/ Prova Brasil, por meio da escala de proficiência em Língua Portuguesa.

Em seguida, é apresentada a metodologia e o *corpus* da pesquisa em que se fundamenta o presente artigo, assim como a classificação dos níveis de complexidade do texto a que se chegou com base nos dados da pesquisa realizada.

Finalmente, são tecidas algumas considerações sobre a relevância de haver maior clareza sobre o nível de complexidade do texto para os processos de elaboração dos testes que avaliam a leitura, assim como para a comunicação dos resultados das avaliações às escolas e aos professores.

A COMPLEXIDADE DOS TEXTOS NA COMUNICAÇÃO DE RESULTADOS DE AVALIAÇÕES INTERNACIONAIS E NACIONAIS

No relatório do Sistema de Avaliação da Educação Básica (SAEB) (BRASIL, 2019) são apresentados os resultados das avaliações de Língua Portuguesa e Matemática realizadas, no ano de 2017, no âmbito da Prova Brasil, em todo o território nacional. Nesse relatório, os desempenhos dos estudantes em ambos os componentes curriculares são apresentados por meio de uma escala de proficiência⁵. Com base na escala de proficiência, esses desempenhos são interpretados e descritos, na avaliação de Língua Portuguesa, leitura, em 9 (nove) níveis. Essa forma de apresentação e divulgação de resultados é aquela que vem sendo utilizada pelo INEP ao longo das diferentes edições da Prova Brasil.

No quadro 1, são apresentadas as descrições dos três primeiros níveis de desempenho da escala de proficiência em Língua Portuguesa, tal como descritos no relatório do SAEB (BRASIL, 2019).

⁵ Escala de proficiência: Conjunto de números ordenados obtido pela Teoria de Resposta ao Item (TRI) que representa a medida da proficiência em uma determinada área de conhecimento. No SAEB, em cada área do conhecimento, há uma escala de proficiência única para todos os anos escolares (séries) e todas as edições da avaliação (Ex.: a escala de proficiência em Língua Portuguesa da edição de 2017 é a mesma das edições anteriores e agrega as proficiências dos estudantes que participaram dos testes do 5º e do 9º ano do ensino fundamental e da 3ª série do ensino médio). Cada escala varia de 0 a 500 pontos, com média de 250 e desvio padrão de 50. Disponível em <http://portal.inep.gov.br/documents/186968/484421/Relat%C3%B3rio+Saeb+2017/e683ba93-d9ac-4c2c-8f36-10493e99f9b7?version=1.0> Acesso em: 10 dez. 2019.

Quadro 1 – Níveis de Desempenho em Língua Portuguesa segundo a escala de proficiência da SAEB (fragmento).

NÍVEL	DESCRIÇÃO DO NÍVEL
Nível 1 Desempenho maior ou igual a 125 e menor que 150	Os estudantes provavelmente são capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios. Identificar o tema de um texto. Localizar elementos, como o personagem principal. Estabelecer relação entre partes do texto: personagem e ação; ação e tempo; ação e lugar.
Nível 2 Desempenho maior ou igual a 150 e menor que 175	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informações explícitas em contos. Identificar o assunto principal e a personagem principal em reportagem e em fábulas. Reconhecer a finalidade de receitas, manuais e regulamentos. Inferir características de personagens em fábulas. Interpretar linguagem verbal e não verbal em tirinhas.
Nível 3 Desempenho maior ou igual a 175 e menor que 200	Além das habilidades anteriormente citadas, os estudantes provavelmente são capazes de: localizar informação explícita em contos e reportagens e em propagandas com ou sem apoio de recursos gráficos. Reconhecer relação de causa e consequência em poemas, contos e tirinhas. Inferir o sentido de palavra, o sentido de expressão ou o assunto em cartas, contos, tirinhas e histórias em quadrinhos, com o apoio de linguagem verbal e não verbal.

Fonte: Relatório SAEB 2017.

A descrição dos demais níveis da escala segue o mesmo padrão da descrição dos três primeiros, trazidos, aqui, a título de exemplificação. Com base nesse exemplo, observa-se que a descrição do desempenho em leitura dos estudantes se faz numa articulação entre a habilidade de leitura avaliada (por exemplo, localizar informação explícita, inferir, reconhecer a finalidade do texto) e o gênero textual utilizado como suporte aos itens que avaliam a habilidade (reportagem, poemas, conto, fábula). A progressão entre os diferentes níveis é dada pelo acréscimo de habilidades àquelas indicadas no nível anterior, entretanto, pela pesquisa que vimos desenvolvendo, constatamos que dessa forma não é possível compreender como se altera a complexidade da tarefa de leitura ou o que torna um gênero mais complexo que outro, do ponto de vista da interação do leitor com o texto. A título de exemplificação tem-se, no nível, o mais elementar, a descrição de que os estudantes, provavelmente, “são capazes de: localizar informações explícitas em textos narrativos curtos, informativos e anúncios.” Já no nível 2, no qual os estudantes teriam uma proficiência maior do que aqueles de nível 1, lê-se que os estudantes provavelmente são capazes de, além de realizar as tarefas de leitura do nível anterior “localizar informações explícitas em contos”. Os contos podem ser textos narrativos curtos. Além disso, não existem informações sobre a posição da informação a ser localizada e nem como ela deve ser recuperada pelo leitor, com base no tipo de situação problema construída no item. Desse modo, quanto à habilidade de localizar informações, não é possível perceber o que distingue um estudante com proficiência compatível com o nível 1 daqueles com proficiência compatível com o nível 2. Esse mesmo problema se repete com relação a outros descritores e na passagem entre outros níveis.

Conclui-se, a partir do exposto, que narrativas, como um conto, por exemplo, podem se materializar em textos com diferentes níveis de complexidade, consequência do tipo de sintaxe utilizada, do léxico empregado, do tema abordado, dentre outras possibilidades. Além disso, no caso da habilidade de localizar informações explícitas, a posição da informação solicitada e, ainda, a existência de maior ou menor quantidade de informações no texto também são fatores a considerar, além do modo como a informação é recuperada na situação problema proposta pelo item: se por meio de uma repetição, de uma quase repetição, ou por uma paráfrase simples, por exemplo. Desse modo, a descrição das habilidades de leitura tal como se apresenta no relatório do SAEB é pouco elucidativa da natureza das habilidades de leitura desenvolvidas por estudantes que se encontram nos diferentes níveis da escala de proficiência.

No relatório do PISA, o desempenho dos estudantes em leitura também é apresentado por meio de 6 (seis) níveis. No quadro 2, são apresentados os três primeiros níveis da escala de proficiência do PISA (OECD, 2019).

Quadro 2 – Níveis de desempenho em *reading literacy proficiency* do PISA (fragmento).

Level 1 B	Readers at level 1b can locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text with a familiar context and text type, such as a narrative or a simple list. Texts in level 1b tasks typically provide support to the reader, such as repetition of information, pictures or familiar symbols. There is minimal competing information. Level 1b readers can interpret texts by making simple connections between adjacent pieces of information.
Level 1B	Readers at Level 1a can locate one or more independent pieces of explicitly stated information; they can recognise the main theme or author's purpose in a text about a familiar topic, or make a simple connection between information in the text and common, everyday knowledge. Typically, the required information in the text is prominent and there is little, if any, competing information. The student is explicitly directed to consider relevant factors in the task and in the text.
Level 2	Readers at Level 2 can locate one or more pieces of information, which may need to be inferred and may need to meet several conditions. They can recognize the main idea in a text, understand relationships, or construe meaning within a limited part of the text when the information is not prominent and the reader must make low level inferences. Tasks at this level may involve comparisons or contrasts based on a single feature in the text. Typical reflecting tasks at this level require readers to make a comparison or several connections between the text and outside knowledge, by drawing on personal experience and attitudes.

Fonte: Assessment and Analytical Framework, PISA (OECD, 2019, s/p).

No relatório do PISA, é possível encontrar um maior detalhamento das tarefas de leitura que os estudantes que se encontram nos diferentes níveis de proficiência provavelmente podem realizar. Por exemplo, a habilidade de localizar informações é descrita como uma habilidade que pode ser mais ou menos complexa se o texto apresenta muitas informações que concorrem com aque-

la solicitada pelo item, se a informação solicitada é mais ou menos proeminente no texto. Não há referências ao gênero textual, mas a características gerais do texto, como ao fato de abordar um tema mais ou menos familiar. Neste caso, há uma análise mais detalhada da tarefa de leitura, mas tal detalhamento se faz sem uma definição mais precisa da natureza do texto, suas características e microestrutura, que podem tornar essas tarefas de leitura mais ou menos complexas. Entretanto, há um avanço em relação ao relatório do SAEB, no sentido de tornar mais evidentes as tarefas de leitura que os estudantes dos diferentes níveis têm probabilidade de realizar com êxito.

Os estudos desenvolvidos no âmbito da pesquisa de que trata o presente artigo têm mostrado que as habilidades de leitura desenvolvidas pelos estudantes estão fortemente relacionadas à estrutura do texto e ao tipo de sinalização que ele oferece para que os estudantes possam realizar a tarefa de leitura requerida, evidenciando o desenvolvimento de determinadas habilidades. Um poema, por exemplo, pode apresentar uma linguagem próxima àquela utilizada na vida cotidiana, o que exigirá uma proficiência em leitura relativamente menor, para sua compreensão pelo leitor; ou pode utilizar uma linguagem metafórica, o que demandará uma elevada proficiência em leitura para que seja compreendido pelo leitor.

A DEFINIÇÃO DOS NÍVEIS DE COMPLEXIDADE DO TEXTO NA PESQUISA EM AVALIAÇÃO CAED, ÁREA DE LÍNGUA PORTUGUESA

No âmbito da pesquisa em avaliação, área de Língua Portuguesa, foram analisados, ao longo dos anos de 2017 e 2018, quatro descritores que integram a Matriz de Referência para Avaliação do CAEd, quais sejam: D06 - Localizar informação explícita; D07 - Reconhecer o assunto de um texto; D08 - Inferir o sentido de uma palavra ou expressão; D09 - Inferir informação em um texto verbal. Em todos esses descritores, são utilizados como suporte para a construção da situação problema dos itens textos de diferentes gêneros. Constituíram o corpus da pesquisa 763 textos, que deram suporte a itens que avaliaram esses descritores.

As análises empreendidas no âmbito da pesquisa tinham como objetivo a identificação da progressão da dificuldade que cada uma das habilidades apresenta ao longo da escala de proficiência do SAEB. Para compreender tal progressão, a complexidade dos textos utilizados como suporte para a elaboração dos itens que avaliam cada uma das habilidades se mostrou um ele-

mento decisivo. Isso ocorreu porque a leitura pressupõe uma interação entre leitor e texto que são fatores relevantes à natureza do texto e à experiência do leitor. O texto, objeto central da investigação, configura-se como um conjunto de sinalizações que “guiam” o leitor para que ele produza sentidos para o que lê, e o leitor como alguém que atua nesse processo a partir de uma gama de conhecimentos – sobre o mundo, sobre temas específicos do mundo, sobre os textos, sobre a língua.

Para a análise dos descritores da matriz de Referência para Avaliação em Língua Portuguesa, foi realizada, inicialmente, uma explicitação da natureza linguística do descritor com base nas teorias do texto, sobretudo as teorias de base cognitivista (SMITH 1991; KLEIMAN 1989; KOCH 1998) e na linguística do texto (KOCH, 1989; KOCH, TRAVAGLIA 1989).

A TRI – Teoria da Resposta ao Item – e o modelo de três parâmetros foram aqueles que subsidiaram a análise dos resultados das avaliações. Para identificar os níveis de complexidade dos textos que deram suporte aos itens, foram eliminados: itens com parâmetro A menor que 0,01, uma vez que esses itens apresentam fraco poder de discriminação entre estudantes que desenvolveram as habilidades avaliadas e aqueles que não as desenvolveram e, ainda, itens com parâmetro C superior a 0,025, uma vez que esses itens apresentam forte probabilidade de acerto ao acaso. Foram considerados, na análise, o ponto de ancoragem dos itens na escala de proficiência em leitura do SAEB e a distribuição desses itens em intervalos de 25 pontos.

O segundo passo consistiu numa incursão aos itens que avaliam cada um dos descritores, buscando identificar a gradação da complexidade com que tais habilidades se manifestam ao longo da escala, o que ofereceu indícios da proficiência em leitura dos estudantes nas diferentes etapas de escolarização. Tal identificação se deu pela comparação entre textos que davam suporte a itens que ancoravam em pontos próximos da escala de proficiência, buscando identificar aspectos em comum entre esses textos. As definições das habilidades avaliadas, elaboradas na primeira etapa da pesquisa foram, também, fundamentais para a análise dos textos que davam suporte aos itens.

A análise das tarefas propostas pelos itens considerou tanto a complexidade do texto quanto a natureza das sinalizações que ele oferecia para a realização da tarefa. São exemplos de sinalizações de apoio à leitura: a presença, no texto, de sinônimos que auxiliam a tarefa de inferir o sentido de uma palavra; a presença de marcação tópica que facilita ao leitor a tarefa de reconhecer o assunto de um texto; um enunciado parafrástico que permite acessar uma

informação, dentre outras possibilidades. Por outro lado, o nível de complexidade de um texto – definido a partir de critérios como sintaxe, léxico, tema, extensão – tem, também, impacto sobre a complexidade da tarefa.

A classificação decorrente dos critérios definidos no âmbito da pesquisa considera quatro níveis de complexidade dos textos, em cuja definição a sintaxe se apresenta como elemento bastante definidor da complexidade, quando se consideram os dois primeiros níveis e o léxico e o tema se apresentam como bastante determinantes nos níveis três e quatro. Com base nos critérios tamanho, sintaxe, léxico e tema, definiram-se os seguintes níveis de complexidade dos textos a seguir.

Textos de Nível 1: os textos são pouco extensos (até 15 linhas). A **sintaxe** é simples, em que predominam os períodos curtos e estruturas coordenadas, ou ainda em tópicos; os processos coesivos são simples, como aqueles envolvendo retomadas com referente próximo – por referência através de pronomes pessoais do caso reto ou reiteração, por repetição ou sinonímia; a estrutura dos textos é convencional e as frases apresentadas na ordem canônica. O **léxico** é mais próximo do coloquial e adequado ao interlocutor preferencial do texto, normalmente criança. Os **temas** são mais cotidianos, uma vez que temas menos familiares requerem um tratamento que “facilite” a leitura do texto. Os textos mais comuns neste nível são de estrutura canônica, estando presentes as narrativas e os textos expositivos: sobre os primeiros, podemos destacar parlendas, lendas, narrativas da literatura infantil e, ainda, os poemas de estrutura narrativa; entre os segundos, se encontram as “curiosidades científicas” ou “notas científicas”, como definidos pela BNCC.

Textos de Nível 2: os textos são pouco extensos (até 15 linhas). A **sintaxe** é um pouco mais complexa do que aquela que se observa em textos de nível 1, com pequenas alterações na estrutura sintática canônica – orações reduzidas no início dos períodos e intercalações breves, por exemplo; há o predomínio de períodos um pouco mais longos, com duas linhas em média; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos mais usuais. Há textos que apresentam **léxico** mais usual ou aqueles cujo léxico é próprio do discurso formal, mas com estratégias de facilitação do discurso, como aposto depois de um vocábulo pouco usual ou de um termo científico ou técnico. Os **temas** são relacionados ao cotidiano e, quando menos familiares, o texto oferece pistas para a identificação de seu conteúdo. Os textos mais comuns são as narrativas de estrutura canônica e as argumentações breves e de estrutura mais simples.

Textos de Nível 3: os textos apresentam-se em qualquer **tamanho**, ou seja, tanto há textos curtos quanto muito longos entre os classificados nesse nível. A **sintaxe** é complexa e própria do uso formal da língua, com estruturas subordinadas, intercalações mais longas ou impessoalizações, por exemplo. Os períodos são mais longos, com cerca de três linhas; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos menos usuais, assim como retomadas de referentes mais distantes. O **léxico** é próprio do discurso formal, apresentado em textos diversos, de estrutura menos canônica. Os **temas** são pouco familiares e os suportes mais variados, voltados a interlocutores diversos. As narrativas que aparecem nesse nível são os contos, as crônicas, os fragmentos de romance – textos da literatura brasileira. Encontram-se nesse nível textos expositivos ou argumentativos mais longos, como notícias, reportagens e poemas de elaboração metafórica.

Textos de Nível 4: assim como no nível anterior, os textos de nível 4 apresentam-se em qualquer **tamanho**. A **sintaxe** é complexa e própria do uso formal da língua, com estruturas subordinadas, intercalações mais longas e impessoalizações, por exemplo. Os períodos são mais longos, com três linhas em média; os processos coesivos são mais complexos, como aqueles envolvendo retomadas por elipse, pronomes oblíquos e pronomes relativos menos usuais, assim como retomadas de referentes mais distantes. O **léxico** é próprio do discurso formal, especializado e erudito, apresentado em textos diversos, de estrutura menos canônica. O predomínio é de textos com tratamento literário mais elaborado, afetando sua estrutura e seus recursos, e aqueles que tratam de **temas** especializados, sobretudo do campo da ciência e da vida em sociedade. Como exemplo de textos narrativos, destacam-se textos clássicos: os contos, crônicas e fragmentos de romances, narrativas menos focadas na ação e com estratégias mais complexas, como o discurso indireto livre, descrições que capturam o fluxo de consciência da personagem. Nesse nível encontram-se, ainda, resenhas, notícias e reportagens, textos de opinião com vários argumentos e poemas com tratamento literário mais elaborado, com a presença de metáforas, alegorias e imagens poéticas.

Com base nos quatro níveis de complexidade dos textos aqui apresentados, foram analisados os itens referentes aos descritores referidos anteriormente e aplicados em avaliações realizadas pelo CAEd a alunos 5º ao 12º ano de escolarização. Nos quadros 3, 4, 5 e 6, a seguir, é apresentada a distribuição dos itens analisados, relacionados a cada descritor, por intervalos de 25 pontos da escala de proficiência.

Quadro 3 – Distribuição dos itens que avaliam a habilidade de localizar informações explícitas em textos por níveis de complexidade.

D06- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
100-125	1	0,31						
126-150	3	0,95	5	1,58				
151-175	16	5,07	29	9,20			1	0,31
176-200	6	1,90	38	12,06	5	1,58		
201-225	3	0,95	51	16,19	17	5,39	2	0,63
226-250			33	10,47	18	5,71	5	1,58
251-275			22	6,98	13	4,12	4	1,26
276-300			9	2,85	3	0,95	7	2,22
301-325			3	0,95	2	0,63	7	2,22
326-350			1	0,31	2	0,63	4	1,26
351-375							3	0,95
376-400								
Acima 401					1	0,31	1	0,31
Totais	29	9,18	191	60,59	61	19,32	34	10,74
Total Geral	315 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 4 – Distribuição dos itens que avaliam a habilidade de reconhecer o assunto de um texto por níveis de complexidade.

D07- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	3	1,33	1	0,54				
151-175	4	1,78	16	7,10	1	0,54		
176-200	3	1,33	17	7,58	11	4,91		
201-225			20	8,92	21	9,37		
226-250			14	6,25	35	15,62	3	1,33
251-275			5	2,23	14	6,25	16	7,10
276-300			6	2,67	9	4,01	2	0,89
301-325			1	0,54	2	0,89	3	1,33
326-350					7	3,12	6	2,67
351-375					1	0,54	2	0,89
376-400								
Acima 401					1	0,44		
Totais	10	4,44	80	35,83	102	45,25	32	14,21
Total Geral	224 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 5 – Distribuição dos itens que avaliam a habilidade de inferir o sentido de uma palavra ou expressão por níveis de complexidade

D08- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	1	0,65						
151-175	4	2,80	2	1,40				
176-200	7	4,89	12	8,39				
201-225	3	2,09	20	13,98	3	2,09		
226-250	8	5,59	19	13,29	1	0,65		
251-275	1	0,65	19	13,29	8	5,59		
276-300			3	2,09	13	9,09		
301-325			1	0,65	11	7,69		
326-350			2	1,40	2	1,40		
351-375					1	0,65		
376-400			1	0,65	1	0,65		
Acima 401								
Totais	24	16,67	79	55,14	40	27,81		
Total Geral	143 textos							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Quadro 6 – Distribuição dos itens que avaliam a habilidade de inferir informação em um texto verbal por níveis de complexidade.

D09- QUANTITATIVO DE COMPLEXIDADE DO TEXTO POR INTERVALO								
Intervalos	NÍVEL 1		NÍVEL 2		NÍVEL 3		NÍVEL 4	
	Qtidade	%	Qtidade	%	Qtidade	%	Qtidade	%
125-150	1	1,23	1	1,23				
151-175	2	2,46	1	1,23				
176-200	3	3,70	11	13,58				
201-225			5	6,17	2	2,46	1	1,23
226-250	1	1,23	7	8,64	7	8,64	2	2,46
251-275			2	2,46	6	7,40	3	3,70
276-300			1	1,23	12	14,81	3	3,70
301-325					1	1,23	2	2,46
326-350					3	3,70	2	2,46
351-375					1	1,23	1	1,23
376-400								
Acima 401								
Totais	7	8,62	28	34,54	32	39,47	14	17,24
Total Geral	81							

Fonte: Pesquisa em Avaliação CAEd, 2016-2019.

Como é possível observar pela análise dos quadros 3, 4, 5 e 6, os textos de nível 1 são aqueles predominantes em itens que ancoram entre 125 e 200 pontos da escala de proficiência em leitura do SAEB. Os textos de nível 2 predominam como suporte de itens que ancoram entre 151 e 275 pontos da escala de proficiência. Textos de nível 3 aparecem predominantemente como

suporte de itens que ancoram entre 176 e 300 pontos da escala. Finalmente, os textos de nível 4 aparecem como suporte de itens que ancoram acima de 226 pontos da escala, com maior concentração entre itens que ancoram acima de 250 pontos da escala. A comparação entre a distribuição dos textos, de acordo com seu nível de complexidade, entre os diferentes descritores da matriz de referência, indica uma certa regularidade dessa distribuição. Tal regularidade é indicativa de padrões de desempenho em leitura apresentados pelos estudantes que realizaram os testes e que se encontram em diferentes etapas da educação básica. É possível afirmar, por exemplo, com base na descrição dos níveis de complexidade do texto, que estudantes com proficiência entre 125 e 200 pontos leem textos pouco extensos (até 15 linhas), que apresentam sintaxe simples, em que predominam os períodos curtos e estruturas coordenadas, ou em tópicos, com processos coesivos simples, como aqueles envolvendo retomadas com referente próximo, por referência através de pronomes pessoais do caso reto ou reiteração, por repetição ou sinonímia. A estrutura dos textos que esses estudantes leem é convencional, com frases apresentadas na ordem canônica (sujeito, verbo, objeto), o léxico é mais próximo do coloquial e os temas mais cotidianos. Parlendas, lendas, narrativas da literatura infantil, os poemas de estrutura narrativa, as “curiosidades científicas” ou “notas científicas”, como definidos pela BNCC, são os gêneros com os quais os estudantes com essa proficiência leem.

De acordo com a análise realizada, o que distingue os textos de nível 1 dos textos de nível 2 é fundamentalmente a sintaxe. Pequenas alterações da estrutura sintática canônica (como intercalações e antecipações), bem como a extensão dos períodos, podem tornar um texto mais complexo para leitores iniciantes que não possuem fôlego de leitura. O que se observa é que a extensão dos períodos é um fator mais determinante para a complexidade do texto do que, propriamente, a extensão do texto. Por outro lado, textos de nível 1 e 2 caracterizam-se por exibir léxico mais familiar, próximo do uso coloquial da linguagem.

O léxico se apresentou, nesta investigação, como elemento definidor da complexidade de textos de nível 3 e 4. Textos de nível 3 apresentam, normalmente, léxico próprio do discurso formal, implicando em atividade inferencial mais constante. Textos de nível 4 apresentam léxico especializado, presente em textos cuja temática também se especializa, ou seja, textos que abordam questões de áreas de conhecimento específicas: ciência e tecnologia, vida em sociedade etc.

No entanto, a consideração desses traços do léxico (se mais, ou menos, próximo do discurso formal) não deve ser rígida. Em alguns textos, por exemplo,

usos formais da língua são seguidos de recursos de “facilitação do discurso” que permitem lidar com um léxico desconhecido. Textos que tratam de temas especializados e que são dirigidos ao leitor criança ou adolescente costumam exibir várias estratégias para facilitar sua recepção (como expressões explicativas, sinonímia, exemplos etc.). Léxico e tema são, portanto, critérios importantes para a classificação de textos de nível 3 e 4.

A investigação feita até aqui permite propor que a competência de leitura é fortemente determinada, inicialmente, pelo domínio da sintaxe da língua. Os leitores iniciantes vão, aos poucos, apropriando-se da sintaxe da língua, de seu uso formal, a partir do contato com textos cada vez mais complexos do ponto de vista sintático. O amadurecimento desse leitor se consolida à medida em que ele vai ampliando seu repertório de leitura e, conseqüentemente, seu repertório lexical, familiarizando-se com textos de vários campos do conhecimento, dominando sua estrutura, abordando temas diversos cada vez mais especializados. Essa competência se amplia ainda mais quando se compreende o potencial expressivo da linguagem recriada literariamente.

CONSIDERAÇÕES FINAIS

Os resultados da pesquisa aqui apresentada indicam que a complexidades dos textos é um critério importante a ser considerado na descrição de padrões de desempenho em leitura, definidos com base em avaliações de larga escala. Considerar os níveis de complexidade dos textos que dão suporte aos utilizados nessas avaliações permite avançar para além da descrição das tarefas de leitura propostas por esses itens e vislumbrar competências de leitura mais amplas, possivelmente desenvolvidas pelos estudantes que se encontram em determinados intervalos de proficiência.

Os resultados da pesquisa apontam, ainda, que, na definição dos níveis de complexidade que um texto pode apresentar, é necessário considerar um conjunto de características, que se combinam para que um texto seja mais ou menos desafiador ao leitor e que, ao longo do processo de formação do leitor, esses fatores desempenham um papel diverso. Para leitores com pouca proficiência, a sintaxe do texto é um fator bastante definidor da complexidade do texto. Para leitores mais experientes, o tema desempenha um papel mais determinante para que um texto possa ser considerado mais ou menos complexo. Deduz-se, do exposto, que a categoria gênero e a extensão do texto não são suficientes para uma descrição mais consistente das habilidades de leitura desenvolvidas pelos estudantes. Por exemplo, um poema, a despeito de sua extensão, pode se apresentar como de leitura mais ou menos comple-

xa, a depender de sua estrutura: se for um poema narrativo que faça uso de uma linguagem coloquial, com versos de estrutura simples, ele será de leitura mais fácil do que se apresentar uma linguagem metafórica e versos com uma estrutura menos canônica.

Dado o exposto, conclui-se que a descrição de padrões de desempenho em leitura deva considerar uma análise mais criteriosa dos suportes utilizados na construção das situações problema propostas pelos itens, para além da consideração do tipo de habilidade avaliada. Nesse sentido, é importante que, ao se descrever a habilidade, por exemplo, de localizar uma informação explícita em um texto, se considere de que tipo de texto se está falando e do tipo de apoio oferecido por ele para que o leitor realize a tarefa proposta. Se considerarmos que a leitura é um processo de interação entre leitor e texto, é preciso considerar como o texto conduz o leitor nessa interação, além de considerar o que o leitor traz para ela.

Outra consideração importante com base nos dados da pesquisa é a de que uma maior clareza quanto às características dos textos, obtida por meio da definição de critérios claros para a classificação de seus níveis de dificuldade, favorece a produção de itens que contemplem, de forma mais satisfatória, a avaliação de leitura de estudantes que se encontram em etapas específicas de escolarização, considerando o que está previsto na norma curricular. Assim, uma próxima etapa da pesquisa consistirá em cotejar a classificação dos níveis de dificuldade dos textos, produzida com base na incursão aos itens já utilizados em avaliações de larga escala, para estabelecer uma relação entre esses níveis e as diferentes etapas da educação básica.

Destaca-se, ainda, o fato de que a maior clareza quanto ao nível de dificuldade dos textos é um conhecimento relevante para subsidiar a produção de itens que avaliam outras áreas de conhecimento, como as ciências da natureza e as ciências humanas, vez que nessas áreas também são utilizados textos verbais que dão suporte à elaboração dos itens que as avaliam.

REFERÊNCIAS

BAKHTIN, M. **Estética da criação verbal**. São Paulo: Martins Fontes, 2003.

BRASIL. **Base Nacional Comum Curricular (BNCC)**. Educação é a Base. Brasília, MEC/CONSED/UNDIME, 2017. Disponível em: [568http://basenacionalcomum.mec.gov.br/images/BNCC_publicacao.pdf](http://basenacionalcomum.mec.gov.br/images/BNCC_publicacao.pdf). Acesso em: 10 dez. 2019.

BRASIL. Secretaria de Educação Fundamental. **Parâmetros Curriculares Nacionais (PCNs):** terceiro e quarto ciclos do ensino fundamental: Língua Portuguesa. Brasília: MEC/SEF, 1998.

BRASIL. **Relatório SAEB 2017.** Brasília, 2019. Disponível em: http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/6730262. Acesso em: 10 dez. 2019.

KLEIMAN, A. Compreensão Leitora. In: **Glossário CEALE.** Disponível em: <http://ceale.fae.ufmg.br/app/webroot/glossarioceale/verbetes/compreensao-leitora>. Acesso em: 10 dez. 2019.

_____. **Texto e leitor:** aspectos cognitivos da leitura. Campinas, SP: Pontes, 1989.

KOCH, I. G. V; TRAVAGLIA, C. **Texto e Coerência.** São Paulo: Cortez, 1989.

KOCH, I. G. V. **O texto e a construção dos sentidos.** São Paulo: Contexto, 1998.

_____. **A coesão textual.** São Paulo: Contexto, 1989.

OECD. **PISA 2018.** Assessment and Analytical Framework, PISA, OECD Publishing, Paris, 2019. Disponível em: <https://doi.org/10.1787/b25efab8-em>. Acesso em: 10 dez. 2019.

PRIMI, R. **Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes.** Avaliação psicológica, 2012. Disponível em: <http://pepsic.bvsalud.org/pdf/avp/M1n2/M1n2a15.pdf>. Acesso em: 10 dez. 2019.

SMITH, F. **Compreendendo a leitura:** uma análise psicolinguística da leitura e do aprender a ler. Porto Alegre: Artes Médicas, 1991.

SOUSA, L. B. de; HÜBNER, L.C. Desafios na avaliação da compreensão leitora: demanda cognitiva e leiturabilidade textual. **Revista neuropsicologia latinoamericana**, 2014. Disponível em :https://www.neuropsicolatina.org/index.php/Neuropsicologia_Latinoamericana/article/view/237. Acesso em: 10 dez. 2019.

Recebido em: 10 de dezembro de 2019

Aceito em: 15 de junho de 2020

Publicado em: 30 de junho de 2020