

ARTIGO ORIGINAL

Do grau de dificuldade para um contínuo de complexidade: análise textual de itens do Enem

From the degree of difficulty to a continuum of complexity: textual analysis of items of the Enem

Clemilton Lopes Pinheiro¹ , Jailson Almeida Conceição² 

1Universidade Federal do Rio Grande do Norte, clemiltonpinheiro@hotmail.com

2Universidade Estadual do Piauí, almeidajai@hotmail.com

Como citar o artigo.

Pinheiro, C. L.; Conceição, J. A. Do grau de dificuldade para um contínuo de complexidade: análise textual de itens do Enem. *Revista Horizontes de Linguística Aplicada*, v. 20, n. 1, p. AG4, 2021. Disponível em: <https://periodicos.unb.br/index.php/horizontesla/article/view/36140>.

Resumo

As pesquisas no campo dos estudos da linguagem sobre as provas do Enem compõem um cenário vasto de temas e abordagens. Neste trabalho, particularizamos a dimensão textual dos itens da prova. Nosso objetivo é, portanto, o de propor a análise do item pela perspectiva da complexidade textual, e, dessa forma, contribuir para o debate acerca da sua qualidade e do seu comportamento empírico. Apresentamos indicadores de complexidade de natureza textual, relacionados à dimensão composicional do texto, que permitem observar, nos itens, diferentes níveis dentro de um contínuo de complexidade. Consideramos duas categorias textuais: os tipos de discurso e o tópico discursivo. Para demonstrar o raciocínio de que os itens da prova podem ser analisados segundo o critério da complexidade textual, apoiado nessas duas categorias, analisamos 24 itens da prova do Enem/2016.

Palavras-chave: Complexidade textual. Enem. Tipo de discurso. Tópico discursivo.

Abstract

Research in the field of language studies about Enem compose a vast scenario of themes and approaches. In this paper, we focus on the textual dimension of the items in Enem. Therefore, our goal is to propose the analysis of the item through the perspective of textual complexity, and thus to contribute to the debate about its quality and its empirical behavior. We present indicators of textual-nature complexity, which are related to the compositional dimension of the text, and which allow us to observe different levels within a continuous of complexity in the items. We consider two textual categories: types of discourse and the discursive topic. To demonstrate the reasoning that items can be analyzed according to the textual complexity criterion, supported by these categories, we analyze 24 items of Enem/2016.

Keywords: Discursive topic. Enem. Textual complexity. Types of discourse.

Apoio financeiro: Nenhum.

Recebido em 20 Janeiro, 2021. Aceito em 10 Junho, 2021.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution Non-Commercial No Derivative, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais, sem alterações e que o trabalho original seja corretamente citado.

1 INTRODUÇÃO

O Exame Nacional do Ensino Médio (Enem) foi instituído no Brasil, em 1998, com o objetivo de avaliar o desempenho dos estudantes concluintes do ensino médio, e servir de parâmetro para a melhoria da qualidade do ensino no país. A partir de 2009, o exame passou também a ser usado como instrumento de seleção para o ingresso no ensino superior. A prova é composta de 180 questões (itens) de múltipla escolha, que avaliam o conhecimento em quatro áreas (linguagens, códigos e suas tecnologias, ciências humanas e suas tecnologias, ciências da natureza e suas tecnologias, matemática e suas tecnologias) e uma redação.

Considerando a sua importância para o contexto educacional brasileiro, o rápido e grande interesse acadêmico, em diferentes campos do saber, pelo Enem, não foi surpresa. Há algum tempo, muitos trabalhos têm apresentado discussões e reflexões sobre aspectos variados envolvidos no exame. Se ficarmos no campo apenas dos estudos linguísticos, alguns temas podem ser destacados: gêneros e tipologias textuais, concepções de língua e linguagem, relação de temas linguísticos e literários com os Parâmetros Curriculares Nacionais (PCN), metodologias e práticas de ensino de língua e literatura, matriz de referência para elaboração e correção de itens e da redação. Todos os trabalhos buscam entender como a linguagem é concebida e explorada nesse processo de modo a discutir a sua eficiência em consonância com o conhecimento produzido na área dos estudos da linguagem.

Nesse contexto vasto de temas e abordagens, particularizamos, neste trabalho, a dimensão textual dos itens da prova. Do ponto de vista dos mecanismos/processos de textualidade, um item é um texto constituído por um entrelaçamento de outros textos, verbais e não verbais, correlacionados formal e funcionalmente. Isso pressupõe que as reflexões que orientam a elaboração de um item precisam passar pela reflexão da sua dimensão textual. As regras que orientam a elaboração dos itens do Enem são estabelecidas em diretrizes próprias pelo INEP (BRASIL, 2010, p. 12). No que diz respeito a aspectos da organização textual, essas regras são genéricas e vagas: “utilize termos impessoais como: ‘considere-se’, ‘calcula-se’, ‘argumenta-se’ etc.; não utilize sentenças como: ‘pode-se afirmar que’; paralelismo sintático e semântico, extensão equivalente e coerência com o enunciado”. No âmbito dos estudos acadêmicos, parece-nos que esses aspectos do item ainda têm sido pouco explorados. Além disso, o texto, como suporte linguístico da prova, é negligenciado tanto no processo em si de avaliação do item (na meta-avaliação) como nos estudos sobre esse processo.

Nesse sentido, ao particularizarmos a dimensão textual, buscamos refletir sobre a contribuição dessa dimensão para o debate acerca da qualidade e do comportamento empírico dos itens. Tomamos como ponto de partida a seguinte questão: o que torna um texto, e, por extensão, um item do Enem, mais ou menos difícil? Ao revisitarmos algumas possibilidades de respostas a essa questão, propomos a análise do item pela perspectiva da complexidade textual. A proposta aponta indicadores de complexidade de natureza textual, relacionados à dimensão composicional do texto, que permitem observar, em um item, não um grau definido de dificuldade (fácil, médio, difícil), mas um nível dentro de um contínuo de complexidade (do mais para o menos complexo).

Concebemos a complexidade textual como um fenômeno objetivo e estruturado, ligado ao conjunto de propriedades composicionais e estruturais do texto, que o tornam, como base nesses atributos, mais ou menos complexos. Essa concepção foi construída a partir de trabalhos anteriores sobre o tema e baseada em duas abordagens dos estudos do texto das quais extraímos duas categorias analíticas: tipos de discurso (Interacionismo Sociodiscursivo, ISD), e tópico discursivo (Perspectiva Textual-interativa, PTI). Para demonstrar o raciocínio de que os itens da prova do ENEM podem ser analisados segundo o critério da complexidade textual, apoiado especificamente nessas categorias, realizamos a análise de um pequeno *corpus*: 24 itens da prova do ENEM/2016, divididos em fáceis e difíceis, conforme um critério estatístico.

2. PERSPECTIVAS SOBRE DIFICULDADE/COMPLEXIDADE

A discussão acerca do grau de dificuldade dos itens do Enem é comumente posta em foco, sobretudo, porque está relacionada à atribuição da nota da prova. “Do ponto de vista empírico, a prova é organizada com questões de diferentes níveis de dificuldade para melhor medir e situar os resultados individuais” (BRASIL, 2002, p. 17). O grau de dificuldade é medido estatisticamente através de testes psicométricos. Segundo Trivitz (2017, p. 264), há dois grupos de testes:

Aqueles reunidos sob a alcunha da Teoria Clássica dos Testes – que se baseia em conceitos estatísticos mais genéricos, como correlação e porcentagem –, e os da Teoria da Resposta ao Item – com modelos próprios, desenhados especificamente para analisar testes de inteligência, conhecimento ou outras propriedades que não podem ser diretamente observadas.

A Teoria Clássica dos Testes (TCT), que foi usada na primeira década de aplicação do Enem, aponta, com base na porcentagem de resposta para cada uma das alternativas do item de múltipla escolha, o grau de dificuldade e o grau de discriminação do item. “O ‘grau de dificuldade’ é definido como sendo a porcentagem de acerto de todo o grupo de examinados” (ANDRADE; KLEIN, 2005, p. 108). A Teoria de Resposta ao Item (TRI) é o modelo atualmente utilizado. Segundo Pasquali e Primi (2003, p. 102) a TRI apresenta dois postulados básicos:

1) o desempenho do sujeito numa tarefa (item de um teste) pode ser predito a partir de um conjunto de fatores ou variáveis hipotéticas, ditos aptidões ou traços latentes (identificados na TRI com a letra grega teta: θ); o teta sendo a causa e o desempenho o efeito. Trata-se de modelagem latente (latent trait modeling). Ou seja, comportamento = função (traço latente);

2) a relação entre o desempenho e os traços latentes pode ser descrita por uma equação matemática monotônica crescente, chamada de Curva Característica do Item – CCI

Esse modelo oferece uma série de características de um item entre as quais o parâmetro de dificuldade, que “é o ponto na escala de aptidão no qual a probabilidade de uma resposta correta é de 50% (ou seja, 0,5) (PASQUALI; PRIMI, 2003, p. 107).

Andrade e Klein (2005, p. 108), por exemplo, fazem a análise estatística de um item de uma prova do Enem (Tabela 01), com base na TCT, e constatarem que se trata de um item fácil:

Pode-se concluir que o item foi fácil para o grupo de candidatos e que teve bom desempenho, tendo em vista que quase a totalidade dos candidatos do grupo superior (0,91 ou 91%) e um número bem menor dos candidatos do grupo inferior (0,45 ou 45%) acertaram o item (grau de discriminação igual a 0,46). Dois outros pontos importantes são que os distratores foram igualmente preferidos e que todos os valores do coeficiente bisserial foram negativos para os distratores e positivo e “alto” para a alternativa correta (0,51).

Tabela 01: Estatísticas descritivas de um item

Estatísticas	Alternativas					
	A	B	C	D	E	Outra
Porcentagem de resposta	7	7	8	69	9	0
Coefficiente bisserial	-0,28	-0,17	-0,30	0,51	-0,47	-0,32
Grau de dificuldade = 0,69 ou 69%						
Proporção de acerto do grupo superior = 0,91						
Proporção de acerto do grupo inferior = 0,45						
Grau de discriminação = 0,46						
Coefficiente bisserial = 0,51						

Fonte: Andrade e Klein (2005, p. 108).

Com base em análises estatísticas, um item pode ser, portanto, classificado como fácil, médio ou difícil. Convém, no entanto, destacar a limitação que os testes psicométricos podem apresentar, conforme destaca a conclusão de Travitzi (2017, p. 277) sobre a aplicação da TRI às provas do Enem:

Para uma verificação mais precisa da distribuição do grau de dificuldade dos itens, seria necessária uma definição dos pontos de corte que distinguem os itens fáceis, médios e difíceis no Enem. Tal informação, contudo, não foi encontrada na documentação disponível.

A questão do grau de dificuldade de um item pode também ser abordada com base nas discussões sobre os aspectos que interferem na leitura do texto, embora não existam muitos trabalhos, sobretudo acerca de itens de prova, que a explorem. Coscarelli (2012), por exemplo, aponta alguns dos aspectos do processo da leitura, que facilitam ou dificultam o trabalho do leitor: o processamento lexical, o processamento sintático, a construção da coerência. Especificamente ao processamento lexical, a complexidade da sílaba, o tamanho da palavra, a frequência de uma palavra, a familiaridade do leitor com uma palavra, a ambiguidade lexical, por exemplo, são fatores que interferem no grau de dificuldade da leitura. Dessa forma, a partir da consideração desses fatores, o texto será mais ou menos difícil para um determinado leitor.

A discussão sobre o grau de dificuldade de um texto pelo viés da dificuldade de compreensão de leitura, relacionando-a, portanto, à noção de legibilidade, está relacionada a traços textuais mensuráveis (inerentes ao texto), que, após identificados, compõem os graus de legibilidade de um texto. Silva (2009) desenvolve um trabalho nessa direção com enunciados de exercícios de língua portuguesa, gênero que se aproxima muito dos itens de prova. O estudo investigou a compreensão de enunciados por um grupo de estudantes, e chegou-se à seguinte conclusão:

Os alunos que possuem um conhecimento lexical reduzido não conseguem construir o sentido pretendido pelo professor que produziu o E.E (enunciado), pois demonstram ignorar os novos sentidos que uma palavra pode ganhar em situações diferentes de uso. Podemos verificar, também, que as marcas de subjetividade assinaladas na análise que fizemos, possam constituir-se como elementos facilitadores na compreensão leitora do E.E., por estabelecerem interlocução entre o autor que elaborou o E.E. (locutor) e o discente (alocutário) que irá resolvê-lo. (SILVA, 2009, p. 95-96).

Não pretendemos, aqui, alongar a discussão conceitual sobre o que torna um texto, e, por extensão, um item de prova, mais ou menos difícil. Interessa-nos, de fato, apresentar uma visão alternativa sobre o fenômeno, ou seja, propomos a análise de um item pela perspectiva da complexidade textual.

Ao discutir as definições de complexidade para uma pesquisa sobre complexidade textual e progressão escolar, Martins (2016) destaca a diversidade de abordagens e enquadramentos teóricos sob os quais o fenômeno vem sendo tratado. O autor, então, propõe, a despeito dessa diversidade, “uma linha divisória que permite pôr, de um lado, a complexidade como um fenômeno subjetivo e, de outro, a complexidade como um fenômeno objetivo” (MARTINS, 2016, p. 11).

Como pode sugerir o próprio termo, a complexidade subjetiva é um fenômeno relativo e depende de um sujeito, social e culturalmente situado. Seguindo Fischer, Frey e Lapp (2012) e Kurters (2003), Martins (2016) defende que a complexidade subjetiva é sinônimo de dificuldade. Nessa perspectiva, portanto, complexidade é entendida como a dificuldade/facilidade que pode ser enfrentada por alguém ao realizar determinada tarefa, de forma mais específica, a tarefa de compreender textos. Os estudos que se desenvolvem a partir desse ponto de vista visam, assim, a avaliar o grau de complexidade (leia-se dificuldade/facilidade) de textos, sobretudo escritos, ao serem interpretados por leitores, normalmente, aprendizes (RICHARDS; SCHIMIDT, 2002). Procura-se determinar o nível de adequação de textos em uma escala de dificuldade e se conclui que há textos mais complexos para uns leitores do que para outros.

A complexidade objetiva é vista “como uma propriedade do sistema linguístico e, conseqüentemente, das suas estruturas, e não algo que lhe é exterior” (MARTINS, 2016, p. 13). O princípio subjacente a essa concepção é o de que a complexidade de um objeto está relacionada à quantidade de informação necessária para recriar ou especificar esse objeto, “ou, ainda, para medir a extensão da sua mínima descrição possível”. Essa complexidade estrutural apresenta duas dimensões: a sistêmica e a estrutural.

A complexidade estrutural sistêmica está relacionada, como sugere o próprio nome, às propriedades dos diferentes sistemas linguísticos. Trata-se da complexidade presente nas gramáticas das línguas. Algumas línguas isolantes, por exemplo, são praticamente isentas de morfologia flexional, são simples, nessa perspectiva. Já outras apresentam alto grau de elaboração do léxico e possuem um sistema complexo de tons lexicais (RIDDLE, 2008). Outro exemplo, mais próximo, dessa complexidade sistêmica é também a flexão de gênero e número: mais complexa em português (pequeno garoto/pequena garota, pequenos garotos/pequenas garotas) e mais simples em inglês (small boy/small girl, small boys/small girls).

A complexidade objetiva estrutural também é um fenômeno intrínseco ao objeto, mas está relacionada às escolhas linguísticas que um falante/escritor faz a partir das estruturas disponíveis. A complexidade objetiva estrutural é inerente ao texto, que apresenta determinadas características estruturais (linguísticas e textuais) associadas aos contextos de uso. Em outras palavras, são as estratégias de construção linguístico-textual específicas empregados pelos falantes/escritores em uma dada ocasião, com um propósito específico. Essas estratégias se configuram nas diferentes formas de se dispor das propriedades linguístico-textuais do sistema, e se manifestam nas diferentes escolhas que os escritores instanciam. Essa característica, como já vimos, também integra a complexidade subjetiva. A diferença, na verdade, é o encaminhamento da análise.

A complexidade objetiva não está associada à legibilidade nem a habilidades ou condições de determinados tipos de leitores, ou seja, a complexidade objetiva se desvincula da noção de dificuldade. Em outras palavras, ocorre o deslocamento do quadro da compreensão, e o encaminhamento não é determinar a legibilidade do texto, tendo em vista habilidades de linguagem de um ou de um conjunto de indivíduos. Em resumo, segundo essa perspectiva a complexidade textual pode ser entendida como um fenômeno objetivo e estrutural ligado ao conjunto de atributos composicionais e estruturais de um texto. Um texto será mais ou menos complexo conforme esses atributos, que são os indicadores de complexidade. Esses atributos ou indicadores de complexidade não deixam de constituir escolhas de um dado locutor em uma dada situação comunicativa, mas os aspectos que colaboram para essas escolhas são metodologicamente desconsiderados.

Se considerarmos a diversidade econômica, social e cultural da população de estudantes que se submete a uma avaliação como o Enem, e se pretendermos analisar a complexidade dos itens, a perspectiva subjetiva não daria conta de estabelecer níveis de dificuldade ou de complexidade confiáveis, e ainda se correria o risco de se sustentarem tomadas de posições depreciativas do ponto de vista ideológico. Portanto, entendemos que, se pretendemos abordar a complexidade de textos dessa natureza, ao menos, em um primeiro momento, precisamos assumir uma perspectiva objetiva estrutural, porque excluiria variáveis de difícil controle. Não estamos negando com isso, convém frisar, que o grau de complexidade do texto não está relacionado ao fato de que ele se constitui como um objeto psicológico social e semiótico. Nesse sentido, embora uma abordagem meramente linguística do texto pareça impraticável, isso não impede o desenvolvimento de perspectivas de análise que, sem reduzir o objeto, possam dar conta de aspectos específicos.

Figueiredo-Gomes e Mendes (2018), por exemplo, desenvolveram um “instrumento de avaliação de complexidade” que envolve três dimensões. As duas primeiras dimensões são linguísticas: uma estrutural, que envolve o número de palavras e a estrutura sintática do texto, e outra cognitiva, que envolve a junção, a referência, a relação dado novo/dado velho e a pontuação. A terceira dimensão é não linguística, porque envolve aspectos visuais do texto. Com base nessas categorias, os autores analisaram textos e provas do Enem e chegaram a uma escala de complexidade com quatro níveis: complexo, difícil, mediano, fácil. “Desse modo, esse instrumento permite avaliar textos de língua portuguesa, diferentemente das medidas até então existentes” (FIGUEIREDO-GOMES; MENDES, 2018, p. 439). A proposta que defendemos, aqui, segue nessa direção: pretendemos apontar indicadores de complexidade de natureza textual, relacionados à dimensão composicional do texto, mas que permitam um espectro maior de níveis.

Os estudos linguísticos do texto contam, atualmente, com diferentes perspectivas teóricas, que apesar de divergirem, seja na concepção de texto, seja na forma de compreender seu funcionamento, convergem no sentido de propor recursos analíticos para explicar a composição textual. Algumas dessas perspectivas oferecem categorias de natureza textual, além das exploradas por Figueiredo-Gomes e Mendes (2018) como a junção e a referência, que também podem dar conta de indicadores de complexidade textual. Recorremos aqui a duas dessas perspectivas: ISD (BRONCKART, 1999) e PTI (JUBRAN, 2006). Não nos interessa, neste momento, discorrer sobre os princípios de cada uma dessas perspectivas. Retomaremos de uma forma bastante geral a proposta de cada uma apenas para justificar o motivo de recorrermos a elas para extrairmos categorias relativas à dimensão composicional do texto que podem ser tomadas como indicadores de complexidade textual.

No quadro do ISD, há duas propostas fundamentais de modelos para a análise de textos: o modelo da ação de linguagem e o modelo da arquitetura textual. O primeiro objetiva caracterizar o contexto de produção de um texto. O segundo modelo é o que nos interessa aqui por tratar especificamente da organização interna do texto. Segundo esse modelo, a organização interna de todo texto é concebida como um folhado, constituído por três camadas superpostas: a infraestrutura geral, os mecanismos de textualização e os mecanismos enunciativos. Cada uma dessas três camadas corresponde a um nível da arquitetura interna dos textos, ou seja, diferentes aspectos linguísticos que desempenham funções específicas na organização textual. O termo nível, no modelo, significa o grau de dependência contextual dos fenômenos.

O primeiro nível ou camada do folhado textual é o dos mecanismos enunciativos. É o mais superficial, ou seja, o que mais diretamente depende dos parâmetros da situação. Objetiva descrever os posicionamentos enunciativos, as diferentes vozes presentes no texto e as modalizações, ou seja, as “avaliações formuladas sobre alguns aspectos do conteúdo temático” (BRONCKART, 1999, p. 131). O segundo nível ou camada é o dos mecanismos de textualização. É o nível intermediário, e visa a apreender a coerência temática do texto, através de três mecanismos: coesão nominal, coesão verbal e conexão. A infraestrutura

geral do texto corresponde ao terceiro nível, o mais profundo, ou seja, o que menos diretamente depende da situação. Essa camada comporta o plano geral do texto, o tipo de discurso, as sequências e outras formas de planificação.

Como nosso propósito é buscar categorias de natureza textual diretamente ligadas à dimensão composicional, acreditamos que entre essas categorias previstas no modelo a que pode estar relacionada à complexidade textual são os tipos de discurso. Os tipos de discurso são “formas linguísticas que são identificáveis nos textos e que traduzem a criação de mundos discursivos específicos” (BRONCKART, 1999, p. 149). Mundos discursivos são mundos virtuais criados pelos agentes humanos pela atividade de linguagem. Os mundos discursivos se constroem, então, com base na relação que se estabelece entre dois tipos de operações que envolvem a produção do texto e a atividade humana. A primeira operação explicita a relação entre as coordenadas gerais que organizam o conteúdo temático do texto e as coordenadas do mundo ordinário no qual se desenvolve a ação de linguagem representada por esse texto. A segunda operação explicita a relação entre as instâncias de agentividade e de tempo/espço mobilizadas no texto e os parâmetros físicos da ação de linguagem. A ancoragem ou não do conteúdo temático nas coordenadas do mundo estabelece a oposição entre mundo do narrar (disjunção) e mundo do expor (conjunção). Já a mobilização ou não de parâmetros do ato de produção e as referências dêiticas a esses atos estabelecem a oposição entre mundo implicado e mundo autônomo. O cruzamento das duas oposições define, então, quatro tipos de mundo (expor implicado, expor autônomo, narrar implicado e narrar autônomo), que, por sua vez, são traduzidos nos quatro tipos de discurso: discurso interativo, discurso teórico, relato interativo e narração.

O discurso teórico, por exemplo, é autônomo em relação aos parâmetros da ação de linguagem, ou seja, mobiliza operações cognitivas em um sistema que é genérico, abstrato e depende de circunstâncias particulares. Ao contrário, o relato interativo tem outro tipo de relação com a ação de linguagem e mobiliza operações menos abstratas. Logo, um texto que apresenta discurso teórico tende a ser mais complexo que um texto que apresenta relato interativo. Com base nisso, entendemos que os tipos de discurso podem compor uma propriedade de natureza textual através do qual se pode analisar o nível de complexidade do texto.

A PTI (JUBRAN *et al*, 1992; 2006), como proposta teórica para a análise textual, foi desenvolvida pelo grupo do texto do Projeto Gramática do Português Falado (PGPF), no Brasil (KATO, 1996; CASTILHO, 2007). Em conformidade com nosso objetivo de propor categorias do plano linguístico do texto como indicadores de complexidade textual, tomaremos dessa perspectiva o aporte para a análise dos procedimentos de elaboração textual. O primeiro e principal deles é a organização tópica do texto.

Para que se verifiquem os processos globais de organização textual-interativa, é necessário isolar convenientemente as informações do texto e acompanhar os seus diferentes estágios de desenvolvimento, através de unidades analíticas menores compatíveis com essa forma de concebê-lo. Para cumprir, portanto, com o objetivo de propor uma unidade analítica compatível com a forma de conceber o texto e sua organização, a PTI apresenta a noção de tópico discursivo. A primeira análise do texto requer, portanto, a apreensão da sua organização tópica.

O tópico é uma categoria abstrata, que se manifesta no texto mediante enunciados formulados em torno de um conjunto de referentes explícitos ou inferíveis, concernentes entre si e em relevância num determinado ponto. O tópico, nessa perspectiva, abrange duas propriedades que o particularizam: a centração e a organicidade. A centração abrange ainda a concernência, que é relação de interdependência semântica entre os enunciados; a relevância e a pontualização (JUBRAN, 2006).

As propriedades do tópico se atualizam no texto de diferentes maneiras. Um tópico pode ser expresso diretamente, mas pode também ser expresso indiretamente, pois a relação de concernência pode se dar por inferência ou associação. Assim, entendemos que o tópico discursivo também pode constituir um atributo relacionado à complexidade textual. Se o

uso de expressões que colaboram para a centração e o consequente estabelecimento do tópico é explícito e apresenta relação mais direta, isso deixa o texto menos complexo. Ao contrário, se o estabelecimento da concernência se dá por inferência ou associação, o texto será mais complexo.

A organicidade constitui outra propriedade que também está relacionada à complexidade. Se um texto, por exemplo, apresenta um único tópico com poucos desdobramentos em subtópicos e relação de concernência explícita, ele será menos complexo. Ao contrário, um texto com mais de um tópico, sucessivos níveis de desdobramento e relação de concernência implícita será mais complexo.

Podemos, agora, sistematizar as categorias que podem ser mobilizadas para a análise da complexidade textual tal como a concebemos, ou seja, como um fenômeno objetivo e estruturado ligado ao conjunto de atributos composicionais do texto, que o tornam mais ou menos complexo. Prevemos duas categorias textuais: os tipos de discurso e o tópico discursivo. Em referência aos tipos de discurso, levamos em conta os seguintes atributos: (a) discurso interativo (conjunção/implicação); (b) discurso teórico (conjunção/autonomia); (c) relato interativo (disjunção/implicação); (d) narração (disjunção/autonomia). Com base no tópico, os atributos são: (a) um único tópico; (b) mais de um tópico; (c) tópico sem desdobramento; (d) tópico com desdobramento; (e) relação de concernência explícita; (f) relação de concernência implícita.

As diferentes combinações desses atributos servem para dispor o texto em um contínuo de complexidade, que vai do menos complexo, extrema esquerda, ao mais complexo, extrema direita.

3 O CONTÍNUO DE COMPLEXIDADE EM ITENS DE PROVA DO ENEM

Para demonstrar o raciocínio de que os itens da prova do ENEM podem ser analisados segundo o critério da complexidade textual, apoiado nas duas categorias (tipos de discurso e tópico discursivo), propomos, aqui, a análise de um pequeno *corpus* de itens do Enem. A partir das respostas de mais de 300.000 alunos no aplicativo *Quero Minha Nota!*, a TUNEDUC¹, *startup* dedicada a estudos educacionais, com base em critérios estatísticos, elencou as três questões mais fáceis e difíceis do Enem/2016 por área do conhecimento. Os 24 itens distribuídos nos dois grupos (fácil e difícil) constituem, portanto, o *corpus* objeto da nossa análise (Quadro 01)².

Quadro 01: Características dos itens

Área do conhecimento	Número do item			Classificação
Linguagens, Códigos e suas Tecnologias	105	121	122	Fácil
	100	113	129	Difícil
Matemática e suas Tecnologias	153	154	155	Fácil
	175	178	179	Difícil
Ciências da Natureza e suas Tecnologias	79	87	89	Fácil
	66	70	86	Difícil
Ciências Humanas e suas Tecnologias	31	32	41	Fácil
	14	23	35	Difícil

Fonte: Elaborado pelos autores

¹ <https://www.tuneduc.com.br>

² A prova completa com os itens do Quadro 01 pode ser acessada em: <http://portal.inep.gov.br/web/guest/provas-e-gabaritos>

Os 24 itens foram dispostos em um contínuo, que apresentou 5 níveis em que 1 é o menos complexo e 5 o mais complexo. Esse contínuo foi constituído com base na identificação dos atributos composicionais, conforme explicado na seção anterior. São considerados atributos de baixa complexidade: narração ou relato interativo, um tópico, tópico sem desdobramento, relação de concernência explícita. Os atributos considerados de alta complexidade são: discurso teórico ou discurso interativo, mais de um tópico, tópico com desdobramento, relação de concernência implícita.

No nível 1, estão os itens que apresentam quatro atributos de baixa complexidade. O nível 2 é formado pelos itens que apresentam três atributos de baixa complexidade e um de alta. No nível 3, estão os itens que apresentam dois atributos de baixa complexidade e dois de alta. O nível 4 reúne os itens que apresentam três atributos de baixa complexidade e dois de alta. Finalmente, no nível 5, estão os itens que apresentam dois atributos de baixa complexidade e três de alta. Esse contínuo está representado na Figura 01.



Figura 01: Contínuo de complexidade

Fonte: Elaborado pelos autores

Vejamos, a título de exemplo, um caso para cada um dos níveis. O item 155 (Figura 02) é composto pela narração. O conteúdo temático é ancorado em coordenadas do mundo: “os alunos de uma escola...”, “...uma aula ao ar livre...” e “...ao final da aula...”, o que marca a disjunção, e não apresenta marcas de elementos da situação de produção, o que caracteriza a autonomia. A disjunção autônoma, portanto, caracteriza a narração, que, por sua vez, é um atributo de baixa complexidade. O tópico é “cadeira”. Esse tópico é retomado por expressões referenciais “fechassem as cadeiras”, “guardá-la”, “esboço lateral da cadeira”. A retomada dessas expressões evidencia a relação de concernência explícita. Esse tópico não se desdobra, logo não há instauração de subtópicos. Trata-se, por isso, com base na categoria tópico discursivo, de um texto que apresenta três indicadores de baixa complexidade: um tópico, tópico sem desdobramento e relação de concernência explícita. O item apresenta, portanto, 4 atributos de baixa complexidade (narração, um único tópico, tópico sem desdobramento e concernência explícita) e está no nível 1 de complexidade.

QUESTÃO 155

Os alunos de uma escola utilizaram cadeiras iguais às da figura para uma aula ao ar livre. A professora, ao final da aula, solicitou que os alunos fechassem as cadeiras para guardá-las. Depois de guardadas, os alunos fizeram um esboço da vista lateral da cadeira fechada.



Qual é o esboço obtido pelos alunos?



Figura 02: Item nível 1, Matemática e suas tecnologias

Fonte: Prova Enem – Caderno Azul 7 – 2016, p. 23

No item 154 (figura 03), predomina o discurso teórico (conjunção e autonomia), pois não há unidades que se referem à situação de produção nem termos que marquem a disjunção do texto em relação às coordenadas gerais do mundo. Há predominância de frases declarativas, marcadas pelo uso de verbos no presente genérico “apresentam”, “possui” e oração com voz passiva “Recomenda-se a ingestão do pão que possui a maior concentração de fibras”, que aparecem no texto-base e no enunciado. O tópico é “Marcas de pão integral”. Ele é instaurado pela retomada no texto pelas expressões referenciais “concentração de fibras”, “massa de fibra”, “2-7g de fibras”. Assim sendo, a relação de concernência é explícita, já que essas expressões pertencem ao mesmo campo semântico, além de haver, no decorrer do texto, a repetição da palavra “pão”. Esse é o único tópico do texto, que não se desdobra, não havendo, por isso mesmo, subtópicos. Trata-se, portanto, de um texto que apresenta um atributo de alta complexidade (discurso teórico) e três atributos de baixa complexidade (um tópico, tópico sem desdobramento e relação de concernência explícita). Com essas características, o item se insere no nível 2 de complexidade.

QUESTÃO 154

Cinco marcas de pão integral apresentam as seguintes concentrações de fibras (massa de fibra por massa de pão):

- Marca A: 2 g de fibras a cada 50 g de pão;
- Marca B: 5 g de fibras a cada 40 g de pão;
- Marca C: 5 g de fibras a cada 100 g de pão;
- Marca D: 6 g de fibras a cada 90 g de pão;
- Marca E: 7 g de fibras a cada 70 g de pão.

Recomenda-se a ingestão do pão que possui a maior concentração de fibras.

Disponível em: www.blog.saude.gov.br. Acesso em: 25 fev. 2013.

A marca a ser escolhida é

- A** A.
- B** B.
- C** C.
- D** D.
- E** E.

Figura 03: Item nível 2, Matemática e suas tecnologias

Fonte: Prova Enem – Caderno Azul 7 – 2016, p. 22

O item 122 (figura 04) se insere no nível 3. No segmento do texto-base que vai do início até “podem doar sangue” (os três primeiros parágrafos), o conteúdo está em conjunção com as coordenadas do mundo já que não há nenhum elemento no texto que marca uma ancoragem espacial ou temporal e os agentes do ato de produção não estão implicados. Configura-se, assim, um segmento de discurso teórico. Já o segmento que se inicia em “se você acha que sua saúde ou comportamento” e segue até o final (três últimos parágrafos) é marcado pelo discurso interativo, pois o conteúdo está em conjunção com as coordenadas do mundo, mas um agente do ato de produção está implicado pela mobilização do uso dos pronomes “você” e “seu”. No enunciado, o discurso teórico é retomado. Como os dois tipos de discurso são características de alta complexidade, consideramos que o item apresenta, portanto, um atributo de alta complexidade.

O tópico é “doação de sangue”. Esse tópico é instaurado pela retomada, no texto-base, no enunciado e nas alternativas, das expressões referenciais “segurança do sangue”, “indivíduos saudáveis”, “hábitos de doar”, “doadores”, “transportar”, “estocar”, “testar”, “produto seguro”, “risco de doenças infecciosas”, “transmissíveis”, “hepatite B e C”, “HIV”, “sífilis”, “Chagas”. Há relação de concernência explícita, uma vez que essas expressões pertencem ao mesmo campo de sentido, além da própria repetição do referente “sangue” ao logo do texto. Da mesma maneira, percebemos que o tópico doação de sangue passa por uma especificação e se desdobra em dois subtópicos: segurança do sangue e campanha de doação. Trata-se, por essa razão, com base na categoria tópico discursivo, de um item que apresenta: duas propriedades de baixa complexidade (um tópico e relação de concernência explícita), e um de alta complexidade (tópico com desdobramento). No cômputo geral, são, em resumo, dois atributos de alta complexidade (discurso teórico/discurso interativo e tópico com desdobramento) e dois de baixa (apenas um tópico e relação de concernência explícita).

QUESTÃO 122**Qual é a segurança do sangue?**

Para que o sangue esteja disponível para aqueles que necessitam, os indivíduos saudáveis devem criar o hábito de doar sangue e encorajar amigos e familiares saudáveis a praticarem o mesmo ato.

A prática de selecionar criteriosamente os doadores, bem como as rígidas normas aplicadas para testar, transportar, estocar e transfundir o sangue doado fizeram dele um produto muito mais seguro do que já foi anteriormente.

Apenas pessoas saudáveis e que não sejam de risco para adquirir doenças infecciosas transmissíveis pelo sangue, como hepatites B e C, HIV, sífilis e Chagas, podem doar sangue.

Se você acha que sua saúde ou comportamento pode colocar em risco a vida de quem for receber seu sangue, ou tem a real intenção de apenas realizar o teste para o vírus HIV, NÃO DOE SANGUE.

Cumpra destacar que apesar de o sangue doado ser testado para as doenças transmissíveis conhecidas no momento, existe um período chamado de janela imunológica em que um doador contaminado por um determinado vírus pode transmitir a doença através do seu sangue.

DA SUA HONESTIDADE DEPENDE A VIDA DE QUEM VAI RECEBER SEU SANGUE.

Disponível em: www.prosangue.sp.gov.br. Acesso em: 24 abr. 2015 (adaptado).

Nessa campanha, as informações apresentadas têm como objetivo principal

- A** conscientizar o doador de sua corresponsabilidade pela qualidade do sangue.
- B** garantir a segurança de pessoas de grupos de risco durante a doação de sangue.
- C** esclarecer o público sobre a segurança do processo de captação do sangue.
- D** alertar os doadores sobre as dificuldades enfrentadas na coleta de sangue.
- E** ampliar o número de doadores para manter o banco de sangue.

Figura 04: Item nível 3, Linguagens, Códigos e suas Tecnologias

Fonte: Prova Enem – Caderno Azul 7 – 2016, p. 13

Na composição do item 121 (figura 05), entram dois tipos de discurso: discurso interativo e narração. No texto-base, o discurso interativo é marcado pela presença de elementos do ato de produção, o locutor (escrevo) e o interlocutor (a personificação da poesia pelo uso do pronome “te”), e o conteúdo temático não está situado em nenhuma coordenada do mundo, portanto conjunto. No enunciado, há uma unidade que remete o conteúdo a um espaço (“Nesse fragmento de João Cabral de Melo Neto”) e a um tempo (“geração de 1945, João Cabral de Melo Neto”). Ocorre, assim, a disjunção. Não há, por outro lado, nenhuma marca dos agentes da situação de produção, o que configura a autonomia, e, portanto, a narração. Nesse sentido, o item, no que diz respeito ao tipo de discurso, apresenta um atributo de alta complexidade (discurso interativo), e um de baixa, (narração).

O tópico discursivo é “a palavra”. Esse tópico é instaurado por uma cadeia referencial composta por expressões cujas relações não são explícitas. No texto-base, os referentes “flor”, que é retomado várias vezes, e “poesia”, que aparece no texto-base, são retomados no enunciado. Da mesma forma, uma série de outros referentes sem relações explícitas entre si aparecem nas alternativas: “palavra”, “urinal”, “ave”, “máquina”, “tecido”. Para que se estabeleça o referente “palavra” como tópico do texto, é preciso uma série de inferências, o que permite dizer que as relações de concernência para o estabelecimento do tópico são implícitas. O item, como o próprio enunciado aponta, apresenta, do ponto de vista da criação poética as características da palavra, que se configura como o único tópico, sem desdobramento. Trata-se, portanto, com base na categoria tópico discursivo, de um item que apresenta dois atributos de baixa complexidade (um tópico, tópico sem desdobramento) e um atributo de alta complexidade (relação de concernência implícita). No

cômputo geral, são dois atributos de alta complexidade (discurso interativo e relação de concernência implícita) e três de baixa (narração, um tópico, tópico sem desdobramento), o que caracteriza o nível 4.

QUESTÃO 121**Antiode**

Poesia, não será esse
o sentido em que
ainda te escrevo:
flor! (Te escrevo:
flor! Não uma
flor, nem aquela
flor-virtude — em
disfarçados urinóis).
Flor é a palavra
flor; verso inscrito
no verso, como as
manhãs no tempo.
Flor é o salto
da ave para o voo:
o salto fora do sono
quando seu tecido
se rompe; é uma explosão
posta a funcionar,
como uma máquina,
uma jarra de flores.

MELO NETO, J. C. *Psicologia da composição*. Rio de Janeiro: Nova Fronteira, 1997 (fragmento).

A poesia é marcada pela recriação do objeto por meio da linguagem, sem necessariamente explicá-lo. Nesse fragmento de João Cabral de Melo Neto, poeta da geração de 1945, o sujeito lírico propõe a recriação poética de

- A** uma palavra, a partir de imagens com as quais ela pode ser comparada, a fim de assumir novos significados.
- B** um urinol, em referência às artes visuais ligadas às vanguardas do início do século XX.
- C** uma ave, que compõe, com seus movimentos, uma imagem historicamente ligada à palavra poética.
- D** uma máquina, levando em consideração a relevância do discurso técnico-científico pós-Revolução Industrial.
- E** um tecido, visto que sua composição depende de elementos intrínsecos ao eu lírico.

Figura 05: Item nível 4, Linguagens, Códigos e suas Tecnologias

Fonte: Prova Enem – Caderno Azul 7 – 2016, p. 13

Por fim, o item 23 (figura 06) se insere no nível 5, caracterizado por apresentar três atributos de alta complexidade e dois de baixa. O item é composto integralmente pelo discurso teórico (conjunção e autonomia), pois os temas tratados aparecem de forma conjunta às coordenadas do mundo, visto que não há ancoragem no espaço/tempo. O item apresenta, portanto, como base no tipo de discurso, um atributo de alta complexidade: o discurso teórico.

Os dois fragmentos do texto-base instauram um único tópico: “mutabilidade do ser”. A cadeia referencial desses fragmentos não permite a identificação direta do tópico. É preciso um complexo processo de inferência para se vincular, por exemplo, expressões “banhar duas vezes no mesmo rio”, “substância”, intensidade”, “rapidez”, “mudança”, “dispersa” e “reúne”, “compacto”, “inabalável”, “homogêneo”, “uno”, “contínuo”, “perecer” em um conjunto homogêneo. As relações de concernência que instauram o tópico são, portanto, implícitas. No enunciado, o tópico é outro: “pensamento pré-socrático”. Esse tópico é explicitamente anunciado (“os fragmentos do pensamento pré-socrático...”).

No que diz respeito ao tópico, esse item apresenta características diferentes dos demais: dois tópicos (atributo de alta complexidade), relações de concernência implícitas para instauração de um dos tópicos (alta complexidade) e explícitas para instauração de outro (baixa complexidade), e tópicos sem desdobramento em ambos os casos (baixa

complexidade). O item apresenta, portanto, três atributos de alta complexidade e dois de baixa.

QUESTÃO 23

TEXTO I

Fragmento B91: Não se pode banhar duas vezes no mesmo rio, nem substância mortal alcançar duas vezes a mesma condição; mas pela intensidade e rapidez da mudança, dispersa e de novo reúne.

HERÁCLITO. *Fragmentos (Sobre a natureza)*. São Paulo: Abril Cultural, 1996 (adaptado).

TEXTO II

Fragmento B8: São muitos os sinais de que o ser é ingênito e indestrutível, pois é compacto, inabalável e sem fim; não foi nem será, pois é agora um todo homogêneo, uno, contínuo. Como poderia o que é perecer? Como poderia gerar-se?

PARMÊNIDES. *Da natureza*. São Paulo: Loyola, 2002 (adaptado).

Os fragmentos do pensamento pré-socrático expõem uma oposição que se insere no campo das

- A** investigações do pensamento sistemático.
- B** preocupações do período mitológico.
- C** discussões de base ontológica.
- D** habilidades da retórica sofística.
- E** verdades do mundo sensível.

Figura 06: Item nível 5, Ciências Humanas e suas Tecnologias

Fonte: Prova Enem – Caderno Azul 1 – 2016, p. 8

Feita essa análise, os 24 itens foram distribuídos no contínuo de complexidade: entre os itens fáceis, o item 155 está no nível 1, os itens 105, 153, 154, 79, 87, 89, 31, 32, 41 estão no nível 2, o item 121, no nível 4, e o item 122, no nível 3. Entre os difíceis, os itens 100, 129, 175, 178, 179, 66, 70, 86, 14, 35 estão no nível 2, o item 113, no nível 4, e o 23, no nível 5. A Tabela (02) resume a distribuição quantitativa dos itens conforme o nível de complexidade e o grau de dificuldade, estatisticamente medido. Como se pode observar, com base no quadro, a grande parte dos itens (19 itens, 09 fáceis e 10 difíceis) estão no nível 2. Os outros se distribuem pelos demais níveis: 1 item no nível 1 (fácil), 1 item no nível 3 (fácil), 2 itens no nível 4 (1 fácil e 1 difícil), e 1 item no nível 5 (difícil).

Tabela 02: Disposição dos itens no contínuo

Nível de complexidade	Item fácil	Item difícil	Total
1	01	00	01
2	09	10	19
3	01	00	01
4	01	01	02
5	00	01	01
			24

Fonte: Elaborado pelos autores

Essa simples constatação mostra a falta de pareamento entre item difícil/fácil, segundo o critério estatístico, e item mais complexo/menos complexo, segundo o critério da

complexidade textual. Grande parte dos itens (20, no total) se situam mais à esquerda do contínuo (nível 1 e 2), de baixa complexidade textual, portanto, independentemente de serem fáceis ou difíceis. Dos dois itens que estão no nível 4 de complexidade, ou seja, mais complexo, um é fácil, o outro é difícil. Trata-se de mais um dado que evidencia a falta de pareamento entre os dois critérios. Há apenas dois casos de pareamento: um item, ao mesmo tempo, menos complexo (nível 1) e fácil, e um item, ao mesmo tempo, mais complexo (nível 5) e difícil. Como são casos pontuais, apostamos na coincidência e optamos por aceitar a ideia do não pareamento.

Com base nessa constatação, convém se perguntar: por que razão, por exemplo, um item menos complexo do ponto de vista da composição textual não é, necessariamente, fácil do ponto de vista da dificuldade? E o contrário: por que um item mais complexo não é o mais difícil? A princípio parece óbvio afirmar que um texto mais complexo do ponto de vista da sua composição deve apresentar maior dificuldade. Essa obviedade, na verdade, só existe, se não são consideradas as concepções específicas de complexidade e dificuldade com as quais se opera.

Quando se fala de itens de provas do Enem, dificuldade é concebida como uma das categorias a que se recorre para realizar uma meta-avaliação, ou seja, “o ato de se avaliar uma avaliação, com objetivo de evitar distorções e imprecisões que possam prejudicar o processo educativo de crianças e jovens” (TRAVITZI, 2017, p. 262). Como já dissemos, o grau de dificuldade é um dos parâmetros de um item, ao lado de outros, como a discriminação (capacidade de alguém acertar a resposta em itens mais difíceis do que outras) e probabilidade (a possibilidade de alguém acertar a resposta ao acaso), que podem atestar seu comportamento empírico e verificar sua qualidade psicométrica.

Em outras palavras, isso significa dizer que a falta de pareamento entre grau de dificuldade e nível de complexidade ocorre, porque, de fato, são dimensões não relacionadas, que atuam de forma independente, pelo menos a princípio, na composição do item. Os testes psicométricos que analisam o grau de dificuldade de um item contemplam uma habilidade latente e como essa habilidade influencia a resposta de cada item. No caso dos itens do Enem, são consideradas as 21 habilidades que expressam as cinco competências que estruturam o exame (BRASIL, 2002).

Embora essas competências pressuponham, de forma geral e transversal, a competência textual (ler e produzir textos), as características organizacionais e composicionais do texto não compõem um parâmetro a ser medido na avaliação do item, seja em relação à dificuldade, discriminação ou probabilidade. Em outras palavras, o instrumental linguístico em si (a organização textual do item) usado pelos participantes na mobilização do conhecimento não entra nesses parâmetros. A complexidade textual tal como a concebemos aqui constitui, portanto, a nosso ver, um outro parâmetro a ser considerado na avaliação do item. Trata-se, no caso, como já dissemos, de um parâmetro distinto. A dificuldade, como já vimos, avalia o item do ponto de vista de habilidades relativas à mobilização de conhecimentos de diversas áreas e suas tecnologias, e se atribuem os graus de fácil, médio e difícil. A complexidade, por outro lado, avalia o conjunto de atributos composicionais e estruturais do item, que apresenta níveis de complexidade conforme esses atributos, ou, em outras palavras, conforme os indicadores de complexidade.

4 CONCLUSÃO

Ao retomarmos a proposta de apontar indicadores de complexidade de natureza textual de modo a identificar um contínuo de complexidade (do mais para o menos complexo) nos itens da prova do Enem, podemos chegar a dois tipos de conclusão: uma que diz respeito a considerações teórico-analíticas de uma abordagem textual de um item e outra sobre a aplicação possível desse conhecimento.

No que diz respeito ao primeiro tipo, o trabalho focaliza a concepção de uma complexidade textual estrutural, entendida como um fenômeno intrínseco ao texto e

relacionado às escolhas que o escritor/falante realiza a partir da dimensão estrutural e composicional do texto. A análise da complexidade textual passa, portanto, pela identificação de atributos de natureza textual, que são considerados como indicadores do grau de complexidade do texto. Os atributos de complexidade considerados na análise (tipos de discurso e tópico discursivo) foram tomados emprestados de duas abordagens que oferecem modelos de análise da estrutura do texto: ISD e PTI. No entanto, outras categorias que compõem modelos analíticos de outras abordagens podem ser acrescentadas, e, dessa forma, incrementa-se a análise da complexidade textual, o que possibilita a ampliação da quantidade de atributos e dos pontos do contínuo.

Quanto ao segundo tipo de conclusão, a análise da complexidade aqui proposta pode ser usada como um dos procedimentos a serem levados em conta para criação, pré-testagem, seleção e definição de características psicométricas desejáveis para os itens do Enem. Em outras palavras, o conhecimento sobre a complexidade textual pode ampliar as políticas de elaboração e aplicação da prova. De forma mais pontual, esse conhecimento pode colaborar, diretamente, nas atividades de ensino, em sala de aula. O professor, por exemplo, pode recorrer à análise da complexidade para melhor fundamentar a leitura do item pelos estudantes de forma a melhor chegar à resposta.

Por fim, queremos assinalar a natureza propositiva deste trabalho, e esperamos que as questões aqui apontadas sirvam de inspiração para trabalhos futuros, para que a sua relevância e alcance possam ser discutidos.

REFERÊNCIAS

- ANDRADE, D. F.; KLEIN, R. Aspectos quantitativos da análise dos itens da prova do Enem. In: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Exame Nacional do Ensino Médio (Enem): fundamentação teórico-metodológica**. Brasília: Inep, 2005. p. 107-112.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Exame Nacional do Ensino Médio: documento básico**. Brasília: Inep, 2002.
- BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Guia de elaboração e revisão de itens**. Brasília: INEP, 2010.
- BRONCKART, J-P. **Atividade de linguagem, discurso e desenvolvimento humano**. Campinas: Mercado de Letras, 1999.
- CASTILHO, A. T. Fundamentos teóricos da gramática do português culto falado no Brasil: sobre o segundo volume, classes de palavras e as construções gramaticais. **Alfa**, v. 51, n. 1, p. 99-135, 2007.
- COSCARELLI, C. V. Entendendo a leitura. **Revista de Estudos da Linguagem**, v. 10, n. 1, p. 7-27, 2002.
- FIGUEIREDO-GOMES, J. B.; MENDES, W. V. Busca de medida de complexidade textual numa abordagem funcionalista. **Entrepalavras**, v. 8, n. 2, p. 419-441, 2018.
- FISHER, D., FREY, N.; LAPP, D. **Text complexity: raising rigor in reading**. Newark: International Reading Association, 2012.
- JUBRAN, C. C. A. S. *et al.* Organização tópica da conversação. In: ILARI, R. (Org). **Gramática do português falado**, v. 2. Campinas/SP: Unicamp/São Paulo: FAPESP, 1992. p. 322-384.
- JUBRAN, C. C. A. S. A Perspectiva textual-interativa. In: JUBRAN, C. C. A. S.; KOCH, I. G. V. (Orgs.). **Gramática do português falado culto falado no Brasil: a construção do texto falado**. v. I. Campinas: Editora da UNICAMP, 2006, p. 27-36.
- KUSTERS, W. **Linguistic complexity: the influence of social change on verbal inflection**. Utrecht: LOT, 2003.
- MARTINS, M. G. C. **Complexidade textual e progressão escolar em dois registos: um estudo de correlação baseado em um corpus quase-longitudinal**. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade de Lisboa, Lisboa, 2016.
- PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item. **Avaliação Psicológica**, v. 2, n. 2, p. 99-110, 2003.
- RICHARDS, J. C.; SCHMIDT, R. **Longman dictionary of applied linguistics and language teaching**. Londres: Longman, 2002.

- RIDDLE, E. Complexity in isolating languages: lexical elaboration versus grammatical economy. In: Karlsson, F.; Miestamo, M.; Sinnemäki, K. (Eds.). **Language complexity: typology, contact, change**. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2008. p. 133-151.
- SILVA, É. M. P. **Leitura e compreensão de enunciados de exercícios de Língua Portuguesa**. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Filosofia, Comunicação, Letras e Artes, Pontifícia Universidade Católica de São Paulo, São Paulo, 2009.
- TRAVITZI, R. Avaliação da qualidade do ENEM 2009 e 2011 com técnicas psicométricas. **Estudos de Avaliação Educacional**, v. 28, n. 67, p. 256-288, 2017.

Contribuição dos autores:

Clemilton Lopes Pinheiro contribuiu com conceptualização, escrita (rascunho original, análise e edição).

Jailson Almeida Conceição contribuiu com conceptualização, escrita (rascunho original).