



Propriedades linguísticas da redação do Enem: uma análise computacional

Linguistic properties of Enem essays: a computational analysis

Roberlei Alves Bertucci

Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná / Brasil

bertucci@utfpr.edu.br

<https://orcid.org/0000-0003-4014-5610>

Resumo: Este texto descreve algumas propriedades linguísticas recorrentes em textos nota 1000 do Enem. Parte-se do princípio de que o gênero *redação do Enem* tem características próprias, como a exposição aliada à argumentação, e que elementos como repertório e amplo uso de conectivos e modalizadores podem contribuir para a caracterização do texto. Nesta pesquisa, levando em conta o rigoroso processo de avaliação pelo qual passam, consideram-se as redações nota 1000 como exemplares prototípicos do gênero, ou seja, cumprem todos os requisitos exigidos pela banca. O *corpus* é constituído de 95 redações que alcançaram a nota máxima nos anos de 2014, 2018 e 2019, analisadas por meio do *software Tropes*, uma ferramenta computacional de análise lexical que verifica as recorrências de categorias e repertório. Os resultados mostraram que tais redações apresentam uma estrutura muito próxima àquela que se verifica na literatura sobre o tema, em especial o predomínio da estrutura impessoal (terceira pessoa), a amplitude do repertório (universo de referência), a variedade e alta frequência de conectivos e modalizadores, além da recorrência de verbos estativos, como *ser*. Com isso, conclui-se tanto que a ferramenta contribui para a descrição do gênero em questão quanto que os resultados fomentam um debate em torno da padronização da estrutura do texto do Enem.

Palavras-chave: gênero textual; ferramenta computacional; redação do Enem; propriedades linguísticas.

Abstract: This paper aims to describe some linguistic properties found in Enem essays graded to 1000 points, the maximal score. It considers that this genre has its own characteristics, such as the relation between exposition and argumentative types, a wide repertoire, as well as a large use of connectives and modal elements, all of them

contribute to characterize this genre. In this research, by considering the strict evaluation process, it has been considered that these essays graded to 1000 points prototypically represent the genre, once that they meet the criteria required by the evaluation panel. The *corpus* consists of 95 essays that got 1000 points in 2014, 2018 and 2019 exams. It has been analyzed by means of Tropes, a computational tool that verifies the frequency of lexical items group them in categories and repertoire. The results show that such essays follow what the literature proposed by the genre, especially, the impersonality (third person only), the vast repertoire extension (universe of reference), the range of connectives and modal elements, besides the recurrence of stative verbs, such as *ser* ('to be'). Consequently, one concludes that both the tool contributes to the genre description, and the results put forward the debate around the standardization of the Enem essay structure.

Keywords: textual genre; computational tool; Enem essay; linguistic properties.

Recebido em 01 de outubro de 2020

Aceito em 25 de novembro de 2020

1 Introdução

Trabalhar com textos argumentativos requer do pesquisador a capacidade de refletir sobre estratégias linguísticas que amarram as ideias num todo capaz de defender uma tese. Em outras palavras, exige que ele seja capaz de descrever os elementos textuais que apontam para uma determinada conclusão argumentativa. Assim, a argumentação tem sido colocada em foco nos estudos em linguagem, esta encarada como meio de produção de interações sociais. Levando em conta esse caráter da linguagem, Koch (2011, p. 15) defende o trabalho com o tema por considerar que a argumentação caracteriza a ação linguística “dotada de intencionalidade”. Além disso, para a autora, é pelo trabalho com a argumentação que se pode desenvolver “a capacidade de refletir, de maneira crítica, sobre o mundo”, por meio da linguagem (KOCH, 2011, p. 15). De maneira mais completa, Koch (2011, p. 17, grifos da autora) assinala que

a interação social por intermédio da língua caracteriza-se, fundamentalmente, pela argumentatividade. Como ser dotado de razão e vontade, o homem, constantemente, avalia, julga, critica, isto é, forma juízos de valor. Por outro lado, por meio do

discurso – ação verbal dotada de intencionalidade – tenta influir sobre o comportamento do outro ou fazer com que compartilhe determinadas de suas [*sic*] opiniões. É por esta razão que se pode afirmar que o **ato de argumentar**, isto é, de orientar o discurso no sentido de determinadas conclusões, constitui o ato linguístico fundamental, pois a **todo e qualquer discurso subjaz uma ideologia**, na acepção mais ampla do termo.

Como se vê, Koch (2011) deixa claro seu posicionamento sobre o papel central da argumentação nas ações linguísticas cotidianas, já que os juízos que formamos, em geral, são compartilhados nessas interações, concretizadas nos discursos produzidos.

Nesse sentido, as redações do Exame Nacional do Ensino Médio (Enem) são uma fonte interessante de estudo de argumentação. Realizada por mais de 4 milhões de candidatos anualmente, a redação do Enem é um texto que merece uma atenção por parte de professores e pesquisadores, uma vez que o sucesso na prova pode contribuir para a continuidade dos estudos de muitos estudantes a partir da política educacional estabelecida (OLIVEIRA; CABRAL, 2017). Nessa prova, além da exposição, discussão e análise de um tema, há uma necessidade de articular informações de áreas diversas com coesão e coerência. Além disso, conforme se lê nos enunciados das propostas – e na Cartilha do Participante (BRASIL, 2019) –, o candidato precisa apontar uma proposta de intervenção na situação-problema indicada no tema. Por essa razão, selecionamos para este trabalho alguns exemplares de textos nota 1000 com a intenção de descrever as características linguísticas mais recorrentes nesse gênero.

Nesta pesquisa, consideramos os textos nota 1000 como prototípicos do gênero *redação do Enem*, porque, para a atribuição de tal nota, é necessária uma avaliação criteriosa da banca a partir do que exige o Exame. Sabe-se que a avaliação ocorre às cegas por no mínimo dois profissionais capacitados para a atividade (BRASIL, 2019), mas, no caso das redações que alcançaram a nota máxima, a atribuição é realizada por outros profissionais que precisam ratificar a avaliação. Essa prototipicidade tem levado alguns estudiosos a debaterem sobre um excesso de padronização em detrimento de questões autorais (LIMA, 2019; PAIVA, 2020). Ainda que indiretamente, acreditamos que o presente trabalho pode contribuir para o debate.

Apesar dessa discussão em torno da padronização, dados revelam que o número de textos que recebem a nota máxima vem caindo ao longo dos processos, como se vê no Quadro 1.

QUADRO 1 – Número de redações nota 1000 (últimos 7 anos do Enem)

Ano do Enem	Textos com nota 1000
2013	481
2014	250
2015	104
2016	77
2017	53
2018	55
2019	53

Fonte – Adaptado de Campos (2020).

Isso parece indicar que, apesar de a proposta de produção ser basicamente a mesma ao longo dos anos, os textos parecem estar se distanciando do padrão exigido.¹ A pergunta que primeiramente motivou o olhar para essas redações prototípicas foi: quais seriam as características linguísticas recorrentes da redação do Enem, a partir do que é exigido pelo Exame? Entendemos que a melhor resposta passaria por uma análise dos textos que obtiveram nota máxima no certame.

Várias pesquisas têm sido feitas sobre a redação do Enem, numa tentativa de se compreender melhor o gênero. Do ponto de vista argumentativo, muitos trabalhos relacionam, por exemplo, os aspectos da Nova Retórica com os textos do Enem, como se vê em Magalhães (2013), Barros e Albuquerque (2015), Azevedo (2015), Oliveira, F. (2016), Oliveira e Cabral (2017), Pinheiro e Cortez (2017), Bertucci (no prelo), entre outros.

¹ Apesar de não ser foco desta pesquisa, é igualmente válido considerar a subjetividade da banca de avaliação, conforme sugerem Cançado *et al.* (2020), ou o fato de a mesma banca ter se tornado mais rígida quanto à atribuição da nota máxima. Mesmo assim, esperava-se que mais candidatos tirassem a nota máxima ao longo do tempo, em razão de a preparação para o texto ser beneficiada pela recorrência do gênero no Exame.

Do ponto de vista da estrutura, Bertucci *et al.* (2020) descrevem a presença de recursos anafóricos nesses textos, mostrando como os encapsulamentos contribuem para a coesão e a direção argumentativa do texto. Um pouco mais próximo àquilo que realizamos aqui, Paiva (2020) abordou um conjunto de 32 redações nota 1000 entre os anos de 2014 e 2017. A autora atentou-se para marcas estruturais no texto, as quais são entendidas desta forma: “[...] o modo como o texto é composto e dividido, ou seja, toda a sua teia de escolhas, desencadeamentos e composição (...)”. Mais especificamente, olhou

para a presença ou não de título; quantidade de parágrafos; quantidade de períodos por parágrafos; sequência das ideias, evidenciando de que modo a tese, os argumentos e a proposta de intervenção foram organizados; presença ou não de discurso direto ou indireto; e a forma como os recursos linguísticos expressam esses discursos. (PAIVA, 2020, p. 54.)

Ainda que a autora tenha comentado sobre elementos linguísticos recorrentes, como o uso de verbos modais, o texto não enfoca aspectos quantitativos, como pretendemos. Assim, podemos dizer que as pesquisas mais próximas àquela que aqui desenvolvemos são de Pereira (2018) e Silva (2018). Com um *corpus* bem menor (apenas 16 redações), Pereira (2018) analisou elementos estruturais presentes em redações nota 1000 referentes ao ano de 2016, por meio do aplicativo Tropes. Paralelamente, Silva (2018) descreveu diferenças importantes em relação aos textos nota 1000 (verificados por Pereira, 2018) e aqueles considerados medianos, como o menor número de modalizadores ou conectivos. Sua análise, por meio do mesmo software, foi feita a partir de um *corpus* de textos de um curso preparatório.²

Nesse contexto, este trabalho tem como objetivo central descrever as características linguísticas recorrentes da redação do Enem. Para isso, selecionamos 95 textos de anos distintos, conforme sua disponibilidade: 20 (das 250) redações nota 1000 do processo de 2014, todas elas

² Seguindo a correta orientação de um parecerista anônimo, não faremos uma comparação em relação ao texto de Silva (2018), uma vez que o objetivo não é comparar nossa descrição com a análise da autora, cujo *corpus* é diverso daquele analisado na presente pesquisa (ela analisa textos considerados mediados, com notas abaixo de 700). Por isso, deixamos ao leitor interessado no tema a referência para leitura adicional.

divulgadas pela imprensa; 31 (dos 55) do Exame de 2018 e 44 (das 53) de 2019, divulgadas por Felpi (2019, 2020, respectivamente). Após essa coleta, utilizamos o Tropes, uma ferramenta computacional que faz uma caracterização linguística de textos, baseada especialmente nas ocorrências lexicais, já utilizada em outros trabalhos para análise de textos de gêneros distintos (ARAÚJO, 2017; BERTUCCI, 2020)

Entendemos que a utilização de ferramentas digitais que favoreçam a coleta e a análise de dados é essencial para a pesquisa empírica. É válido dizer que a Linguística de *Corpus* vem aproveitando bem o surgimento de novas tecnologias com o intuito de gravar, armazenar e analisar dados (RASO; MELO, 2012; SARDINHA, 2000). Nesse âmbito, pode-se dizer que houve avanços na descrição de gêneros textuais por meio de recursos computacionais, algo que tem contribuído muito para o entendimento das línguas, tal como se vê em Finatto (2017) – e nos artigos que fazem parte daquele número temático. Apesar disso, muito há ainda por ser feito, sendo importante destacar que o uso de ferramentas como o Tropes em ambientes de ensino e/ou pesquisa pode ajudar.

Por isso, o presente trabalho pode contribuir para pelo menos duas frentes distintas e complementares: a primeira, os estudos de *corpora*, na tentativa de preencher as lacunas desse tipo de trabalho em especial com gêneros escolares e com o uso de ferramentas digitais no Brasil; a segunda, o ensino, uma vez que, bem delineadas as propriedades do gênero em questão, os profissionais da área podem se beneficiar de métodos diversos de análise desse tipo de texto – ou ao menos, como indicamos antes, para um debate acerca da padronização do gênero.

2 Análise de gêneros e tecnologia

2.1 Linguagem e tecnologia

A discussão sobre a importância da tecnologia em nossas vidas vem sendo posta em diferentes perspectivas de análise. Isso certamente ocorre porque, como nos aponta Cupani (2016, p. 9) “a tecnologia nos afeta e desafia qualquer que seja a nossa atividade”. E afeta porque a ela estão relacionadas diferentes práticas cotidianas, inclusive as linguísticas. Na mesma direção, o mesmo autor aponta que as técnicas são “procedimentos sujeito a regras” e, por isso, passíveis de serem aprendidas; são, ao mesmo tempo, “manifestações da capacidade humana

de fazer coisas” (CUPANI, 2016, p. 15). Consequentemente, os produtos (artefatos) decorrem do saber-fazer do homem, que modifica o seu meio também artisticamente, porque está carregado de criatividade para concepção do processo e para chegar ao resultado planejado. Dessa forma, tal procedimento está imbuído de regras relacionadas ao saber-fazer.

Nesse ponto, em que se concebe a tecnologia em relação direta com o planejamento, destaca-se o papel da linguagem. Vieira Pinto (2005) considera a linguagem um requisito essencial para planejamento por seu caráter simbólico: foi por isso que ela permitiu avançar no momento presente e estabelecer possibilidades de mudanças da realidade para além das necessidades imediatas. Para o autor, não fosse isso, o desenvolvimento de tecnologias associadas ao conhecimento, seria mínimo ou até mesmo impossível. Assim, a linguagem também preenche o requisito essencial para o desenvolvimento da própria tecnologia. Para Cassirer (1994, p. 47-48, grifos do autor) o caráter simbólico é a chave para se entender o próprio ser humano:

Entre o sistema receptor e o sistema de reação, que se encontram em todas as espécies animais, encontramos no homem um terceiro elo, que podemos descrever como o *sistema simbólico*; esta nova aquisição transforma toda a vida humana. Em confronto com os outros animais, o homem não vive apenas numa realidade mais vasta; vive, por assim dizer, numa nova dimensão da realidade.

Como se lê acima, o sistema simbólico, ofereceu ao homem condições claras de diferenciação dos outros animais, mas, sobretudo, de superar dimensões mais próximas e desejar, planejar e até mesmo realizar outras, trazendo-as à dimensão real. A linguagem, portanto, exerceu um papel decisivo, por ser modeladora dessa capacidade.

Pode-se dizer que é graças a isso que as inovações modificam o ambiente humano. Barton e Lee (2015) consideram que toda mudança tecnológica causa alguma mudança na vida social; ou seja, à medida que novas técnicas, processos e produtos aparecem, a vida das pessoas se modifica. Como a linguagem é parte essencial no processo, não podemos desconsiderar a relação plena entre texto/linguagem e tecnologias: linguagem modificando práticas sociais (e tecnológicas) e tecnologias modificando/influenciando práticas linguísticas.

Com isso, é possível se falar sempre em uma valoração da tecnologia, seus processos e produtos. É por isso que Cupani (2016,

p. 12) argumenta que “aquilo que denominamos tecnologia se apresenta, pois, como uma realidade polifacetada: não apenas em forma de objetos e conjuntos de objetos, mas também como sistemas, como processos, como modo de proceder, como uma certa localidade”. Em outras palavras, se o mundo está cada vez mais permeado de tecnologia e a linguagem é a grande mediadora da sociedade com essa realidade, lança-se o desafio aos estudiosos da área de identificarem objetos, processos e modos de proceder na coleta e análise de dados.

Nesse contexto, a escola é um espaço propício para a discussão da ciência e da tecnologia, pois tenta apresentar aos alunos de que forma o homem chegou aos conhecimentos que transformaram criativamente seu meio ao longo do tempo. Isso também se aplica aos gêneros textuais: espera-se que as aulas de línguas levem o aluno a compreender e a aplicar um “saber-fazer” específico para cada situação de interação linguística por meio de gêneros. Quanto ao objeto desta pesquisa, há inclusive uma série de aulas, vídeos e livros que apresentam supostas técnicas aos alunos para conquistar a (quase impossível!) nota 1000 (ALVES; BESSA, 2018).

Por isso, é inconcebível se pensar em uma discussão de linguagem e tecnologia sem se associar também a capacidade do homem de pensar sobre o próprio fazer da linguagem, ou seja, sem desenvolver elementos metalinguísticos específicos para ampliação das suas esferas de saber (AUROUX, 2014). Dessa forma, os gêneros podem ser compreendidos como parte fundamental dos processos de ensino-aprendizagem de língua, os instrumentos pelos quais as trocas linguísticas se realizam.

Acrescenta-se a essa reflexão que a possibilidade de analisar os gêneros por meio de ferramentas tecnológicas é uma das maneiras de contribuir para o entendimento das línguas naturais e dos processos relacionados à escrita. Isso nos leva a concluir que pesquisas de *corpus* são essenciais para se atingir esse objetivo.

Quando tratamos das pesquisas em *corpora*, podemos observar que o uso de tecnologias é um fato; afinal, ferramentas que contribuem para coleta, armazenamento e análise de *corpus* de grande extensão têm beneficiado as pesquisas, em especial com relação a textos escritos no Brasil (RASO; MELO, 2012). Ao mesmo tempo, o uso da internet tem facilitado também o acesso a gêneros que poderiam ficar restritos a seu ambiente de produção, como é o caso da redação do Enem. Assim, ações como de Felpi (2019, 2020), que compilou um bom número das redações nota 1000 do Enem de 2018 e 2019, contribuem para que

pesquisadores e estudantes analisem padrões de textos caracterizados de forma homogênea em uma situação, como é o caso da redação nota 1000.

O presente trabalho segue a tendência de análise com *corpora* de textos escritos e para estudos relacionados ao gênero, com a possibilidade de estudos relativos a *corpora* com propósitos específicos. Com isso, acreditamos que a utilização de textos produzidos em ambientes formais e monitorados (como o exame do Enem) pode fornecer elementos importantes para o estudo da língua, seja para a análise do discurso, o estudo de texto/gêneros, seja para a variação linguística, entre outros.

Nesse contexto, Finatto (2017) destaca a importância do apoio computacional quer para a formação de *corpora* de diferentes modos e fontes, quer para sua análise textual, em especial para a descrição de gêneros textuais/discursivos. Concordamos com a autora, pois tal apoio, especialmente para os linguistas, mostra que a tecnologia é um meio que ajuda a explicar diferentes fenômenos da língua natural. Por isso, este artigo quer colaborar para enfatizar o apoio dado de recursos de tecnologia na área da linguagem, em especial dos gêneros textuais: a partir da coleta de dados, vamos analisar o *corpus* por meio do Tropes. Nesse sentido, destaca-se que a contribuição dos estudos com *corpora* é oferecer aos pesquisadores a possibilidade de fazer análise de dados reais e, para tal, a aplicação de recursos de coleta, armazenamento e análise de dados é importante (SARDINHA, 2000).

Sobre a extensão de nosso *corpus*, nos baseamos na ideia de representatividade exposta em Sardinha (2000). Para o autor, o tamanho do *corpus* é menos importante que os dados, em si, que devem ser agrupados conforme critérios do pesquisador. Nesse aspecto, o que é de fato necessário é sua representatividade linguística em algum contexto. Logo, não se deve definir um tamanho mínimo para um *corpus*, ainda que seja fundamental o fato de ser entendido como suficientemente representativo. Além disso, embora não tenha extensão pré-definida, a representatividade é, logicamente, melhor caracterizada quanto maior for o *corpus*. Quando específicos (e não abertos), os *corpora* são de acesso exclusivo do pesquisador, que os produzem para uma finalidade pré-determinada, não sendo disponíveis para outros pesquisadores e, conseqüentemente, acabam não sendo “verificáveis, o que compromete a pesquisa em termos de sua replicabilidade e generabilidade” (SARDINHA 2000, p. 348). Acreditamos que os 95 textos do nosso *corpus* sejam suficientemente representativos para uma compreensão

geral do gênero em questão, inclusive porque, só do ano de 2019, há 44 textos no *corpus*, dos 53 possíveis.

De qualquer modo, se as ferramentas tecnológicas são meio para análise dos gêneros, poderíamos pensar, igualmente, que os próprios gêneros (em especial os escritos) são também uma tecnologia. Se tomarmos a perspectiva da língua humana como “instrumento” de interação e ação do homem no mundo, podemos atribuir-lhe um caráter tecnológico: como instrumento, a língua é um meio para se alcançar um determinado objetivo, exigindo dos falantes um amplo conhecimento de suas capacidades para cumprir com suas funções.

Nesse contexto, assumimos que os gêneros escritos são um tipo de tecnologia, não só por exigirem a escrita (também uma tecnologia, como sustenta Auroux (2014)), como pelo caráter de planificação (próprio da tecnologia, como defende Vieira Pinto (2005)), uma vez que pesquisas têm demonstrado que o planejamento é essencial para construção de textos escritos (CABRAL, 2013; OLIVEIRA, 2018).

A respeito disso, Campos e Ribeiro (2013, p. 25) afirmam que a natureza do gênero é ser

[...] um instrumento semiótico que é apropriado por sujeitos, os quais, uma vez instrumentalizados, podem adquirir novos conhecimentos e saberes. Em outras palavras, instrumentalizado em alguns gêneros, o sujeito é capaz de alçar voo de forma independente porque desenvolve habilidades que lhe permitem usar a língua nas mais variadas práticas sociais, que são mediadas pelos gêneros.

Para tal instrumentalização, a escola tem um papel essencial, já que deveria possibilitar aos estudantes refletirem sobre o fazer linguístico em situações diversas, especialmente aquelas de maior complexidade, que exigem graus de letramento mais aprofundados (como a escrita). Por isso, Wachowicz (2010) defende que cabe à escola promover práticas em que os gêneros sejam os instrumentos por meio dos quais os alunos se relacionem com a cultura letrada. Por exemplo, o gênero artigo de opinião de um jornal é o instrumento socialmente elaborado que o jornal utiliza para estabelecer sua relação com o leitor; caberia, assim, à escola instrumentalizar o aluno para a ação sobre esse gênero, seja na leitura, seja na produção, seja na interação com o jornal que produziu esse artigo. Da mesma maneira que compreendemos a técnica da fabricação

de artefatos como mediação das ações humanas no mundo do trabalho, e tecnologia como um pensar sobre essas ações, tornando-as mais eficientes, compreendemos os gêneros como mediadores da produção de conteúdo linguístico, que permitem aos homens agir no mundo letrado. Acreditamos, portanto, que a relação entre linguagem e tecnologia se dá na medida em que ambas se utilizam do aprimoramento de instrumentos para se concretizarem como práticas sociais.

2.2 O gênero *redação do Enem*

A redação do Enem é chamada de texto dissertativo-argumentativo, uma mescla de tipos textuais. Cantarin *et al.* (2017, p. 83) consideram que dissertar é “[...] fazer uma reflexão teórica sobre um assunto”, o que pressupõe que o autor lance mão de seu repertório sociocultural e utilize recursos de diferentes naturezas na exposição. Já a argumentação, numa perspectiva retórica, por exemplo (FIORIN, 2017; PERELMAN; OLBRECHTS-TYTECA, 2014), está ligada à capacidade de convencer um determinado auditório, a partir de relações possíveis, prováveis e plausíveis (e não por demonstrações lógicas, por exemplo). O candidato usa o seu conhecimento para expor uma questão; mas, ao articular as informações escolhidas, acaba por construir um argumento em torno de um ponto de vista.

Em textos argumentativos, como o que analisamos nesta pesquisa, o planejamento pode contribuir para o projeto do autor na defesa de um ponto de vista, já que um produtor “instrumentalizado” no gênero seria capaz de traçar um caminho argumentativo claro em seu texto. Coroa (2017, p. 52) sugere que a escrita desse tipo de texto segue um mapeamento por parte do autor: “a cada marca ou pista, ele avança, recua ou reorienta seu caminho.” No caso dos textos nota 1000, considerando a avaliação criteriosa, o planejamento se revela na completa coerência do texto com relação ao ponto de vista defendido, aos argumentos que o sustentam e às propostas de intervenção sugeridas pelo candidato. Assim é essa materialização que se realiza no gênero e que propomos pode ser de alguma forma descrita pelo Tropes.

Ao tratar dos tipos textuais, Garcez (2017a, p. 45) ressalta que eles “se definem pela natureza linguística intrínseca de sua composição. As escolhas lexicais, os aspectos sintáticos, o emprego de tempos verbais, as relações lógicas estabelecidas definem o tipo textual.” Por isso,

levantamos a hipótese de que haja recorrência de elementos linguísticos nos textos prototípicos do Enem (nota 1000).

De início, um ponto importante a ser reforçado é o fato de o texto do Enem ser considerado um gênero misto (dissertativo-argumentativo), o que significa dizer que a argumentação, ali, exige uma exposição do tema – como se disse, é aqui que o repertório vai contribuir para a dissertação. Coroa (2017) e Cantarin *et al.* (2017) mostram que a dissertação, entendida como exposição ou aprofundamento do tema, é essencial no texto do Enem. Nesse sentido, a redação também requer uma descrição de diferentes aspectos e relativamente aprofundada do tema, de tal forma que os argumentos se sustentem nesse contexto. De modo bem específico sobre a exposição, Cantarin *et al.* (2017, p. 85) afirmam que

linguisticamente, há elementos básicos para a exposição, como a utilização de verbos no presente do indicativo (“exercem”, “surge”, “é”), especialmente os de ligação (“ser” e “estar”, por exemplo), e marcadores de qualidade e classe, como adjetivos (“diferentes”, “grande”, “infantil”) e substantivos (“processo”, “poder”, “hábitos”), entre outros elementos.

O uso de verbos estativos (“ser” e “estar”, por exemplo) ou de adjetivos, por exemplo, são importantes nesse gênero e devem indicar que o texto está dissertando sobre o tema, e ao mesmo tempo, argumentando, pela escolha de tais itens, conforme sugerem os mesmos autores.

Passando para a argumentação, é preciso dizer que o tipo argumentativo é bastante complexo. Coroa (2017), ao comparar a possibilidade de caracterização dos tipos, sustenta que as características de organização textual são de difícil identificação para o tipo argumentativo, embora recorra,

com muita frequência às relações lógicas para demonstrar a verdade daquilo que diz, como as de causa e consequência e as de condição. Comumente tais relações são expressas por conectivos de finalidade, de causa, de justificativa, como em “por causa das múltiplas realidades” ou em “a fim de que, não importa quais sejam os resultados de nossos estudos, nosso compromisso”. Mas também a seleção lexical, como “complexa”, “múltiplas realidades”, “elementos significativos”, deve estar a serviço do objetivo. (COROA, 2017, p. 61.)

Como se vê, a autora indica aspectos importantes para a caracterização do tipo argumentativo no texto do Enem, em especial o cuidado com as relações lógicas e a seleção lexical, que pode indicar o repertório do candidato. Para Peixoto (2017, p. 163), são exatamente os conectivos os elementos responsáveis pela introdução da argumentação no texto: “argumentos com diferentes forças e orientações discursivas são colocados na cena textual pelos chamados operadores e conectivos argumentativos.”

Outro fator importante é a análise com relação aos modalizadores. Autores como Koch (2011) defendem que esses elementos são formas importantes de mostrar um posicionamento do autor em relação ao conteúdo que expressa. Por isso, além de verbos modais, como “dever” e “poder”, por exemplo, é importante que se verifique a ocorrência de outros marcadores de modalidade no texto, em especial advérbios. Neste trabalho, procuramos descrever a frequência de modalizadores presentes no *corpus*, imaginando que muitos deles serão aqueles adequados para caracterizar a argumentação no texto (uma vez que é prototípico do gênero).

Um último ponto de destaque vem da (im)personalidade no gênero sob análise. Garcez (2017b, p. 276, grifos nossos) afirma que “o Enem, tradicionalmente, exige um texto dissertativo-argumentativo, ou seja, aquele que apresenta ideias e informações e é impessoal (**preferencialmente deve-se evitar a primeira pessoa**) (...). Um bom modelo desse gênero é o editorial jornalístico”. Essa “preferência” pela impessoalidade não significa, no entanto, que o candidato que usar primeira pessoa possa ser prejudicado pela escolha de registro, como afirmam Sandoval *et al.* (2017). Ainda assim, o uso de pronomes de alguma pessoa que não a terceira é um fator a ser investigado também neste trabalho.

Tudo isso vai ao encontro do que se lê em uma cartilha do participante do Enem (BRASIL, 2019), a qual detalha que o candidato precisa defender um ponto de vista com argumentos consistentes e elencar uma proposta de intervenção para o problema discutido. Concluímos que o conjunto de textos aqui analisados cumpriu de modo exemplar essas indicações, já que foram avaliados com a nota máxima pela banca de correção.

O mesmo material detalha aos candidatos as cinco competências avaliadas no processo, indicando inclusive uma matriz de referência de notas, em que se apontam os motivos pelos quais os avaliadores atribuem

notas de 0 a 200 para cada competência. O Quadro 2 a seguir resume os objetivos de cada competência e a descrição do modo como se avalia um texto com nota máxima em cada uma delas.

QUADRO 2 – Competências avaliadas na redação do Enem

Competência	Objetivo	Para nota máxima (200 pontos)
1	Demonstrar domínio da modalidade escrita formal da língua portuguesa.	Demonstra excelente domínio da modalidade escrita formal da língua portuguesa e de escolha de registro. Desvios gramaticais ou de convenções da escrita serão aceitos somente como excepcionalidade e quando não caracterizarem reincidência.
2	Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo em prosa.	Desenvolve o tema por meio de argumentação consistente, a partir de um repertório sociocultural produtivo, e apresenta excelente domínio do texto dissertativo-argumentativo.
3	Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista.	Apresenta informações, fatos e opiniões relacionados ao tema proposto, de forma consistente e organizada, configurando autoria, em defesa de um ponto de vista.
4	Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação.	Articula bem as partes do texto e apresenta repertório diversificado de recursos coesivos
5	Elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos.	Elabora muito bem proposta de intervenção, detalhada, relacionada ao tema e articulada à discussão desenvolvida no texto.

Fonte: Adaptado de Brasil (2019)

Novamente, podemos sustentar que o conjunto em análise na presente pesquisa é composto de textos que cumpriram com excelência todos os requisitos indicados pelas competências. Sendo exemplares prototípicos, levantamos a hipótese de que devem ser formados por itens recorrentes, os quais podem estar intimamente ligados às características do gênero textual em questão. É o que pretendemos verificar a partir da próxima seção, em que apresentamos a metodologia e análise dos dados. Vamos descrever a ferramenta Tropes, os textos selecionados e os resultados obtidos.

3 Metodologia e análise dos dados

3.1 Dados, ferramenta e hipóteses

Para esta pesquisa, selecionamos um conjunto de 95 textos nota 1000 do Enem, sendo divididos desta forma: 20 textos do exame de 2014, disponíveis na imprensa;³ 31 redações de 2018, disponibilizados por Felpi (2019); e 44 redações de 2019, disponibilizadas por Felpi (2020). O objetivo era verificar os elementos recorrentes no conjunto que pudessem descrever o gênero redação do Enem.

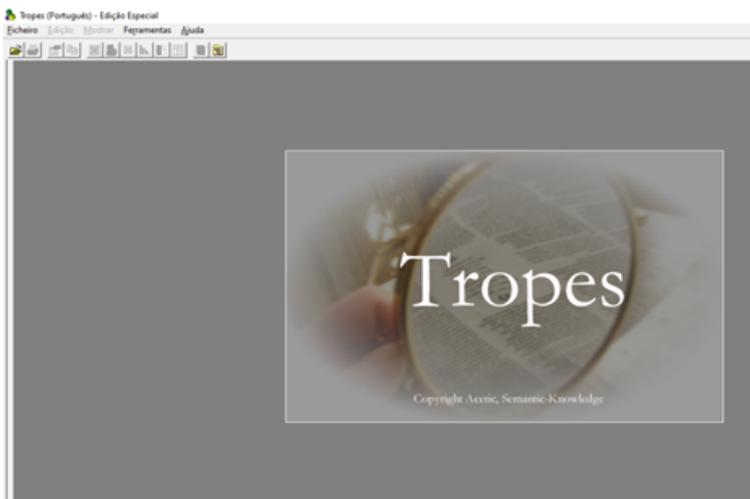
A análise foi realizada por meio do Tropes, uma ferramenta de análise textual em versão gratuita. Criado nos anos 90, pela Semantic-Knowledge (Acetic), com sede em Paris, o Tropes é uma das ferramentas que a empresa desenvolveu para análise textual. Por meio da incorporação de conhecimentos advindos da área de Processamento de Linguagem Natural, o objetivo do *software* era servir para estudos de áreas diversas, como de Sistemas de Informação, Sociologia e Linguística, por exemplo. Araújo (2017, p. 300) observa que

[este] *software* destaca-se pelo processamento semântico de textos em línguas naturais. Para descrever as características dos enunciados em análise, o Tropes 7.2.3 vale-se de critérios linguísticos pré-programados e os associa às estruturas linguísticas encontradas nos textos processados.

³ Entre os exemplos está o portal G1: Redações nota 1000: Disponível em: <<http://g1.globo.com/educacao/enem/2015/noticia/2015/05/leia-redacoes-do-enem-que-tiraram-nota-maxima-no-exame-de-2014.html>> Acesso em: 17 jul. 2018.

Por suas características de análise lexical, o programa faz um processamento do qual decorrem informações como o “estilo textual”, o contexto básico relativo ao texto, denominado de “Universo de referências” (sendo uma espécie de mapeamento do repertório mobilizado no texto), e dados quantitativos referentes a classes lexicais. A Figura 1 apresenta a tela de abertura da ferramenta.

FIGURA 1 – Tela inicial do Tropes



Fonte: Tropes.

Neste trabalho, vamos enfatizar dois pontos: o universo de referência, pela relação com o repertório e o tema; e a frequência de classes lexicais, que tem relação com os elementos linguísticos que constituem o gênero. No primeiro caso, o software relaciona palavras de mesmo campo semântico, criando conjuntos de referência. Para a referência de “sociedade”, por exemplo, ele agrupa expressões como *família*, *qualidade de vida* e *população* entre muitas outras. Nesse caso, acreditamos que o aplicativo poderá indicar o universo mais próximo do tema proposto bem como o repertório utilizado nas redações.⁴

⁴ O aplicativo oferece subdivisões do Universo de referência; aqui, apresentaremos a mais geral, intitulada “Universo de referência 1”.

No segundo caso, verificaremos a ocorrência das classes. Entre os conectivos, o Tropes divide a classe em *condição, causa, finalidade (escopo), adição, disjunção, oposição, comparação, tempo e lugar*; os modalizadores são divididos em *tempo, lugar, modo, afirmação, negação, dúvida e intensidade*; os verbos em *factivos, estativos, declarativos e performativos*; os adjetivos entre *subjetivos, objetivos e numéricos*; os pronomes entre *primeira, segunda, terceira pessoa* e o clítico *se*.

Considerando as funcionalidades do *software*⁵ e as características da redação do Enem, apresentadas anteriormente, levantamos as seguintes hipóteses a serem testadas no conjunto de dados por meio do Tropes:

1. Espera-se um volume grande de repertório (universo de referência) mobilizado pelos autores, o que favorece a exposição e, articulado, leva à argumentação – Cantarin *et al.* (2017).
2. Haverá uma variedade nos tipos de conectivos, já que as relações presentes no texto argumentativo os exigem – Coroa (2017) e Peixoto (2017).
3. Haverá variedade de modalizadores, com enfoque para aquelas classes que contribuem mais para processos argumentativos, como os de modo e intensidade – Koch (2011).
4. Espera-se uma alta frequência de estativos pelo caráter de exposição (*ser* e *estar*, por exemplo) e modalização (*poder* e *dever*, por exemplo) – Cantarin *et al.* (2017).
5. Haverá alta frequência de adjetivos pela exigência da dissertação/descrição – Cantarin *et al.* (2017).
6. Espera-se um predomínio da terceira pessoa – Garcez (2017b).

A seguir, apresentamos os resultados e verificamos se as hipóteses foram corroboradas ou não pelos dados analisados no *corpus* com as 95 redações nota 1000.

⁵ Para mais informações sobre o aplicativo, sugerimos Araújo (2017) e Bertucci (2020).

3.2 Resultados e análise

3.2.1 Universo de referência/repertório

Os primeiros dados que colhemos do conjunto de redações diz respeito ao repertório utilizado nos textos. A especificação “Universo de referência” no aplicativo apresentou os seguintes resultados, conforme os Quadros 3 a 5.

QUADRO 3 – Principais classes no universo de referência (2014)

Referência/Classe	Ocorrências	Exemplos
vida_humana	475	idade, crianças, propaganda
conceitos_gerais	377	padrões, necessidade, produto
negócios	147	consumo, publicidade, mercado
sociedade	78	pais, filhos, família

Fonte: Elaboração própria com dados do Tropes.

QUADRO 4 – Principais classes no universo de referência (2018)

Referência/Classe	Ocorrências	Exemplos
vida_humana	761	liberdade, indivíduo, manipulação
conceitos_gerais	691	comportamento, conteúdo, perfil
comunicação_e_mídia	170	linguagem, internet, comunicação
educação_e_ensino	79	educação, ensino, escola

Fonte: Elaboração própria com dados do Tropes.

QUADRO 5 – Principais classes no universo de referência (2019)

Referência/Classe	Ocorrências	Exemplos
conceitos_gerais	1278	inclusão, desenvolvimento, acesso
vida_humana	757	poder, conhecimento, construção
arte_e_cultura	557	cinema, cultura, arte
sociedade	217	classes sociais, sociedade, população

Fonte: Elaboração própria com dados do Tropes.

Como se percebe nos dados, os itens ‘vida_humana’ e ‘conceitos_gerais’ são os mais frequentes nos três conjuntos de redações nota 1000 analisados. Isso parece indicar que os textos tratam de um conjunto de elementos relativos à sociedade (o que pode ser reforçado pela presença do item “sociedade” em dois conjuntos), cuja abordagem é de caráter teórico/conceitual. É uma sugestão de leitura dos dados e que, na prática, pode ser confirmada por uma análise qualitativa adicional. Outro ponto de destaque é que o universo de referência pode ser considerado amplo nos três casos: verificamos que, em média, 35 conjuntos de relações no universo de referência foram mobilizados nas redações. Sem dúvida, o ideal seria uma análise individual de cada redação para uma média por redação, tal como fez Pereira (2018). Analisando redações nota 1000, ela observou um total de 33 classes no universo de referência em seu *corpus* – o que se aproxima do nosso número. Além disso, verificou que, em média, cada candidato mobilizou cerca de 16,5 universos distintos. No nosso caso, quando se observam os dados de forma mais específica, verificam-se distintas categorias tais como *filosofia*, *história*, *economia*, *ciência*, entre muitas outras, o que revela um amplo repertório mobilizado. Esse é um ponto que parece ser decisivo para a categorização das redações nota 1000.

Com isso, acreditamos que a Hipótese 1, relativa ao repertório, foi parcialmente corroborada pelos dados – mas apenas uma análise individual das redações poderia dar uma noção mais precisa. A variedade observada favorece, portanto, a exposição, um dos elementos essenciais para a construção do gênero. Vale acrescentar ainda que a frequência observada nos Quadros 3 a 5 sugere uma relação próxima dos textos com a temática proposta, ou seja, os candidatos se mantiveram no escopo do tema solicitado: em 2014, o tema foi “A publicidade infantil em questão no Brasil”, por isso *negócios* e *sociedade* aparecem também entre os itens mais frequentes do universo de referência; em 2018, o tema solicitado foi “A manipulação do comportamento do usuário pelo controle de dados da internet”, por isso *comunicação_e_mídia* foi bastante recorrente; finalmente, o tema de 2019 foi “Democratização do acesso ao cinema no Brasil”, o que explica a recorrência do item *arte_e_cultura*.

Finalmente, cabe a observação de que a mobilização de um amplo repertório parece estar ligada, primeiramente, às articulações expositivas no texto. Para Coroa (2017), o tipo dissertativo privilegia a exposição de uma ideia, enquanto, enquanto o argumentativo, o modo como se

defende um ponto de vista. Logo, no tipo dissertativo-argumentativo, essa exposição é mobilizada em favor de uma argumentação. Nesse sentido, o repertório estaria ligado à noção de legitimação e produtividade que aparecem nos materiais referentes à avaliação da redação do Enem (INEP, 2020). Assim, como os dados apontam para uma recorrência do fato, é possível que os candidatos treinem essa mobilização de diversos universos de referência com o objetivo de dar credibilidade à sua redação dissertativo-argumentativa.

3.2.2 Conectivos

A Tabela 1 resume o resultado obtido com a análise das 95 redações nota 1000.

TABELA 1 – Ocorrência de conectivos

Conectivo	2014		2018		2019		Geral	
	Total	%	Total	%	Total	%	Total	%
Adição	190	59	294	58,3	361	51,5	845	55,4
Causa	8	2,5	33	6,6	49	6,9	90	6
Comparação	22	6,9	49	9,7	97	13,9	168	11
Condição	30	9,4	42	8,4	78	11,1	150	9,9
Disjunção	15	4,6	16	3,2	23	3,3	54	3,5
Escopo (fim)	15	4,6	31	6,2	21	3	67	4,4
Lugar	2	0,6	0	0	8	1,2	10	0,6
Oposição	27	8,4	30	6	52	7,4	109	7,1
Tempo	12	3,7	8	1,6	12	1,8	32	2,1
TOTAL	321	100	503	100	701	100	1525	100

Fonte: Elaboração própria com dados do Tropes.

Os dados revelam que a categoria de adição é a mais frequente, com mais de 55% das ocorrências entre os conectivos. Isso é fruto, certamente, da alta frequência da conjunção em português, em especial pela possibilidade de ser usada em posição interior ao sintagma e não

apenas para ligação de orações, períodos ou parágrafos.⁶ Outras conjunções que merecem destaque são as de comparação (11%), de condição (9,9%), oposição (7,1%) e causa (6%), exemplificadas respectivamente por *como*, *mesmo que*, *no entanto* e *uma vez que*. Todas elas são essenciais para a construção do teor argumentativo, para o direcionamento que o autor pretende dar no texto. Além disso, cabe ressaltar ainda dois pontos: primeiro, que todas as categorias foram contempladas na análise *corpus*, sendo as de lugar e tempo as menos frequentes – o que é de se esperar em textos de teor argumentativo. De fato, conforme observa Coroa (2017), espera-se mais elementos de referência temporal na narração e de referência espacial na descrição. Em segundo lugar, que, dado o total de 1.525 conectivos, tem-se uma média de 16 categorias distintas para cada uma das 95 redações analisadas. Pereira (2018) havia observado um total de 308 conectivos nas 16 redações analisadas, o que dá uma média próxima de 19 por texto. Nesse caso, também se observou uma alta frequência do conectivo de adição (acima de 54%) e porcentagens similares àquelas encontradas aqui para as outras categorias.

Isso parece indicar que a quantidade de conectivos no texto é tão relevante quanto sua variedade. E, de fato, isso está de acordo os materiais sobre a Competência 4 para avaliadores (GARCEZ; CORRÊA, 2017; INEP, 2020), que orientam sobre a diversidade e sobre a frequência dos elementos de coesão (como os conectivos) na redação do Enem. Em geral, tais materiais sugerem que uma redação do Enem deve apresentar pelo menos dois conectivos interparágrafos e pelo menos um intraparágrafos, o que já poderia contabilizaria um número de 6 conectivos (no mínimo) em uma redação de 4 parágrafos. Mas, quando se olha para a especificação das competências (INEP, 2020), percebe-se uma orientação para atribuição de nota máxima apenas aos casos de presença expressiva, com adequação e sem repetições, de elementos coesivos, o que deve explicar a recorrência nos textos nota 1000. Neste caso, reforçamos, como se trata de textos avaliados com a nota máxima, pode-se dizer que a quantidade e a qualidade do emprego foram importantes para a argumentação, sobretudo pela necessidade do uso dos conectivos para a

⁶ Uma rápida pesquisa no site “Corpus do Português”, com a entrada “e”, devolveu mais de 34 milhões de ocorrências no termo (num *corpus* de mais de 1 bilhão de palavras). A título de comparação, a conjunção “mas” apareceu em mais de 3 milhões de casos, o que corresponde a apenas 10% de “e”.

construção da argumentação. Portanto, podemos concluir que os dados corroboram a Hipótese 2, sobre a variedade com relação aos tipos de conectivos exigidos no gênero.

3.2.3 Modalizadores

A Tabela 2 resume o resultado sobre os modalizadores.

TABELA 2 – Ocorrência de modalizadores

Modalizador	2014		2018		2019		Geral	
	Total	%	Total	%	Total	%	Total	%
Afirmção	08	3	12	3,2	19	3	39	3
Dúvida	02	0,7	2	0,5	0	0	04	0,3
Intensidade	91	33,7	98	26,2	238	35	427	32,3
Lugar	07	2,6	6	1,5	17	2,6	30	2,3
Modo	65	24,1	142	38,1	204	30	411	31,1
Negação	52	19,3	44	12,4	107	15,5	203	15,4
Tempo	45	16,6	66	18,1	96	14	207	15,6
TOTAL	270	100	370	100	681	100	1321	100

Fonte: Elaboração própria com dados do Tropes.

Os dados mostram um fenômeno interessante: modalizadores de modo (como advérbios terminados em *-mente*) e intensidade (como *sobretudo*) predominam no tipo de texto em análise, revelando um posicionamento do autor em relação ao fato apresentado. Ainda que não esperada, a alta frequência da categoria de negação é explicada pelo amplo uso de *não* nos textos, outra palavra frequente na língua (quase 10 milhões indicadas no site “Corpus do Português”). Igualmente, a frequência da categoria de tempo parece estar relacionada com o uso de advérbios como *atualmente* e *hoje*. No entanto, uma pesquisa mais específica, com cada uma das redações, seria necessária para uma explicação mais minuciosa.

Com relação aos dados totais, temos uma média de 13,9 modalizadores para cada uma das 95 redações analisadas, um número

próximo dos 15,7 encontrados por Pereira (2018). Novamente, percebe-se que a maior frequência aliada a uma variedade de modalizadores parece constituir o gênero do ponto de vista da sua prototipicidade. Assim, tais dados revelam que os modalizadores, além de indicarem a capacidade do candidato em articular elementos linguísticos importantes para a construção da sua argumentação, se constituem como indícios de autoria importantes no texto dissertativo-argumentativo: é por meio do uso de elementos como esses que o autor modaliza as vozes/discursos que apresenta no texto (COSTA; GUEDES, 2017). Portanto, os dados apresentados corroboram para a Hipótese 3 a respeito da variedade de modalizadores, com enfoque para modo e intensidade.

De modo adicional, podemos dizer que o emprego dos modalizadores estaria ligado mais ao caráter argumentativo que dissertativo, na diferenciação dos tipos proposta por Coroa (2017). São esses elementos que vão contribuir para que o texto tenha um caráter de defesa de ponto de vista sobre o tema, a partir das informações e fatos mobilizados pelos candidatos. Ou seja, são eles que contribuem para a formação das estratégias argumentativas que sustentarão o texto.

3.2.4 Verbos

Na Tabela 3 são apresentados os dados relativos à frequência dos tipos de verbos.

TABELA 3 – Frequência de verbos

Verbo	2014		2018		2019		Geral	
	Total	%	Total	%	Total	%	Total	%
Factivo	727	69,2	1225	73	1642	68,9	3594	70,3
Estativo	302	28,7	418	24,9	689	28,9	1409	27,5
Outros	23	2,1	36	2,2	52	2,2	111	2,2
TOTAL	1.052	100	1679	100	2383	100	5114	100

Fonte: Elaboração Própria com dados do Tropes.

Os números relativos ao *corpus* em análise demonstram que os verbos classificados como *factivos* (verbos de ação) são predominantes. Ainda que o índice de 27,5% de estativos seja significativo, o resultado

geral parece ir na contramão da Hipótese 4, que afirmava sobre a alta frequência dos estativos. Pereira (2018), que também encontrou uma alta frequência de factivos em seu *corpus*, recorre a Araújo e Cunha (2009, p. 33-34) para explicar a relação dessa classe de verbos com a argumentação: os autores consideram que essa categoria “é aquela em que um sujeito intencional e animado realiza uma ação e essa ação afeta (ou efetua) um objeto paciente”. Para Pereira (2018), essa ação com intencionalidade que afeta um paciente é percebida no encadeamento lógico (de ações/reflexões que se sucedem), exigido na redação do Enem, e que culmina com a proposta de intervenção. Vê-se como adequada, portanto, a presença de verbos factivos. Apesar disso, decidimos verificar no *corpus* quais verbos eram os mais frequentes desse total de 5114. O resultado se vê na Tabela 4.

TABELA 4 – Verbos mais frequentes

Verbo	Geral	
	Total	% (do total de 5114)
ser	792	15,5
dever	142	2,8
poder	141	2,8
ter	111	2,2
tornar	92	1,8
TOTAL	1278	25

Fonte: Elaboração Própria com dados do Tropes.

À primeira vista, os dados parecem contradizer a alta frequência de verbos factivos observada nos dados anteriores, uma vez que os cinco verbos mais frequentes no *corpus* são todos estativos: *ser*, *dever*, *poder*, *ter* e *tornar*, sendo que apenas o primeiro agrega mais de 15% do total de 5114 verbos etiquetados pelo Tropes. No entanto, a verdade é que tais números atestam a hipótese levantada a respeito da alta frequência de estativos esperada no *corpus*, sobretudo pelo papel que esses verbos têm na língua. Por isso, três outros pontos podem ser relacionados aos dados. O primeiro sobre o uso de modalizadores, como *poder* e *dever*.

Assim como se viu antes, a modalização é um processo esperado em textos argumentativos e, além dos advérbios, o uso de verbos modais é uma marca para essa ação linguística. Tal fato explica que esses verbos tenham uma frequência alta nos textos. O segundo ponto é com relação ao uso de *ser*. Bertucci (no prelo) mostra que o uso do argumento de definição é muito recorrente em textos do Enem e, para a construção desse tipo de argumento, um dos verbos mais recorrentes é exatamente o verbo *ser*. Finalmente, Pereira (2018) também observou índices gerais muito parecidos com o que apresentamos aqui, seja das classes verbais, seja dos exemplos em si. Assim, os números que apresentamos vão ao encontro da proposta de que as redações do Enem se caracterizam pela alta frequência de estativos específicos e de uma predominância de factivos.

Nesse sentido, podemos sugerir que, apesar de percentualmente inferiores, os verbos estativos são essenciais tanto para o caráter expositivo do texto quanto para o teor argumentativo pretendido pelo autor, em especial no uso de verbos modais e na formação de estruturas opinativas, como aquelas formadas com adjetivos (PADILHA, em preparação). Logo, no contexto de ensino, por exemplo, não se deve concluir que um tipo de verbo deve predominar: o que se espera é um emprego adequado (e não quantificado), ainda que os dados revelem pontos interessantes sobre a redação do Enem como gênero textual.

3.2.5 Adjetivos

O resultado da análise sobre adjetivos pode ser visto na Tabela 5.

TABELA 5 – Frequência de adjetivos

Adjetivo	2014		2018		2019		Geral	
	Total	%	Total	%	Total	%	Total	%
Objetivo	301	48,6	418	46,3	670	45,7	1389	46,5
Subjetivo	300	48,5	400	44,7	644	43,8	1344	45
Numérico	18	2,9	81	9	155	10,6	254	8,5
TOTAL	619	100	899	100	1469	100	2987	100

Fonte: Elaboração própria com dados do Tropes.

No que diz respeito à subcategorização de adjetivos, os dados revelam um relativo equilíbrio entre adjetivos classificados como objetivos (*infantil, social, cultural* entre outros) e aqueles marcados como subjetivos (*necessário, grande, importante* entre outros). Observa-se que o número de adjetivos utilizados (2987) é bastante alto, comparado ao de verbos (5114), ainda mais se se considerar que o verbo é um elemento essencial numa oração, mas o adjetivo não. O que se percebe é que os adjetivos têm um papel relevante em textos como do Enem, uma vez que são um meio também de revelar o posicionamento do autor. Isso fica claro, sobretudo, com os subjetivos, que são articulados em torno de um ponto de vista em relação a um fato apresentado, tal como observa Padilha (em preparação). Para a autora, a forte presença dos subjetivos, em especial aqueles de caráter modal e gradual, contribui para a noção argumentativa no texto, muito mais do que para uma simples exposição do tema. Voltando à Hipótese 5, podemos constatar que a média de adjetivos para cada uma das 95 redações é de 31,4, o que daria cerca de um adjetivo por linha, na redação do Enem. Isso mostra a importância da categoria nesse gênero e corrobora a referida hipótese.

3.2.6 Pronomes

Finalmente, a última categoria sob análise é dos pronomes. O Tropes destacou uma única referência à primeira pessoa do singular e observamos nos textos que se referia ao trecho seguinte, presente na redação 3 da cartilha referente ao Exame de 2018.

Em segundo lugar, o ser humano perde a sua capacidade de escolha. Conforme o conceito de “Mortificação **do Eu**”, do sociólogo Erving Goffman, é possível entender o que ocorre na internet que induz o indivíduo a ter um comportamento alienado. (FELPI, 2019, p. 11, grifos nossos)

Nesse caso, o uso do pronome é feito como citação/menção e não como uso por parte do candidato. Vale destacar ainda que o aplicativo destacou 276 ocorrências do clítico *-se*, que tem a função de impessoalizar a ação verbal (ou seja, é também um item de terceira pessoa). Pereira (2018), com o já referido *corpus* de 16 redações nota 1000, fez observações semelhantes às nossas. Tudo isso, então, corrobora a Hipótese 6, sobre a predominância da terceira pessoa em textos do Enem.

Nesse ponto, é importante destacar que os materiais relativos à avaliação da redação do Enem não excluem a possibilidade de escrita do texto em primeira pessoa (BRASIL, 2019; GARCÊZ; CORREA, 2017; INEP, 2020; SANDOVAL *et al.*, 2017). No entanto, o que se vê nos dados é que os textos avaliados com a nota máxima não apresentam pronomes dessa categoria, o que é um indício de que a avaliação se dá num nível bastante restrito de impessoalização do texto, valorizando a escrita em terceira pessoa. Não nos parece justo dizer que a argumentação exija, em si, uma impessoalização; pelo contrário: pessoalizar pode ser uma forma de persuasão bastante válida, como mostram Freitas e Marra (2016) a respeito de anúncios publicitários. Este ponto, aliás como todos os já comentados aqui, merece mais espaço de estudo em português brasileiro, em especial no que tange à caracterização dos gêneros e, por consequência, suas possibilidades de reflexão e ensino.

4 Considerações finais

Os resultados da presente pesquisa nos permitem defender que ferramentas computacionais contribuem na descrição de gêneros textuais/discursivos, porque são capazes de analisar volumes de dados maiores do que aqueles que se faz manualmente (impensáveis para um pesquisador ou professor). Nesse sentido, pensamos ter justificado a importância da relação entre linguagem, tecnologia e um estudo de *corpus* com gêneros escolares, uma vez que pode ter impacto no modo como se analisam e se produzem os textos em ambientes de letramento.

Isso não significa, obviamente, que o ensino de produção textual deva ser comparado a uma receita pronta, uma vez que a ação linguística é extremamente complexa. No entanto, compreender as recorrências de um gênero pode contribuir para ampliar a capacidade de reflexão de professores e estudantes a respeito do fazer linguístico, especialmente de textos que se mostram bastante formatados, como a redação do Enem. Por isso, tentamos discutir os dados à luz dos materiais oficiais referentes à prova, além de outras pesquisas, como uma forma de vincular os dados ao modo como se interpreta (ou avalia) o texto escrito no Exame.

A partir da proposta de utilizarmos o Tropes para análise dos dados, podemos concluir que ele apresentou um panorama geral relativo às propriedades linguísticas das redações nota 1000 que são coerentes

com aquilo que se vê na literatura sobre esse tipo de texto. Com isso, concordando com Araújo (2017) e Bertucci (2020), pensamos que o Tropes contribui para a identificação de elementos linguísticos que caracterizam um gênero. Deixamos em aberto, no entanto, pesquisas que possam analisar suas limitações e possíveis inconsistências de análise. Entre as que identificamos, há a marcação dos artigos *o* e *a* como pronomes oblíquos de terceira pessoa, o que, no entanto, não prejudica as conclusões gerais apresentadas, em especial porque o número de pronomes de primeira e segunda pessoa foi nulo no *corpus*.

No presente trabalho, também não tratamos de questões a respeito do debate sobre a padronização do gênero, nem de como essa descrição realizada pode contribuir para o ensino de produção textual, ou seja, para a discussão do “saber-fazer” que a tecnologia metalinguística abarca na escola, uma vez que não era nossa intenção fornecer um manual de ensino do gênero, mas analisar um conjunto significativo de textos prototípicos do Enem e mostrar que a ferramenta Tropes pode contribuir para a descrição de textos. Ao leitor interessado nesse tópico, no entanto, a divulgação do material de apoio para avaliadores, divulgado em 2017 pelo Inep (GARCEZ; CORRÊA, 2017), e dos materiais para formação de avaliadores (INEP, 2020), recentemente disponibilizados, podem auxiliar a entender as minúcias desse gênero textual/discursivo exigido no Enem, com ênfase na produção e avaliação dos textos.

Por fim, entendemos que será uma grande contribuição, tanto para a área acadêmica, quanto para a pedagógica, que profissionais se debrucem sobre questões como as apresentadas aqui e desenvolvam estratégias que contribuam para um ensino da língua que, realmente, ajudem os alunos a fazer escolhas linguísticas relevantes na construção de seus textos. Reiteramos que isso não significa haver uma receita, mas pode ser um olhar diferente para a constituição do gênero. Igualmente relevantes são pesquisas com aplicativos como o aqui usado e com *corpora* similares, a fim de se contribuir para o entendimento de gêneros e para sua instrumentalização.

Referências

ALVES, W. M.; BESSA, J. C. R. Orientações para escrita da redação do Enem em vídeos do Youtube. *Hipertextus*, Recife, v. 19, n. 1, p. 1-23, 2018. Disponível em: <https://periodicos.ufpe.br/revistas/hipertextus/article/view/247974/36463>. Acesso em: 1 out. 2020.

ARAÚJO, L. S. de. O gênero entrevista radiofônica em comunidades hispânicas: um aporte da Análise Textual Automática. *Domínios de Linguagem*, Uberlândia, v. 11, n. 2, p. 289-312, 2017. Disponível em: <https://doi.org/10.14393/DL29-v11n2a2017-2>. Acesso em: 1 out. 2020.

ARAÚJO, F. de C.; CUNHA, M. A. F. da. A estrutura argumental dos verbos de ação. *Publica*, Natal, v. 3, n. 1, p. 28-35, 2009. Disponível em: <https://periodicos.ufrn.br/publica/article/view/106>. Acesso em: 1 out 2020.

AUROUX, S. *A revolução tecnológica da gramatização*. 3. ed. Campinas: Editora da UNICAMP, 2014.

AZEVEDO, I. C. M. Organização de textos dissertativo-argumentativos em prosa: o que se percebe em dez anos de realização do Enem? In: SILVA, L. R. da; FREITAG, R. M. K. (org.). *Linguagem, interação e sociedade: diálogos sobre o Enem*. João Pessoa: Editora do CCTA, 2015. p. 33-50.

BARROS, M.; ALBUQUERQUE, M. G. As técnicas argumentativas e a construção de sentidos em redações do Enem. In: SEMINÁRIO DE ESTUDOS SOBRE DISCURSO E ARGUMENTAÇÃO, 2., 2015, Belo Horizonte. *Anais...* Belo Horizonte: Editora Fale, 2015. p. 545-559.

BARTON, D.; LEE, C. *Linguagem online: textos e práticas digitais*. Trad. Milton Camargo Mota. 1. ed. São Paulo: Parábola Editorial, 2015.

BERTUCCI, R. A. Aplicação de ferramentas para coleta e análise de dados em Linguística. *Diacrítica*, Braga, Portugal, v. 32, n. 3, p. 129-155, 2020. DOI: <https://doi.org/10.21814/diacritica.576>

BERTUCCI, R. A. Análise do argumento por definição em redações do Enem. *Acta Scientiarum*, Maringá, PR. (no prelo).

BERTUCCI, R. A.; MALHEIROS, A. J.; LOPES, W. de S. Ocorrências de anáforas encapsuladoras em redações do Enem. *Filologia e Linguística Portuguesa*, São Paulo, v. 22, n. 1, p. 81-102, 2020. DOI: 10.11606/issn.2176-9419.v22i1p81-102. Disponível em: <https://www.revistas.usp.br/flp/article/view/164142>. Acesso em: 23 nov. 2020.

BRASIL. *Redação no Enem 2019*: cartilha do participante. Brasília: Daeb/Inep/MEC, 2019. Disponível em: http://download.inep.gov.br/educacao_basica/enem/downloads/2019/redacao_enem2019_cartilha_participante.pdf. Acesso em: 1. out. 2019.

CABRAL, A. L. T. O conceito de plano de texto. *Linha d'Água*, São Paulo, n. 26, v. 2, p. 241-259, 2013. Disponível em: <https://www.revistas.usp.br/linhadagua/article/download/64266/71562/>. Acesso em: 1 ago. 2019.

CAMPOS, L. V. Mais de 143 mil participantes tiraram zero na redação do Enem 2019. *Brasil Escola, Site Uol*, 17 jan. 2020. Disponível em: <https://vestibular.brasilecola.uol.com.br/enem/mais-143-mil-participantes-tiraram-zero-na-redacao-enem-2019/347183.html>. Acesso em: 28 set. 2020.

CAMPOS, C. M.; RIBEIRO, J. Gêneros. In: COSTA, I. B.; FOLTRAN, M. J. (org.). *A tessitura da escrita*. São Paulo: Contexto, 2013. p. 23-44.

CANÇADO, M.; AMARAL, L.; AMORIN, E.; VELOSO, A.; MELLO, H. Subjetividade em correções de redações: detecção automática através de léxico de operadores de viés linguístico. *Linguamática*, Braga, v. 12, n. 1, p. 63-79. 2020. DOI: <https://doi.org/10.21814/lm.12.1.313>.

CANTARIN, M.; BERTUCCI, R. A.; ALMEIDA, R. C. de. A análise do texto dissertativo-argumentativo. In: GARCEZ, L. H.; CORRÊA, V. R. (org.). *Textos dissertativo-argumentativos*: subsídios para a qualificação de avaliadores, Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017. p. 81-91.

CASSIRER, E. *Ensaio sobre o homem*: introdução a uma filosofia da cultura humana. São Paulo: Martins Fontes, 1994.

COROA, M. L. O texto dissertativo-argumentativo. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativo-argumentativos*: subsídios para a qualificação de avaliadores. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017. p. 59-71.

CORPUS do Português. Disponível em: <https://www.corpusdoportugues.org/>. Acesso em: 30 set. 2020.

COSTA, J. de R. O.; GUEDES, M. A. A avaliação dos indícios de autoria. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativo-argumentativos: subsídios para qualificação de avaliadores* Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017. p. 101-108.

CUPANI, A. *Filosofia da tecnologia: um convite*. Florianópolis: Editora da UFSC, 2016.

FELPI, L. (org.). *Cartilha redação a mil*. 2019. Disponível em: <https://www.lucasfelpi.com.br/redamil>. Acesso em: 1 out. 2020.

FELPI, L. (org.). *Cartilha redação a mil 2.0*. 2020. Disponível em: <https://www.lucasfelpi.com.br/redamil>. Acesso em: 1 out. 2020.

FINATTO, M. J. B. Apresentação: descrição dos gêneros textuais/discursivos com apoio computacional. *Domínios de Linguagem*, Uberlândia, v. 11, n. 2, p. 282-288. 2017. DOI: <https://doi.org/10.14393/DL29-v11n2a2017-1>.

FIORIN, J. L. *Argumentação*. São Paulo: Contexto, 2017.

FREITAS, H. C.; MARRA, M. N. A. Tipos de argumentos utilizados nos anúncios publicitários das Havaianas. *Domínios de Linguagem*, Uberlândia, v. 10, n. 1, p. 304-329, 2016. DOI: <https://doi.org/10.14393/DL21-v10n1a2016-16>.

GARCEZ, L. H. do C. Gênero e tipo de texto. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativos-argumentativos: subsídios para qualificação de avaliadores*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017a. p. 51-58.

GARCEZ, L. H. do C. O ensino de redação. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativos-argumentativos: subsídios para qualificação de avaliadores*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017b. p. 275-279.

GARCEZ, L. H. C.; CORRÊA, V. R. (org.). *Textos dissertativo-argumentativos: subsídios para a qualificação de avaliadores*. Brasília: Inep, 2017. Disponível em: <http://portal.inep.gov.br/documents/186968/484421/TEXTOS+DISSERTATIVO+ARGUME>

NTATIVOS/7809ef0d-5a4a-4c24-9a03-9db15e0bdacf?version=1.0. Acesso em: 28 set. 2020.

INEP disponibiliza material inédito sobre correção da redação do Enem. *Portal do MEC*, Brasília, 2020. Disponível em: <http://portal.mec.gov.br/pronatec/oferta-voluntaria/418-noticias/enem-946573306/90611-inep-disponibiliza-material-inedito-sobre-correcao-da-redacao-do-enem>. Acesso em: 28 set. 2020.

KOCH, I. G. V. *Argumentação e linguagem*. 13. ed. São Paulo: Cortez, 2011.

LIMA, L. I. *Mapeamento semântico da construção de autoria no Ensino Médio*. 2019. 163f. Tese (Doutorado em Letras) - Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, 2019. Disponível em: <http://www.prppg.ufpr.br/signa/visitante/trabalhoConclusaoWS?idpessoal=29931&idprograma=40001016016P7&anobase=2019&idtc=1533>. Acesso em: 01 out. 2020.

MAGALHÃES, M. M. A argumentação em redações escolares. In: SILEL – SIMPÓSIO NACIONAL E INTERNACIONAL DE LETRAS E LINGUÍSTICA, 2013, Uberlândia. *Anais...* Uberlândia: EDUFU, 2013, p. 1-13. Disponível em: http://www.ileel.ufu.br/anaisdosilel/wp-content/uploads/2014/04/silel2013_645.pdf. Acesso em: 8. Jun. 2020.

OLIVEIRA, F. C. C. de. *Um estudo sobre a caracterização do gênero redação do ENEM*. 2016. 166f. Tese (Doutorado em Linguística) – Centro de Humanidades, Universidade Federal do Ceará, 2016.

OLIVEIRA, W. R. de. *Planejamento de escrita em meio digital e analógico*. 2018. 175f. Dissertação (Mestrado em Estudos de Linguagens) - Departamento Acadêmico de Linguagem e Comunicação, Universidade Tecnológica Federal do Paraná, 2018. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/3319>. Acesso em: 1 out. 2020.

OLIVEIRA, M. I. S.; CABRAL, A. L. T. Política de Língua Portuguesa para o ensino de Redação no nível médio da educação brasileira: o texto argumentativo dos PCN's à redação do Enem. *Verbum*, São Paulo, v. 6, n. 2, p. 6-30. 2017. Disponível em: <https://revistas.pucsp.br/verbum/article/view/30274>. Acesso em: 1 out. 2020.

PADILHA, M. P. *O papel modalizador dos adjetivos em redações do Enem*. Dissertação (Mestrado em Estudos de Linguagens) – Departamento Acadêmico de Linguagem e Comunicação, Universidade Tecnológica Federal do Paraná. Em preparação.

PAIVA, R. I. de S. *Redações nota mil no ENEM: um estudo analítico da massificação se sua estrutura e conteúdo*. 2020. 132f. Dissertação (Mestrado em Estudos de Linguagem) – Centro de Ciências Humanas, Letras e Artes, Universidade Federal do Rio Grande do Norte, 2020. Disponível em: <https://repositorio.ufrn.br/handle/123456789/29929>. Acesso em: 1 out. 2020.

PEIXOTO, J. dos S. A avaliação do emprego de operadores e conectivos argumentativos. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativos-argumentativos: subsídios para qualificação de avaliadores*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017. p. 163-171.

PEREIRA, K. A. P. L. *A contribuição de um analisador automático para a caracterização de gêneros textuais*. 2018. 83f. Trabalho de Conclusão de Curso (Graduação em Letras) - Universidade Tecnológica Federal do Paraná, 2018. Disponível em: <http://repositorio.roca.utfpr.edu.br/jspui/handle/1/11580>. Acesso em: 1 out. 2020.

PERELMAN, C.; OLBRECHTS-TYTECA, L. *Tratado da argumentação: a nova retórica*. 3. ed. São Paulo: Martins Fontes, 2014.

PINHEIRO, C.; CORTEZ, J. Teorias da argumentação na prova de redação do ENEM. *Linguagem & Ensino*, Pelotas, v. 20, n. 1, p. 61-80. 2017. Disponível em: <https://doi.org/10.15210/rle.v20i1.15215>. Acesso em: 1 out. 2020.

RASO, T.; MELLO, H. (org.). *C-ORAL-BRASIL I: corpus de referência do português brasileiro falado informal*. Belo Horizonte: Editora UFMG, 2012.

SANDOVAL, A. N.; ALCÂNTARA, S. S.; ZANDOMÊNICO, S. C. M. de R. Notas sobre a avaliação de desvios de registro. In: GARCEZ, L. H. do C.; CORRÊA, V. R. (org.). *Textos dissertativos-argumentativos: subsídios para qualificação de avaliadores*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017. p. 29-35.

SARDINHA, T. B. Corpus Linguistics: History and problematization. *DELTA*, São Paulo, v. 2, n. 16, p. 323–367, 2000. Disponível em: <http://dx.doi.org/10.1590/S0102-44502000000200005>. Acesso em: 26 abr. 2018.

SILVA, N. S. *Análise textual mediada por ferramenta computacional: um estudo sobre redações estilo Enem*. 2018. 73f. Trabalho de Conclusão de Curso (Graduação em Letras) - Universidade Tecnológica Federal do Paraná, 2018.

TROPES. *Semantic-Knowledge: text analysis, qualitative analysis & text mining*. [s.d.]. Disponível em: <https://www.semantic-knowledge.com/tropes.htm>. Acesso em: 30 jun. 2018.

VIEIRA PINTO, Á. *O conceito de tecnologia*. Rio de Janeiro: Contraponto, 2005.

WACHOWICZ, T. C. *Análise linguística nos gêneros textuais*. Curitiba: Intersaberes, 2010.