

ORIENTAÇÕES DE ATIVIDADES DE ANÁLISE LINGUÍSTICA BASEADAS EM DIAGNÓSTICOS EXTRAÍDOS DE UM CORPUS DESENVOLVIMENTAL

Taiza de Oliveira Santos¹

RESUMO

O presente trabalho tem o objetivo de apresentar orientações baseadas em diagnósticos para a preparação de atividades de análise linguística extraídas de um corpus desenvolvimental no contexto da escrita escolar de crianças e adolescentes da Educação Básica. É uma pesquisa que se enquadra na área dos estudos do desenvolvimento linguístico em idade escolar, em que se defende que o desenvolvimento é contínuo, havendo uma consolidação e aperfeiçoamento do conhecimento já existente das estruturas linguísticas, em especial o sintático, e pelo aparecimento de outras construções (SIM-SIM, 1998; CHOUPINA, 2010). Neste estudo, com o intuito de constituir evidências que substanciem as orientações propostas, que abrangem os campos sintáticos, morfológicos e lexicais da língua, foram extraídos dados do corpus desenvolvimental Doeste (MARTINS *et al.*, 2020). O Doeste é um repositório de textos escritos por crianças e adolescentes portugueses e brasileiros em idade escolar, sendo os textos brasileiros escritos por alunos do 5º e 9º anos do Ensino Fundamental e do 3º do Ensino Médio de instituições públicas de três cidades do estado do Rio Grande do Norte. Os diagnósticos coletados no corpus referem-se ao uso dos antecedentes das orações relativas, dos advérbios e as conjunções. Espera-se que este trabalho propicie ao professor refletir sobre a elaboração de estratégias de ensino baseadas em evidências, estimulando-o a ele próprio tornar-se investigador e fazer uso de dados extraídos de um corpus nas suas aulas.

Palavras-chave: Desenvolvimento Linguístico; escrita escolar; linguística de corpus; ensino baseado em dados.

¹ Graduanda em Letras-Português pela Universidade Federal Rural do Semi-Árido (UFERSA). E-mail: taiza_oliveira@hotmail.com

Trabalho orientado pelo professor Dr. Mário Martins (UFERSA). E-mail: mario.martins@ufersa.edu.br

Considerações iniciais

A língua, enquanto sistema semiótico muito bem articulado, organizado e diversificado, constitui-se num recurso que nos permite interagir socialmente. Essa interação materializa-se por meio de textos, sejam eles falados, escritos ou ainda multimodais. Todo esse processo de interação está situado em um contexto também ele carregado de significado, pelo que interferem diretamente na seleção dos elementos linguísticos que são acionados pelos interlocutores. Por isso, a língua – e consequentemente os elementos que a constroem – são maleáveis, sofrendo alterações à medida que o indivíduo que a usa sente necessidades.

Os elementos constitutivos do sistema linguístico podem ser entendidos como um conjunto de regras e princípios que os falantes aprendem (ou adquirem, numa perspectiva inatista), convertendo-se em seu conhecimento linguístico. Esse conhecimento, em princípio, é implícito, ou seja, um conhecimento irrefletido, imbuído da intuição, em que as estruturas da língua são desenvolvidas de forma espontânea a partir da interação. Embora implícito, esse conhecimento linguístico sustenta-se num sistema considerado altamente complexo, que envolve diversos estratos de composição.

Assumimos, porém, que, nas sociedades modernas, em que se demanda dos indivíduos conjuntos cada vez mais refinados de competências, inclusive relacionadas aos usos da língua, é necessário que a escola transforme de forma eficaz esse conhecimento linguístico implícito – próprio de qualquer indivíduo com desenvolvimento típico – em consciência linguística. Isso implica, como ponderam Costa *et al.* (2017), um domínio deliberado que o falante passa a ter dos diferentes estratos da língua (lexical, fonológico, morfológico, sintático e semântico-discursivo e pragmático), tornando-se capaz de controlar o seu uso baseado em reflexão linguística, em quaisquer processos de produção e compreensão de textos.

Para que essa tomada de consciência linguística ocorra, é papel da escola propor estratégias de aprendizagens formais e sistêmicas, o que permitirá qualificar os alunos como usuários proficientes da língua. Defendo, por isso, a necessidade de se formar indivíduos capazes de transitar em todas as esferas da sociedade através da linguagem, competindo à escola “contribuir para o crescimento linguístico de todos os alunos, estimulando-lhes o desenvolvimento da linguagem e promovendo a aprendizagem das competências que não decorrem do processo natural de aquisição” (SIM-SIM, 1998, p. 35).

Em um contexto escolar, é reconhecido que o trabalho de reflexão consciente sobre a língua permite o aumento da capacidade dos aprendizes de “identificar e nomear as unidades da língua (por exemplo, fonemas, sílabas, morfemas, palavras, grupos sintáticos, frases), de caracterizar as suas propriedades, as suas regras de combinação e os processos que atuam sobre as estruturas formadas.” (DUARTE, 2008, p. 17). Nesse sentido, é premente admitir a importância que os estudos linguísticos fundamentados no uso real da língua exercem na descoberta e na análise de elementos que são inerentes à língua e aos usos que fazemos dela, sendo possível apontar e descrever padrões existentes.

Faz-se necessário, pois, que o professor tenha conhecimento sobre esses padrões, inclusive padrões produzidos no percurso do aprendizado escolar, e possa trabalhá-los de forma eficiente e aplicada em sala de aula. Propõe-se a esse professor um duplo papel: primeiro, que seja um utilizador competente da língua e, segundo, que forme falantes competentes (DUARTE, 2008). Ainda com esse pensamento, compreendo que os estudos linguísticos, em especial os estudos com corpora² de textos, têm assumido uma importância fundamental no momento histórico em que vivemos, sendo uma fonte valiosa de informações, que nos oferece um entendimento sobre a língua por meio da descrição das suas regularidades e das frequências com que essas regularidades estão presentes nos usos, fazendo-nos observar o real comportamento linguístico dos falantes de uma dada comunidade, a partir de dados precisos e isentos de opiniões e pré-julgamentos dos fatores de uma língua (TRASK, 2004).

Através dessa compreensão, percebemos que o corpus se torna uma ferramenta de pesquisa, mas também uma ferramenta pedagógica, que o professor pode usar em sala de aula, acionando problemas linguísticos reais que podem ser analisados, garantindo que o processo de ensino-aprendizagem dos alunos seja cientificamente rigoroso, contribuindo para uma melhor compreensão dos elementos que constituem a língua e para a formação do pensamento científico.

Portanto, o propósito deste estudo se justifica por percorrer caminhos ainda em exploração no nosso país, objetivando encontrar possíveis potencialidades para o uso de um corpus desenvolvimental, principalmente de como este pode ser de grande ajuda para a elaboração de aulas de Língua Portuguesa para a Educação Básica. Para isso, fez-se uso de dados extraídos do Doeste (MARTINS, et al., 2020)³, um corpus desenvolvimental de

² Neste texto, para fins de simplificação, não se destaca com itálico a palavra corpora e corpus, em linha com o que fazem McEnery e Hardie (2012) e Martins (2016).

³ Para uma melhor compreensão, o [DOESTE \(ufersa.edu.br\)](http://DOESTE.ufersa.edu.br) será apresentado na seção metodológica.

que foi possível extrair evidências baseadas em dados que permitem sugerir orientações para a elaboração de atividades a nível sintático, morfológico, semântico, lexical e gramatical da língua. Contribui-se, assim, com a prática pedagógica do professor, auxiliando-o no processo de ensino-aprendizagem, em que se permite ao aluno desenvolver o seu conhecimento sobre a língua que ele próprio produz, suas estruturas linguísticas e o seu repertório lexical.

2. Aporte teórico

Esta seção apresenta um olhar pedagógico para a Linguística de Corpus, com o intuito de entender como a sua natureza empírica pode auxiliar na construção de estratégias voltadas para o ensino de línguas. Portanto, divide-se em duas subseções, em que se fala, respectivamente, sobre a Linguística de Corpus e o corpus como ferramenta de ensino. Nessa seção, portanto, tem-se o objetivo de demonstrar o alicerce desta pesquisa, em que se utiliza das ferramentas disponibilizadas pela Linguística de Corpus para se pensar em orientações de atividades de análise linguística baseadas em diagnósticos a serem aplicadas nas aulas de Língua Portuguesa da Educação Básica.

2.1 Definição de corpus e técnicas de extração de informações

A Linguística de Corpus é apresentada como uma área da Linguística em que se utiliza de corpora de textos, tratados por programas de computador, para a exploração de línguas e de variedades linguísticas. A natureza empírica de seus dados auxilia em uma análise mais objetiva dos fenômenos linguísticos, pois o estudo da língua é baseado em exemplos do uso da língua da “vida real” (MCENERY; WILSON, 1996). Antes de debruçar em específico sobre o objeto de estudo da Linguística de Corpus, o corpus, é preciso reconhecer, que, embora alguns trabalhos baseados em corpus tenham sido realizados nos anos 60, como é o caso do “Survey of English Usage”, de Quirk *et. al* (1960), e do “Brown Corpus”, de Francis e Kučera (1964), é só nas últimas três décadas, com os avanços na disponibilidade de computadores institucionais e pessoais, que a Linguística de Corpus começou a crescer e a ser aceita como um meio válido de investigação linguística.

No que tange o âmbito nacional, as pesquisas nessa área ainda caminham a passos tímidos, o que difere do cenário internacional, em especial da Europa, em que essa metodologia exerce grande influência nas pesquisas de cunho linguístico, utilizando-se desse aparato metodológico para descrever variados aspectos da linguagem, como

explicita Berber-Sardinha (2004). Portanto, a história da Linguística de Corpus está diretamente vinculada à tecnologia, permitindo não somente o armazenamento de corpora, mas também a sua exploração. Esse processo só se tornou possível devido à disponibilidade de ferramentas computacionais, softwares, para análise de corpus, em que se manipulam os dados linguísticos, mostrando, pelo menos, a sua frequência, a frase estruturada e a colocação dos elementos (HUNSTON, 2002).

Na literatura, muitas são as definições de corpus, porém opto pela que melhor se ajusta ao escopo deste trabalho. Em um passado não tão distante, o corpus estava associado a um conjunto de obras, como, por exemplo, todos os textos escritos de um determinado autor. Com o surgimento dos computadores e a possibilidade de armazenamento de grandes quantidades de textos, o termo passou a ser entendido como uma coleção de textos eletrônicos, na modalidade escrita ou falada da língua, ou de ambas, podendo ser acrescidas de elementos de multimídia, sendo normalmente armazenados num computador, em que se utiliza de ferramentas de software específicas para análise, como descrevem O’Keeffe *et al.* (2007).

Uma outra definição de corpus, apresentada por Sinclair (2004), considera o corpus como uma coleção de peças de textos linguísticos em formulário eletrônico, que são fonte de dados para uma investigação linguística. Esses textos são selecionados a partir de critérios externos específicos para representar, na medida do possível, uma língua ou variedade linguística. Resumidamente, Friginal (2018) define o corpus como um sistema eletrônico de concepção sistemática de coleção de textos naturais, em que se apresentam as construções existentes das modalidades orais ou escritas de uma língua e que são relevantes para auxiliar a entender as tendências linguísticas usadas pelos falantes.

A nomenclatura empregada na Linguística de Corpus para se definir qual o conteúdo e o propósito de um corpus é muito extensa. Vejam-se, portanto, algumas tipologias citadas por Berber-Sardinha (2004) e agrupadas segundo critérios estabelecidos pelo autor: modalidade, se é falado ou escrito; tempo, se é diacrônico, sincrônico, contemporâneo, histórico; seleção, se é um corpus de amostragem, monitor, dinâmico, estático etc.; conteúdo, se é de tipos textuais, regional, multilíngue; autoria, de aprendizes ou de nativos; finalidade, se é de estudo, de referência, de teste.

A verdade é que existe uma gama de possibilidades para a construção e a escolha de corpus, entretanto o que se deve ter em mente ao selecionar um corpus é que, através dele, seja possível responder a questões específicas de investigação, havendo uma

correspondência entre o conteúdo de um corpus e os propósitos que se deseja investigar (MCENERY; HARDIE, 2012). Em outras palavras, um corpus deve ser representativo dos fatos linguísticos que se pretende estudar, eis o motivo pelo qual a maioria dos corpora disponíveis on-line se apresenta construído de vários critérios particulares de compilação, como é o caso do Doeste.

Uma outra característica de um corpus, como Biber, Conrad e Reppen (1998) expõem, é que se trata de uma coleção de textos disponíveis para análises quantitativas e qualitativas, o que permite um conhecimento mais apurado e minucioso da língua, em que se observam as técnicas de análise de frequência e concordância. A análise de frequências permite a obtenção de valores de ocorrência, sejam absolutos ou normalizados, de um determinado termo, permitindo a sua observação ou a comparação entre termos num dado corpus, sendo possível observar quais deles têm ocorrências mais expressivas.

Ao utilizar um corpus de textos narrativos escritos por alunos sobre as suas férias, como é o caso de um dos estímulos presentes no corpus Doeste, pode-se examinar, por exemplo, quantas vezes os lemas “amigo” e “família” ocorrem (239 e 45 ocorrências, respectivamente). São resultados quantitativos que podem ser comparados à frequência de outros elementos no mesmo corpus ou em outros corpora. Como a observação e análise da frequência de um termo é relativa, testes estatísticos podem garantir a fiabilidade dos dados.

No entanto, o corpus permite também um olhar qualitativo para a forma como uma palavra, frase ou texto é utilizado. Para isso, é preciso olhar para além da frequência da ocorrência da palavra e olhar para o modo como determinado elemento é inserido nos seus contextos mais imediatos (contexto linguístico ou cotexto), o que se denomina como concordância, que é entendida como um procedimento de obtenção de listas de ocorrências de um termo específico presente num corpus.

O resultado, como explicam Baker *et. al* (2006), são linhas em que o termo pesquisado, conhecido como nó, é demarcado pelo contexto em que ele ocorre. O contexto é tido como um conjunto, que é regulável, de palavras à direita e à esquerda do termo consultado, o que permite observar um dado elemento linguístico e chegar a determinadas conclusões acerca dele, como se exemplifica abaixo.

Imagem 1: Linhas de concordância com o nó *nós* e contexto à esquerda e à direita de cinco palavras, retiradas do corpus Doeste.

contexto	uma casa de praia para nós passa a noite, eu
contexto	casa da Praia, quando nós chegou em casa, combinamos
contexto	tinha chegado o Parque quando nós viu ficamos muito alegre,
contexto	no restaurante, de noite nós foi para o parque chegamos
contexto	au parque logo fomos brinca nós brinquedo no último brinquedo foi

Fonte: Corpus Doeste (MARTINS *et al.*, 2020)

Com base nesses exemplos retirados do Doeste, podemos observar o que foi exemplificado anteriormente sobre as linhas de concordância. Nesse caso, representa a construção da pesquisa do pronome “*nós*” seguido de *verbo*, podendo ser feita uma análise de coocorrência entre este elemento e o verbo que o segue. No exemplo, que foi extraído de um corpus de escrita escolar, evidenciamos que o pronome e o verbo não estão em concordância, se considerarmos a variedade padrão da língua portuguesa.

As linhas de concordância se apresentam no contexto escolar como uma ferramenta importante na construção das aulas do professor, permitindo a seleção prévia das linhas que melhor se adequam ao que será trabalhado em sala, já que tal seleção não pode ser feita de modo aleatório, mas de modo que permita a clareza dos dados, que os alunos saibam o que está sendo pedido e que evidenciem o uso do fato linguístico a ser estudado, tornando-se, assim, úteis para a compreensão e análise.

2.2 O corpus em contexto de aprendizagem

Num contexto pedagógico, as ferramentas e técnicas disponíveis pela Linguística de Corpus podem e devem ser usadas em uma rotina que promova o aprendizado. Ou seja, um corpus, para além de pesquisas linguísticas, como foi visto, pode ser utilizado no ensino de línguas, seja de língua materna ou de língua adicional, com o intuito de propiciar ao professor a construção de estratégias que possibilitem o ensino e a imersão do aluno frente a língua. Duas são as possibilidades de uso de corpus em contexto de aprendizagem: o ensino-aprendizagem baseado em dados ou o ensino-aprendizagem movido por dados.

Para auxiliar o professor nesse processo, Johns (1991, 1994) apresenta uma metodologia que recebe o nome de “*Data-Driven Learning*” (DDL), em uma tradução para o português, Berber-Sardinha (2004) propõe como um “Aprendizado Movido por Dados” (AMD). Essa metodologia de ensino auxilia na motivação dos aprendizes, especialmente por lidar com questões diretamente relacionadas à língua e à forma como

ela se apresenta nos seus contextos de uso, promovendo um aprendizado que confronta de forma direta o aprendiz com os dados que dispõe, proporcionando um contexto de aprendizado centrado no uso, e não apenas na transcrição de regras.

Os estudos de natureza empírica sobre os benefícios da aprendizagem da AMD ainda são relativamente poucos, mas têm mostrado resultados satisfatórios, apontando que o uso pedagógico dessa abordagem reúne teorias da aprendizagem construtivistas, além de proporcionar a aquisição de novo vocabulário através de encontros contextuais, sendo possível reconhecer que a metodologia em questão pode desempenhar um papel na racionalização e aceleração da aprendizagem, principalmente na aquisição de uma nova língua, fornecendo uma rica fonte de contextos a partir do novo vocabulário.

A abordagem destacada tem um olhar especial para o ensino de gramática, revalorizando o seu lugar na aprendizagem de línguas, não a excluindo, mas usando-a com novos propósitos, pois os métodos tradicionais ainda são baseados em suposições sobre como se aprende a gramática de uma língua e o que se deve aprender, apresentando um conjunto de regras e padrões já estabelecidos, o que não se torna viável, podendo ser falhos porque as línguas, como demonstra Martins (2016), incluindo o português, não estão demasiadamente estudadas para que se tenham uma ampla noção dos fenômenos que a cercam, carecendo ainda de uma descrição sobre os modos como estão fundamentadas e estruturadas.

O ensino-aprendizagem baseado em dados, “corpus-based”, norteia a construção de materiais pedagógicos. Esses materiais utilizam-se de um corpus como fonte de exemplos que serão utilizados em sala de aula, assim é possível que o professor verifique o que está sendo usado e como deve ser usado, auxiliando na aprendizagem dos alunos. É válido destacar que o professor não levará diretamente o corpus para a sala de aula, como no método “*corpus-driven*”, mas irá basear as suas aulas a partir dos elementos encontrados no corpus, em que examina a frequência e a concordância de fenômenos previamente escolhidos e consolida estratégias de ensino a partir do que foi encontrado (TOGNINI-BONELLI, 2001).

A escolha por um ou outro desses tipos de metodologia vinculados a corpora faz-se necessária no contexto atual do ensino de línguas. Numerosos estudos têm mostrado que a língua apresentada nos livros didáticos escolares ainda é frequentemente baseada em intuições sobre a forma como a língua é utilizada, apresentando o uso de frases soltas e isoladas, as quais estão fora de um contexto de uso, além da presença de uma transcrição de regras que não permite refletir sobre o uso, de acordo com Friginal (2018). Um olhar

atento para essa forma de ensino coloca em xeque a viabilidade de centrar o ensino apenas no método tradicional da gramática, portanto fazemos uso dos dados observados no corpus para orientar as formas de se trabalhar os conteúdos programáticos da disciplina de Língua Portuguesa de acordo com a necessidade dos alunos, para que esses desenvolvam competências linguísticas e consigam transitar livremente nos discursos da sociedade em que se encontram.

3. Procedimentos metodológicos

Neste estudo, propõe-se a construção de orientações baseadas em diagnósticos para a preparação de atividades que objetivam o desenvolvimento linguístico dos alunos, adentrando nos campos semântico, sintático, morfológico e lexical, em que se observam, através de um corpus desenvolvimental, as regularidades da língua produzidas pelos próprios alunos da Educação Básica, sendo possível conhecer lacunas no processo de ensino-aprendizagem e, assim, apresentar orientações específicas com o intuito de colmatá-las. Portanto, o teor desta pesquisa é baseado em diagnósticos, em que a identificação de situações problemáticas fundamenta orientações de ensino.

O corpus usado para a extração de dados e produção das atividades é o Doeste, que apresenta textos argumentativos e narrativos escritos por crianças e adolescentes monolíngues de português europeu e português brasileiro, o que significa uma coleção de 619 textos (120.721 formas distribuídas em 3.129 frases). Esses textos foram escritos a partir de estímulos semelhantes, o que auxilia numa comparação mais eficiente das mudanças linguísticas entre alunos do mesmo ano escolar e entre alunos de diferentes anos escolares. É importante destacar que os subcorpora estão tokenizados e anotados morfológicamente, além de apresentar os transcritos originais (com as devidas anonimizações).

Nesta pesquisa, opto por utilizar apenas os textos escritos por alunos do português brasileiro, os quais estão sendo coletados desde 2017, em escolas públicas de três cidades do estado do Rio Grande do Norte (Caraúbas, Mossoró e Umarizal). Os textos foram escritos por estudantes do quinto e do nono ano do Ensino Fundamental e do terceiro ano do Ensino Médio, representando os anos finais de cada ciclo da educação brasileira. As necessidades evidenciadas no corpus desenvolvimental legitimam orientar as formas como os conteúdos devem ser trabalhados em sala de aula

Para se obter os dados coletados, fundamentais para a definição das orientações a seguir apresentadas, foi necessária a utilização de construções de pesquisas feitas

diretamente no corpus para facilitar a busca pelos dados requeridos, opção que restringe a busca pelos requisitos desejados. Abaixo, segue uma imagem da tela de construção de pesquisa, dispondo de todas as opções disponíveis pelo corpus, desde a escolha do tipo de texto até a escolha por etiquetas morfológicas específicas, além das construções de pesquisas utilizadas em cada proposta.

Imagem 2: tela da opção de *construção de pesquisa*.

Fonte: Corpus Doeste (MARTINS, 2020)

A partir desse recurso do Doeste, foram extraídos dados para três orientações, adequadas ao escopo deste trabalho. No que tange a Orientação 1, com que se pretende trabalhar com os antecedentes das orações relativas e a importância que essas orações exercem na construção do texto, foi necessário saber o número total de relativas presentes nos textos, pelo que se buscou por: *[pos = "PR.*"] :: match.text_language = "Brazilian Portuguese"*. Já para saber essa quantidade por tipos de textos, acrescentou-se a opção de “tipo de texto”, a saber: *[pos = "PR.*"] :: match.text_language = "Brazilian Portuguese" & match.text_purpose = "Argumentative" within text*. Para especificar o ano escolar, acrescentou-se a opção com nome “ano escolar”, restando a seguinte construção: *[pos = "PR.*"] :: match.text_language = "Brazilian Portuguese" & match.text_education = "5" & match.text_purpose = "Narrative" within text*. Para a construção da busca por nomes genéricos, apenas digitou-se o nome que se desejava pesquisar *[form = "pessoa"] :: match.text_language = "Brazilian Portuguese"*, além de restringir a pesquisa para o tipo de texto e ano escolar. Para se obter a frequência dos nomes genéricos “pessoa(s), coisa(s) e gente” em função de antecedentes do pronome

relativo, buscou-se por: *[form = "pessoa" & pos = "|PR.*"] :: match.text_language = "Brazilian Portuguese" & match.text_purpose = "Argumentative" within text* .

Para a Orientação 2, que tem por objetivo fundamentar atividades sobre os usos dos advérbios, em específico os terminados em -mente, as construções de busca usadas se centraram na etiqueta morfológica de advérbios e a língua portuguesa brasileira, *[pos = "R.*"] :: match.text_language = "Brazilian Portuguese" within text* . No intuito de restringir ainda mais a busca e observar apenas os que possuíam a terminação -mente, a busca construída foi a seguinte: *[form = ".*mente" & pos = "R.*"] :: match.text_language = "Brazilian Portuguese"*. Para suprir a necessidade de se saber a quantidade desses elementos por ano de ensino, buscou-se por: *[form = ".*mente" & pos = "R.*"] :: match.text_language = "Brazilian Portuguese" & match.text_purpose = "Argumentative" within text* .

A fim de construir os dados presentes na Orientação 3, relativa aos usos das conjunções nos textos argumentativo e narrativos, usaram-se distintas construções de pesquisas. Como forma de identificar o valor total de elementos conjuntivos presentes, apliquei a busca *[pos = "C.*"] :: match.text_language = "Brazilian Portuguese" within text*. Para pesquisar pelos dois tipos de conjunções (coordenativas e subordinativas), tomei por base a construção *[pos = "CC.*"] :: match.text_language = "Brazilian Portuguese" within text* . Para restringir a busca das conjunções por “ano escolar” e “tipo de texto”, a construção resultou do seguinte modo: *[pos = "CC.*"] :: match.text_language = "Brazilian Portuguese" & match.text_education = "5" & match.text_purpose = "Argumentative" within text* . Por fim, para procurar por uma conjunção em específico, apliquei a seguinte construção *[form = "mas" & pos = "|C.*"] :: match.text_language = "Brazilian Portuguese" within text* .

As construções de pesquisas são essenciais para que se possa fazer um panorama sobre o desenvolvimento linguístico, uma vez que auxiliam o pesquisador na construção quantitativa dos elementos presentes nos textos, demonstrando estatisticamente o que se pode encontrar. Para o professor, elas são necessárias para que busque de maneira mais rápida e eficiente pelo que se pretende trabalhar em sala, auxiliando-o a identificar lacunas gramaticais que estão presentes no ensino da produção textual. No Doeste, portanto, as construções de pesquisas foram essenciais para que pudesse identificar de forma rápida e eficiente os déficits apresentados na escrita das crianças e adolescentes.

4. Orientações baseadas em evidências

Nesta seção, apresentaremos orientações sobre fatos da língua que podem ser abordados em sala de aula a partir dos dados extraídos do corpus, o que demonstra um contorno de como tais fatos têm se apresentado na escrita de crianças e adolescentes, refletindo o seu desenvolvimento linguístico. Para isso, aplicamos os procedimentos de extração presentes na seção anterior em elementos que consideramos marcadores do desenvolvimento linguístico, sendo: i) as orações relativas e termos antecedentes restritos ao uso de formas genéricas, como *pessoa*, *coisa* e *gente*; ii) o uso dos advérbios, com um olhar especial para os terminados em *-mente*, pois são essenciais na construção de sentido modais em textos argumentativos e narrativos; e, por fim, iii) o uso das conjunções coordenativas e subordinativas, as quais contribuem para a complexificação de ideias de modo coeso.

4.1. Antecedentes de orações relativas

Quando a criança atinge a idade dos cinco ou seis anos, pode-se dizer que ela já adquiriu a sua língua materna e sabe usá-la de modo satisfatório, mostrando um domínio das estruturas básicas da língua, sejam elas sintáticas, morfológicas ou lexicais. Nesse período, os enunciados das crianças já demonstram a presença das “orações relativas completas e todas as formas de subordinadas conjuntivas, sejam elas temporais, finais, causais, entre outras” (CHOUPINA, 2010, p. 50). Partindo disso, observei como o uso das orações relativas⁴ e a escolha lexical do antecedente retomado pelo constituinte relativo, o que representa um indicador do desenvolvimento linguístico (CHOUPINA, 2010; NIPPOLD, 2004), tem se apresentado no corpus desenvolvimental Doeste.

Aplicada a construção de busca, foi possível identificar o número total de ocorrências das orações relativas presentes nos textos, totalizando um número de 817 ocorrências, das quais 510 são em textos argumentativos (5EF=83, 9EF=182 e 3EM=245) e 307 em textos narrativos (5EF=44, 9EF=102 e 3EM=161). Com isso, podemos perceber que há um crescimento expressivo do 5º ano Ensino Fundamental ao 3º ano do Ensino Médio, porém menos expressivo entre o 9º do Ensino Fundamental e 3º Ensino Médio. Esse crescimento pode ser justificado pela noção de que o domínio das orações relativas na escrita ocorra em estágios de desenvolvimento mais tardios, o que significa um grau

⁴ Compreende-se, numa visão mais geral, que as orações relativas exercem uma função de um adjunto adnominal de um substantivo ou pronome antecedente (CUNHA; CINTRA, 2017). Ou seja, as relativas são um tipo de oração com valor sintático-semântico de um adjetivo, que sofre um processo de transposição da sua condição de independente para uma condição de subordinada, sendo esse processo realizado pelo uso do pronome relativo (BECHARA, 2009).

de maturidade linguística maior (PERERA, 1984). No que tange o crescimento pouco expressivo dos textos que abarcam o 9º do Ensino Fundamental ao 3º Ensino Médio, pode ser justificado pelo domínio ainda restrito das estruturas desse tipo de oração e as escolhas lexicais que esses alunos estão fazendo.

Abaixo, seguem alguns exemplos extraídos do corpus das orações relativas (em itálico) e a presença de palavras genéricas, como *pessoas* e *gente*, funcionando como antecedentes. Os códigos identificativos em parênteses correspondem às iniciais do nome do aluno (dmss), o estímulo a partir do qual o texto foi produzido, se narrativo (1) ou argumentativo (2), o ano escolar, 5º ano (5) e 9º ano (9) do Ensino Fundamental ou 3º ano Ensino Médio (12) e, por fim, as iniciais da escola de coleta (elgo).

- (1) “... o lado ruim é gente *que fica falando mau das pessoas...*” (dmss_2_5_ejo)
- (2) “as parte ruim das redes sociais são a pessoas *que faz racismo com pessoas morenas...*” (mpsy_2_9_elgo)
- (3) “em que eu conheci aquela pessoa *que me mostrou como aproveitar...*” (vn ds_1_12_esg)

Nos três anos escolares escolhidos para a análise, é possível observar um expressivo uso de nomes genéricos, como *pessoa*, *coisa* e *gente* que acompanham o pronome relativo, em especial o “que”, apresentando um total de 150 ocorrências, das quais 135 são em textos argumentativos. Além de acompanhar o pronome relativo, essas formas genéricas ocorrem 1161 vezes (5EF= 208, 9EF= 343 e 3EM= 610) em vários momentos dos textos e, em alguns deles, dificultam a clareza daquilo que se deseja expressar. Abaixo, temos um exemplo retirado do corpus de um aluno do 5EF, que apresenta necessidade urgente em demonstrar possíveis escolhas lexicais que auxiliem esses alunos a enriquecer seus textos, além de aumentar o grau de complexidade das orações em questão.

“As *pessoas* hoje usam muito as redes sociais para se comunicarem, postar coisas, fazer amizades novas e encontrar *pessoas*, as redes sociais também servem para ajudar *pessoas*, doar coisas, saberem oque as *pessoas* passam, e as *pessoas* que tem doenças para poderem ajudar, tabém serve para saber coisas a venda para poder comprar As *pessoas* hoje usam mais para se comunicar e fazer várias amizades, as redes sociais são muito usadas hoje

em dia, mais de 20 milhões de *pessoas* usam as redes sociais para as *pessoas* são muito importantes...”

As orações relativas se apresentam como um conteúdo complexo, que implica estudar não só os níveis morfológicos, o uso dos pronomes e os antecedentes, mas também estudar o campo sintático, observando como as orações se apresentam em conexão com o restante da estrutura sentencial. Além disso, é preciso adentrar o texto no que diz respeito ao valor coesivo apresentado pelo pronome relativo (CHOUPINA, 2010). Por isso, orienta-se que, ao apresentar as orações relativas em sala de aula, o professor estude-a pensando nos níveis da palavra, da frase e do texto, e que essa inserção seja feita ainda no início dos anos finais do Ensino Fundamental, uma vez que a Base Nacional Comum Curricular (BRASIL, 2018) sugere que o estudo dos pronomes e dos recursos coesivos iniciem no 5º ano, fechamento do ciclo dos anos iniciais.

Portanto, quando se pensa no estudo da palavra, entendemos que nesse momento é preciso que o professor trabalhe o pronome relativo, ensinando o aluno a identificá-lo, conhecer a função que ele exerce na oração e qual a sua relação com o antecedente, além de compreender que esse pronome tem uma função específica dentro da oração relativa e que pode ser diferente do seu antecedente. No que tange o estudo da frase, é importante que se reflita sobre as frases complexas, pois isso levará “o aluno a distinguir com clareza uma oração que é um acrescento ao SN e uma oração que é um argumento seleccionado por um verbo” (CHOUPINA, 2010, p. 62). Estudar a palavra é também estudar a frase, devido a relação que existe com o antecedente.

Ao se refletir sobre o caráter coesivo do pronome relativo, trabalha-se esse recurso ao nível do texto. É importante que essa reflexão seja contínua, começando, também, nos anos iniciais do Ensino Fundamental até o Ensino Médio, como explicita a BNCC, nas habilidades em que se trabalham os recursos de coesão pronominal. Para que se alcance esse propósito, tem-se como alternativa as atividades de reescrita, pois auxiliam no aperfeiçoamento do texto, permitindo abordar o conteúdo que abrange o uso das orações relativas, uma vez que seu uso impede a repetição lexical e, conseqüentemente, enriquece o texto, apresentando uma variedade lexical maior. Nesse tipo de atividade, são trabalhados processos morfológicos, sintáticos, semânticos, lexicais e também referenciais. Dominar esses processos no ato da escrita significa dizer, como aponta Zayas (1997), que o aluno será capaz de revisar seus textos e identificar possíveis problemas de coesão e corrigi-los.

As orações relativas devem ser abordadas em sala de aula numa perspectiva do texto, de modo que o estudo das relativas siga as etapas propostas com o intuito de auxiliar os alunos a identificarem os elementos que estão presentes nesse tipo de oração, desde os pronomes relativos até a função que essas orações exercem dentro do contexto de produção do texto. Ao fazer isso, os alunos serão capazes de desenvolver habilidades linguísticas, evitando desvios que fogem da norma padrão da língua, além de aumentar a densidade semântica desejável em determinadas produções dessas estruturas, demonstrando sua competência linguística enquanto produtor e usuário da língua.

4. 2 O uso de advérbios

As classes morfológicas estão previstas, como planeja a BNCC, para serem introduzidas no ensino das crianças ainda nos anos iniciais do Ensino Fundamental, a partir do 3º ano, em especial o uso dos substantivos, verbos e pronomes, para que os estudantes desenvolvam até o fim do ciclo dos anos iniciais as habilidades da concordância verbal e nominal e saibam reconhecer e utilizar em seus textos os recursos de coesão pronominal. Entretanto, os estudos com a classe dos advérbios só são inseridos no ensino a partir do 7º ano, com o objetivo de que os alunos identifiquem advérbios e locuções adverbiais que ampliam o sentido do verbo. A partir do 8º ano do Ensino Fundamental até os anos do Ensino Médio, os advérbios aparecem como habilidades requeridas, mas de forma apenas que os alunos percebam, com o auxílio da identificação desse elemento, o teor ideológico presente nos textos.

Estudar esse grupo de palavras na sala de aula está para além de classificar a sua posição na frase, significa entender e analisar as suas multifunções, principalmente a nível sintático e semântico (NEVES, 2000; ILARI, 2007). Aqui, cabe dizer que a gramática tradicional entende os advérbios no que tange o nível morfológico, tendo-o como invariável, sintático e semântico, quando se diz que ele pode modificar ou alterar o sentido de um grupo de palavras, estabelecendo uma relação sintática entre termo modificador e termo modificado, e no critério semântico classifica-os de acordo com as circunstâncias ou ideias acessórias que expressam (CUNHA; CINTRA, 2017; BECHARA, 2009).

Em particular quanto aos advérbios terminados em -mente, foco desta orientação baseada em evidências, a gramática tradicional também se atenta a eles, porém de forma superficial, entendendo-os e classificando-os como palavras que indicam modo. No entanto, os estudos linguísticos têm apontado um comportamento heterogêneo, em que esse elemento pode assumir funções variadas em uma sentença, a nível sintático-

semântico (CASTILHO, 2000). Os advérbios terminados em -mente possibilitam ao falante articular recursos variados da língua, permitindo que se expressem diferentes relações no enunciado, desde afirmar ou negar um posicionamento, ou relativizá-lo. Por isso, espera-se que em textos de base argumentativa, por exemplo, o uso deste elemento seja recorrente.

No que tange o uso de advérbios terminados em -mente no corpus Doeste, identifiquei 221 ocorrências (de um total de 3673 advérbios presentes no corpus), dos quais 134 ocorrem em textos argumentativos e 87 em textos narrativos. Dessas ocorrências, por ano escolar, tem-se que 13 estão no 5º ano e 56 ocorrências no 9º ano do Ensino Fundamental. Já no 3º do Ensino Médio, há 152 ocorrências. Com base nesses dados, esperava-se que o número de ocorrências geral fosse maior, uma vez que os dois estímulos de construção do corpus, o narrativo e o argumentativo, suscitam o surgimento de estratégias textuais em que tal classe de palavras cabe. Os textos narrativos se caracterizam por se codificarem com elementos que qualifiquem as ações narradas, enquanto que os textos de cunho argumentativo se materializam com elementos que tecem comentários e confrontam ideias. Portanto, para expressar essas ideias, é enriquecedor o uso dos advérbios terminados em -mente.

Um olhar mais atento percebe que o valor pouco expressivo desses elementos presentes nos textos do Ensino Fundamental, principalmente do 5º ano, dá-se pelo tardio encontro dessas crianças com esse grupo de palavras, já que, como explicitado anteriormente, esses elementos começam a ser trabalhados a partir do 7º ano, o que explica o uso mais expressivo no 9ºEF e 3ºEM, porém ainda deficitário para as produções exigidas, uma vez que, ao se utilizar desses elementos, se poderia ter textos argumentativos mais fluídos no desenvolvimento das ideias e textos de base narrativa mais expressivos e robustos, demonstrando um desenvolvimento linguístico satisfatório para os níveis escolares desses alunos.

Partindo desse entendimento, é clara a necessidade de se repensar a forma como os advérbios têm sido tratados em sala de aula da Educação Básica, em especial os terminados em -mente. Nesse sentido, é sugerível que os advérbios sejam introduzidos ainda mais cedo nas salas de aula de Língua Portuguesa, ainda no 5º ano, sendo uma fase do desenvolvimento cognitivo em que os alunos já conseguem compreender os valores sintático e semântico que esse elemento exerce nos enunciados. Para além desse ano escolar, é preciso que os professores trabalhem esses elementos com seus alunos de forma

contínua, perpassando por todos os ciclos, uma vez que configuram elementos importantes na construção dos textos.

Ainda como orientação, é preciso que o professor leve os alunos a uma reflexão da funcionalidade desses advérbios, expondo-os a situações práticas de uso, em que se apresenta o advérbio em diferentes contextos comunicativos e gêneros textuais. Com isso, é provável que os alunos percebam que a função exercida pelo advérbio não é apenas a de modo, como indica a gramática tradicional, mas que há matizes de significado, cabendo ao professor mostrar aos alunos o poder de maleabilidade do advérbio, que pode assumir funções semânticas variadas no enunciado a depender do contexto, além do seu caráter de transitar em diferentes posições do enunciado. Ao entender essas noções, os alunos se tornarão ainda mais competentes linguisticamente.

4.3 O uso de conjunções

Para esta orientação, observei o uso das conjunções no corpus, entendendo-a como elemento necessário na construção das ideias de um texto, responsáveis por garantir a compreensão e a unidade do texto. Entendo que as crianças, desde cedo, são capazes de realizar enunciados complexos, com o recurso a conjunções, que têm o objetivo de enriquecer e de tornar claro aquilo que se deseja expressar. Para que isso seja feito de forma eficaz, é preciso dar condições para que o conhecimento linguístico sobre essa classe de palavras tenha pleno emprego em todas as faixas escolares, colmatando as lacunas presentes em cada nível do desenvolvimento.

As conjunções são elementos gramaticais responsáveis por encadear orações ou termos de uma mesma função sintática. Neves (2010) classifica as conjunções em coordenativas ou subordinativas adverbiais. As conjunções coordenativas atuam em construções aditivas, adversativas e alternativas. No que tange às conjunções subordinativas, estabelecem diferentes sentidos, desde relações temporais, causais, condicionais, concessivas até modais. Esses marcadores fazem ligações significativas e específicas entre elementos ou orações do texto. Porque os conectores são fundamentais na construção de informações mais complexas, compreender o uso real desses elementos implica diretamente na construção do texto.

O uso consciente e variado das conjunções pressupõe usuários competentes e textos bem produzidos do ponto de vista linguístico, discursivo e gramatical. A falta de conjunções, enquanto elementos coesivos e mantenedores da coerência num texto, pode não somente implicar desconhecimento das possibilidades que o sistema linguístico do

português oferece, mas implica também que os estudantes deixam de materializar distintos significados textuais em que a diversidade de conjunções poderia ser benéfica. Sendo assim, é preciso que se criem estratégias que facilitem a compreensão sobre o que são as conjunções e qual a sua importância semiótica dentro do texto, já que se configuram como um elemento essencial na construção de sentido complexos.

No Doeste, foram identificadas 3090 ocorrências de elementos conjuntivos, das quais 2135 são de conjunções coordenativas (1041 em textos argumentativos distribuídos da seguinte forma: 5EF=223; 9EF=348; 3EM=470; nos textos narrativos, têm-se 1094 ocorrências, sendo: 5EF= 324; 9EF= 332 e 3EM= 448). As conjunções subordinativas têm um total de 955 ocorrências (532 em textos argumentativos, sendo 5EF=89; 9EN=168 e 3EM=275; e 423 em textos narrativos, 5EF=76; 9EF=148 e 3EM=199). Dentre o número de ocorrências, percebe-se que a maioria corresponde ao conectivo “e”, conjunção coordenativa que aparece em 2265 ocorrências. Esse elemento, como aponta Neves (2010), evidencia a ligação entre dois segmentos, os quais se articulam, em que um segundo elemento é acrescentado ao primeiro de forma recursiva e determinado por variações de contexto, sendo um elemento comumente usado por crianças. Essa conjunção marca uma relação de adição, porém pode assumir características de contraste e causa-consequência.

(4) “... em nosso destino nos organizamos *e* curtimos as arvores o ceu...”

(kayt_1_9_eliv)

(5) “O nome dela era Cassão *e* ele gostava gostava de gostava de jogar bola...”

(jvsl_1_9_eliv)

(6)

(7) “... muito medo chegamos em casa *e* não contamos pra ninguém esse...”

(maln_1_9_eliv)

Outra ocorrência que chama a atenção é o uso do “mas”, que aparece em 259 ocorrências. Sua função discursiva é marcar uma desigualdade entre os segmentos, colocando “o segundo segmento como de algum modo diferente do primeiro, especificando-se essa desigualdade conforme as condições contextuais” (NEVES, 2010, p. 756). Abaixo seguem exemplos retirados do Doeste, em que também se identificam equívocos ortográficos no uso do “mas” e do “mais”, apontando que esses alunos

entendem o seu valor, mas não o grafam de acordo com a convenção da norma escrita, sendo mais um ponto a ser trabalhado em sala de aula.

(8) “... eu venci, *mas* Franjinha quase que me passava...” (eacx_1_9_eliv)

(9) “... levado para o hospital, *mais* havia chegado sem vida e...” (glan_1_12_esg)

(10) “... sim usar as redes sociais *mais* usar por a sua melhoria...”
(esox_2_12_esg)

Os números representam a realidade do corpus. É perceptível que as escolhas dos alunos, nos três níveis escolares, mesmo diante de uma gama de opções de termos conjuntivos, ainda sejam apenas por uma ou outra das conjunções referidas, restringindo a potencial complexidade linguística de seu texto e consequentemente podendo comprometer a compreensão do seu leitor, pois as escolhas pelas unidades conjuntivas do texto, como se viu, interferem diretamente na coesão e coerência do texto. Assim, pensar no ensino das conjunções, implica pensar sobre a melhoria na escrita dos alunos e isso significa inserir a tomada de consciência da existência de recursos linguísticos conjuntivos na produção textual de forma mais recorrente em sala de aula.

Portanto, como proposta para auxiliar o professor em sala de aula, entendo que uma maneira de sanar e diversificar o uso dos elementos conjuntivos seja por meio da (re)escrita de textos, não se tendo em conta apenas o conceito e a classificação das conjunções, mas a observação de como esses elementos são construídos no texto, o que permitirá ao aluno pensar estratégias que agreguem mais significados aos seus textos. Nesse sentido, o aluno conseguirá desenvolver as habilidades que são pretendidas pela BNCC e conseguirá estabelecer sentidos diversificados no texto, enriquecido por valores expressivos ditados pelas conjunções, garantindo mais fortemente a coesão sequencial.

Considerações finais

Os textos das crianças e adolescentes têm o poder de mostrar a complexidade dos recursos linguísticos de que eles fazem uso. Se, aos cinco anos, uma criança, em seu pleno desenvolvimento, é capaz de produzir enunciados simples, mas que contemplem linguisticamente o que se deseja dizer, quanto mais é possível que ao longo do desenvolvimento escolar, esse indivíduo aprenda a produzir enunciados que demandem grande desempenho linguístico. Quando, por algum motivo, o desenvolvimento linguístico, aquele que se aprende por meio de ambiente monitorado, é permeado por

entraves em seu ensino, ter-se-á indivíduos que apresentam um baixo grau de desenvolvimento, no que tange aspectos relacionados à modalidade escrita e oral, uma vez que essas não são independentes, mas configuram como partes de um mesmo sistema.

Neste trabalho, apontaram-se as finalidades pedagógicas de um corpus desenvolvimental, entendendo que um corpus está para além de meros dados estatísticos, mas que é possível se pensar em estratégias de ensino através dos números demonstrados, possibilitando ao professor entender quais as necessidades mais urgentes dos alunos. Ao propor rotinas baseadas em dados, o que consistiu em levantar os dados sobre fatos da língua específicos, apontar as lacunas relativas a eles e transpô-los para a sala de aula, estabelece-se apenas uma proposta de estímulos para os professores, não sendo rotinas fechadas e limitadas, podendo ser moldadas às necessidades do professor e de sua sala de aula.

Nos contextos específicos desta pesquisa, foi possível observar lacunas significativas no que tange o desenvolvimento linguístico dos escritores, observando-se situações que estão em desencontro com o que as convenções da variedade padrão indicam, o que sugere que habilidades que são propostas pela BNCC para cada ciclo de ensino não estão sendo satisfeitas. A culpa dessas lacunas não é do professor. Inúmeras situações, estruturais, sociais e cognitivas, cooperam para que os nossos alunos não desenvolvam completamente a sua capacidade linguística e tornem-se competentes linguisticamente.

Os déficits linguísticos apresentados, a meu ver, tiram a oportunidade desses alunos de se manifestar socialmente, uma vez que dominar as estruturas convencionadas que regem a variedade padrão da língua é sinônimo de ascensão social. Ignorar esses déficits, portanto, é negar aos nossos alunos o direito de transitar livremente em diferentes contextos comunicativos. Portanto, o que se tentou fazer neste trabalho, foi estimular o professor a enxergar em dados que se aproximam de sua realidade novas possibilidades de ensino.

REFERÊNCIAS

- BAKER, P.; HARD, A.; MCENERY, T. **A glossary of corpus linguistics**. Edimburgo: Edinburgh University Press, 2006.
- BECHARA, Evanildo. **Moderna Gramática Portuguesa**. Rio de Janeiro: Nova Fronteira, 2015.
- BERBER-SARDINHA, T. **Linguística de corpus**. São Paulo: Manole, 2004.

BIBER, D.; CONRAD, S.; REPPEN, R. **Corpus Linguistics: Investigating Language Structure and Use**. New York: Cambridge University Press, 1998. Disponível em: <<http://dx.doi.org/10.1017/CBO9780511804489>> Acesso em 06 de outubro de 2021.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, 2018.

CASTILHO, A. T. **O modalizador realmente no português falado**. Revista Alfa. São Paulo, v. 44, 2000, p. 147-169.

CHOUPINA, Celda M. G. M. **Orações Relativas: Como e Quando inseri-las na sala de aula?** Exedra, Coimbra, p. 49-68, mar. 2010.

COSTA, A. L.; COSTA, A.; GONÇALVES, A. Consciência linguística: aspetos sintáticos. Em Maria João Freitas & Ana Lúcia Santos (eds.), **Aquisição de língua materna e não materna: Questões gerais e dados do português**. Berlin: Language Science Press, 2017, p. 409-438.

CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. 5.ed. Rio de Janeiro: Lexikon, 2012.

DUARTE, Inês. **O conhecimento da língua: Desenvolver a consciência linguística**. Lisboa: DGIDC - Ministério da Educação, 2008.

FRIGINAL, Eric. **Corpus Linguistics for English Teachers: New Tools, Online Resources, and Classroom Activities**. New York and London: Routledge, Taylor & Francis, 2018.

HUNSTON, Susan. **Corpora in Applied Linguistics**. Cambridge University Press, 2002.

ILARI, R. **A categoria advérbio na gramática do português falado**. Revista Alfa. São Paulo, v. 51, 2007, p. 151-174.

JOHNS, Tim. **Should you be persuaded: Two samples of data-driven learning materials**. Lexically, 1991. Disponível em: <https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and%20DDL.pdf> Acesso em: 17 de abril de 2021.

JOHNS, Tim. From printout to handout: Grammar and vocabulary teaching in the context of datadriven learning. In: ODLIN, T. (ed.). **Perspectives on pedagogical grammar**. Cambridge: Cambridge University Press, 1994.

KUCERA, H; FRANCIS, W. N. **Computational Analysis of Present-Day American English**. Providence R.I.: Brown University Press, 1967.

MARTINS, Mário; JANSSEN, Maarten; SANTOS, Taiza; et al., 2020, **DOESTE v0.5**, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics

(ÚFAL), Faculty of Mathematics and Physics, Charles University, <http://hdl.handle.net/11234/1-3262>.

MARTINS, Mário. O aprendizado movido por dados na sala de aula de línguas adicionais. In: MARTINS, Mário; ZONI, Martha (org.). **Experiências e reflexões sobre a educação de línguas adicionais**. Macapá: UNIFAP, 2016, p. 107-121.

MCENERY, T.; HARDIE, A. **Corpus linguistics: Method, theory, and practice**. Cambridge: Cambridge University Press, 2012.

MCENERY, T; WILSON, A. **Corpus linguistics**. Edinburgh: Edinburgh University Press, 1996.

MCCARTHY, M. J. **Issues in Applied Linguistics**. Cambridge: Cambridge University Press, 2001.

NEVES, M. H. M. **Gramática de usos do português**. São Paulo: UNESP, 2000.

NIPPOLD, M. A. **Research on later language development**. Disponível em: <[\(PDF\) Later Language development: The school-age and adolescent years \(researchgate.net\)](#)>

Acesso em 30 de outubro de 2021.

O'KEEFFE, A.; MCCARTHY, M.; CARTER, R. **From Corpus to Classroom: language use and language teaching**. Inglaterra: Cambridge University Press, 2007.

PERERA, Katharine. **Children's writing and reading: analysing classroom language**. Oxford: Blackwell, 1984.

QUIRK, R., GREENBAUM, S. LEECH, G.; SVARTVIK, J. **A Comprehensive Grammar of the English Language**. London: Longman, 1985. Disponível em: <SIM-SIM, I. **Desenvolvimento da linguagem**. Lisboa: Universidade Aberta, 1998.

<https://silo.pub/qdownload/a-comprehensive-grammar-of-the-english-language.html>>

SINCLAIR, J. **Trust the text: Language, corpus and discourse**. Nova Iorque: Routledge, 2004.

SUSMICKAT, I. S.; AGUIAR, T. P.; KANTHACK, G. S. **O uso multifuncional de advérbios terminados em -mente: uma análise em textos de alunos do ensino fundamental e médio**. Sergipe: Interdisciplinar, IX, v.20, p. 09-20.

TOGNINI-BONELLI, E. **Corpus Linguistics at Work**. Amsterdam: John Benjamins, 2001.

TRASK, R. L. **Dicionário de Linguagem e Lingüística**. São Paulo: Contexto, 2004

ZAYAS, F. Las actividades gramaticales desde una perspectiva textual. In C. Lomas & A. Osoro (Org.), **El enfoque comunicativo de la enseñanza de la lengua**. Barcelona: Paidós, 1997.