

A CO-EXTENSIVIDADE E ORGANIZAÇÃO LEXICAL NO PORTUGUÊS BRASILEIRO: UMA INTRODUÇÃO DESCRITIVA A PARTIR DE UMA ABORDAGEM SISTÊMICO-FUNCIONAL

Júlia Santos Nunes RODRIGUES*
Giacomo Patrocínio FIGUEREDO**
Francieli Silveria OLIVEIRA***

- RESUMO: Tendo por finalidade identificar princípios de organização do léxico segundo a Abordagem sistêmico-funcional (HALLIDAY; MCINTOSH ; STREVENS, 1964; MARTIN, 1992; HALLIDAY, 2002; FIGUEREDO, 2011), subsidiada por metodologia da Linguística de *Corpus*, o presente artigo explica dois processos de transformação pelos quais o léxico se integra aos sistemas da língua, quando relacionado ao eixo paradigmático nos estratos gramatical e do registro por meio da co-extensividade. Tem-se aqui como motivação o fato de que as concepções acerca do léxico modular – modelado como preenchimento de estruturas gramaticais – ou léxico contíguo – modelado como a gramática delicada – necessitam de maior elaboração quanto aos processos internos relativos à dimensão do léxico, bem como às relações do léxico, para além da gramática, com o discurso e o contexto. Neste artigo, especificamente, investigam-se os processos de transformação da gramática e do registro, incluindo principalmente o desenvolvimento de protocolos para análise de processos internos do léxico, assim como a caracterização de como itens lexicais são demandados por um determinado registro.
- PALAVRAS-CHAVE: funcionalismo; léxico; co-extensividade; transformação da gramática; transformação do registro.

Introdução

Partindo da concepção de léxico das abordagens funcionalistas da linguagem – caracterizada por relações de Lexicalização/Gramaticalização (BASÍLIO, 1991; NEVES, 1997; BORBA, 2003; CASTILHO, 2010), contiguidade (HASAN, 1987;

* Universidade Federal de Minas Gerais (UFMG), Belo Horizonte – MG – Brasil. juliasrodrigues@ufmg.br. ORCID: 0000-0002-7673-1833

** Universidade Federal de Ouro Preto (UFOP), Mariana – MG – Brasil. giacomo.figueredo@ufop.edu.br. ORCID: 0000-0002-0967-244X

*** Universidade Federal de Minas Gerais (UFMG), Belo Horizonte – MG – Brasil. francielioliveira@ufmg.br. ORCID: 0000-0002-4895-1919

HALLIDAY, 2008) e contextualização (HASAN, 1985; ROSE, 2006; HAO, 2015) – e dos subsídios metodológicos da Linguística de *Corpus* (PEARSON, 1998; SINCLAIR, 2004; BIBER; CONRAD; CORTES, 2004; MOON, 2010), o presente artigo investiga a relação complementar entre as dimensões linguísticas do léxico e da organização sistêmica da língua (MATTHIESSEN, 1995; HALLIDAY, 2002; MARTIN, 2013). Tomando com fenômenos a serem investigados (i) a organização interna do léxico e (ii) a co-extensividade entre o léxico e a dimensão da estratificação do sistema linguístico, o artigo apresenta, mais especificamente, um conjunto de mecanismos que caracteriza o modo pelo qual o léxico interage com a organização sistêmica, ao mesmo tempo que identifica as características próprias da constituição interna da dimensão do léxico.

A relação entre léxico e as outras partes do sistema linguístico permite a compreensão de fenômenos complexos de Lexicalização (BASÍLIO, 1991), Gramaticalização (KURYLOWICZ, 1975; HOPPER; TRAUGOTT, 1993), delicadeza (HASAN, 1987) e realização (MARTIN; WHITE, 2005; MATTHIESSEN; TERUYA; LAM, 2010). No entanto, esta compreensão se vê majoritariamente restrita ao estrato gramatical e, muito particularmente, à uma unidade desse, que é a palavra. Essa restrição possibilita inclusive a hipótese da relação unívoca entre um item lexical e uma palavra (cf. SINCLAIR, 2004; CASTILHO, 2010). Por outro lado, quando a produção de significado em contextos sociais é considerada, conforme pressupostos funcionalistas, o aspecto complementar entre a dimensão do léxico e a organização sistêmica passa a ser motivado pela produção de sentido, o qual inclui a criação, organização e distribuição de significados nos textos segundo os domínios experienciais das atividades humanas e as relações sociais (HAO, 2015; ROSE, 2019).

Por conseguinte, um item lexical sempre está relacionado à gramática e, da mesma forma, à semântica discursiva, o que torna a interpretação de um item lexical sempre inter-estratos (HALLIDAY, 2008). Isto pode ser verificado pelo papel importante que o léxico ocupa conjuntamente com sistemas discursivos, tais como a IDENTIFICAÇÃO realizada na gramática pela DÊIXIS NOMINAL, quando ambas são co-extensivas a itens lexicais (ROSE, 2019). O mesmo acontece em determinados subsistemas da COESÃO LEXICAL e de algumas opções de SUBSTITUIÇÃO OU CONJUNÇÃO, que se estabelecem por meio do léxico, quando existe um padrão covariante entre opções desses sistemas e escolhas de itens lexicais (MARTIN, 1992; HAO, 2015). É nesse sentido que tanto a organização interna do léxico quanto sua co-extensividade à dimensão da estratificação se tornam os fenômenos investigados neste artigo.

Além disso, existe sempre uma correlação entre um item lexical e o contexto no qual ele é instanciado, conforme o modo pelo qual os significados são criados, organizados e distribuídos. Esse tipo de correlação é amplamente investigado pela Linguística de *Corpus*, a qual aponta que unidades lexicais ganham significado quando se relacionam com outras unidades, quer no mesmo texto (i.e., cotexto), quer na sua frequência distribuída em diferentes textos (i.e., contexto) (PEARSON, 1998; SINCLAIR, 2004; MOON, 2010). Do ponto de vista sistêmico-funcional o contexto é concebido como um estrato dos sistemas linguísticos responsável por alinhar a variação das situações

sociais à variação linguística. Assim, a funcionalidade encerra também um caráter probabilístico que abordamos de forma quantitativa. É nesse sentido que a metodologia da Linguística de Corpus vai ao encontro da investigação funcionalista.

Essa correlação permite, por conseguinte, que se considerem variáveis contextuais de registro e a distribuição na investigação do léxico. Ela, de maneira semelhante, também pode ser verificada a partir do registro, quando cadeias de itens lexicais são particulares a determinados tipos de campo, o que condiciona a forma pela qual as sequências de atividades são apresentadas e estabelecidas no gênero (HASAN, 1985; PEARSON, 1998; HAO, 2020).

Assim, o artigo admite dois objetos adjacentes, a saber, a dimensão do léxico e sua interação com a dimensão da estratificação do sistema linguístico, incluindo: (i) o estrato da gramática, por meio do contínuo de Lexicalização/Gramaticalização, a formação de itens lexicais e o compartilhamento com unidades da gramática tais como a palavra, o grupo e o morfema, além dos processos de transformação de gramática; e (ii) o estrato do registro, revelando como os itens lexicais são, efetivamente, criados a partir de demandas do campo relacionado à construção da experiência humana no processo de transformação de registro.

No que tange à organização interna do léxico, o artigo, ainda, contribui com a elucidação de como se dá a distribuição dos diferentes conjuntos de itens lexicais por tipo de texto, relacionando a dimensão lexical não apenas ao registro em termo de sua configuração, mas também ao emprego dos itens – o que, por sua vez, também se constitui como outra forma de produzir significado.

Em um primeiro momento, a relação entre o léxico (que aqui é tratado como mais uma dimensão do sistema linguístico; sendo as outras dimensões: estratificação, instanciação, metafunção, eixo e individualização) e a estratificação (no escopo deste artigo, limitada aos estratos da gramática e do registro) pode ser caracterizada a partir de duas perspectivas, às quais denominam-se modular e contígua (cf. HASAN, 1985; HALLIDAY, 2008). É importante destacar como as concepções de léxico e item lexical são tratadas no artigo a partir de uma perspectiva sistêmico-funcional. Assim, “léxico” é a dimensão, ao passo que ‘item lexical’ é uma instância do léxico.

A perspectiva modular trata o léxico como um módulo separado do restante da estratificação – em particular, referindo-se à gramática (cf. SINCLAIR, 2004) – de forma que a gramática responde pela organização estrutural da língua, ao passo que ao léxico cabe preencher a estrutura com sentido por meio dos itens lexicais (cf. MARTIN, 1992).

A perspectiva contígua, por sua vez, entende que léxico e estratificação – novamente, concentrada na gramática – são, na verdade, o mesmo estrato denominado lexicogramática e que, portanto, compreendem o mesmo fenômeno de produzir significado. A diferença visível entre um item lexical e outro gramatical, segundo essa perspectiva, se deve apenas ao ponto de observação e ao nível de delicadeza em um sistema a partir do qual a observação é conduzida (HASAN, 1987; HALLIDAY, 2008).

Do ponto de vista funcional e da produção de significado, a perspectiva modular apresenta alguns problemas que, em face do escopo do presente artigo, incluem: (i)

o mecanismo pelo qual as estruturas são preenchidas pelos itens lexicais – ou, em que momento da produção isso acontece; as regras lexicais (e não gramaticais) que determinam este preenchimento; os princípios gerais de Lexicalização/Gramaticalização nessa perspectiva; (ii) a correlação entre um item lexical e suas propriedades sintáticas (assumindo-se que os itens lexicais possuem, de alguma forma, propriedades gramaticais); (iii) a possibilidade de um item lexical possuir significado mesmo estando isolado das relações linguísticas e contextuais.

A perspectiva contígua, por sua vez, apresenta problemas que, restritos ao interesse de pesquisa deste artigo, incluem: (i) a limitação da relação entre o léxico e a gramática, desconsiderando-se a pesquisa sobre os outros estratos do sistema linguístico como semântica discursiva e registro; (ii) os mecanismos de organização interna do léxico relativamente àquela dos sistemas (a qual inclui hierarquia composicional, organização axial, realização, meta-redundância e diversidade metafuncional).

Dessa forma, o artigo busca averiguar em que medida os fenômenos da organização interna do léxico e da sua co-extensividade à estratificação podem ser compreendidos segundo as perspectivas modular e contígua de modo a responder a esses problemas. O artigo traça, assim, os objetivos: (a) apresentar uma alternativa à dicotomia módulo/contiguidade, mostrando mecanismos pelos quais a dimensão do léxico é co-extensiva aos sistemas gramaticais, (b) indicar a forma como os itens lexicais são distribuídos na língua a partir da sua co-extensividade ao registro (contexto), ajudando a compreender de que modo o léxico contribui para a produção de significado.

O léxico e a gramática: perspectiva do módulo

A definição de léxico, seu processo de formação e sua relação com a gramática foram discutidos pela literatura em vários momentos do desenvolvimento dos estudos linguísticos (HALLIDAY; HASAN, 1976; NEVES, 1997; SINCLAIR, 2004, entre outros). Uma das formas mais frequentes de abordagem do item lexical é concebê-lo como uma unidade modular, estando assim associada a um conjunto de morfemas (BASÍLIO, 1991; BORBA, 2003; CASTILHO, 2010). Contudo, essa concepção implica na relação direta entre o léxico e um estrato específico – a gramática – e, dentro desse estrato, uma unidade gramatical particular – a palavra. Dessa maneira, apesar de compartilharem recursos de produção de significado, o item lexical e a palavra operam papéis distintos na organização da língua.

Uma das formas significativas de conceber o módulo é por meio da vertente funcional-cognitivista que descreve o léxico, primeiramente, pelo viés psicológico/individual. Ele é compreendido como “um inventário (i) de categorias e subcategorias cognitivas; e (ii) de traços semânticos inerentes” (CASTILHO, 2010, p.110). Esse inventário integra propriedades que utilizamos para a criação de palavras, portanto, a palavra é formada a partir do léxico de uma língua (CASTILHO, 2010).

O item lexical é distinto da palavra morfológica (+núcleo; +afixos), fonológica (+sílabas tônicas; + sílabas átonas) ou gráfica (espaços em branco ao redor de um elemento gráfico, cf. *token*¹). Ele pode constituir-se de mais uma palavra em sua totalidade, como também, pode ser formado por uma única palavra (CASTILHO, 2010). Isso pode ser verificado por meio da análise de colocados que têm relação direta com significado do item lexical (SINCLAIR, 2004; VIANA; TAGNIN, 2010).

Outra forma pertinente de compreender o módulo é mediante a perspectiva coletiva/social da Linguística de *Corpus*. Por essa vertente, o significado construído pelo léxico é abordado na sua determinação funcional, uma vez que o item lexical “realiza uma unidade de significado, que implica função desse item tanto em seu cotexto quanto em seu contexto” (SINCLAIR, 2004, p.121, our translation).²

Mais especificamente, o plano autônomo da língua, segundo a concepção da língua em uso de forma dinâmica, é o modelo responsável pela organização textual. Esse plano responde pela organização linguística da experiência humana que é registrada na forma de textos. Com isso, é nesse plano que se estabelece uma relação complexa entre o item lexical e a forma como ele ganha significado. A relação é complexa porque, sendo a estrutura do texto o modo como a língua se desenvolve enquanto um repositório de experiência humana, um item lexical do texto depende de outros elementos (i.e., cotexto) desse repositório para fazer sentido (PEARSON, 1998). Da mesma maneira, depende desses mesmos itens estarem registrados em outros textos do repositório (i.e., contexto) para serem aceitos, compreendidos e compartilhados pela comunidade de falantes (SINCLAIR, 2004).

Portanto, Sinclair (2004) aponta a relação entre a Linguística de *Corpus* e a Linguística Funcionalista, visto que esta última ocupa-se da investigação estrutural juntamente com as funções linguísticas que operam em um determinado contexto. Essa é a abordagem utilizada neste estudo, em particular, a junção entre a metodologia utilizada pela Linguística de *Corpus* em confluência com as categorias teóricas e descritivas da Linguística Sistêmico-Funcional.

Processos de expansão lexical: Lexicalização e Gramaticalização

A partir desse ponto, o item lexical (i) se relaciona às outras unidades da gramática, sendo concebido de forma psicológica e individual, e (ii) depende igualmente da semântica discursiva, do cotexto e do contexto para ganhar sentido, sendo compreendido de maneira coletiva e social. Dessa forma, podem-se identificar duas perspectivas para o processo de expansão lexical: Lexicalização e Gramaticalização.

¹ *Token* é um conjunto de caracteres entre espaços em branco que, independentemente de seu significado semântico, é reconhecido e contabilizado pelo computador (VIANA; TAGNIN, 2010).

² No original: “realize an element of meaning which is the function of the item in its cotext and context” (SINCLAIR, 2004, p.121).

Pela perspectiva da Lexicalização, o processo de expansão lexical é iniciado por impulsos mentais transformados em ondas sonoras. Ela pode ter início a partir da motivação individual dos falantes em um primeiro momento que, na sequência, se torna um processo negociado no decorrer das interações humanas com o objetivo de suprir uma necessidade comunicativa (CASTILHO, 2010). Esse processo é subclassificado como composição e derivação (cf. BASÍLIO, 1991).

Em contraponto, a Gramaticalização, em sua abordagem clássica, é concebida como um processo de mudança que leva itens lexicais a assumirem funções gramaticais ou, se já gramaticais, a assumirem funções ainda mais gramaticais, por alterações em suas estruturas morfológica e semântica (cf. KURYŁOWICZ, 1975; HOPPER; TRAUGOTT, 1993). Portanto, a criação de novos itens linguísticos por meio desse processo é subcategorizada conforme o processo de criação, que se dá pela perspectiva da morfologia lexical ou da morfologia flexional. À primeira, morfologia lexical, cabe estruturar e enriquecer o léxico de uma língua; à segunda, morfologia flexional, compete a imputação de valores gramaticais aos itens (cf. BORBA, 2003). A distinção entre esses dois tipos de morfologia torna-se um dos critérios formais mais importantes para a divisão dos significados em dois grupos: lexical e gramatical (BORBA, 2003).

Dessa forma, o processo de criação de itens lexicais é descrito mediante duas perspectivas: (i) do ponto de vista do impacto sobre o próprio léxico (i.e., Lexicalização); e (ii) do ponto de vista do impacto sobre a gramática (i.e., Gramaticalização). A morfologia lexical, classificada como subcategoria da Gramaticalização, possibilita a formação de novos itens linguísticos que são organizados pelo mecanismo da derivação e composição. No entanto, esses mecanismos compõem o processo da Lexicalização, ocorrendo, assim, um paralelo entre os processos de Lexicalização e de Gramaticalização.

Por conseguinte, a formação de itens lexicais percorre o âmbito lexical, como também o gramatical. Esse processo pode ocorrer partindo tanto de um item lexical como de outro gramatical; conseqüentemente, o processo Lexicalização-gramaticalização é apresentado através do viés da contigüidade (KURYŁOWICZ, 1975; HOPPER; TRAUGOTT, 1993).

O léxico e a gramática: perspectiva da contigüidade

Nessa perspectiva, léxico e gramática são compreendidos como conjuntos abertos lexicais e sistemas fechados gramaticais (HALLIDAY; MCINTOSH; STREVEN, 1964; HALLIDAY, 2002). A relação estabelecida entre esses dois pólos seria de contínuo, uma vez que as construções linguísticas não são escolhas polarizadas – lexicais ou gramaticais, mas podem pertencer a qualquer ponto no contínuo em que em uma extremidade encontra-se o léxico e na outra a gramática.

Dessa maneira, léxico e gramática são interpretados como sendo parte do mesmo estrato linguístico – a lexicogramática – que possui a delicadeza como seu princípio sistêmico de organização (HALLIDAY; MCINTOSH; STREVEN, 1964; HALLIDAY;

MATTHIESSEN, 2014). Por esse princípio, as unidades da lexicogramática são descritas através da rede de sistemas que é organizada pelos eixos paradigmático e sintagmático. O sistema pode ser detalhado por meio de subsistemas, e quanto mais subsistemas são utilizados na descrição, maior é o seu nível de delicadeza. Assim, o sistema é desenvolvido por níveis de delicadeza. No ponto mais delicado do sistema, localiza-se o léxico e, no ponto menos delicado, a gramática, compondo o contínuo lexicogramatical.

A co-extensividade do item lexical na estratificação

Uma das principais dimensões da organização da língua é a estratificação (HALLIDAY; 2002; HALLIDAY; MATTHIESSEN, 2014). Essa dimensão é formada pelos estratos da expressão: fonologia e grafologia, e pelos estratos do conteúdo: lexicogramática, semântica discursiva. Além da língua, o contexto: registro e gênero³ também compõe parte do modelo em virtude de ser modelado como um sistema semiótico que, portanto, produz significado. O princípio que organiza os estratos linguísticos é a realização, responsável pela relação conjunta entre um padrão linguístico abstrato com um padrão mais complexo.

O léxico e a gramática compõem um estrato linguístico que se localiza no nível mais abstrato que a expressão e no nível menos abstrato que a semântica discursiva. A lexicogramática é constituída por quatro ordens que se relacionam por meio da hierarquia composicional: morfema, palavra, frase/grupo e oração (HALLIDAY; MATTHIESSEN, 1999, 2014).

A palavra é uma ordem da escala de ordens que se encontra entre o morfema e a frase/grupo. Ela constitui uma ordem composta por um ou mais morfemas e possui uma única função na ordem superior (HALLIDAY; MATTHIESSEN, 2014). Desse modo, um elemento da ordem palavra é caracterizado por uma única função que desempenha na ordem frase/grupo.

Nem sempre um elemento da ordem palavra é coextensivo a um item lexical. Dois ou mais elementos podem funcionar como um único item lexical, como, por exemplo, em doenças crônicas e controle metabólico. Cada exemplo apresenta dois elementos da ordem palavra exercendo função de substantivo e adjetivo, todavia, compõem um item lexical, logo, o item lexical pode ser formado por um ou mais elementos da ordem palavra (EGGINS, 2004; MARTIN, 1992). Portanto, um elemento da ordem palavra e item lexical não são termos sinônimos, o item lexical corresponde às palavras de conteúdo ou também chamadas de léxico (MARTIN; ROSE, 2007).

O item lexical integra-se diretamente ao estrato da semântica e, da mesma forma, ao da gramática (HALLIDAY, 2008). No estrato gramatical, por exemplo, relaciona-se não apenas à ordem da palavra, conforme os diferentes parâmetros que a definem,

³ O gênero é o estrato linguístico mais abstrato e consiste na configuração de significados que são frequentes e importantes para uma cultura (MARTIN; ROSE, 2008).

mas com outras ordens – estendendo-se aos elementos do morfema, da palavra e do grupo e, a algumas expressões fixas da língua, até mesmo à oração (Quadro 1). A relação estabelecida entre o item lexical e as ordens da gramática é denominada uma propriedade co-extensiva (HALLIDAY, 2002).

Quadro 1 – Co-extensividade entre itens lexicais e ordens da gramática

ordem gramatical	item lexical	significado
morfema	autocuidado	‘auto-’ = próprio
palavra/composta	açúcar	= alimento
grupo/frase	pé diabético	= sintoma
oração/complexo	se deus quiser	= provavelmente

Fonte: Elaboração própria.

A co-extensividade de um item lexical também pode se expandir aos demais estratos linguísticos. Na co-extensividade com a semântica discursiva, por exemplo, a coesão lexical e a colocação apresentam uma relação entre itens lexicais concebida no texto. A coesão lexical é apresentada pelo sistema semântico discursivo de **RELAÇÕES TAXONÔMICAS** e a colocação é dependente do sistema de **RELAÇÕES NUCLEARES**. Tanto as **RELAÇÕES TAXONÔMICAS** quanto **RELAÇÕES NUCLEARES** pertencem ao primeiro nível de delicadeza do sistema semântico da **IDEAÇÃO** (MARTIN; ROSE, 2007).

Na co-extensividade com o registro, o conjunto de itens lexicais determina o tipo de registro, orientando o desenvolvimento do seu campo, sintonia e modo. O campo apresenta como o registro concebe a experiência humana, sendo subdividido em domínio experiencial e processos sociossemióticos, a sintonia mostra como as relações humanas são encenadas e o modo apresenta como campo e sintonia são configurados de forma a configurar simbolicamente o registro (HALLIDAY; MATTHIESSEN, 2014). Na co-extensividade com o campo, uma frequência de itens lexicais auxilia a determinar um domínio experiencial, assim, o conjunto de itens lexicais aliado à estrutura sistêmica organiza e amplia o registro (NEUMANN, 2013).

A mesma relação ocorre, ainda relativamente ao campo, com os tipos de processos sociossemióticos, os itens lexicais do processo sociossemiótico “explorar” são distintos se comparados com o processo sociossemiótico “relatar”. Os itens lexicais do primeiro apresentam maior densidade lexical (cf. URE, 1971) com relação ao segundo. Isso também é evidenciado no trabalho de Figueredo *et al.* (2019), por meio do Protocolo de Investigação Lexical: Processo de Generalização, cujo enfoque recai sobre a co-extensividade entre tipos de itens lexicais e processos sociossemióticos.

Na co-extensividade com a sintonia, os itens lexicais auxiliam no desenvolvimento da relação de proximidade/distanciamento entre os participantes. Quanto mais lexicalmente densa é estabelecida a construção de significados, maior é o distanciamento entre os participantes, determinando a seleção do tipo de sintonia (LIMA, 2013). Desse modo, a

co-extensividade entre item lexical, campo e sintonia apresentada por meio da densidade lexical dos registros estudados em Figueredo *et al.* (2019) revela uma mudança conjunta, ou seja, uma covariação entre o tipo de processo sociossemiótico e o tipo de sintonia.

Na co-extensividade com o modo, os itens lexicais escolhidos para elaboração de textos colaboram na determinação do tipo de modo ao qual pertencem. Itens lexicais nominalizados ou metaforizados compõem textos mais tipicamente do modo escrito em maior número quando comparados com textos configurados no modo falado (HALLIDAY, 2002), uma vez que esses últimos possuem densidade lexical mais reduzida, no entanto, apresentam uma estrutura textual mais intrincada (HALLIDAY, 2002).

Da mesma forma, a co-extensividade entre item lexical e o gênero reside na formação do próprio gênero por meio da seleção de conjuntos de itens lexicais. A escolha de determinados itens lexicais indica o tipo de gênero desenvolvido – podendo ser histórico, descritivo, entre outros (ROSE; MARTIN, 2012; ROSE, 2019), como também, determina a forma que os estágios do gênero são elaborados juntamente com seus propósitos comunicativos (MARTIN; ROSE, 2007). A escolha de um único item lexical, por sua vez, pode indicar um gênero. Por exemplo, o item lexical “aula de gramática” mostra o tipo de aula, como também uma configuração de sequência de atividades específica, que seria diferente da sequência de atividades para o gênero em aula de física.

Problematizando a perspectiva da contiguidade

A perspectiva da contiguidade concebe léxico e gramática como parte do mesmo contínuo. Todavia, os sistemas da chamada lexicogramática privilegiam a descrição gramatical, não mantendo uma relação equilibrada entre ambos os pólos (HASAN, 1987; MARTIN, 1992). O léxico, visto pelo ponto mais delicado do sistema, é sempre interpretado pelo olhar da gramática, e não há um modelo que descreva o léxico em seu âmbito de significação.

Ademais, a gramática é realizada pela estrutura, organizada pela escala de ordens e pelo princípio da abstração, já a organização do léxico está relacionada à probabilidade de ocorrência de um item lexical em determinados registros. Enquanto a gramática é determinista, o léxico é probabilístico (HALLIDAY, 2002), portanto, léxico e gramática possuem organização e natureza distintas, o que também vai de encontro ao princípio da contiguidade, na forma do contínuo lexicogramatical.

A relação entre léxico e gramática pode ser definida como complementaridade, visto que estruturas gramaticais são construídas por itens lexicais, bem como, itens lexicais apresentam colocações em estruturas gramaticais (MATTHIESSEN, 1995; HALLIDAY, 2008). Estrutura e probabilidade trabalham complementarmente na construção e compreensão de qualquer produção linguística. Em outras palavras, um significado lexical sempre é formado por uma porção lexical e uma porção gramatical.

O léxico, além de estar associado à gramática, também é co-extensivo (HALLIDAY, 2002) aos demais estratos da língua, a semântica discursiva, o contexto, registro e gênero. O léxico co-extensivo à gramática mostra que um item lexical não está somente associado à ordem da palavra, mas estabelece relações de significado com as demais ordens. A co-extensividade à semântica discursiva, registro e gênero aponta como certos conjuntos de itens lexicais podem (i) criar cadeias de significação e de coocorrência; (ii) organizar e ampliar domínios experienciais; (iii) diferenciar processos sociossemióticos; (iv) colaborar na construção da relação de proximidade/distanciamento entre os participantes; (v) auxiliar na elaboração do modo de um registro e (vi) indicar o gênero e o desenvolvimento de seus estágios.

Vale acrescentar que um item é lexical quando for passível de generalização (cf. FIGUEREDO *et al.*, 2019). A generalização compõe um dos protocolos que são realizados mediante a aplicação de relações taxonômicas de identificação, classificação e generalização. Sendo assim, a generalização aponta uma diferença de funcionamento entre o item lexical e o item gramatical, enquanto o primeiro é passível de ser generalizado, o segundo é passível de ser expoente de categorias mais abstratas.

Consequentemente, a co-extensividade aos estratos linguísticos e a generalização mostram que o item lexical possui organização própria, uma vez que gera significados distintos da gramática, como também, dos demais estratos. A investigação dessa organização permite que o item lexical seja descrito e mapeado de forma ampla, com o objetivo de compreender, mais precisamente, sua natureza.

Metodologia

Segundo os pressupostos preconizados pela Linguística Sistêmico-funcional (HALLIDAY; MCINTOSH; STREVENSON, 1964; HALLIDAY, 2002), o desenvolvimento dos chamados protocolos de investigação lexical pode ser considerado uma maneira de aprofundamento dos estudos acerca do léxico como conceito teórico no âmbito da língua portuguesa do Brasil (FIGUEREDO *et al.*, 2019). Nesse sentido, torna-se necessário delimitar o conjunto de textos a ser utilizado para a aplicação dos protocolos de investigação lexical, bem como o seu domínio experiencial.

A primeira parte da metodologia deste artigo diz respeito ao *corpus* compilado para aplicação dos protocolos de investigação lexical. A segunda parte explica o desenvolvimento de dois protocolos de investigação lexical. O primeiro protocolo aborda a co-extensividade entre item lexical e o estrato da gramática, e o segundo lida com a co-extensividade entre item lexical e o estrato do registro.

O Corpus

Os textos selecionados para a aplicação dos protocolos de investigação lexical compreendem artigos acadêmicos que retratam pesquisas desenvolvidas no âmbito

do Diabetes Mellitus tipo II. O domínio experiencial (HALLIDAY; MATTHIESSEN, 1999; HAO, 2015) desses textos é, portanto, essa condição crônica, especificamente, o autocuidado em Diabetes Mellitus.⁴

Esses artigos acadêmicos foram extraídos de publicações da área das Ciências da Saúde a partir da busca da palavra-chave “autocuidado em diabetes mellitus” na parte de artigos acadêmicos disponíveis na aba de busca do Google Acadêmico.⁵ Foram coletados 40 artigos originalmente escritos em PB. Esses 40 artigos correspondem à quantidade mínima suficiente para obtenção de resultados significativos em função dos protocolos de investigação propostos. Os artigos foram publicados entre 2010 e 2019, gerando, ao todo, 133.232 *tokens*, sendo que a seleção por artigos desse período de tempo é justificada pela necessidade de textos atuais, capazes de retratar padrões produzidos na última década. A escolha por esse tipo de texto para compor o corpus se deu pela necessidade de lidar com uma linguagem utilizada entre pares, ou seja, o(s) autor(es) e o(s) leitor(es) desse tipo de texto está(ão) no ambiente acadêmico e pertence(m) a uma determinada área do conhecimento, cuja expertise de ambos está localizada. Além disso, por se tratar de uma linguagem utilizada por pares considerados expertos na área de domínio dos artigos, esperava-se encontrar uma grande variedade de itens lexicais, passível de ser utilizado na aplicação dos protocolos de investigação lexical, detalhados adiante.

Cada artigo foi salvo separadamente em um arquivo no bloco de notas e nomeado como Título da Revista_ano de publicação, sendo que para dois ou mais artigos de um determinado periódico publicados no mesmo ano, a nomeação se deu da seguinte forma: Título da Revista_ano de publicação.

É importante destacar que os *abstracts*, quadros, gráficos, anexos, figuras, tabelas e referências de todos os artigos selecionados para o *corpus* foram excluídos no arquivo do bloco de notas, por não tomarem parte do objetivo do trabalho em questão, uma vez que não apresentavam uma quantidade de itens lexicais suficientes para os protocolos de investigação lexical.

O desenvolvimento de um protocolo de investigação lexical

A criação dos protocolos de investigação lexical se dá por meio do processo de produção de significado particular do item lexical no PB, contrapondo-se ao chamado item gramatical. Nesse sentido, a literatura existente acerca do léxico, como, por exemplo, os tipos de coesão lexical, sobretudo o conceito de colocação (cf. HALLIDAY;

⁴ É importante mencionar que este artigo faz parte de uma pesquisa de doutorado em desenvolvimento. Tal pesquisa está localizada na área de Estudos Linguísticos no escopo do Programa de Pós-graduação em Estudos Linguísticos da Faculdade de Letras da UFMG e também no âmbito do Projeto Empoder@ - Protótipo conceitual e metodológico para avaliação de intervenções orientadas ao autocuidado em diabetes, uma parceria entre o Laboratório Experimental de Tradução (LETRA) da FALE/UFMG, a Escola de Enfermagem da UFMG e o Departamento de Estatística do ICEx/UFMG.

⁵ Disponível em: <https://scholar.google.com.br/?hl=pt>

HASAN, 1976; HALLIDAY; MATTHIESSEN, 2014), bem como os termos técnicos conjunto lexical e série apresentados em Halliday, McIntosh e Strevens (1964) e Halliday (2002) são considerados. A seguir, dois dos protocolos de investigação lexical desenvolvidos são detalhados.

Do Protocolo de Investigação Lexical: Processo de Transformação da Gramática - ProtinLex: PTG

O processo metodológico que engloba o ProtinLex: PTG é formado por duas fases, o desenvolvimento e a aplicação do protocolo. No desenvolvimento, observam-se as distinções entre os itens considerados gramaticais e aqueles considerados lexicais, isto é, as características que fazem com que determinados itens do PB se agrupem no conjunto dos itens gramaticais; e outros itens possam ser inseridos no conjunto dos itens lexicais. No que diz respeito à aplicação, os itens considerados lexicais são extraídos do *corpus*, e o protocolo de investigação lexical em questão é testado, a fim de se verificar a relevância de determinado protocolo para a caracterização do item lexical no PB.

Para o desenvolvimento do ProtinLex: PTG, os textos do *corpus* foram importados em formato .txt para o *software* concordanciador AntConc (ANTHONY, 2019), e uma lista de palavras-chave foi gerada pelo software, a partir da inserção da parte escrita monológica do CALIBRA (cf. FIGUEREDO, 2011) como *corpus* de referência. É importante mencionar que como somente o recurso de palavras-chave do software concordanciador foi utilizado, não foi necessário excluir do *corpus* os itens denominados como funcionais.⁶

Após a geração das listas de palavras-chave, a morfologia dos itens considerados lexicais, entendidos, até aquele momento, como aqueles de conteúdo, que estão em oposição aos itens funcionais, foi analisada. Essa análise revelou que os itens considerados lexicais apresentam duas porções (i.e., um conjunto de morfemas). A primeira porção não sofre alterações mesmo quando o item lexical gera outros itens lexicais; essa porção foi chamada de porção lexical, porque compreende a parte que dá significado ao item. A outra porção que sofre alterações conforme as transformações pelas quais aquele item considerado lexical passa é chamada de porção gramatical e refere-se à parte morfológica do item lexical. Assim, diferentemente dos itens entendidos como gramaticais (funcionais), os itens considerados lexicais apresentam a capacidade de desmembramento em porções – a porção lexical e a porção gramatical.

Para a aplicação do ProtinLex: PTG, a lista de palavras-chave gerada na fase de desenvolvimento do protocolo em questão foi utilizada. 12 itens que estavam relacionados ao escopo do domínio experiencial do Diabetes Mellitus tipo II e que apresentavam ocorrências capazes de identificar a existência do Processo de Transformação da Gramática (PTG) foram selecionados para a aplicação do protocolo em questão.

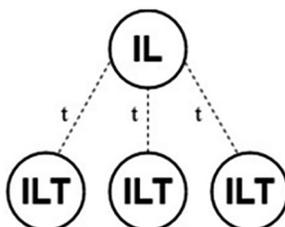
⁶ Itens funcionais são entendidos, nesta pesquisa, como aqueles que estão no escopo dos sistemas fechados, como, por exemplo, preposições, conjunções, artigos e advérbios (cf. HALLIDAY; MCINTOSH; STREVEN, 1964).

Para serem selecionados, os itens devem: (i) aparecer como item lexical e como item lexical transformado (entendido como o item lexical que passou por um processo de transformação⁷ na sua porção gramatical). Isso é caracterizado pelo número de morfemas dos itens, o item lexical deve contar com menos morfemas que o item lexical transformado.

Entretanto, se houver apenas um item lexical transformado para um dado item lexical, a transformação entre esse item lexical e o respectivo item lexical transformado não pode estar relacionada somente aos morfemas de gênero, número e/ou pessoa. A partir da co-extensividade entre item lexical e a ordem da palavra, o acréscimo desses morfemas aos itens lexicais não envolve uma mudança de classe de palavra, ou seja, um item lexical que está em co-extensão com um substantivo, por exemplo, quando adicionado algum desses morfemas – gênero, número e/ou pessoa – continuará em co-extensão com um substantivo.

Nesse caso, o item lexical não deve ser selecionado para a aplicação do ProtinLex: PTG. (i) Apesar de ser uma transformação possível é insuficiente para caracterizar um item lexical como lexical, de fato. (ii) O item lexical deve ter pelo menos um item lexical transformado para assegurar a eficácia do ProtinLex: PTG na caracterização dos itens lexicais no PB. (iii) O item lexical e o item lexical transformado devem estar presentes no *corpus* utilizado e pelo menos um deles deve aparecer na lista de palavras-chave (ver Fig. 1).

Figura 1 – Representação do ProtinLex: PTG



Fonte: Elaboração própria.

A Fig.1 é formada por um círculo localizado na parte superior, onde encontra-se o item lexical (IL) com menor número de morfemas. Em seguida, na parte inferior, tem-se outros círculos, onde estão os itens lexicais transformados (ILT), aqueles provenientes do item lexical e que devem apresentar maior número de morfemas que o respectivo item lexical. O círculo na parte superior é ligado por cada círculo da parte inferior por

⁷ É importante destacar que a noção de transformação empregada neste artigo não está relacionada à concepção de transformação da gramática transformacional ou gerativista de Chomsky (1957). A percepção de transformação adotada no presente trabalho tem a função de evidenciar como o item lexical no português brasileiro pode ser formado por uma porção lexical e outra porção gramatical, indicando que gramática e léxico operam juntas para construção de itens lexicais no PB.

linhas pontilhadas e conta ainda com a letra t minúscula localizada à direita das linhas, indicando que entre os itens apresentados no círculo superior e no inferior ocorreu o PTG (cf. Fig. 3).

É relevante explicar que a seleção pela linha pontilhada para retratar o PTG se dá para indicar a não linearidade de tal processo. Isso significa que esse processo pode ser variável conforme as necessidades do contexto de uma dada língua, no caso o PB. Já os itens lexicais e os itens lexicais transformados foram inseridos dentro de círculos para representar a ciclicidade desses itens. Mesmo sendo itens que podem ser alterados conforme o registro dos textos do *corpus*, eles continuariam sendo considerados itens lexicais, por isso a escolha da mesma forma (círculo) para representar os itens lexicais e os itens lexicais transformados.

Do Protocolo de Investigação Lexical: Processo de Transformação do Registro - ProtinLex: PTR

O ProtinLex: PTR compreende o segundo protocolo apresentado neste artigo. Trata-se de um desdobramento do protocolo anterior, uma vez que o PTG é uma parte inerente ao Processo de Transformação do Registro (PTR).

O desenvolvimento do ProtinLex: PTR acontece a partir de uma análise dos resultados obtidos pelo ProtinLex: PTG. Mediante essa análise, o conceito de registro (cf. HALLIDAY, 1978; MATTHIESSEN; TERUYA; LAM, 2010; HALLIDAY; MATTHIESSEN, 2014) é adicionado ao PTG descrito anteriormente. Para além disso, as distinções mencionadas por Taverniers (2008) acerca dos significados conotativos e denotativos preconizados por Barthes (1957) e Hjelmslev (1963), sobretudo a especificação feita por Barthes sobre o significado conotativo de um signo estar atrelado a uma situação em particular, contribuíram para a criação do ProtinLex: PTR.

A aplicação do ProtinLex: PTR é constituída de três fases. A primeira diz respeito à análise dos resultados encontrados através da aplicação do ProtinLex: PTG. Como o PTG é visto como parte inseparável do PTR, uma vez que, sem a realização linguística pelas funções estruturais da gramática não é possível depreender o registro, apenas aqueles itens que foram considerados lexicais de fato pelo PTG estão aptos a serem utilizados no PTR.

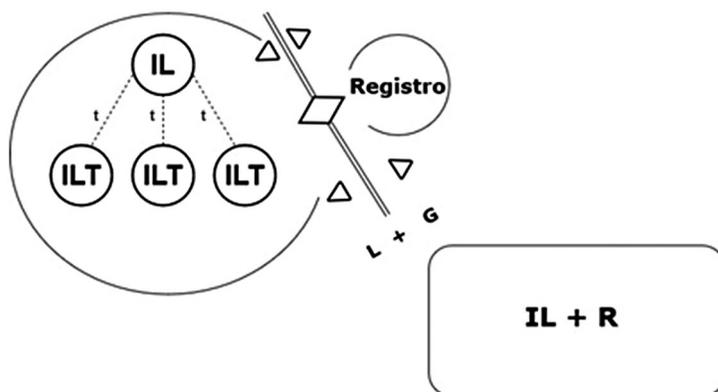
Na segunda fase, após a seleção dos itens lexicais que podem ser úteis para a aplicação do ProtinLex: PTR, as linhas de concordância de cada um dos itens lexicais selecionados foram exploradas, a fim de se encontrar padrões de coocorrência. Esses padrões precisavam ter uma regularidade de frequência que fosse capaz de atestar a confiabilidade de tais padrões, por isso o item lexical x precisa ter no mínimo 3 coocorrências com o item y ou y + z, dependendo do número de componentes do item lexical observado.

Na terceira e última fase da aplicação do ProtinLex: PTR, os padrões de coocorrência são isolados e analisados para cada um dos itens lexicais, verifica-se: (i) a quantidade

de componentes que formam os padrões, sendo que o pressuposto é que exista dois ou três componentes coocorrendo com o item lexical selecionado na primeira fase; (ii) como são esses componentes que coocorrem com o item lexical selecionado anteriormente, por exemplo, qual a classe de palavra que pertence, qual a função que exerce nesse item lexical formado por mais de um componente, é influenciável pelo tipo de registro dos textos do *corpus*; e (iii) se esses itens lexicais formados por mais de um componente são co-extensivos à gramática, ou seja, se são itens lexicais que contribuem para a construção de significado também na co-extensividade relacionada à ordem do grupo nominal.

A partir deste processo de co-extensividade, esses itens lexicais passam a ser itens lexicais que emergiram nos textos do *corpus*, por mudanças nas variáveis do registro – campo, sintonia e modo – que dizem respeito ao tipo de texto (ver Fig. 2).

Figura 2 – Representação do ProtinLex: PTR



Fonte: Elaboração própria.

A Fig. 2 traz a representação do esquema proposto para o ProtinLex: PTR. O semicírculo maior que está na parte central corresponde à etapa inerente ao PTR que é o PTG, explicado previamente. O semicírculo menor localizado na parte superior à direita refere-se ao registro, o qual pelas mudanças nas variáveis do registro pode influenciar a ocorrência de determinados itens lexicais que emergiram dessas mudanças. Na representação, esses itens emergentes, cuja representação se dá pela fórmula IL (item lexical) + R (registro), são apresentados dentro do retângulo com pontas arredondadas que está abaixo da inscrição L + G. Essa inscrição, por sua vez, simboliza a co-extensividade existente entre léxico (L) e gramática (G), uma vez que os itens lexicais emergentes da influência do registro precisam ser formados por dois ou mais componentes, marcando a co-extensividade desses itens lexicais com a ordem do grupo nominal.

A intervenção do registro no aparecimento de determinados itens lexicais acontece em condições indefinidas de tempo. Esse fato indica que alguns itens lexicais podem emergir de maneira mais rápida, significando que as mudanças nas variáveis do registro aconteceram de forma mais intensa, mas outros itens lexicais podem demorar mais para emergirem, sugerindo que as alterações nas variáveis do registro ocorreram de forma espaçada.

Por esse motivo, não há uma simetria na localização dos triângulos na representação acima, ou seja, pode haver influências do registro sobre um componente do item lexical co-extensivo ao grupo nominal, por isso há triângulos que estão mais próximos ou mais distantes à primeira barra. Da mesma forma, há triângulos que também estão mais próximos ou mais distantes à segunda barra, mostrando, nesse caso, que o(s) outro(s) componente(s) desse item lexical exerce(m) uma força maior. Quando a influência do registro no ProtinLex: PTG termina, chega ao fim também o processo de co-extensividade entre léxico e gramática, emergindo um item lexical influenciado pelo registro, na Fig. 2, a co-extensividade é representada pela forma de um losango.

Por fim, é importante salientar que a escolha por semicírculos, que aparecem em paralelo na imagem, pode ser justificada pela necessidade de se indicar que no protocolo em questão os dois conjuntos – PTG e registro – estão interligados, ou seja, quando há mudanças na(s) variável(eis) de um tipo de registro, o ProtinLex: PTG também sofre alterações. Por outro lado, como essas mudanças ocorrem em um prazo de tempo indefinido, os semicírculos não estão unidos.

Existe, no entanto, um espaço entre eles que é responsável por representar a relação léxico (L) e gramática (G), sendo que a forma que simboliza os componentes que constituem os itens lexicais emergidos pela influência do registro nos textos do *corpus* é o triângulo, pois quando dois triângulos são unidos, isto é, quando ocorre o processo de co-extensividade⁸ entre o item lexical a ser emergido e a ordem do grupo nominal, uma nova forma é gerada, o losango, que demonstra que a co-extensividade entre o item lexical emergido e o grupo nominal.

Resultados

Da aplicação do ProtinLex: PTG – *corpus* de artigos acadêmicos

Como mencionado na Metodologia, 12 itens do *corpus* de artigos acadêmicos que estão no escopo do domínio experiencial do Diabetes Mellitus tipo 2 foram selecionados para a aplicação do ProtinLex: PTG (ver Quadro 2).

⁸ É importante mencionar que o ProtinLex: PTR lida com o conceito de co-extensividade em relação ao item lexical e a dimensão estrutural (eixo sintagmático) – escala de ordens da gramática – oração, grupo, palavra e morfema, particularmente a co-extensividade entre item lexical e a ordem do grupo (nominal). Mas há indícios que essa co-extensividade também pode ocorrer entre o item lexical e a estratificação – contexto, semântica, gramática, fonologia e fonética.

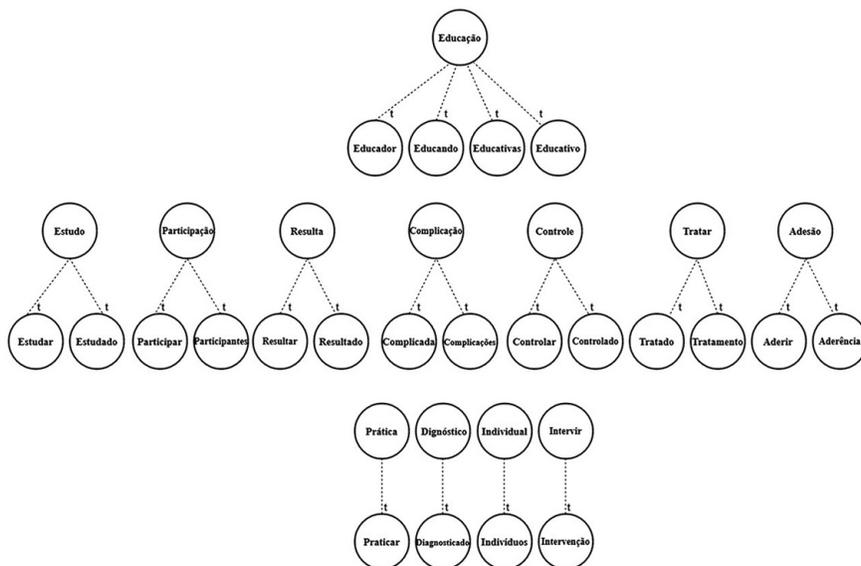
Quadro 2 – Lista de itens selecionados para a aplicação do ProtinLex: PTG – *corpus* de artigos acadêmicos

Número de ocorrências	Chavicidade	Itens retirados da lista de palavras-chave: <i>Corpus</i> de artigos acadêmicos
645	233,91	Estudo
397	149,26	Controle
288	128,81	Educação
276	121,53	Participantes
346	116,14	Tratamento
286	115,86	Resultados
256	112,71	Complicações
225	99,05	Indivíduos
194	85,4	Prática
184	80,99	Diagnóstico
183	80,55	Adesão
175	77,03	Intervenção

Fonte: Elaboração própria.

O Quadro 2 destaca alguns itens retirados da lista de palavras-chave gerada pelo concordanceador a partir da inserção do *corpus* de artigos acadêmicos e da utilização da parte escrita monológica do CALIBRA como como *corpus* de referência. A seleção desses itens é explicada pelo fato de serem itens frequentes no *corpus* utilizado, o que é explicado pela chavicidade detalhada na segunda coluna do Quadro 2, bem como por apresentarem ocorrências que contribuem para reforçar a capacidade de transformação desses itens (ver Fig. 3).

Figura 3 – Representação do ProtinLex: PTG – *corpus* de artigos acadêmicos



Fonte: Elaboração própria.

A Fig. 3 mostra a representação do Protocolo de Investigação Lexical: ProtinLex: PTG com as 12 ocorrências selecionadas a partir da lista de palavras-chave do *corpus* de artigos acadêmicos. A partir da representação, todos os itens selecionados – Estudo, Participantes, Resultado, Complicações, Controle, Tratamento, Adesão, Educação, Prática, Diagnóstico, Indivíduos e Intervenção – apresentam pelo menos um outro item relacionado àquele que foi selecionado. Isso indica que todos os itens, seja o item lexical, seja o item lexical transformado, são itens lexicais pela aplicação do protocolo.

O número de itens lexicais transformados pode ser diferente para um dado item lexical. Há um item lexical – Educação – localizado na parte superior da Fig. 3 – que obteve quatro itens lexicais transformados. Os itens lexicais com dois itens lexicais transformados – Controle, Tratar, Adesão, Estudo, Participação, Resulta e Complicação – encontram-se na parte central da representação. Os itens lexicais com somente um item lexical transformado – Prática, Diagnóstico, Individual e Intervir – estão na parte inferior da Fig. 3.

Esse número de itens lexicais transformados pode ser variável conforme o tamanho do *corpus* utilizado, pois, como descrito na Metodologia, pelo menos um dos itens, o lexical ou o lexical transformado, deve aparecer na lista de palavras-chave. Condicionalmente, aquele que não estiver na lista de palavras-chave deve estar presente no *corpus*. Portanto, dependendo do tamanho do *corpus*, o número de itens lexicais transformados a partir de um determinado item lexical pode ser maior ou menor.

A representação demonstra também que metade dos itens selecionados mediante a lista de palavras-chave – Participantes, Resultado, Complicações, Tratamento, Indivíduos e Intervenção (cf. Quadro 2) – não são os itens lexicais (IL) com o menor número de morfemas – isto é, não ocupam a posição do círculo superior, sendo considerados itens lexicais transformados (ILT), localizados nos círculos inferiores.

Ressalta-se que esse resultado se restringe ao *corpus* desta pesquisa. Assim, em outro *corpus*, os itens que são ILTs podem ser ILs, uma vez que esses itens também podem ser entendidos como itens lexicais dependendo do processo de transformação a partir do qual o texto constrói os significados lexicais. A seguir, o Quadro 3 traz a segmentação em porções lexicais e em porções gramaticais de todos os itens lexicais da Fig. 3.

Quadro 3 – Identificação da porção lexical e da gramatical das ocorrências extraídas a partir da aplicação do ProtinLex: PTG – *corpus* de artigos acadêmicos

Item Lexical	Item Lexical Transformado		Porção Lexical	Porção Gramatical
Prática	Praticar		Pratic	a/ar
Diagnóstico	Diagnosticado		Diagnostic	o/ado
Individual	Indivíduos		Individ	ual/uos
Intervir	Intervenção		Interv	ir/enção
Estudo	Estudar	Estudado	Estud	o/ar/ado
Participação	Participar	Participantes	Particip	ação/ar/ntes
Resulta	Resultar	Resultado	Result	a/ar/ado
Complicação	Complicada	Complicações	Complic	ação/ada/ações
Controle	Controlar	Controlado	Control	e/ar/ado
Tratar	Tratado	Tratamento	Trat	ar/ado/amento
Educação	Educador	Educando	Educ	ação/ador/ando/ativas/ativo

Fonte: Elaboração própria.

Os dados do Quadro 3 mostram que todos os itens selecionados para a aplicação do ProtinLex: PTG no *corpus* de artigos acadêmicos apresentam uma porção lexical, realizada pelas raízes: Pratic-, Diagnostic-, Individ-, Interv-, Estud-, Particip-, Result-, Complic-, Control-, Trat- e Educ-, bem como uma porção gramatical, realizada pelos morfemas: -a/ar, -o/ado, -ual/uos, -ir/enção, -o/ar/ado, -ação/ada/ações, -e/ar/ado, -ar/ado/amento e -ação/ador/ando/ativas/ativo.

Os dados do Quadro 3 evidenciam que a raiz (porção lexical) dos itens lexicais, destacados na terceira coluna do Quadro 3, não sofrem alterações com a realização do ProtinLex: PTG, por exemplo, a raiz Educ- permanece a mesma para todos os itens lexicais transformados – Educador, Educando, Educativas e Educativo – a partir do

item lexical Educação. As partes que sofrem modificações com o ProtinLex: PTG são os morfemas (porção gramatical) dos itens lexicais transformados.

Essa percepção sugere que para o *corpus* de artigos acadêmicos escritos em PB a porção lexical de cada item lexical utilizado nas aplicações do ProtinLex: PTG é responsável por realizar o significado dos itens lexicais, uma vez que se trata de uma parte permanente mesmo após a transformação dos itens lexicais em itens lexicais transformados. Por outro lado, os morfemas (porção gramatical) de cada item lexical variam a cada item lexical transformado encontrado no *corpus*.

A aplicação do protocolo em questão revela que para um item ser considerado um item lexical no PB, precisa apresentar uma porção lexical e outra gramatical, não havendo itens lexicais constituídos somente de uma porção lexical, ou formados apenas por uma porção gramatical. Esse pressuposto assinala que léxico e gramática podem ser concebidos como duas porções que operam juntas para a construção de itens lexicais no PB, sendo porções complementares. A porção gramatical, por sua vez, pode ser entendida como a responsável pelo processo de transformação de um item lexical em outros itens lexicais, uma vez que é essa porção que admite mudanças ao longo do processo de transformação.

Nesse sentido, a análise dos dados obtidos a partir da aplicação deste protocolo indica ainda que, quando o ProtinLex: PTG gera um item lexical transformado, a gramática insere as suas porções gramaticais, os morfemas, em porções lexicais, desenvolvendo um item lexical (ou mais de um item lexical) para suprimir uma necessidade de alguma (s) variável (is) do registro. A partir disso, a quantidade de itens lexicais transformados em um texto aumenta, gerando significados antes inexistentes. Assim, ao longo do tempo e das necessidades emergentes, a quantidade de itens lexicais transformados na própria língua, no caso, o PB, também pode aumentar.

Da aplicação do ProtinLex: PTR – *corpus* de artigos acadêmicos

Analisando as linhas de concordância do *corpus* de artigos acadêmicos para cada um dos itens lexicais a serem utilizados para a aplicação do ProtinLex: PTR, há padrões de coocorrência que se repetem com um mesmo item lexical (ver Quadro 4).

Quadro 4 – Itens selecionados para aplicação do ProtinLex: PTR – *corpus* de artigos acadêmicos

Item encontrado na lista de palavras-chave	Itens que coocorrem com aquele da lista de palavras-chave
Estudo	descritivo – transversal – multicêntrico – qualitativo – prospectivo – experimental
Educação	em diabetes – em DM – em grupo – em saúde – alimentar – individual – física – nutricional – terapêutica
Participantes	do estudo – do grupo – da pesquisa – da intervenção – diabéticos
Tratamento	do diabetes – do DM – da doença – medicamentoso – farmacológico – hemodialítico – dialítico – dietético – ambulatorial – conservador
Resultados	do estudo – da hemoglobina glicada – dos exames laboratoriais satisfatórios – semelhantes – favoráveis – positivos – clínicos – efetivos
Complicações	agudas e crônicas – autorreferidas do diabetes – crônicas do diabetes/DM – da doença – em membros inferiores – microvasculares
Indivíduos	com diabetes – com diagnóstico de DM2 – com DM – diabéticos
Prática	de atividade física – de/do autocuidado – clínica – assistencial – educativa – regular de atividade física
Diagnóstico	do diabetes – de diabetes – do DM – de DM – da doença – médico
Adesão	ao autocuidado – ao tratamento – dos pacientes – terapêutica
Intervenção	educativa – educacional – em grupo – individual – telefônica – nutricional

Fonte: Elaboração própria.

Os dados do Quadro 4 mostram que os itens da primeira coluna, que podem ser caracterizados como itens lexicais pela aplicação do ProtinLex: PTG apresentam um padrão de coocorrência⁹ com os itens da segunda coluna. O item lexical Estudo, por exemplo, coocorre no *corpus* de artigos acadêmicos com os itens: descritivo – transversal – multicêntrico – qualitativo – prospectivo – experimental, formando os seguintes itens lexicais: Estudo descritivo – Estudo transversal – Estudo multicêntrico – Estudo qualitativo – Estudo prospectivo e Estudo experimental. O Quadro 5 a seguir

⁹ Por padrão entende-se que os grupos nominais formados a partir das coocorrências com os itens lexicais – Estudo, Educação, Participantes, Tratamento, Resultados, Complicações, Indivíduos, Prática, Diagnóstico, Adesão e Intervenção – não contam com apenas uma ocorrência no *corpus* de artigos acadêmicos. Para ser considerado padrão, portanto, é necessário que haja três ou mais ocorrências dos itens lexicais que são co-extensivos aos grupos nominais, isto é, os itens lexicais da primeira coluna do Quadro 4 precisa apresentar no mínimo três ocorrências com cada um dos outros itens destacados na segunda coluna do mesmo quadro.

separa as funções do item lexical Estudo a partir de suas coocorrências no *corpus* que estão em co-extensão à ordem do grupo nominal.

Quadro 5 – Separação das funções lexicais e das funções gramaticais a partir da co-extensividade do item lexical Estudo à ordem do grupo nominal

Funções lexicais		Funções gramaticais		
Item encontrado na lista de palavras-chave	Itens que coocorrem com aquele da lista de palavras-chave	Co-extensividade com o grupo nominal	Segmentação do item lexical em co-extensividade com o grupo nominal	
			Ente	Classificador
Estudo	descritivo	Estudo descritivo	Estudo	descritivo
	transversal	Estudo transversal		transversal
	multicêntrico	Estudo multicêntrico		multicêntrico
	qualitativo	Estudo qualitativo		qualitativo
	prospectivo	Estudo prospectivo		prospectivo
	experimental	Estudo experimental		experimental

Fonte: Elaboração própria.

A primeira parte, as duas colunas da esquerda, do Quadro 5 exhibe as funções lexicais que são realizadas pelo item lexical Estudo e pelos outros itens lexicais que coocorrem com ele no *corpus*. A segunda parte do Quadro 5, as duas colunas na direita, traz as funções gramaticais dos grupos nominais que são co-extensivos aos itens lexicais da primeira parte.

Todos os itens lexicais co-extensivos com os grupos nominais retratados na terceira coluna do Quadro 4 apresentam duas funções: um Ente realizado por uma palavra da classe dos nomes, um substantivo: Estudo, e um Classificador realizado também por uma palavra da classe dos nomes, os adjetivos: descritivo – transversal – multicêntrico – qualitativo – prospectivo – experimental.

A segmentação apresentada no Quadro 5 também é adequada para os outros itens lexicais que são co-extensivos à ordem do grupo nominal detalhados no Quadro 4, pois todos têm um Ente, que é co-extensivo aos itens da primeira coluna do Quadro 4 – Estudo, Educação, Participantes, Tratamento, Resultados, Complicações, Indivíduos, Prática, Diagnóstico, Adesão e Intervenção. E, na maior parte das ocorrências, esses itens lexicais também têm um Classificador, que é co-extensivo aos itens da segunda coluna do Quadro 4, sendo que esse Classificador pode ser realizado por um elemento da ordem da palavra (cf. Quadro 5) ou por uma frase preposicional, como, por exemplo, Educação em Diabetes e Participantes do estudo.

É interessante observar que esses resultados corroboram os dados encontrados em Halliday (2002) acerca do fato de que, em inglês, o item lexical pode ser coextensivo

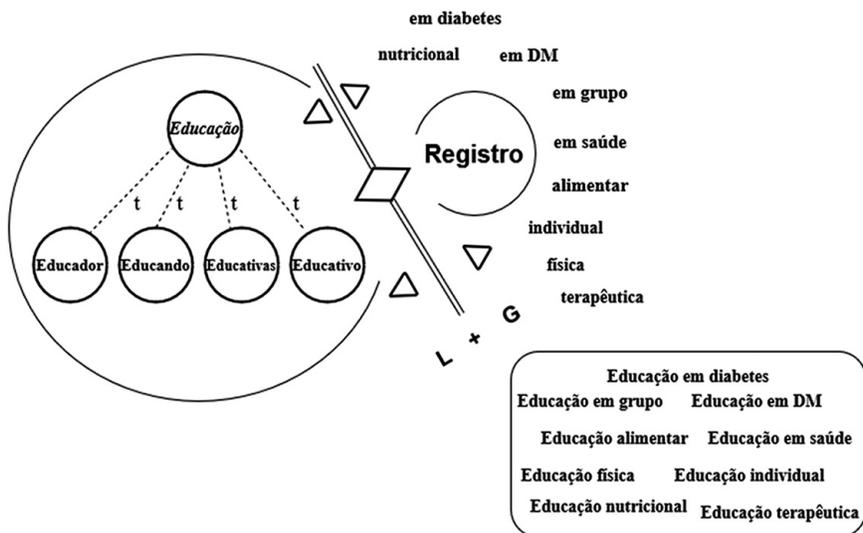
a uma das três ordens da escala de ordens da gramática – grupo, palavra e morfema. Essa perspectiva apresentada em Halliday (2002) também pode ser uma característica possível para os itens lexicais selecionados para a aplicação do protocolo em questão. Contudo, para o PB, os itens lexicais, que são co-extensivos à ordem do grupo (nominal), também estão relacionados às variáveis do registro – campo, sintonia e modo (cf. MARTIN; ROSE, 2008).

Quando os itens lexicais – Estudo, Educação, Participantes, Tratamento, Resultados, Complicações, Indivíduos, Prática, Diagnóstico, Adesão e Intervenção – selecionados para a aplicação do ProtinLex: PTR coocorrem com outros itens, como aqueles apresentados na segunda coluna do Quadro 4, passam a operar juntos no texto em que aparecem, constituindo um novo significado e sendo entendidos como um único item lexical, que funciona de forma co-extensiva à ordem do grupo nominal. Portanto, os itens lexicais selecionados para a aplicação do ProtinLex: PTR constituem outros itens lexicais co-extensivos à ordem do grupo mediante modificações do registro.

Esses resultados sugerem ainda que as coextensões no PB podem fazer com que emergjam itens lexicais na língua em razão de alterações no registro de um conjunto de textos, fazendo com que haja a necessidade de criação de novos significados. Isso significa dizer que a co-extensividade, entendida nesta pesquisa como o item lexical que é extensivo à mais de uma ordem gramatical e a mais de um estrato, é resultado da influência do registro e pode ser responsável por aumentar o potencial de produção de significado da língua (ver Fig. 4).

Além disso, os resultados revelam como a relação entre o processo de produção de significado na língua está imediatamente relacionado às configurações do registro, indicando um aspecto da funcionalidade da língua em seu contexto. Esse fato reforça a maneira pela qual o trabalho quantitativo com o corpus contribui para uma visão funcionalista da co-extensividade entre léxico e registro.

Figura 4 – Representação do ProtinLex: PTR – *corpus* de artigos acadêmicos



Fonte: Elaboração própria.

Na Fig. 4, o semicírculo localizado à esquerda apresenta o ProtinLex: PTG, que é inerente ao ProtinLex: PTR. No caso da Fig. 4, o item lexical utilizado para a representação é Educação, que aparece dentro do círculo superior destacado em negrito e itálico. Esse item lexical foi selecionado para a representação do protocolo em questão, pelo fato de ter sido caracterizado como lexical, após a aplicação do ProtinLex: PTG.

Seguindo os passos metodológicos estabelecidos para o ProtinLex: PTG, o item que tem menor número de morfemas é considerado o item lexical, e aparece no círculo superior, enquanto que os itens que tem maior número de morfemas são entendidos como itens lexicais transformados, e aparecem nos círculos inferiores. Por esse motivo, o item lexical Educação está localizado no círculo superior.

Ressalta-se que o item lexical a ser utilizado para a aplicação do ProtinLex: PTR não precisa necessariamente estar no círculo superior do ProtinLex: PTG. Tal item lexical também pode ser classificado como item lexical transformado, e, ainda assim, ser utilizado para a aplicação do protocolo, uma vez que, de acordo com os parâmetros desenvolvidos para o ProtinLex: PTG, tanto o item lexical como o item lexical transformado podem ser caracterizados como itens lexicais no PB.

Ainda com relação à representação da Fig. 4, os itens que aparecem ao redor do semicírculo do registro compreendem aqueles que coocorrem com o item lexical Educação no *corpus* de artigos acadêmicos. Esses itens variam conforme as variáveis do registro. Isso indica que os itens lexicais formados pela coocorrência de Educação com: em diabetes – em DM – em grupo – em saúde – alimentar – individual – física –

nutricional – terapêutica – podem ser diferentes se o item lexical Educação, por exemplo, aparecer em um outro registro.

As coocorrências, por sua vez, referem-se aos itens lexicais formados por mais de um componente. Em outras palavras, as coocorrências que aparecem dentro do retângulo com pontas arredondadas na representação da Fig. 4 dizem respeito à co-extensividade entre léxico (L) e gramática (G), uma vez que os itens lexicais formados por dois ou mais componentes estão em co-extensividade à ordem do grupo.

É importante salientar, por fim, que os grupos nominais não são analisados nesta pesquisa, porque pertencem a uma ordem da escala da gramática, enquanto que este artigo aborda o léxico como conceito teórico. Para além disso, os grupos nominais são utilizados neste trabalho apenas para mostrar que há no PB itens lexicais formados por mais de um componente, os quais podem ser co-extensivos a mais de uma ordem e que, para tanto, necessitam ser co-extensivos a mais de um estrato (aqui ilustrado pelo registro).

Conclusão

Tendo como base principal a visão funcional sobre o modo pelo qual a língua se organiza – levando assim em conta a interdependência da gramática com o discurso, e deste com o registro (contexto) – este artigo buscou complementar o entendimento dos fenômenos de Lexicalização/Gramaticalização com a funcionalidade do texto em seu contexto, por meio da descrição dos processos de transformação no PB.

Motivado pelos problemas que as visões modular e contígua apresentam na interpretação do léxico, a pesquisa ocupou-se de apontar como o significado de um item lexical não está isolado do restante do sistema ou restrito apenas ao estrato da gramática e à ordem da palavra.

Partindo da necessidade de uma maior compreensão de como o léxico, na condição de dimensão do sistema linguístico, estabelece relações de co-extensividade com o restante do sistema linguístico do PB, bem como da forma como o léxico se organiza internamente, apresentaram-se aqui os mecanismos de co-extensividade do léxico, em particular com os estratos da gramática e do registro.

A observação da coextensividade do léxico no PB foi realizada a partir da compilação de um *corpus* e do desenvolvimento de dois protocolos de investigação lexical, um para a gramática (PTG) e outro para o registro (PTR). Uma vez que o significado do item lexical é construído no texto e na relação entre textos (cotexto e contexto, respectivamente), foi selecionado um *corpus* pertencente a um único domínio experiencial, artigos da área da Ciência da Saúde sobre Diabetes Mellitus.

A aplicação dos protocolos possibilitou, então, a descrição de dois processos de transformação, os quais indicam mecanismos de co-extensividade do léxico com a gramática (ProteinLex: PTG) e com o registro (ProteinLex: PTR). O PTG revelou que o léxico opera junto com a gramática na criação de itens lexicais. Em PB, os itens

possuem uma porção lexical e outra gramatical e, por esse motivo, são capazes de produzir transformações. Este processo gera novos itens lexicais, o que poderia explicar a criatividade linguística demonstrada pela literatura sobre a Lexicalização.

Por sua vez, o PTR possibilitou a observação de como itens lexicais operam em conjunto para formar novos itens lexicais para um registro específico. Esse se constitui como um mecanismo análogo ao processo de Lexicalização/Gramaticalização para o registro, que podemos denominar de lexicalização/contextualização. Para o PB, uma característica da transformação do registro é sua preferência pela realização na ordem do grupo. Assim, a co-extensividade da transformação com o campo é realizada pela co-extensividade da transformação com o grupo nominal.

Os resultados da pesquisa indicam que o registro no PB tem, exatamente como a gramática, um papel fundamental na emergência de itens lexicais da língua em virtude das especificidades para a produção de significados para novos contextos. Dessa forma, o artigo oferece contribuições tanto para os estudos da complementaridade entre léxico e gramática, em dois aspectos.

Primeiramente, mostra como ela pode ser modelada de maneira co-extensiva, o que evita os problemas das visões modular e contígua. Em seguida, demonstra como a co-extensividade se expande para além do estrato gramatical, indo até o registro. O artigo abre ainda a possibilidade para, no futuro, investigar-se a co-extensividade do léxico com o estrato da semântica discursiva e a ampliação para outras configurações de registro.

Embora seja uma pesquisa ainda em desenvolvimento, as seções dos resultados e das conclusões do presente artigo apontam que o item lexical no português brasileiro apresenta uma propriedade chamada de co-extensividade, que é mencionada primeiramente por Halliday (2002), com a escala de ordens da gramática, principalmente, no que diz respeito às ordens da palavra e do grupo. Para além da relação de co-extensividade entre o item lexical e o estrato da gramática, essa propriedade ainda pode ser estabelecida na relação entre o item lexical e o estrato do registro, uma vez que os itens lexicais no português brasileiro podem indicar demandas que surgem de contextos de situação específicos, no caso, o autocuidado em Diabetes Mellitus. Esses fatos caracterizam as relações existentes entre o item lexical (= dimensões linguísticas do léxico) e os estratos da gramática e do registro (ambos organizações sistêmicas da língua).

Agradecimentos

Agradecemos o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq no desenvolvimento desta pesquisa.

RODRIGUES, J.; FIGUEREDO, G.; OLIVEIRA, F. Towards a description of coextensiveness in lexis organization in Brazilian Portuguese. *Alfa*, São Paulo, v.66, 2022.

- *ABSTRACT: This paper aims at identifying principles of lexis organization under a Functional Approach based on corpus evidence. Problems with the modular and delicacy approaches to lexis – i.e., lexis as filling for grammatical structures or most delicate grammar – motivate the paper as they need further elaboration to describe the internal processes of lexis as a dimension of language, as well as lexis relations to discourse and context. The paper goes on to explain transformation processes integrating lexis and language systems, as lexis is coextensive to system in the grammar and register strata. Transformation processes of grammar and register are investigated and lead to the development of protocols of processes internal to lexis. Results show how the protocols can reliably account for the generation of lexical items to a given register. This, in turn, sheds light into lexis internal organization.*
- *KEYWORDS: functional linguistics; lexis; coextensiveness; grammar transformation; register transformation.*

REFERÊNCIAS

ANTHONY, L. **Laurence Anthony Website**. Version 3.5.8. Tokyo: AntConc Homepage, 2019. Disponível em: <https://www.laurenceanthony.net/software/antconc/>. Acesso em: 30 abr. 2020.

BARTHES, R. **Mythologies**. Paris: Seuil, 1957.

BASÍLIO, M. Produtividade, função e Fronteiras Lexicais. In: NEVES, M. H. D. M. **Descrição do Português II**. Araraquara: FCLAR, 1991. p.28-41. (Série Encontros).

BIBER, D.; CONRAD, S.; CORTES, V. If you look at . . . : lexical bundles in university teaching and textbooks, **Applied Linguistics**, Oxford, n.25, p.371–405, 2004.

BORBA, F. D. S. **Introdução aos estudos linguísticos**. 13.ed. Campinas: Pontes, 2003.

CASTILHO, A. T. D. **Nova gramática do Português Brasileiro**. São Paulo: Contexto, 2010.

EGGINS, S. **An introduction to systemic functional linguistics**. 2.ed. London: New York: Continuo, 2004.

FIGUEREDO, G. **Introdução ao perfil metafuncional do português brasileiro: contribuições para os estudos multilíngues**. Tese (Doutorado em Linguística Aplicada) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

FIGUEREDO, G. P. *et al.* O léxico como um recurso linguístico para a produção de significado no texto: um estudo de caso com protocolos de investigação. **Estudos da Língua(gem)**, Vitória da Conquista, v.17, n.3, p.37-59, 2019.

- HALLIDAY, M. A. K. **Complementarities in language**. Beijing: The Commercial Press, 2008.
- HALLIDAY, M. A. K. **On grammar**. London: New York: Bloomsbury Publishing, 2002. v. 1. (Collected Works of M. A. K. Halliday).
- HALLIDAY, M. A. K. **Language as social semiotic The social interpretation of language and meaning**. London: Edward Arnold, 1978.
- HALLIDAY, M. A. K.; HASSAN, R. **Cohesion in English**. London: New York: Longman, 1976.
- HALLIDAY, M. A. K.; MATTHIESSEN, C. M. I. M. **An introduction to functional Grammar**. 3.ed. London: Routledge, 2014.
- HALLIDAY, M. A. K.; MATTHIESSEN, C. **Construing experience as meaning: a language based approach to cognition**. London: Cassell, 1999.
- HALLIDAY, M. A. K.; MCINTOSH, A.; STREVENS, P. **The linguist sciences and language teaching**. London: Longmans, 1964.
- HAO, J. **Analysing scientific discourse from a systemic functional linguistic perspective: a framework for exploring knowledge-building in biology**. New York: Taylor & Francis, 2020.
- HAO, J. **Construing biology: an ideational perspective**. Tese (Doutorado em Filosofia) -University of Sydney, Sydney, 2015.
- HASAN, R. Lending and borrowing: from grammar to lexis. **Beiträge zur Phonetik und Linguistik**, Trier, v.48, p.56–67, 1985.
- HASAN, R. The grammarian's dream: lexis as most delicate grammar. *In*: HALLIDAY, M. A. K.; FAWCETT, R. P. **New developments in systemic Linguistics**. London: New York: Frances Pinter, 1987. v. 1: Theory and Description, 1987.
- HOPPER, P. J.; TRAUGOTT, E. **Grammaticalization**. 1.ed. Cambridge: Cambridge University, 1993.
- HJELMSLEV, L. **Prolegomena to a Theory of Language**. Madison: University of Wisconsin Press, 1963.
- KURYŁOWICZ, J. The evolution of grammatical categories. *In*: KURYŁOWICZ, J. **Esquisses linguistiques II**. [S. l.: s. n.], 1975. p.38-54.
- LIMA, K. C. S. D. **Caracterização de registros orientada para a produção textual no ambiente multilíngue: um estudo baseado em corpora comparáveis**. Tese (Doutorado em Linguística Aplicada) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

- MARTIN, J. R. **English text system and structure**. Philadelphia: John Benjamins, 1992.
- MARTIN, J. R. **Systemic functional grammar: a next step into the theory - axial relations**. Beijing: Higher Education, 2013.
- MARTIN, J. R.; ROSE, D. **Working with discourse meaning beyond the clause**. London: New York: Continuum, 2007.
- MARTIN, J. R.; ROSE, D. **Genre relations: mapping culture**. London: Equinox, 2008.
- MARTIN, J. R.; WHITE, P. R. R. **The language of evaluation: appraisal in English**. New York: Palgrave Macmillan, 2005.
- MATTHIESSEN, C. **Lexicogrammatical cartography: English systems**. Tokyo: International Language Science, 1995.
- MATTHIESSEN, C.; TERUYA, K.; LAM, M. **Key terms in systemic functional linguistics**. London: New York: Continuum, 2010.
- MOON, R. What can a corpus tell us about lexis? *In*: O'KEEFFE, A.; MCCARTHY, M. (ed.). **The Routledge Handbook of Corpus Linguistics**. Abingdon: Routledge, 2010. p.197-211.
- NEUMANN, S. **Contrastive register variation: a quantitative approach to the comparison of English and German**. [*S. l.*]: De Gruyter Mouton, 2013.
- NEVES, M. H. D. M. **A Gramática funcional**. São Paulo: Martins Fontes, 1997.
- PEARSON, J. **Terms in context**. Philadelphia: Amsterdam: John Benjamins, 1998. v.1.
- ROSE, D. A systemic functional approach to language evolution. **Cambridge Archaeological Journal**, Cambridge, v.16, n.1, p.73-96, 2006.
- ROSE, D. **Reading to learn: accelerating learning and closing the gap**. Sydney: Reading to Learn, 2019.
- ROSE, D.; MARTIN, J. R. **Learning to write, reading to learn: genre, knowledge and pedagogy in the Sydney school**. [*S. l.*]: Equinox Publishing, 2012.
- SINCLAIR, J. **Trust the text: language, corpus, and discourse**. London: Routledge, 2004.
- TAVERNIERS, M. Hjelmslev's semiotics model of language: an exegesis. **Semiotica**, Berlin, n.171, p.367-394, 2008.
- URE, J. Lexical density and register differentiation. *In*: PERREN, G. E.; TRIM, I. L. M. **Applications of linguistics**. Cambridge: Cambridge University, 1971. p.443-452.

VIANA, V.; TAGNÍN, S. E. O. **Corpora no ensino de línguas estrangeiras**. São Paulo: HUB Editorial, 2010.

Recebido em 5 de maio de 2020

Aprovado em 23 de setembro de 2020