



# ENSAIOS EM PORTUGUÊS COMO SEGUNDA LÍNGUA OU LÍNGUA ESTRANGEIRA

*Papers in Portuguese  
as a second or foreign language*

Inovações Descritivas para o  
Predicador: Novas Possibilidades de  
Análise do Português como Segunda  
Língua a partir da Exploração de  
Grandes Corpora

Lucas Rezende Almeida

Número 33

## **Inovações descritivas para o predicator: novas possibilidades de análise do Português como Segunda Língua a partir da exploração de grandes corpora**

Lucas Rezende Almeida  
[luks.almeida@hotmail.com](mailto:luks.almeida@hotmail.com)

**Resumo:** Baseando-nos na recorrência de dados retirados de uma plataforma digital com corpus, procuramos investigar padrões linguísticos do uso dos verbos mais comuns em língua portuguesa com o interesse de ensinar aos alunos estrangeiros as estruturas predicativas que são, portanto, mais preponderantes na fala brasileira. Para tanto, partimos de uma reflexão sobre a relação bilateral entre linguística de corpus e descrição da linguagem, levando o caráter empírico dessa ciência. Dessa forma, foram comparados os moldes predicativos dos cinco verbos mais recorrentes no corpus “Museu da Pessoa” com a descrição linguística proposta por quatro gramáticas diferentes de português para estrangeiros. Concluímos que boa parte do que as gramáticas propõem para o ensino nas aulas de PLE ainda se distancia da realidade encontrada nas ocorrências nos dados investigados, o que impõe necessidade de atualização da descrição privilegiada por esses instrumentos didáticos.

**Palavras-chave:** Português para Estrangeiros; Linguística de Corpus; Moldes predicativos.

### **Descriptive innovations for the predicator: new possibilities of analysis of Portuguese as a Second Language from the exploration of corpora**

**Abstract:** Based on the recurrence of data taken from a digital platform with corpus, this research investigates the most common linguistic patterns of verbs in Portuguese in order to teach foreigner students what those predicatives are in Brazilian speech. Therefore, it starts with a debate on a bilateral relationship between linguistic of corpus and language description, taking the empirical character of this science. In this way, the predicative patterns of the five most recurrent verbs in the corpus “Museu da Pessoa” were compared with a linguistic description proposed by four grammars different from Portuguese for foreigners. According to the data base, it concludes that the grammars used to teach in PLE classes are still distant from reality, what impose a need to update the privileged description by these didactic methods.

**Keywords:** Portuguese as a Second/Foreign Language; Corpus linguistics; Predicative molds.

## **1 Introdução**

Qualquer pesquisa atual, de cunho teórico ou não, deve preocupar-se com os dados que a constituem. A fim de embasar um autor ou de (co)construir novas questões fundamentais, um corpus serve como elemento primordial no pensamento linguístico contemporâneo. Neste trabalho, mostramos as formas como a utilização de grandes corpora eletrônicos podem auxiliar na descrição linguística, com ênfase na área de português para estrangeiros.

O avanço do ensino da língua portuguesa torna urgente uma prática mais reflexiva e pedagógica para apresentar nossa cultura, nossos hábitos e nossos comportamentos linguísticos. Não cabe mais embasar os estudos em abstrações que pouco refletem a prática do falante contemporâneo. O presente trabalho propõe uma reflexão sobre esses pontos que separam o universo formalista e funcionalista da ciência empírica da linguagem.

## **2 Pressupostos teóricos**

### **2.1 Linguística de Corpus e descrição linguística: meandros entre suas relações**

A Linguística de Corpus é uma abordagem que permite lidar com a língua de forma a estudá-la, descrevê-la e testá-la, criando hipóteses linguísticas, observando fenômenos raros ou ainda não constatados por meio de grandes volumes de dados que serão processados por ferramentas eletrônicas. Corpus, por sua vez, consiste em uma coleção de documentos que são produzidos naturalmente com uma dimensão considerável (entre 30 e 40 mil) em formato eletrônico, podendo ter informações linguísticas associadas (corpus anotado) ou não (corpus não anotado) com um objetivo específico a cargo do pesquisador.

É expressivo o impacto que a Linguística de Corpus trouxe para a ciência da linguagem por meio da utilização de ferramentas tecnológicas para auxiliar na descrição. Somente por meio do uso da tecnologia é possível a utilização de grandes volumes de dados com fins de uma descrição mais empírica sobre o fenômeno linguístico. A partir de então, o estudo de corpora pôde servir como exemplo e motivação para a discussão ou validação de teorias. Por exemplo, gramáticas funcionalistas modernas utilizam exemplos retirados de corpus para validar sua teoria, conforme observamos na “Gramática do Uso do Português”, escrita por Neves (2011), que se baseia na pesquisa lexicográfica desenvolvida pela Universidade Estadual Paulista. O processo inverso, por sua vez, também pode ocorrer: diante da análise dos dados e do processo de anotação morfossintática, constrói-se uma gramática que funciona como documentação desse material, como é o caso da gramática criada para o ensino de português para estrangeiros pela professora Diana Santos, disponível na plataforma da Linguateca.

A utilização de corpus e a descrição da língua não se esgotam apenas na possibilidade de contribuição de uma área para a outra. Segundo Beaugrande (2002), para que essa relação seja possível, é preciso que o pesquisador esteja atento a três fenômenos recorrentes nesse tipo de análise:

- (i) quantidade e qualidade do corpus: Qual o tamanho de um corpus para que a pesquisa descritiva seja de qualidade? A resposta para essa pergunta está no objetivo do pesquisador ao fazer uma descrição linguística. Esse mesmo objetivo deve resolver o impasse da uniformidade ou diversidade dos dados. Dependendo do interesse do pesquisador, devem-se selecionar dados que sejam uniformes, referentes a gêneros semelhantes, ou dados diversificados, com múltiplos gêneros;
- (ii) abrangência e profundidade do corpus: Deve o pesquisador fazer uma descrição que seja ampla e detalhada. Partindo do pressuposto de que os dados são finitos por natureza, o linguista deve equalizar a quantidade de dados a descrever (amplitude) e o grau de detalhamento (profundidade) desses dados;
- (iii) regularidade e irregularidade linguística do corpus: Entendendo que os dados de um corpus observado são por natureza acidentais e que é dever da ciência lidar com as suas irregularidades, como estabelecer o que é regular e o que é irregular nesse corpus? Na linguística de corpus, não se separam regularidades como objeto e as irregularidades como elementos do discurso, conforme feito nas teorias formalistas. Todo elemento linguístico é passível de conceitualização e descrição.

## **2.2 Desafios e limites no trabalho com corpus**

A Linguística de Corpus permite que a língua seja analisada tanto no eixo paradigmático, por meio das observações das frequências e lemas das sentenças retiradas dos respectivos textos, quanto no eixo sintagmático, através da investigação das linhas de concordância presentes no corpus. Essa interação entre os eixos, destacados por Saussure no Curso de Linguística Geral (1916), revolucionou a visão da linguagem modular, em que léxico e gramática eram vistos como pares indissociáveis.

In this respect, it can be said that the corpus revolution has introduced a new theoretical perspective on linguistic structuring: one in bold contrast to the mainstream paradigm of Chomsky (e.g. Chomsky 1965:84-88) whereby grammar and lexicon are two clearly distinct components. It also challenges a tradition long established in language study, whereby grammars and dictionaries provide distinct kinds of information about a language, and are published in separate covers. (LEECH, 2010, p. 12).

Por conta dessa revolução, atualmente a pesquisa com corpus enfrenta obstáculos com as terminologias solidificadas pelo funcionalismo em relação, por exemplo, à ideia da sintaxe sendo representante da estabilidade desse sistema abstrato linguístico, e à ideia de semântica sendo representante da sua instabilidade. Na linguística de corpus, o léxico pode e é tão estruturado e passível à categorização quanto a sintaxe, conforme vemos nos estudos apresentados por Beaugrande (2002) sobre as colocações (combinações lexicais) e as coligações (combinações gramaticais):

Aqui, podemos destacar a relação entre gramática e léxico. A teoria linguística tem há muito tempo considerada "gramática" como o epicentro da uniformidade e regularidade de toda uma linguagem, instituindo-se como um universo linguístico (compare Saussure 1966 [1916]: 133, 152; Sapir 1921: 38; Bloomfield 1933: 163; Chomsky 1957: 56). Em troca, os linguistas há muito concordaram que o léxico é uma mera "lista de irregularidades básicas" (Bloomfield 1933: 274; cf. doce 1913: 31; Saussure 1966 [1916]: 133; Chomsky 1965: 86f, 142, 214, 216). Em menor escala, essa dicotomia reencena a dicotomia entre a ordem da linguagem e da desordem do discurso (1.11), e novamente linguística escolheram pela ordem: muito trabalho em gramática, pouco no léxico (3.6) (BEAUGRANDE, 2002, p. 12).

### 2.3 Por uma relação dialética entre teoria e prática

No pensamento formalista e funcionalista, a função da teoria era predizer e explicar a prática. Nessa perspectiva, a teoria tem o poder de, simultaneamente, generalizar e ser capaz de explicar a linguagem. Porém, através do corpus, a prática é aquela capaz de especificar a teoria a partir do seu detalhamento. Com a prática, instancia-se a teoria para detalhá-la, saindo, assim, da abstração linguística.

Para Beaugrande (2002), a linguagem é uma teoria geral do conhecimento e da experiência humana, ou seja, uma representação do nosso mundo e de como nos relacionamos com ele. Nessa concepção, teoria e prática constituem-se em uma relação dialética, em um círculo interativo:

Se a linguística é restabelecer a linguagem como um objeto empírico de estudo, devemos reafirmar sua herança descritiva e recuperar a interação dialética entre teoria e linguagem como discurso como prática. Esses dois lados devem ser vistos como constituindo um ciclo dinâmico entre dois modos distintos. A ordem da linguagem deve ser orientada para a prática e expressamente concebida para apoiar a ordem orientada pela teoria do discurso, sem predeterminá-lo totalmente. (BEAUGRANDE, 2002, p. 2).

Conforme percebemos pela citação, a ideia estruturalista e formalista baseada nas dicotomias “discurso e sistema”, “competência e desempenho” deixa de existir em prol de uma contribuição bilateral entre teoria e dados por meio do uso de corpus. A língua é considerada um tipo único de teoria devido ao seu caráter metalinguístico, isto é, o objeto do estudo é o próprio instrumento metodológico para analisá-lo. Dessa forma, qualquer teoria sobre a linguagem nada mais é do que uma teoria sobre a teoria, uma metateoria.

## 2.4 Objeto e objetivo

O objeto deste trabalho são os predicadores e suas relações com os predicados. Consideramos “moldes predicativos”, uma definição cunhada por Rebelo (1999) baseada na gramática funcional de Dik (1981), as representações dos elementos constituintes do enunciado: predicador, argumento, função dos argumentos e os rótulos em forma de paráfrases explicativas.

O objetivo geral é averiguar que contribuições o uso de grandes corpora pode trazer para o ensino dos verbos mais recorrentes da língua portuguesa para estrangeiros. Para isso, é preciso delimitar diversos pontos:

a) Baseando-se em uma busca por lema de todos os corpora juntos da plataforma Languateca, selecionaram-se cinco dos verbos encontrados entre os mais recorrentes na língua portuguesa: **ser, estar, dar, ir e ter**;

b) Em seguida, selecionaram-se dois cenários sintáticos específicos em que são analisados os moldes predicativos:

b1) Quando o predicado se inicia por sintagmas preposicionais: selecionaram-se apenas as quatro primeiras preposições que iniciam esse sintagma em comparação com as demais;

b2) Quando o predicado se inicia por verbo ou por nome: o interesse é observar qual desses usos é mais recorrente; por exemplo, o verbo “ser” desempenha, na posição posposta, a ocorrência maior de nomes ou verbos? Nosso trabalho não procura entender semanticamente o significado desse nome ou desse verbo, mas quais são os tipos de estruturas possíveis de serem formuladas.

## 2.5 Motivação e justificativa

Diante da complexidade na descrição dos verbos no português para alunos estrangeiros e da defasagem dos materiais didáticos para explicá-los, este trabalho é relevante por demonstrar de que forma a utilização de grandes corpora pode apresentar novas ocorrências de uso desses verbos. Além disso, comprovam ou questionam práticas de ensino que não revelam a sua funcionalidade na linguagem diária da sociedade brasileira.

Por conta disso, foi selecionado para essa pesquisa o corpus “Museu da Pessoa”, disponível na Linguateca, por meio do projeto AC/DC (SANTOS & BICK, 2000), por ser considerado um corpus de fala informal e espontânea. Trata-se de um corpus morfossintaticamente anotado pelo *Parser PALAVRAS*, com acesso livre e documentado, constituído das variantes do português brasileiro e de Portugal. Formado pelo gênero entrevista, existe uma preferência nos tempos verbais pelo uso do pretérito. Entretanto, o mesmo corpus não é documentado quanto aos fatores extralinguísticos como gênero, idade e escolaridade, o que impede a pesquisa de apresentar esses resultados sociolinguísticos.

## 2.6 Revisão literária

Entendendo que o nosso trabalho pretende destacar de que forma o uso de grandes corpora pode auxiliar na descrição da língua portuguesa, destacamos três trabalhos da área de português para estrangeiros que utilizam diferentes formas de dados para caracterizar alguns verbos da língua portuguesa.

O primeiro trabalho que inspirou a produção deste artigo foi a dissertação de Ida Maria da Mota Rebelo (1999), intitulada “Os sentidos do verbo ficar: uma formalização semântico-funcional voltada para o português como língua estrangeira”. Em seu trabalho, a linguista identifica 15 moldes predicativos para o verbo “ficar” por meio da criação de um corpus cujos textos foram retirados da imprensa escrita, de páginas de divulgação on-line, de correios eletrônicos, de artigos científicos, de publicações de cunho literário e de conversas na internet. A diversidade dos textos reflete o seu interesse de encontrar o maior número de ocorrências do verbo em situações comunicativas diversas. Não se trata, aqui, de analisar texto, mas de investigar uma determinada estrutura recorrente nos mais variados textos. Por essa razão, o corpus reunido pela autora é formado por textos bastante variados, como pode ser confirmado na página das abreviações (REBELO, 1999, p. 131).

Para operacionalizar os textos selecionados, a autora utilizou um programa do tipo *concordancer* chamado *WordPilot* na sua versão 1.1.1999. A respeito do interesse na utilização dessa ferramenta, a autora diz:

Uma observação que deve ser feita é que, entre os objetivos deste estudo, não se encontra a análise prévia de corpora. A análise quantitativa apresenta-se, apenas, como instrumento para a busca de se confirmarem hipóteses levantadas durante a análise dos dados. Explica-se, assim, a ausência, nos capítulos anteriores, de procedimentos que estão por vezes envolvidos na análise de corpora como, por exemplo, a normatização de frequências. (REBELO, 1999, p. 131).

Percebemos, portanto, que a utilização do corpus no trabalho apresentado tem como intuito averiguar uma teoria e uma descrição já estabelecida. Como se trata de um trabalho feito há mais de dez anos, encontramos muitas diferenças teóricas e metodológicas no que tange às reflexões atuais sobre a interação entre corpus e descrição. Uma revisão desse trabalho certamente proporcionaria novas descobertas e novas perspectivas sobre o verbo em análise.

O segundo trabalho, ao contrário de Rebelo (1999), é extremamente recente. Trata-se da tese de doutorado de Bruno de Andrade Rodrigues (2014): “Ser ou Estar: Eis a questão! Um estudo descritivo de usos voltado para o português como segunda língua para estrangeiros”. No trabalho de Rodrigues (2014), temos um estudo qualitativo, baseado na transcrição de episódios escolhidos na série “Os Normais”. Em seu estudo, Rodrigues (2014) considera seus dados como um corpus, com o qual pretende investigar as ocorrências dos verbos “ser e estar” articulados aos sintagmas adjetivais, nominais predicativos, bem como preposicionais, além de estabelecer as diferenças de uso entre os dois verbos. Não é feita em seu trabalho uma pesquisa de cunho quantitativo, explorando grandes corpora com o fim de averiguar ou embasar os resultados com os verbos. No presente trabalho, fazemos, de forma bastante resumida, uma investigação quantitativa envolvendo os sintagmas preposicionais e selecionando as três preposições mais recorrentes com o verbo “ser”, além das quatro mais recorrentes com o verbo “estar”. Diferentemente de Rodrigues (2014), não apresentamos todos os moldes predicativos para cada uma das ocorrências desse sintagma preposicional; apenas estabelecemos significados desses moldes que não encontramos nas gramáticas de português para estrangeiros aqui analisadas.

Um terceiro trabalho que dialoga com a nossa proposta é a dissertação de Sheila Mejlachowicz (2003), intitulada “Uma análise semântico-aspectual dos verbos de ligação”, em que ela compara os verbos de ligação, considerados na gramática tradicional, com os verbos



suportes, discutidos pelos linguistas. Ao invés de fazer uma pesquisa formal, assim como encontramos nos trabalhos anteriores, a linguista se preocupa com os traços semânticos que diferenciam essa classe de verbos dos seus complementos. Em seu trabalho, analisa os verbos “ser” e “estar”, também aqui estudados. Na sua dissertação, entretanto, destacam-se os complementos do tipo morfológico **nome** que seguem esses verbos. No presente artigo, damos outro enfoque para esses verbos, destacando seus usos como sintagmas preposicionados e as possíveis paráfrases. A respeito do corpus utilizado, a única informação dita por Mejlachowicz (2003) é a seguinte:

Com relação ao corpus a que nos referimos, retiramos exemplos de algumas edições completas de jornais, como o Jornal do Brasil, em que para extrair os verbos de ligação usamos uma ferramenta de concordância denominada INTEX (Silberstein, 1993). Utilizamos ainda exemplos retirados do dia-a-dia (outros jornais, como O Globo, mas não edições completas) e outros coletados pessoalmente. (MEJLACHOWICZ, 1993, p. 68).

Diante da sua afirmação, percebemos que sua preocupação não está em caracterizar com precisão o corpus selecionado. Nos anexos, notamos também apenas o aparecimento das linhas de concordâncias selecionadas para o trabalho, sem conter todo o texto de onde elas foram retiradas. Assim, o leitor fica à mercê dos dados utilizados, não podendo conferi-los.

Nesta revisão, verificamos a urgência de um trabalho descritivo para o ensino de português para estrangeiros, baseado em um corpus bem delimitado e que esteja disponível publicamente para outros pesquisadores também poderem acessar e discutir os resultados obtidos. Nossa pesquisa demonstra, dessa forma, quais são as contribuições e os desafios que essa proposta pode trazer para o estudo do português como segunda língua.

### **3 Metodologia**

Este trabalho foi elaborado em três partes diferentes: inicialmente, procurou-se identificar quais eram os sintagmas que sucediam os verbos mais recorrentes dentro de um corpus de entrevista chamado “Museu da Pessoa”. Optou-se por elencar aqueles sintagmas que desempenhavam a função sintática de acusativo ou dativo e a função morfológica de nome e de verbo. Ao escolhermos a função morfológica de nome, estamos verificando a recorrência de moldes semânticos (REBELO, 1999) e predicativos do sujeito (MEJLACHOWICZ, 1993); enquanto, ao escolhermos a função morfológica de verbo, estamos à procura de investigar as formas compostas dos tempos verbais e dos casos de verbo suporte (TAGNIN, 1989).

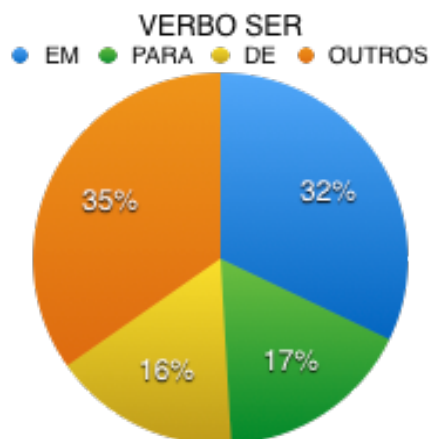
A partir dos resultados quantitativos obtidos, foram elaborados gráficos, com o intuito de melhor observarmos as inovações descritivas ali encontradas - assim nomeadas em quatro tipos de gramática de Português para Estrangeiros. Em um terceiro momento, selecionaram-se as quatro tipos de gramáticas diferentes voltadas para o ensino de português para estrangeiros: (a) “Modern portuguese: a reference grammar”, (b) “Gramática Ativa 1”, (c) “Como está o seu português” e (d) “Gramática Ativa 2”. A escolha por essas gramáticas não foi feita de forma aleatória. A opção pela primeira justifica-se pela necessidade de averiguar de que forma uma gramática descritiva, feita por um linguista brasileiro em outro idioma, apresenta ou não os resultados compatíveis com aqueles descobertos nesta pesquisa. As demais gramáticas foram selecionadas por serem consideradas em nossa investigação bibliográfica como as mais modernas. “Como está o seu português?” foi lançada em 2014 para o ensino de português brasileiro, sendo uma proposta que integra exercícios e teoria; “Gramática Ativa 1” e “Gramática Ativa 2” (dois volumes de uma gramática) foram atualizadas para uma versão do português brasileiro em 2012 e 2014, respectivamente, e apresentam a teoria separada dos exercícios, que vêm subsequentemente.

Analisamos inicialmente os sintagmas preposicionados para cada verbo isoladamente e, em seguida, os predicados do tipo “VERBO + NOME” e do tipo “VERBO + VERBO” comparativamente por meio de tabelas que demonstram as suas ocorrências. Todas as buscas realizadas no site <http://www.linguateca.pt/acesso/corpus.php?corpus=MUSEUDAPESSOA> estão disponíveis em anexo, como também o contexto em que elas foram realizadas.

#### **4 Análise de sintagmas preposicionados**

Nesta etapa, para o estudo comparativo, relacionamos os resultados encontrados na nossa pesquisa quantitativa referentes aos sintagmas preposicionais que acompanham os verbos SER, ESTAR, IR, DAR e TER, com a descrição desses verbos nas gramáticas selecionadas. Os casos inovadores encontrados na pesquisa, em comparação com as gramáticas selecionadas para averiguação, não representam, em alguns casos, todas as ocorrências das estruturas das linhas de concordância. Porém, para este trabalho, foi feito um recorte e foram analisados apenas aqueles pontos não abordados nos volumes compêndios gramaticais anteriormente citados.

#### 4.1 O verbo SER



O gráfico ao lado ilustra os sintagmas preposicionados encontrados para o verbo SER. Foram localizados 3447 casos, com 43 valores diferentes de lemas. Destes, conforme observado, mais da metade (65%) são representados pelas preposições “para”, “de” e “em”, com maior recorrência da preposição “em”.

É interessante observar que o sintagma preposicionado ligado ao verbo SER não é considerado e explicado em nenhuma das gramáticas analisadas. Elas se preocupam em explicar prioritariamente o uso do verbo SER como predicativo do sujeito, como verbo suporte, como formador de tempos compostos – voz passiva – ou como opositor ao verbo ESTAR.

Em “Modern Portuguese: reference grammar”, Perini (2002) trata o verbo SER em oposição ao verbo ESTAR dentro do capítulo “Anotações do uso de certos verbos”. Além disso, mais à frente, quando explica sobre a concordância verbal sujeito-verbo, o autor acrescenta um tópico exclusivo para o verbo SER, apresentando o seu uso como verbo de ligação, ou seja, no que consideramos nesta pesquisa como verbo seguido do tipo morfológico nome.

Em “Gramática Ativa 1 e 2”, o verbo SER é explorado apenas no primeiro volume, que destaca a diferença entre SER e ESTAR, priorizando o seu sentido como caracterizador do sujeito. Existe uma unidade para tratar das estruturas SER + PARTICÍPIO, com enfoque na voz passiva.

Em “Como está o seu português”, o verbo SER é tratado dentro do modo indicativo em duas seções: “Verbos de Ligação” e “Usos específicos de ser, estar, fazer, haver”. Na primeira seção, destaca-se o uso do verbo SER seguido do tipo morfológico **nome**; na segunda seção, destaca-se a diferença entre os verbos SER e ESTAR no português, como também a sua importância na constituição da voz passiva.

A valência do verbo SER, diferente do que é explorado nas gramáticas, varia de acordo com o sintagma preposicionado que o segue. Na análise dos lemas dos nomes que seguem as preposições acima, observamos como a estrutura [ser + para + nome] pode ser substituída pela paráfrase “serve para” ou “é feito para”, conforme observamos nas linhas de concordância a seguir:

E137-BR-1080: Então eu acho que é o projeto básico, serve para todo mundo, é útil para governo, para empresa.

E167-BR-473: Depois veio o Flagil, que era para as meninas, para uso vaginal, ginecológico, que também não ia para as escolas.

E188-BR-1155: Porque vai enxergar que a forma mais barata de cuidar da nossa água, que essa água é para consumo, pra energia etc.

Observamos também como a estrutura [ser + de + nome] pode ser substituída pela paráfrase de pertencimento ou de origem, segundo as linhas de concordância.

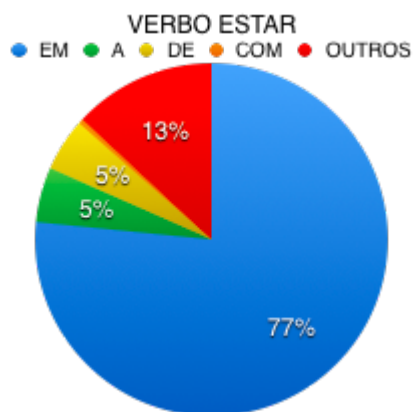
E001-PT-289: Recordo-me da primeira vez que fui para Oia porque foi a primeira vez que vi o mar, que eu sou do interior.

E144-BR-1410: Turma isso aqui é dos asiáticos, é da população local, mas é da humanidade.

E146-BR-32: Para o seu pai ter vindo da Itália e depois ter ficado um tempo na Argentina, era um pouco de italiano e espanhol.

Vale destacar aqui, como grau de curiosidade, que alguns livros didáticos já tratam da valência do verbo SER, principalmente da valência da estrutura SER DE como exemplo da proposta feita pela UFMG: “Terra Brasil: curso de língua e cultura”. (2008).

## 4.2 O verbo ESTAR



O gráfico ao lado representa também os sintagmas preposicionados do verbo ESTAR, nos quais foram encontrados 2.333 casos com 39 valores diferentes de lema, isto é, com 39 tipos diferentes de contrações de preposições. Esse gráfico nos revela a maior recorrência da estrutura ESTAR EM. Uma das funções prototípicas atribuídas à preposição EM refere-se ao seu caráter de localização espaço-temporal. No corpus, verificamos que a estrutura ESTAR A também tem essa paráfrase de localização, incluindo a ideia de posicionamento, com exceção do ESTAR COM, que significa “acompanhado de”.

E083-PT-10: O meu pai estava em Moçambique na altura em que eu nasci e a minha mãe ficou cá.

E200-BR-289: A CIN estava ao lado da produtora do Gilberto Martins, que também foi outro ingrediente de muita sorte, que a gente trabalhava tanto com a CIN quanto com Humberto Martins, que tinha o melhor equipamento para jingles.

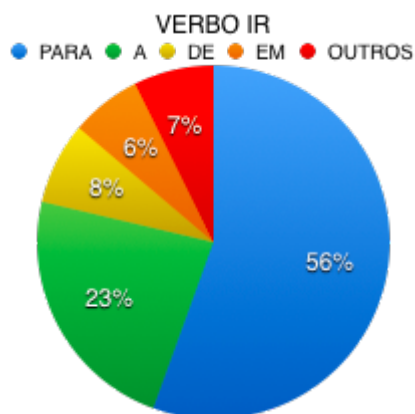
Nas gramáticas observadas, percebemos o mesmo movimento descritivo ocorrido com o verbo SER: há preferência pelos casos em que se explicam os usos do ESTAR, em contraposição com os usos do SER. Além disso, o verbo ESTAR é apresentado como formador de tempos verbais contínuos. “Como está o seu português” não apresenta os tempos verbais do tipo contínuo, restringindo o uso do verbo ESTAR unicamente aos seus usos específicos entre “ser, fazer e haver”. Por outro lado, a “Gramática Ativa 1 e 2” apresenta diversos usos do verbo ESTAR; explica o verbo junto com suas preposições como “em + local” e “com + substantivo” e também quando tem sentido da estrutura “ter + substantivo”. Destaca e diferencia a estrutura “estar + gerúndio” em duas unidades intituladas “estar + gerúndio” e “estou fazendo e faço”, em que diferencia a forma contínua do presente. Há ainda uma unidade para tratar das estruturas “estar + particípio”, com enfoque na voz passiva.

Em “Modern portuguese: reference grammar”, o verbo ESTAR é tratado dentro do presente progressivo e das formas progressivas do passado. Na sua comparação com o verbo SER, uma parte merece ser destacada nos “Casos particulares” (pág. 273 a 275):

The first observation is that *estar* tends to have a more extended usage than the general rule strictly allows, therefore, in some contexts it may be used interchangeably with *ser*. For instance, the general rule applies locational expressions: one uses *estar* to introduce a person's location but *ser* is used when referring to a city or a geographic accident. This makes sense in terms of the general rule, since the location of a city is something inherent to it, unlike the location of a person. But with geographic terms one can also use *estar*. (PERINI, 2002, p. 273).

Neste trecho, percebemos a diferença feita entre os verbos ESTAR e SER. O que, entretanto, a pesquisa quantitativa nos mostra difere-se um pouco: é muito mais recorrente o uso do verbo ESTAR como localizador do que o verbo SER. Percebemos isso na recorrência da preposição “em” no gráfico 02 (77%) em comparação com o gráfico 01 (32%). Perini (2002) também destaca o fato da estrutura ESTAR COM em oposição ao verbo FICAR, relacionando o aspecto durativo que cada estrutura possui.

### 4.3 O verbo IR



O gráfico apresenta os sintagmas preposicionais do verbo IR, mostrando maior recorrência do tipo IR PARA e menor recorrência do tipo IR DE, tão explorado no ensino de português como segunda língua para estrangeiros. Nesta pesquisa, foram encontrados 3.771 casos, divididos em, pelos menos, 31 tipos de lemas diferentes.

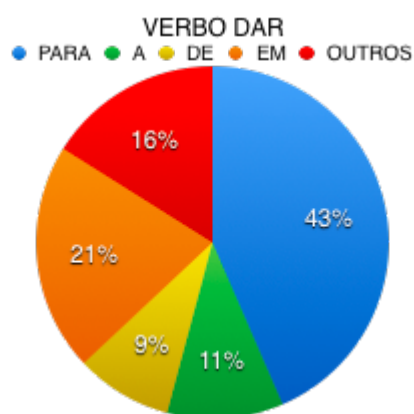
O verbo IR é ensinado como formador do tempo composto considerado “futuro imediato” no ensino de português para estrangeiros. Muito comum também é o ensino da estrutura IR DE apresentada junto a outras como GOSTAR DE, PRECISAR DE, quando referindo-se à regência dos verbos. Ao tratar a estrutura IR DE, por exemplo, os livros didáticos acrescentam o nome com traço [+meio de transporte] e se esquecem de algumas expressões, como “ir de férias” ou “ir de Salvador a Pernambuco”. Como nos casos anteriores, não se chama a atenção do aluno para as outras valências do verbo IR, com a preposição PARA, A e EM, destacadas neste trabalho.

Perini (2002), em “Modern portuguese: reference grammar”, apresenta o verbo IR na formação do futuro composto do indicativo, diferenciando-o da forma simples do futuro. Segundo ele, o futuro imediato é um futuro que logo irá acontecer, enquanto o futuro simples pode tratar de uma ação futura, sem previsão e longínqua, em relação ao momento da enunciação. Oberg (2012), na versão brasileira da “Gramática Ativa 1 e 2”, não trata do futuro imediato com o verbo IR, entretanto apresenta a estrutura “ir + gerúndio” como indicação de realização gradual no seu primeiro volume. Ainda nesse volume, outra unidade intitulada “verbo + preposição” apresenta o verbo IR ligado apenas à preposição COM. No segundo volume, o verbo IR é tratado como verbo suporte, seguido de um elemento morfológico verbo. Otuki (2014) trata o verbo IR unicamente como formador de tempo composto do futuro imediato do indicativo.

### 4.4 O verbo DAR

Com o verbo DAR, percebemos que as frequências das preposições mais recorrentes representam mais de 70% dos seus usos. Nos dados analisados, foram registrados 1.228 casos, com 32 valores diferentes de lema. A predominância da preposição PARA é um forte indício

da estrutura argumental “dar (algo) para (alguém)”, tão conhecida e ensinada aos alunos de



português como segunda língua. Entretanto, a busca por nomes após as preposições revelou outros resultados sobre o verbo DAR.

A pesquisa feita com os verbos anteriores prioriza o sintagma preposicional do tipo PREPOSIÇÃO + NOME. Entretanto, percebemos que um outro tipo de estrutura é possível com o verbo DAR: DAR + PREPOSIÇÃO + VERBO. Observamos essa ocorrência com a preposição PARA e a preposição DE. A tabela a seguir compara o resultado dessa busca. A ocorrência com o sintagma preposicional com o nome no seu núcleo é mais recorrente, fato, entretanto, que não serve como motivo para desconsiderar a estrutura com núcleo verbal.

Diante de uma investigação qualitativa que parte da observação das linhas de concordância, é possível propor uma paráfrase para essas estruturas. Observemos o caso da estrutura DAR + PARA + VERBO.

E020-PT-22: Hoje não, hoje são precisos quinhentos escudos e depende da isca, mas naquele tempo com dez tostões ou quinze tostões já dava para comer.

E020-PT-88: Até naquele tempo, com aquele dinheiro deu para comprar coisas que hoje não se compra.

E020-PT-373: E depois tem outro terreno ao lado que dá para fazer outra casa, tem poço já e tudo porque o meu pai, quando ficou com aquela casa, comprou o terreno ao lado.

Diante das linhas selecionadas, percebemos que essa estrutura indica possibilidade de ocorrência de um fenômeno marcado pelo verbo principal da frase. Uma paráfrase possível seria a substituição do DAR PARA por SER POSSÍVEL.

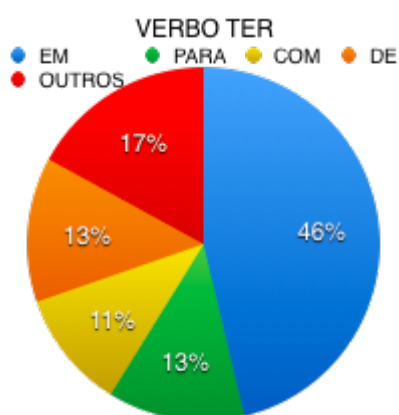
Agora, observemos o caso da estrutura DAR + DE + VERBO:

E093-PT-170: O meu pai passou anos a dar de comer às minhas filhas.

E009-PT-20: Eu costumo dizer: Serviços são poucos e bugios são muitos. Temos que dar de fazer a muitas pessoas, independentemente de que têm que ser pessoas mais ou menos adequadas e por si só manifestar o interesse que têm de fazer, o gostar de fazer isto ou aquilo, porque as pessoas, quer os guias, quer os rabos, quer o velho, estão responsáveis por gerir todos os papéis que competem à Bugiada.

Houve apenas um caso (E009-PT-20) que se aproximou mais do registro informal do uso dessa estrutura em nosso corpus, representando uma ação inesperada que começa a acontecer. A maior parte dos casos encontrados restringiu-se ao uso da expressão cristalizada DAR DE COMER. Perini (2002) destaca a função do verbo DAR como verbo suporte de sintagmas preposicionados sucedidos de nomes ou de verbos. Ele registra o caso da estrutura DAR + PARA + VERBO, como também a estrutura DAR + POR + VERBO. Contudo, o autor não diferencia quais as tipologias verbais podem ocupar o lugar após a estrutura preposicional, como também não diferencia as estruturas que possuem o sintagma preposicionado, seguido de nome, daqueles seguidos de verbo. Oberg (2014) só apresenta o sintagma preposicional seguido de DAR como modificador semântico do verbo no segundo volume da sua coleção, considerando, portanto, essa estrutura, como um elemento avançado na língua. Mesmo no primeiro volume, havendo uma seção chamada “verbo” + “preposição”, trata-se prioritariamente das predicções que não mudam o sentido do seu predicador. No segundo volume, em uma unidade chamada “Dar, ficar e passar seguidos de preposição”, a autora seleciona 5 casos de sintagmas preposicionais seguidos de nome, não havendo nenhum caso seguido de verbo, como ressaltado nesta pesquisa. Otuki (2014) não apresenta nem um caso específico do verbo DAR, que não encabeça nenhuma das frases-modelo das unidades do livro.

#### 4.5 O verbo TER



Diante do gráfico, observamos maior recorrência da estrutura TER + EM. Foram registrados 3.727 casos com 44 valores diferentes de lema. Vale destacar, nesse resultado quantitativo, o fato de o número de ocorrências do verbo TER seguido de um sintagma preposicional ser mais recorrente do que os casos do verbo TER acompanhado de um outro verbo, formando os tempos considerados compostos do subjuntivo e do indicativo.

Esse fenômeno demonstra, baseado no corpus, como a aplicabilidade do verbo TER como formador dos tempos compostos é menor do que o seu papel de predicador. Destaca-se, portanto, a predominância dos tempos simples frente aos compostos do subjuntivo, visto que todos eles são formados pelo verbo TER.

Com o verbo TER também é possível a construção [verbo + preposição + verbo] restritamente – entre as preposições destacadas neste trabalho – quando é a preposição DE,



funcionando, dessa forma, como um modalizador verbal de “compromisso” ou de “dever”, conforme observado nas linhas a seguir:

E188-BR-1143: Eu vejo uma mata que nem essa aqui, puxa. É um prazer que eu tenho de estar lá dentro.

E188-BR-134: Onda eu cheguei a pegar por dois anos, mas logo me dei conta que pra ser um bom surfista tinha de virar vagabundo -- em um bom termo.

E030-PT-27: Mas aconteceu muita vez termos de acender uns paus de urzeira secos e às vezes com «abrotigas», que depois de acesos espetávamos na parede para vermos.

E025-PT-72: Tinha de ir com as cabras.

“Modern Portuguese: a reference grammar” registra o caso modalizador do verbo TER, como também seu papel na formação dos tempos compostos do subjuntivo. Entretanto, não há explicação semântica sobre os casos dos sintagmas preposicionais que seguem esse verbo, como ocorre com o verbo DAR. A “Gramática Ativa 1 e 2” e a gramática “Como está o seu português” só apresentam a estrutura do verbo TER ligada à formação do composto do subjuntivo, sem levar em conta os casos de modalização e dos sintagmas preposicionais que o seguem.

## 5 Análise dos elementos morfológicos do tipo nome e verbo

Nesta segunda etapa, analisamos a recorrência dessas estruturas por meio de uma tabela comparativa entre os verbos e seus argumentos do tipo morfológico nome e do tipo morfológico verbo. Procuramos sustentar, dessa forma, a importância de adotar novas perspectivas no ensino e no preparo de exercícios para esses tipos de estrutura.

### 5.1 Do tipo morfológico nome

Neste tópico, demonstramos a recorrência com que o verbo vem seguido por um nome. Optamos por chamar o termo que segue o verbo pela sua classificação morfológica em vez de sua função sintática, a qual pode ser identificada como complemento verbal, predicativo do sujeito, moldes predicativos (REBELO, 1999), etc.

SER	TER	ESTAR	IR	DAR
4562	4594	126	49	900

Na tabela, percebemos como o tipo de estrutura VERBO + NOME é recorrente com o verbo SER, o que demonstra a sua produtividade como verbo de ligação na construção da caracterização do elemento considerado sujeito. Diferentemente do que se encontra nas gramáticas de português para estrangeiros, o verbo ESTAR não tem essa produtividade tão acentuada. Analisado muitas vezes junto ao verbo SER, ESTAR é ensinado seguido por um número de nomes aos quais ele se comportaria semanticamente de maneira mais propícia do que o verbo SER, por exemplo. Entretanto, a recorrência desse uso foi muito menor do que entre os verbos TER e DAR. Isso ressalta a nossa hipótese de que o verbo ESTAR funciona muito mais como produtor de localização na frase do que como caracterizador nos discursos orais. A proximidade de ocorrência do verbo TER com o verbo SER não reflete a mesma função sintática do argumento nominal. O nome nem sempre funciona como predicativo do sujeito nas orações a seguir:

E078-PT-113: Temos caixa, reforma, temos, se for preciso, fundo de desemprego; se o barco naufragar **temos direito** a uma indenização, temos direito a roupa de pesca e a seguro.

E127-BR-827: Então tem características de região.

E128-BR-188: Não tem problema, imagina.

E129-BR-187: Não tem perigo não, é seguro.

No excerto 113, o verbo TER estabelece uma ideia de posse com o seu argumento, da mesma forma que no excerto 827. É interessante notar a impossibilidade da ideia de caracterização no excerto 827, mesmo com o aparecimento da palavra “característica”, que significa algo que a região possui. Já nos excertos 188 e 187, temos casos diferentes: o verbo TER, no excerto 187, não tem uma paráfrase significativa se não considerado junto com o nome perigo (vale ressaltar que essa estrutura também poderia ser substituída por uma relação de predicação com “Não é perigoso não, é seguro”); no 188, a estrutura “ter problema” nos parece multivocabular (GARRÃO, 2006), ainda mais fechada, pois não pode ser substituída por uma relação de predicação, além de já ter um significado específico quando encontrada interligada. Temos a possibilidade de modificar elementos no excerto 187, mantendo certo paralelismo sintático: “ter entusiasmo, ter coragem” pode ser substituído por “é corajoso, é entusiasmado”,

enquanto esse tipo de paralelismo não é possível em 188. Esses fenômenos retratam a complexidade das estruturas multivocabulares, tema que pode ser ampliado na leitura da tese de Garrão (2006), mas que não será detalhado neste artigo.

Com o verbo DAR, a complexidade dos elementos multivocabulares fica ainda mais evidente, visto que a maior parte das ocorrências da estrutura DAR + NOME são estruturas semelhantes às analisadas anteriormente com o verbo TER, como percebemos com a tabela a seguir, que mostra os lemas mais recorrentes.

**PESQUISA: [lema="dar"] @[pos="N.\*"]**

aula	138	risada	15	condição	10
conta	56	apoio	14	resposta	9
dinheiro	36	resultado	14	oportunidade	9
continuidade	23	prejuízo	11	informação	9
trabalho	23	origem	11	aumento	8
tempo	20	espaço	10	suporte	8

Diante dos dados, os casos do verbo DAR seguido de NOME representam majoritariamente estruturas multivocabulares.

Em “Modern Portuguese: reference grammar”, na seção em que Perini (2002) explora os verbos suportes, ele acrescenta as estruturas multivocabulares junto com aquelas estruturas que vêm acompanhadas de uma preposição e um verbo (DAR + PARA + VERBO). Não existe uma diferenciação entre esses tipos de estrutura, o enfoque é a paráfrase significativa gerada pela predicação verbal. Em “Como está o seu português?”, Otuki (2014) destina uma seção específica para lidar com os verbos de ligação em que é estabelecida a diferença entre ESTAR e SER. Em “Gramática Ativa 1 e 2”, Oberg (2012, 2014) apresenta os derivados do verbo TER e a diferença entre a estrutura SER + NOME e ESTAR + NOME, quando ambos funcionam como verbo de ligação.

## 5.2 Do tipo morfológico verbo

SER	TER	ESTAR	IR	DAR
4741	2363	4500	7019	51

Novamente, registra-se a alta ocorrência do verbo SER acompanhado de outro verbo. Ao solicitarmos a distribuição do tempo verbal e do caso pronominal, vemos que o verbo SER é formador prioritariamente da voz passiva.

No caso de TER, o segundo verbo é um elemento formador dos tempos compostos do indicativo ou uma forma nominal do tipo particípio passado, utilizado nos tempos compostos do subjuntivo e indicativo. A ocorrência do verbo TER no modo subjuntivo é menor do que na forma do imperfeito do indicativo (formador do tempo composto do mais-que-perfeito do indicativo), mostrando a predominância dos tempos compostos do indicativo sobre o subjuntivo com o verbo TER.

O verbo ESTAR é formador dos casos progressivos, independentemente do tempo ou modo verbal predominante.

O verbo IR registra o maior índice da tabela. Na consulta à distribuição de “temcarg”, averiguamos que as estruturas que combinam com o verbo IR respectivamente são o presente, o gerúndio e o infinitivo, o que comprova o seu potencial formador de futuro imediato.

O verbo DAR, diferente de todos os casos anteriores, apresentou um baixo índice de ocorrência nesse tipo de estrutura. Ao olharmos a distribuição do lema do tempo verbal e do caso pronominal que seguem esse predicador, observamos que boa parte desse grupo se caracteriza por formas nominais no particípio passado que podem ser consideradas nomes. A estrutura “[lema=”dar”] @[lema=”errar”]” foi a mais recorrente neste caso:

E185-BR-1026: A gente ia desmanchar tudo, abrir os trabalho da máquina que deram errado.

E176-BR-801: E aí dos projetos que nós desenvolvemos lá, coisas que deram certo, coisas que deram errado, de brigas que nós tivemos com os prefeitos ali da região.

Elas reforçam o indício da recorrência do verbo DAR com estruturas morfologicamente [+ NOME].

Como se pode perceber neste capítulo, em todas as gramáticas, os verbos são tratados como formadores de tempos compostos. Em algumas delas, há uma explicação mais profunda, como a de Perini; em outras, a explicação é mais didática, de acordo com o interesse na produção de cada material.

## 6 Considerações finais

Muitas foram as considerações que viemos apontando ao longo deste artigo. Pretendemos, nesta etapa, uni-las e acrescentar alguns resultados quanto ao procedimento desta pesquisa.

a) A frequência das ocorrências dos verbos em suas estruturas predicativas nos fornece informações importantes no que aparenta ser mais recorrente na língua, norteando, de certa forma, o que deve ser ensinado a um aluno estrangeiro.

b) A utilização de grandes corpora, portanto, funciona sobretudo ao professor de português para estrangeiros como um guia do que se deve ressaltar no universo multiforme, que é a língua.

c) Como exemplos desse direcionamento que a linguística de corpus pode oferecer, trouxemos neste trabalho os moldes predicativos dos verbos encontrados como de maior recorrência no corpus “todos juntos”, da Linguateca.

d) Sobre o verbo SER: descobrimos sua produtividade quando seguido de sintagmas preposicionais, aspecto pouco abordado no ensino de português para estrangeiros.

e) Sobre o verbo ESTAR: constatamos a sua maior recorrência como localizador.

f) Sobre o verbo IR: vimos que, diferente da prioridade dada para o sintagma preposicionado com DE no ensino de português como segunda língua, sua ocorrência é muito maior com PARA e A.

g) Sobre o verbo DAR: encontramos estruturas muito produtivas e pouco ressaltadas no ensino de português para estrangeiros, como DAR + PARA + VERBO e DAR + DE + VERBO, além do alto índice de estruturas multivocabulares, que merecem um melhor estudo.

h) Sobre o verbo TER: percebemos sua recorrência ligada ao sintagma preposicionado maior do que ligada a outro verbo, o que demonstra necessidade de maior descrição desse estudo nas gramáticas de PL2E.

Baseado nas considerações que ponderamos na lista acima, podemos concluir que o ensino de PL2E ainda tem muito a avançar sobre o estudo descritivo do seu esqueleto programático: o próprio verbo. Usado como guia nas preparações dos materiais didáticos e também como nivelador de proficiência, o verbo ganhou o *status* de qualificador na competência deste aluno em relação à língua portuguesa, entretanto, muito ainda se distancia a realidade do seu uso do contexto ensinado no espaço institucional sala de aula. Por este motivo, esta pesquisa, por meio da linguística de corpus, pretende demonstrar a necessidade de levarmos ao espaço sala de aula situações de uso que condizem mais com a realidade linguística a qual esse aluno irá enfrentar, procurando, por meio da descrição dos dados, revelar quais são as amostragens contextuais mais relevantes para os diferentes níveis de aprendizado no processo de ensino do PL2E.

### **Referências Bibliográficas**

CAMACHO, Roberto Gomes. *Da gramática funcional à gramática funcional do discurso*. Signótica Especial, n. 2, pp. 167-180, 2006.

BEAUGRANDE, Robert de. (2002). Descriptive Linguistics at the Millennium: Corpus Data as Authentic. Language. In: *Journal of Language and Linguistics* 1(2), 91-131.

DELL'ISOLA, Regina Lúcia Péret & ALMEIDA, Maria José Aparecida de. *Terra Brasil: Curso de língua e cultura*. Belo Horizonte: Editora UFMG, 2008.

LEECH, Geoffrey Neil. Frequency, corpora and language learning. In: *A Taste for Corpora: In honour of Sylviane Granger, Meunier, Fanny, Sylvie De Cock, Gaëtanelle Gilquin and Magali Paquot* (Eds), 2010.

MEJLACHOWICZ, Sheila. *Uma análise semântico-aspectual dos verbos de ligação*. Dissertação de mestrado: Orientadora: DIAS, Maria Carmelita Padua, PUC/RIO, 2003.

NEVES, Maria Helena de Moura. *Gramática de usos do português*. SP: Editora Unesp, 2 ed. 2011.

OTUKI DE PONCE, Maria Harumi. *Como está o seu português?* São Paulo: 1 ed, HUB Editora, 2014.

BERG, Lamartine Bião. *Gramática Ativa 1*. Lisboa: Lidel, 2 ed. Versão Brasileira, 2012.

BERG, Lamartine Bião & FERNANDES, Alice Ferreira. *Gramática Ativa 2*. Lisboa: Lidel, Versão Brasileira, 2014.

REBELO, Ida Maria da Mota. *Os sentidos de ficar: uma formalização semântico-funcional voltada para o português como língua estrangeira*. Dissertação de mestrado: Orientadora: MEYER, Rosa Marina de Brito, PUC/Rio, 1999.

RODRIGUES, Bruno de Andrade. *Ser ou Estar: Eis a questão! Um Estudo Descritivo de Usos voltado para o Português como Segunda Língua para Estrangeiros*. Tese de doutorado: Orientadora: MEYER, Rosa Marina de Brito, PUC/ RIO, 2014.

PERINI, Mário Alberto. *Modern Portuguese: a reference grammar*. Yale University: Yale language series, 2002.

SANTOS & BICK. *Museu da Pessoa: corpus*. 2000. Disponível em 10 de dezembro de 2014: <http://www.linguateca.pt/acesso/corpus.php?corpus=MUSEUDAPESSOA>

TAGNIN, Stella Esther Ortweiler. *Expressões idiomáticas e convencionais*. São Paulo: Ática, 1989.

## **Anexos**

Neste anexo, apresentamos todas as buscas realizadas no corpus com fins de validar os resultados apresentados nessa pesquisa, seja por meio de tabela, seja por meio de gráficos.

### **1. Análise dos sintagmas preposicionais**

#### **O VERBO SER**

Para calcularmos os gráficos, foram feitos dois tipos de pesquisas separadas:

Na primeira, utilizamos a seguinte busca, com fins de descobrir as preposições que sucediam o verbo ser: [lema="ser"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*"]

Diante da necessidade de encontrar todos os casos das preposições mais recorrentes, realizamos a seguinte busca, com os seguintes códigos:

PREPOSIÇÃO EM: [lema="ser"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="em|em\+.\*"]

TOTAL ENCONTRADO: 1108 CASOS ENCONTRADOS (32,14%)

PREPOSIÇÃO PARA: [lema="ser"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="para|para\+.\*"] (16,94%)

TOTAL ENCONTRADO: 584 CASOS ENCONTRADOS

PREPOSIÇÃO DE: [lema="ser"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de|de\+.\*"] (16,15%)

TOTAL ENCONTRADO: 557 CASOS ENCONTRADOS

RESTANTE DAS PREPOSIÇÕES 34, 76%

TOTAL DE CASOS ENCONTRADOS COM TODAS AS PREPOSIÇÕES: 3447

## O VERBO ESTAR

Para calcularmos os gráficos, foram feitos dois tipos de pesquisa separadamente:

Na primeira, utilizamos a seguinte busca, com fins de descobrir as preposições que sucediam o verbo ser: [lema="estar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*"]

Diante da necessidade de encontrar todos os casos das preposições mais recorrentes, realizamos a seguinte busca, com os seguintes códigos:

PREPOSIÇÃO EM: [lema="estar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="em|em\+.\*"]

TOTAL ENCONTRADO: 1787 CASOS ENCONTRADOS (76,59 %)

PREPOSIÇÃO A: [lema="estar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="a|a\+.\*"]

TOTAL ENCONTRADO: 118 CASOS ENCONTRADOS (5,05%)

PREPOSIÇÃO DE: [lema="estar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de|de\+.\*"]

TOTAL ENCONTRADO: 113 CASOS ENCONTRADOS (4,84%)

PREPOSIÇÃO COM: [lema="estar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="com|com\+.\*"]

TOTAL ENCONTRADO: 102 CASOS ENCONTRADOS (0,43%)

RESTANTE DAS PREPOSIÇÕES 13,09%

TOTAL DE CASOS ENCONTRADOS COM TODAS AS PREPOSIÇÕES: 2333

## O VERBO IR

Para calcularmos os gráficos, foram feitos dois tipos de pesquisas separadamente:

Na primeira, utilizamos a seguinte busca, com fins de descobrir as preposições que sucediam o verbo ser: [lema="ir"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*"]



Diante da necessidade de encontrar todos os casos das preposições mais recorrentes, realizamos a seguinte busca, com os seguintes códigos:

PREPOSIÇÃO PARA: [lema="ir"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="para|para\+.\*"]  
 TOTAL ENCONTRADO: 2095 CASOS ENCONTRADOS (55,55 %)  
 PREPOSIÇÃO A: [lema="ir"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="a|a\+.\*"]  
 TOTAL ENCONTRADO: 872 CASOS ENCONTRADOS (23,12%)  
 PREPOSIÇÃO DE: [lema="ir"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de|de\+.\*"]  
 TOTAL ENCONTRADO: 283 CASOS ENCONTRADOS (7,50%)  
 PREPOSIÇÃO EM: [lema="ir"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="em|em\+.\*"]  
 TOTAL ENCONTRADO: 244 CASOS ENCONTRADOS (6,47%)  
 RESTANTE DAS PREPOSIÇÕES 7,35%  
 TOTAL DE CASOS ENCONTRADOS COM TODAS AS PREPOSIÇÕES: 3771

## O VERBO DAR

Para calcularmos os gráficos, foram feitos dois tipos de pesquisas separadamente:

Na primeira, utilizamos a seguinte busca, com fins de descobrir as preposições que sucediam o verbo ser: [lema="dar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*"]

Diante da necessidade de encontrar todos os casos das preposições mais recorrentes, realizamos a seguinte busca, com os seguintes códigos:

PREPOSIÇÃO PARA: [lema="dar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="para|para\+.\*"]  
 TOTAL ENCONTRADO: 534 CASOS ENCONTRADOS (43,48 %)  
 PREPOSIÇÃO A: [lema="dar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="a|a\+.\*"]  
 TOTAL ENCONTRADO: 117 CASOS ENCONTRADOS (10,66%)  
 PREPOSIÇÃO DE: [lema="dar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de|de\+.\*"]  
 TOTAL ENCONTRADO: 109 CASOS ENCONTRADOS (8,87%)  
 PREPOSIÇÃO EM: [lema="dar"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="em|em\+.\*"]  
 TOTAL ENCONTRADO: 257 CASOS ENCONTRADOS (20,92%)  
 RESTANTE DAS PREPOSIÇÕES 16,07%  
 TOTAL DE CASOS ENCONTRADOS COM TODAS AS PREPOSIÇÕES: 1228

## O VERBO TER

Para calcularmos os gráficos, foram feitos dois tipos de pesquisas separadamente:

Na primeira, utilizamos a seguinte busca, com fins de descobrir as preposições que sucediam o verbo ser: [lema="ter"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*"]

Diante da necessidade de encontrar todos os casos das preposições mais recorrentes, realizamos a seguinte busca, com os seguintes códigos:

PREPOSIÇÃO EM: [lema="ter"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="em|em\+.\*"]

TOTAL ENCONTRADO: 1723 CASOS ENCONTRADOS (46,23%)

PREPOSIÇÃO PARA: [lema="ter"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="para|para\+.\*"]

TOTAL ENCONTRADO: 467 CASOS ENCONTRADOS (12,53%)

PREPOSIÇÃO COM: [lema="ter"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="com|com\+.\*"]

TOTAL ENCONTRADO: 405 CASOS ENCONTRADOS (10,86%)

PREPOSIÇÃO DE: [lema="ter"] [pos!="V.\*"]\* @[func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de|de\+.\*"]

TOTAL ENCONTRADO: 494 CASOS ENCONTRADOS (13,25%)

RESTANTE DAS PREPOSIÇÕES 17,03%

TOTAL DE CASOS ENCONTRADOS COM TODAS AS PREPOSIÇÕES: 3727

Quando o pesquisador teve o interesse de apresentar as linhas de frequência dos nomes que seguiam a preposição nos sintagmas preposicionais, foi realizada a seguinte busca:

- Nos casos em que havia um *nome* após a preposição: [lema="\_\_\_\_\_"] [pos!="V.\*"]\* [func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de.\*"] [pos="SPEC.\*|DET.\*"]\* [pos="ADJ.\*"]\* @[pos="N"]

- Nos casos em que havia um *verbo* após a preposição: [lema="\_\_\_\_\_"] [pos!="V.\*"]\* [func="<PIV|<ADV.\*" & pos="PRP.\*" & lema="de.\*"] @[pos="V"]

## 2) Análise dos elementos morfológicos do tipo nome e verbo.

- Para cada verbo foi realizada a seguinte busca por lemas, quando se tratava de **nomes**, respectivamente:

[lema="ser"] @[pos="N.\*"]

[lema="estar"] @[pos="N.\*"]

[lema="ir"] @[pos="N.\*"]

[lema="dar"] @[pos="N.\*"]

[lema="ter"] @[pos="N.\*"]

- Para cada verbo foi realizada a seguinte busca por lemas, quando se tratava de **verbos**, respectivamente:

[lema="ser"] @[pos="V.\*"]

[lema="estar"] @[pos="V.\*"]

[lema="ir"] @[pos="V.\*"]

[lema="dar"] @[pos="V.\*"]

[lema="ter"] @[pos="V.\*"]