



Universidade de Brasília

Instituto de Letras

Departamento de Linguística, Português e Línguas Clássicas

**COLOCAÇÕES LEXICAIS EM PORTUGUÊS E FRANCÊS:  
ocorrências em *corpora* e obras lexicográficas**

Elisa Maiby Carvalho Augusto

Brasília - DF

2020



Universidade de Brasília

Instituto de Letras

Departamento de Linguística, Português e Línguas Clássicas

Elisa Maiby Carvalho Augusto

**Colocações lexicais em português e francês:**

ocorrências em *corpora* e obras lexicográficas

Trabalho apresentado à disciplina “Projeto de Curso: Elaboração de Multimeios”, do Departamento de Linguística, Português e Línguas Clássicas do Instituto de Letras da Universidade de Brasília, como requisito parcial para obtenção de grau de licenciatura em Letras – Português do Brasil como Segunda Língua, sob orientação da Profa. Dra. Orlene Lúcia de Saboia Carvalho.

Brasília - DF

2020

## RESUMO

Esta pesquisa tem como embasamento teórico a Linguística de *Corpus* (KUHNS; FERREIRA, 2020; BERBER SARDINHA, 2004) e a Lexicografia (POLGUÈRE, 2018; CARVALHO, 2001), que motivaram a observação das possibilidades de combinatória lexical restrita, conforme o discutido por Carvalho (2017) e Williams (2001). O presente trabalho é pautado pelos seguintes percursos metodológicos: i) revisão bibliográfica sobre o fenômeno das colocações lexicais, que compõem o objeto de estudo; ii) identificação e análise do fenômeno estudado em *corpora* das línguas românicas português e francês, tendo como foco as ocorrências de combinação entre verbo + substantivo/adjetivo/advérbio (colocações verbais); iii) investigação da amostra de ocorrências em diferentes seções dos verbetes constantes de obras lexicográficas monolíngues e bilíngues. Como resultado, verificou-se a alta frequência de uso das colocações lexicais nas ferramentas de *corpora*, que se opõe à presença relativamente baixa destas combinações nos dicionários sob análise; discorreu-se, também, sobre a possibilidade de consulta às ferramentas analisadas, de forma mais consciente sobre o fenômeno colocacional, no processo de ensino-aprendizagem de léxico.

**Palavras-chave:** Colocações lexicais. Combinatória lexical. Lexicografia.

## ABSTRACT

The present research is based on the theoretical framework of *Corpus* Linguistics (KUHN; FERREIRA, 2020; BERBER SARDINHA, 2004) and Lexicography (POLGUÈRE, 2018; CARVALHO, 2001). These fields motivated the observation of the possibilities of restricted lexical combinatorics, as discussed by Carvalho (2017) and Williams (2001). Our work is guided by the following methodological paths: i) bibliographic review on the phenomenon of lexical collocations, which compose the object of study; ii) identification and analysis of such phenomenon in *corpora* of the Portuguese and French languages, focusing on the occurrences of a combination of verb + noun / adjective / adverb (verbal collocations); iii) investigation of the sample of occurrences in different sections of entries in monolingual and bilingual dictionaries. As a result, we verified a high frequency of use of lexical collocations in *corpora* tools, in contrast to the relatively low presence of such combinations in the dictionaries under analysis. We also discussed the possibility of consulting the analyzed tools in a more conscious way about the collocational phenomenon, in the lexicon teaching-learning process.

**Keywords:** Lexical collocations. Lexical combinations. Lexicography.

*Palavra puxa palavra, uma ideia traz outra. [...]*

(Machado de Assis)

## SUMÁRIO

<b>Introdução</b>	<b>7</b>
<b>1 Colocações e demais unidades linguísticas convencionais</b>	<b>8</b>
1.1 Tipologias de colocações	9
<b>2 Corpora de colocações verbais em português e em francês</b>	<b>14</b>
2.1 <i>O corpus Portuguese Web 2011</i>	15
2.2 Contraste entre os <i>corpora</i> de português e francês	20
<b>3 As colocações em obras lexicográficas monolíngues e bilíngues</b>	<b>23</b>
3.1 Dicionários monolíngues padrão	23
3.2 Dicionários bilíngues	28
3.2.1 Informatizados	29
3.2.2 Impressos	32
<b>Considerações finais</b>	<b>36</b>
<b>Referências</b>	<b>38</b>
<b>Anexo I – <i>Wordlist</i>: verbos e substantivos mais comuns da língua portuguesa, por Kilgarriff e Rychlý (s.d.)</b>	<b>41</b>

## INTRODUÇÃO

Os falantes de português brasileiro e das demais línguas naturais materializam atos comunicativos dotados de significado por meio de construções textuais de natureza simples, como conversas espontâneas ou narrativas, e complexas, como aulas ou documentos oficiais. Essas construções mobilizam diferentes recursos linguísticos e discursivos, e um dos recursos mais utilizados para atribuir significado às interações, ainda que não seja suficientemente explorado, é o de combinação lexical. Trata-se de uma relação sintagmática que permite associar as palavras que são usadas, normalmente, “em conjunto com outras palavras que as completam, ou no nível da construção de discurso que faz eco delas quando substituídas ou definidas” (TRÉVILLE; DUQUETTE, 1996 *apud* BINON; VERLINDE, 2000a, p. 124).

Tomando por base esta constatação, apoiada inicialmente nos trabalhos de Carvalho (2017; 2001), Polguère (2018) e Williams (2001), e orientada metodologicamente pela pesquisa bibliográfica e documental (GIL, 2008), procedemos ao trabalho monográfico de pesquisa sobre combinações lexicais, também denominadas colocações lexicais ou, simplesmente, colocações, nos contextos gerais de uso das línguas românicas português e francês. Estas línguas foram escolhidas tanto pelo número expressivo e similar de falantes ao redor do globo, quanto pelo vínculo da pesquisadora como falante nativa da variedade brasileira do português (daqui em diante, PB) e aprendiz de longa data da variedade europeia do francês.

A partir das inquietações decorrentes das reflexões sobre língua e linguagem, materializadas na pergunta “quais são as principais possibilidades de combinatória lexical em português e francês, em uso corrente?”, o presente trabalho foi organizado de forma a evidenciar as tipologias e critérios lexicais, sintáticos e semânticos que subdividem as colocações lexicais no primeiro capítulo, para, a partir desta subdivisão, destacar as ocorrências de uso das colocações verbais e possíveis implicações no processo de ensino-aprendizagem do léxico de PB, no segundo capítulo. Estas mesmas colocações lexicais serão analisadas, de igual modo, em obras lexicográficas monolíngues e bilíngues, no terceiro capítulo. Ao final, anexamos uma lista com os substantivos e verbos mais frequentes no PB, gerada a partir das ocorrências analisadas no segundo capítulo.

## 1 COLOCAÇÕES E DEMAIS UNIDADES LINGUÍSTICAS CONVENCIONAIS

As interações humanas que se estabelecem no mundo são mediadas por diferentes construções linguísticas e discursivas. Nosso estudo tratará da forma como essas construções mobilizam recursos de combinação lexical, que são um dos padrões recursivos que as línguas naturais oferecem (CARVALHO, 2017, p. 82). Esses padrões são parte importante do processo criativo dos falantes de determinada língua, pois é a partir deles que geramos “expressões estruturadas sem limites” (BERWICK; CHOMSKY, 2011 *apud* NEGRÃO, 2013, p. 88), que não são escolhidas de forma aleatória, mas obedecem a critérios sintáticos e semânticos (BERBER SARDINHA, 2004, p. 200; COSERIU, 1977, p. 143-161).

Em realidade, quando adquirimos ou aprendemos uma língua, não nos apropriamos das palavras de forma isolada, mas, sim, no âmbito de interações que demandam a mobilização de diferentes recursos: a maneira como falamos com nossos familiares e amigos próximos não costuma ser a mesma de que nos valem para interagir com colegas de trabalho; difere-se também da maneira como um autor se dirige aos seus leitores. Nessas diferentes interações, espera-se que saibamos o que dizer, quando dizer e como dizer (TAGNIN, 2013, p. 21) e, para facilitar este processo, temos à disposição alguns “sintagmas pré-construídos” (CARVALHO, 2017, p. 81), que são usados com grande frequência pelos interlocutores nativos e, idealmente, pelos não nativos em imersão. Esses sintagmas, no PB, podem ser:

- a) Fórmulas de rotina: expressões de função comunicativa, como saudações, agradecimentos e desculpas;
- b) Expressões idiomáticas: combinação de itens lexicais e gramaticais que funcionam em uma unidade de sentido, como “dor de cotovelo”;
- c) Binômios: estruturas coordenativas, como “noite e dia”, “ir e vir”, “cão e gato”. Alguns binômios se assemelham ao item (b), como “gato e sapato”;
- d) Coligações: combinação de um item lexical com um item gramatical, como “fugir de [algo/alguém]”. Na gramática tradicional, os fenômenos de coligação são categorizados como regência verbal ou nominal;

e) Colocações: combinação de dois ou mais<sup>1</sup> itens lexicais em coocorrência regular, como “linha tênue” (CARVALHO, 2017, p. 82-83; TAGNIN, 2013, p. 53 et seq.; CRYSTAL, 2008, p. 51, 104), entre outros.

Como os itens lexicais são palavras que possuem referentes no mundo extralinguístico, como substantivos, verbos e adjetivos, enquanto os itens gramaticais possuem referentes apenas no mundo linguístico, como preposições, conjunções e artigos (VILARINHO, 2019, p. 152, 165; CRYSTAL, 2008, p. 222), interessa-nos analisar, neste primeiro momento, o quanto as colocações lexicais perpassam nossas interações como falantes de PB.

### 1.1 Tipologias de colocações

Os estudos de combinatória lexical, bem como a utilização do termo “colocação” na linguística, foram difundidos por John R. Firth (1890-1960), para quem as colocações lexicais eram observáveis por meio de um “teste de colocabilidade”. Este teste, que revela aspectos semânticos das palavras, foi demonstrado de forma simples pelos pares “noite/dia” e “claro/escuro”, em que se percebe uma maior regularidade de combinação entre “noite” e “escura”, “dia” e “claro” (FIRTH, 1951 *apud* ROBINS, 2013, p. 65; BARNBROOK et al., 2013, p. 37).

Como o grau de dependência entre os itens lexicais na construção no significado pode variar, atualmente, os exemplos de Firth poderiam ser considerados como sintagmas livres, mais do que como colocações, uma vez que se pode substituir “noite” por “casa”, “óculos” e outros tantos; e “escura” por “fria”, “longa” etc. (CARVALHO, 2017). Em realidade, em uma colocação, é necessário que ambos os itens lexicais estejam envolvidos em uma relação de solidariedade semântica, normalmente manifesta no eixo sintagmático, para conferir às palavras escolhidas uma especificidade de sentido, como “montar a cavalo” (COSERIU, 1977, p. 144, 153).

Assim, dentre as propostas tipológicas mais recentes para as colocações lexicais – daqui em diante, também denominadas apenas “colocações” – destacamos as de Deradra

---

<sup>1</sup> Apesar da discussão quanto ao número ideal de itens lexicais em uma colocação (CARVALHO, 2017, p. 84), adotaremos dois itens como unidade padrão em nosso trabalho, sobretudo porque os sintagmas tendem a ser de natureza binária (cf. Mikus, 1947).

(2009a; 2009b), Tagnin (2013) e Carvalho (2017). Tais propostas têm em comum o fato de perceberem os itens lexicais em uma relação assimétrica de solidariedade: Deradra (2009a, p. 240, tradução nossa) fala de “lexia-base” e “constituente/colocado”, enquanto Tagnin (2013, p. 64-65) trata de “base” e “colocado/elemento convencional”, de forma similar à “base/nódulo” e “colocado”<sup>2</sup> de Carvalho (2017, p. 84). Como as semelhanças se situam, basicamente, apenas neste âmbito, passamos à discussão das diferenças de cada proposta.

A pesquisa de mestrado de Deradra (2009a; 2009b) visa apresentar “o contexto em que se emprega determinada colocação” a partir de um *corpus* literário: o romance *Il était une fois un vieux couple heureux*, de Mohammed Khaïr-Eddine (DERADRA, 2009a, p. 239, tradução nossa). Para a autora, critérios de ordem sintático-semântica, como testes de movimentação dos constituintes (*ibid.*, p. 241), e critérios semânticos, como os de binaridade, assimetria e coocorrência restrita/seleção lexical, orientam o “grau de fixidez” das colocações, que podem ser

[...] opacas: quando o sentido do colocado em coocorrência com a base é diferente do sentido que tem fora dessa combinação, [...] [enquanto] o sentido da base se mantém interpretável;  
transparentes: quando o sentido do colocado é facilmente compreensível [...];  
regulares: [...] combinações de palavras ou de sequências de palavras nas quais o sentido do todo é, geralmente, dedutível e parece previsível [...] (*ibid.*, p. 240, tradução nossa)<sup>3</sup>.

Além desses critérios, Deradra (2009b) emprega o consagrado critério morfossintático de divisão por classes de palavras<sup>4</sup> para categorizar as colocações como:

- a) colocações nominais, compostas por: substantivo (posição 1) + preposição [posição opcional] + substantivo (posição 2)  
**la nuit des temps / tempos remotos;**
- b) colocações verbais, compostas por: (verbo) + [preposição] + (substantivo ou adjetivo)  
**crever de faim / morrer de fome;**

---

<sup>2</sup> Estes termos também são aplicáveis à busca por ocorrências linguísticas em pesquisas da área de Linguística de *Corpus*, como veremos no segundo capítulo.

<sup>3</sup> Original: “[Collocations] opaques : lorsque le sens du collocatif en cooccurrence avec la base est différent du sens qu’il a en dehors de cette association, [...] [tandis que] le sens de la base reste interprétable ; transparentes : lorsque le sens du collocatif est facilement compréhensible [...] ; régulières, autrement dit des associations de mots ou de suites de mots dans lesquelles le sens du tout est généralement déductible et semble prédictible [...].”

<sup>4</sup> Cf. Tagnin (2013, p. 63).

- c) colocações adjetivas, compostas por: (adjetivo) + (substantivo ou advérbio)  
***armé jusqu'aux dents / armado até os dentes;***
- d) colocações adverbiais, compostas por: (advérbio ou locução adverbial) + (adjetivo ou verbo)  
***en même temps / ao mesmo tempo;***
- e) colocações prepositivas e conjuntivas, compostas por: (preposição ou conjunção) + (substantivo)  
***à défaut de / na falta de***<sup>5</sup> (DERADRA, 2009b, p. 52 et seq., tradução nossa).

Observamos que, para a autora, vários dos sintagmas pré-construídos se fundem em um só conceito. “Colocações”, segundo os dados de Deradra (2009a; 2009b), abrangem tanto a combinação de itens lexicais (item *b*), quanto a combinação de itens lexicais e gramaticais (item *e*) e, ainda, expressões idiomáticas, por ela denominadas “colocações opacas” (itens *a* e *c*). Assim, consideramos que os trabalhos de Tagnin (2013) e Carvalho (2017) delimitam mais o escopo deste fenômeno.

Tagnin (2013) apresenta, em seu livro *O jeito que a gente diz: combinações consagradas em inglês e português*, as possibilidades de combinação lexical das duas línguas em contraste. A autora sinaliza, como Deradra (2009a), que as colocações lexicais podem apresentar um maior ou menor nível de fixidez de acordo com a categoria a que pertencem, mas difere da linguista argelina ao considerar que a opacidade semântica (ou “idiomaticidade”) não diz respeito às colocações, mas, sim, às expressões idiomáticas (TAGNIN, 2013, p. 22, 99). Para a autora, há:

- a) colocações adjetivas, compostas por (adjetivo) + (substantivo):  
***foreign policy / política externa;***
- b) colocações nominais, compostas por (substantivo) + [preposição] + (substantivo):  
***credit card / cartão de crédito;***
- c) colocações verbais, compostas por (verbo) + [preposição] + (substantivo ou adjetivo):  
***take care / tomar cuidado, go wrong / dar errado;***
- d) colocações adverbiais, compostas por (advérbio) + [preposição] + (adjetivo ou verbo):  
***strictly prohibited / expressamente proibido, fully agree / concordar plenamente*** (*ibid.*, p. 63-73).

---

<sup>5</sup> Nas traduções de Deradra (2009a; 2009b) e Tagnin (2013), consideramos que, por não haver “isomorfia lexical entre duas línguas” (CARVALHO, 2001, p. 111), algumas colocações em português brasileiro podem ser formadas por palavras de classes diferentes ou demandar diferentes preposições.

Tagnin (2013) tipifica, em seguida, as expressões especificadoras de unidade, que acompanham “substantivos que designam objetos incontáveis ou abstratos”, e os coletivos, “convencionalmente usados para se referir a um grupo de coisas” (TAGNIN, 2013, p. 74-75). A autora termina o capítulo com a exposição de quadros comparativos em que amplia sua análise contrastiva, pois adiciona equivalentes de colocações em alemão, espanhol, francês e italiano (*ibid.*, p. 77-78).

Na esteira do proposto por Tagnin (2013), Carvalho (2017, p. 82-84) adota os critérios de “autonomia sintática e transparência semântica” para subdividir as colocações lexicais do PB em:

- a) colocações nominais, compostas por (substantivo ou adjetivo) + [preposição] + (substantivo ou adjetivo):  
**suaves prestações, modéstia à parte;**
- b) colocações adjetivas, compostas por (advérbio) + (adjetivo ou particípio):  
**gravemente ferido, redondamente enganado;**
- c) colocações verbais (verbo ou substantivo) + [preposição] + (substantivo, verbo, adjetivo ou advérbio):  
**prestar atenção, levar a sério, mentir descaradamente** (CARVALHO, 2017, p. 82-87).

Carvalho (2017) analisa, ainda, o quanto os traços semânticos positivos e negativos se estendem de um item lexical a outro, mediante o conceito de “prosódia semântica” (STUBBS, 1995 *apud* CARVALHO, 2017, 87-88): o verbo oferecer, por exemplo, apresenta prosódia semântica positiva, pois se associa a “cursos, serviços, vagas, oportunidades, condições, subsídios, vantagens [...]” (*ibid.*, p. 90). A autora também revisa dados de *corpora* de português e de obras lexicográficas, para propor, ao final, um processo de ensino-aprendizagem de PB como língua estrangeira mais atento às questões de combinatória lexical nos diferentes gêneros textuais que integram o processo.

Desta feita, pode-se dizer que uma colocação é uma combinação lexical, geralmente binária, cujos itens se relacionam em regime de solidariedade e transparência semântica. Podemos dizer, de igual modo, que as colocações não são sintagmas abertos (**casa amarela, anexar território**), uma vez que ambos os itens lexicais podem ser trocados sem prejuízo da constituição do sintagma; tampouco são expressões idiomáticas (**quebrar um galho, virar a**

**página**), pois ainda que haja uma relação explícita em nível sintático, a semântica dos itens não é transparente<sup>6</sup>.

Continuaremos a análise, no capítulo seguinte, de colocações verbais do português, porque, como discutido por Tagnin (2013) e Carvalho (2017), esta categoria de colocações evidencia as dificuldades que ocorrem no ensino-aprendizagem do léxico de PB como língua estrangeira: aprender os tipos de complementos que são demandados por um verbo, sem ter noção de quais são os complementos verbais de maior frequência, pode resultar em equívocos durante situações comunicativas concretas. Tais dificuldades decorrem, dentre outros fatores, da escassez de obras lexicográficas dedicadas à descrição das colocações em geral, conforme discutiremos no último capítulo. Carvalho (2017, p. 92), aliás, constata que “não há dicionários monolíngues direcionados para aprendizes de português brasileiro como língua estrangeira, seja semasiológico, de frequência ou de colocações”, e Tagnin (2013, p. 70) reforça que não há dicionário de colocações, “mesmo de língua materna”.

Assim, com base na metodologia de coleta de dados adotada por Carvalho (2017), discutiremos as colocações verbais da língua em uso no próximo capítulo, e contrastaremos essas colocações em dicionários de português, nas variedades brasileira e europeia, e francês europeu no capítulo seguinte, adotando a tipologia de colocações proposta pela mesma autora (CARVALHO, 2017, p. 80-87).

---

<sup>6</sup> Lembramos que, conforme Williams (2001), ainda não há uma definição de colocação considerada consensual, posto que esse tipo de combinação se situa num *continuum* entre sintagmas abertos, locuções e expressões idiomáticas.

## 2 *CORPORA* DE COLOCAÇÕES VERBAIS EM PORTUGUÊS E EM FRANCÊS

Feita a diferenciação entre as categorias de colocações, lançamo-nos à análise de colocações verbais, conforme o levantamento de dados que fizemos no *corpus* linguístico de português *Portuguese Web 2011*. Antes, porém, é necessário discorrer sobre a constituição de um *corpus* (pl. *corpora*) e quais são as possibilidades de trabalho que ele oferece.

Os *corpora* de texto são ferramentas da maior relevância para a Linguística de *Corpus*, uma corrente de análise linguística inaugurada nos anos 1960 cujo principal objetivo é explorar a “linguagem por meio de evidências empíricas, extraídas [de gêneros textuais diversos] por computador.” (BERBER SARDINHA, 2004, p. 3). Na era da sociedade da informação, existem *corpora* com bilhões de palavras que oferecem possibilidades de análise, por exemplo, dos contextos de ocorrência (concordâncias), das palavras mais frequentes, das relações sintáticas e semânticas, entre outros.

Os *corpora* utilizados na atualidade disponibilizam conhecimento sobre uma determinada língua ou variedade linguística por meio de uma compilação de dados textuais, que são, normalmente, coletados de documentos *on-line*, como *sites*, documentos nos formatos HTML, PDF etc., e são acessados de acordo com as ferramentas disponíveis. Para o português, destacamos os *corpora* Brasileiro<sup>7</sup>, coordenado por Tony Berber Sardinha, *Corpus do Português*<sup>8</sup>, coordenado por Mark Davies, e *Portuguese Web 2011*<sup>9</sup>, coordenado pela equipe do Sketch Engine. Este último *corpus* foi o escolhido para nossa pesquisa por sua maior disponibilidade de ferramentas de análise e de dados compilados.

Ressaltamos que os *corpora* foram inicialmente pensados para linguistas, mas sua utilização não se restringe ao contexto acadêmico. Profissionais e curiosos das áreas do jornalismo, audiovisual, licenciaturas, mídias sociais, entre outras, podem se beneficiar dessas compilações de dados textuais já existentes e até mesmo elaborar seus próprios *corpora*, posto que as ocorrências do uso permitem a exploração atenta, em nível quantitativo e qualitativo, da língua que se fala, observa e aprende (cf. KUHN; FERREIRA, 2020).

---

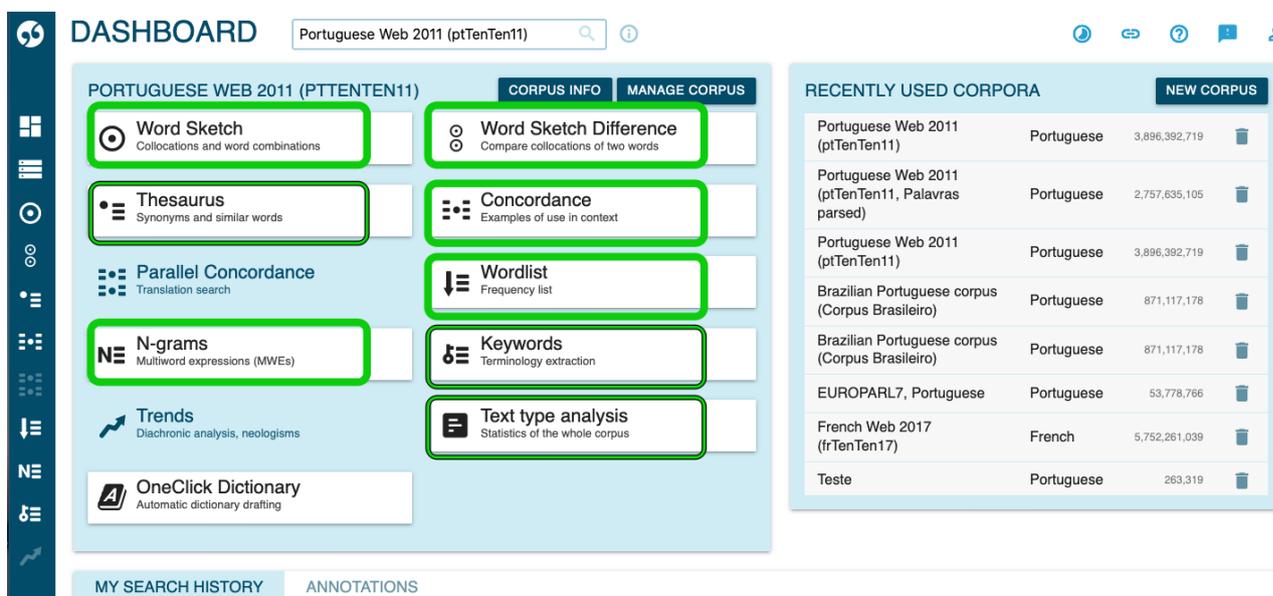
<sup>7</sup> O acesso ao *corpus* Brasileiro e a outros *corpora* de língua portuguesa pode ser feito de forma gratuita em: <https://www.linguateca.pt/ACDC/> (LINGUATECA).

<sup>8</sup> Também gratuito. Disponível em: <https://www.corpusdoportugues.org/xp.asp> (DAVIES).

<sup>9</sup> Disponível em: <https://app.sketchengine.eu> (KILGARRIFF; RYCHLÝ).

## 2.1 O corpus Portuguese Web 2011

O corpus *Portuguese Web 2011* (ou *ptTenTen11*) integra a família *TenTen*, que conta com 41 *corpora* gerenciados pelo Sketch Engine, programa pago da empresa britânica Lexical Computing. Desde a última atualização, de 2018, o *ptTenTen11* armazena mais de 10 milhões de documentos coletados de *sites* científicos e acadêmicos, com 3.896.392.719 palavras, em português brasileiro e europeu, cujo acesso se dá por meio das ferramentas do Sketch Engine destacadas a seguir:



The screenshot displays the Sketch Engine dashboard for the Portuguese Web 2011 corpus. The main area is titled 'DASHBOARD' and shows the corpus name 'Portuguese Web 2011 (ptTenTen11)'. Below this, there are two columns of tool cards. The left column includes: Word Sketch (Collocations and word combinations), Thesaurus (Synonyms and similar words), Parallel Concordance (Translation search), N-grams (Multiword expressions (MWEs)), Trends (Diachronic analysis, neologisms), and OneClick Dictionary (Automatic dictionary drafting). The right column includes: Word Sketch Difference (Compare collocations of two words), Concordance (Examples of use in context), Wordlist (Frequency list), Keywords (Terminology extraction), and Text type analysis (Statistics of the whole corpus). A sidebar on the left contains navigation icons. A table on the right lists 'RECENTLY USED CORPORA' with columns for corpus name, language, and word count. The table includes entries for Portuguese Web 2011 (ptTenTen11), Portuguese Web 2011 (ptTenTen11, Palavras parsed), Brazilian Portuguese corpus (Corpus Brasileiro), and others.

Corpus Name	Language	Word Count
Portuguese Web 2011 (ptTenTen11)	Portuguese	3,896,392,719
Portuguese Web 2011 (ptTenTen11, Palavras parsed)	Portuguese	2,757,635,105
Portuguese Web 2011 (ptTenTen11)	Portuguese	3,896,392,719
Brazilian Portuguese corpus (Corpus Brasileiro)	Portuguese	871,117,178
Brazilian Portuguese corpus (Corpus Brasileiro)	Portuguese	871,117,178
EUROPARL7, Portuguese	Portuguese	53,778,766
French Web 2017 (frTenTen17)	French	5,752,261,039
Teste	Portuguese	263,319

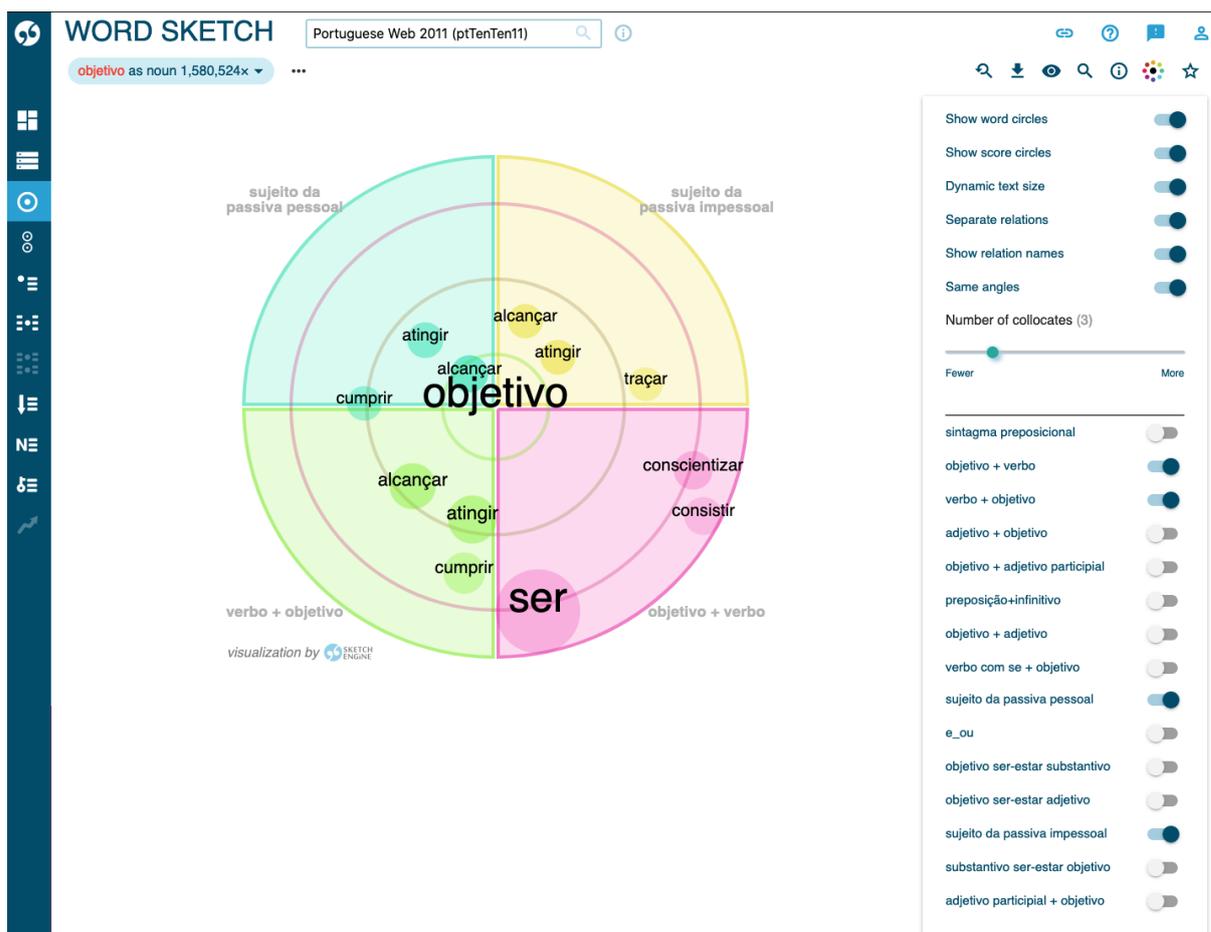
Figura 1 – Área inicial do Sketch Engine

Fonte: Kilgarriff; Rychlý (s.d.).

No corpus *Portuguese Web 2011*, lançamos mão das ferramentas *Wordlist* (lista das palavras mais frequentes no corpus), *Word Sketch* (lista de combinações lexicais e gramaticais) e *Word Sketch Difference* (comparação entre os itens lexicais que constituem uma colocação) para coletar os dados. As demais possibilidades de trabalho para este corpus são: compilar sinônimos e palavras similares (*Thesaurus*), exemplos de uso em contexto (*Concordance*), expressões com até 5 palavras (*N-grams*), extrair termos de áreas de

especialidade (*Keywords*) e estatísticas gerais dos documentos que compõem o *corpus* (*Text type analysis*), em uma interface de *site* apresentada em inglês, e dados na língua original.

Eis as seguintes combinações das quais a palavra **objetivo**, 55º substantivo mais usado no português<sup>10</sup>, é base/núcleo. Selecionamos os contextos de ocorrência de **Verbo + Substantivo** e **Substantivo + Verbo**, nas vozes ativa e passiva.



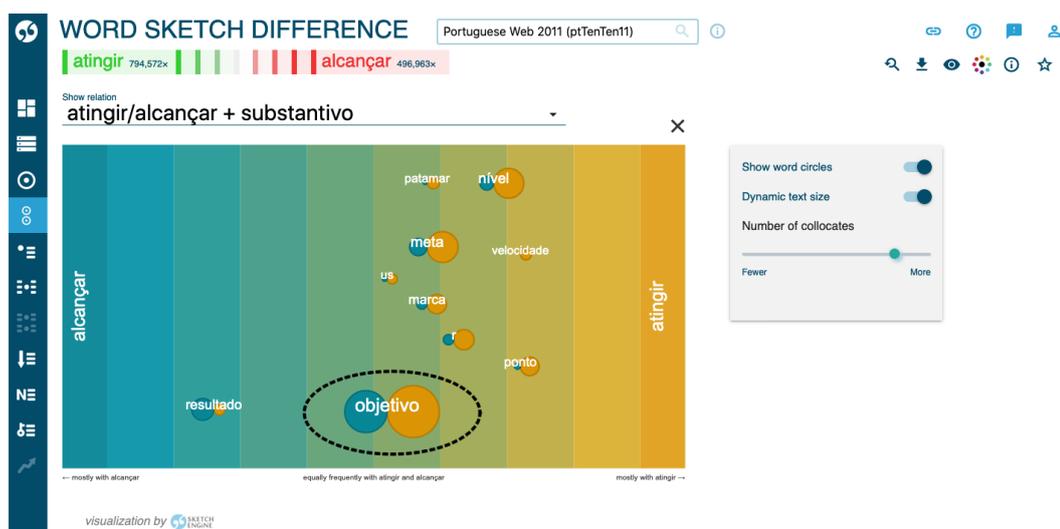
**Figura 2** – Ferramenta *Word Sketch*: “objetivo”

Fonte: Kilgarriff, Rychlý (s.d.).

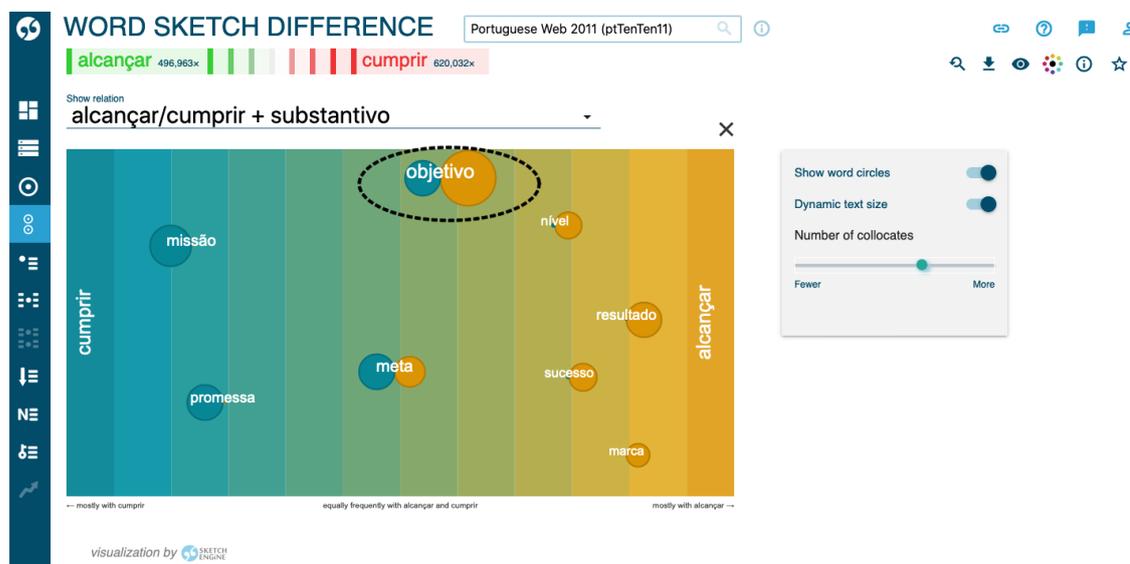
Nas categorias de colocações sob investigação, percebe-se uma maior frequência dos colocados **atingir**, **alcançar** e **cumprir** em coocorrência com **objetivo**, sendo **atingir + objetivo** e **alcançar + objetivo** as colocações de frequência mais alta. Isso é confirmado pelos

<sup>10</sup> Cf. anexo I.

scores<sup>11</sup>, índices numéricos de frequência das ocorrências nos contextos apontados pela Figura 1: a média simples do *score* para **alcançar** + **objetivo**, é de 10.97, para **atingir** + **objetivo**, de 10.82, e para **cumprir** + **objetivo**, de 9.2. A comparação do *Word Sketch Difference* também demonstra estas nuances, como demonstrado nas figuras a seguir:

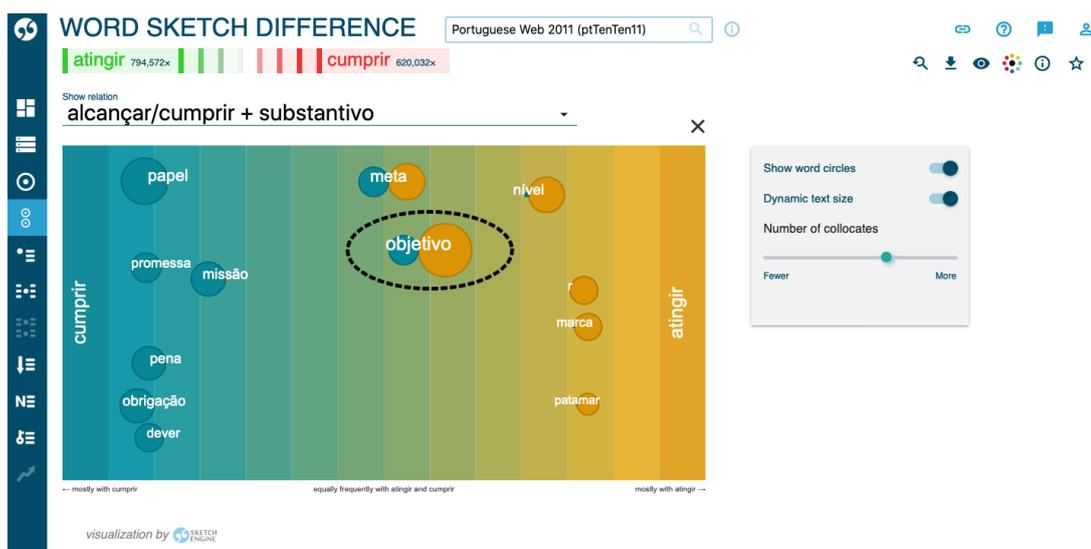


**Figura 3** – Ferramenta *Word Sketch Difference*: alcançar vs. atingir  
 Fonte: Kilgarriff; Rychlý (s.d.).



**Figura 4** – Ferramenta *Word Sketch Difference*: cumprir vs. alcançar  
 Fonte: Kilgarriff; Rychlý (s.d.).

<sup>11</sup> Os critérios de cálculo do *score* podem ser conferidos na página “Statistics used in Sketch Engine” (SKETCH ENGINE, 2015).



**Figura 5** – Ferramenta *Word Sketch Difference*: cumprir vs. atingir  
 Fonte: Kilgarriff; Rychlý (s.d.).

Seguindo estes passos, compilamos as colocações verbais mais frequentes nos usos do português como língua geral<sup>12</sup> e as dispomos no quadro a seguir. Algumas dessas combinações já foram discutidas por Tagnin (2013) e Carvalho (2017) – com base nessas autoras, aliás, mantivemos o destaque dos itens lexicais considerados como base/nódulo em negrito.

**Quadro 1** – Colocações verbais no português brasileiro

Verbo	Substantivo / advérbio	Média simples do score
alcançar / atingir	<b>objetivo</b>	10.97 10.82
chamar	<b>atenção</b>	7.59
dar o	<b>exemplo</b>	7.21
dar	<b>continuidade</b>	6.86
falar / dizer	<b>besteira</b>	5.06 3.95

<sup>12</sup> Distinguimos língua geral de linguagem de especialidade, no sentido de a primeira materializar-se nas interações espontâneas, enquanto a segunda media as interações de especialistas em áreas diversas. Foram detectadas algumas colocações utilizadas no Direito e na Economia (**decretar estado de emergência, deixar legado, tributar serviço**), mas a discussão sobre esses dados é mais pertinente aos estudiosos da Terminologia, área da Linguística que estuda as linguagens de especialidade (cf. Vilarinho, 2019).

firmar / assinar	<b>acordo</b>	10.56 9.1
juntar	<b>esforços</b>	6.86
levar a	<b>sério</b> (adv.)	7.56
levar em	<b>consideração</b>	5.34
mentir	<b>descaradamente</b> (adv.)	7.0
pedir	<b>desculpas</b>	7.4
prestar	<b>serviço</b>	10.32
promover	<b>desenvolvimento</b>	6.2
ranger	<b>dentes</b>	7.97
tirar / levar	<b>vantagem</b>	5.79 5.64
tomar	<b>consciência</b>	5.99
travar	<b>batalha</b>	10.61
(não) vir ao	<b>caso</b>	5.98

Fonte: Kilgarriff; Rychlý (s.d.).

Chamamos atenção a alguns fatos linguísticos, para além da evidente relação sintático-semântica existente entre as colocações listadas. Para começar, pode-se observar que o padrão Verbo + Substantivo, na voz ativa, é o mais comum dentre as colocações verbais. No entanto, em colocações como **travar batalha**, **prestar serviço** e **firmar/assinar acordo**, há grande frequência de ocorrências na voz passiva.

Percebe-se, de igual modo, que o item lexical considerado como base/nódulo é o de maior carga semântica. Na categoria das colocações verbais, a base/nódulo costuma ser o substantivo, uma vez que os verbos considerados suportes, como **dar**, **falar**, **tomar** “não têm semântica plena – não são, no entanto, completamente vazios de significado” (CARVALHO, 2017, p. 86)<sup>13</sup>.

<sup>13</sup> Na pesquisa, também encontramos combinações lexicais com verbos de carga semântica mais forte, cujo fator de restrição semântica suplanta o de solidariedade semântica, que julgamos ser essencial em uma colocação (cf. capítulo 1.1). Por essa razão, não incluímos os pares **folhear livro** (verbo ligado a objetos com folhas) e **fidelizar cliente** (verbo restrito a relações comerciais) em nossas análises, ainda que seus elementos ocorram com frequência nos usos de PB.

Em colocações lexicais como **alcançar/atingir objetivo**, **falar/dizer besteira** ou **tirar/levar vantagem**, é possível que o colocado possa ser trocado por outro de significado similar, sem prejuízo da relação de solidariedade semântica, como apontado por Tagnin (2013, p. 64) e pelos resultados das ferramentas do *corpus*. A variação dos índices de frequência entre uma e outra colocação é baixa: em **alcançar/atingir um objetivo**, há uma lacuna de *score* de 0.15, sendo **alcançar objetivo** a mais frequente (média de 10.97); em **tirar/levar vantagem**, a lacuna de *score* também é de 0.15, com maior frequência de **tirar vantagem** (média de 5.79); em **falar/dizer besteira**, o *score* varia 1.11 e a colocação mais frequente é **falar besteira** (média de 5.06) (KILGARRIFF; RYCHLÝ).

Finalmente, ressaltamos que algumas colocações incorporam em si uma função tradicionalmente atribuída às coligações, ao demandarem preposições e artigos para se concretizarem nas situações comunicativas. Colocações como **dar o exemplo**, **levar a sério / em consideração**, [não] **vir ao caso** tornaram-se tão consolidadas na língua em uso que são, inclusive, consideradas como locuções em alguns dos dicionários analisados, como se verá no terceiro capítulo desta pesquisa.

## 2.2 Contraste entre os *corpora* de português e francês

A seguir, listam-se as equivalências das colocações analisadas em francês. Consultamos as equivalências tanto no *corpus* francês da família *TenTen*, quanto no *Dictionnaire des combinaisons de mots (Le Robert, 2007)*<sup>14</sup>. Salientamos que, embora estas línguas pertençam à mesma família, o processo histórico de formação de cada uma delas faz com que poucos itens lexicais apresentem equivalência plena: a formação linguística do português e do francês é perpassada pela relativa distância geográfica entre os territórios de Portugal e França, assim como pelo contato linguístico decorrente das ocupações dos mouros – no caso de Portugal –, normandos – no caso da França – e outros povos (BASSO; GONÇALVES, 2014).

---

<sup>14</sup> Agradecemos, ainda, ao sujeito nativo de francês, plurilíngue e amigo Yann Amoussou, que ratificou e esclareceu as nuances de algumas das colocações registradas.

**Quadro 2** – Colocações verbais no português e equivalências em francês

Colocação em português	Colocação em francês
atingir / alcançar <b>objetivo</b>	<i>atteindre un <b>objectif</b></i>
chamar <b>atenção</b>	<i>attirer l'<b>attention</b></i>
dar o <b>exemplo</b>	<i>donner l'<b>exemple</b></i>
dar <b>continuidade</b>	<i>donner <b>suite</b></i>
falar / dizer <b>besteira</b>	<i>dire / raconter des <b>bêtises</b></i>
firmar / assinar <b>acordo</b>	<i>conclure un <b>accord</b></i>
juntar <b>esforços</b>	<i>unir ses <b>efforts</b></i>
levar a <b>sério</b> (adv.)	<i>prendre au <b>sérieux</b></i>
levar em <b>consideração</b>	<i>prendre en <b>compte</b></i>
mentir <b>descaradamente</b>	<i>mentir <b>effrontément</b></i>
pedir <b>desculpas</b>	<i>présenter ses <b>excuses</b> / demander <b>pardon</b></i>
prestar <b>serviço</b>	<i>rendre <b>service</b></i>
promover <b>desenvolvimento</b>	<i>promouvoir le <b>développement</b></i>
ranger <b>dentes</b>	<i>grincer des <b>dents</b></i>
tirar / levar <b>vantagem</b>	<i>profiter / <b>bénéficier</b></i>
tomar <b>consciência</b>	<i>prendre <b>conscience</b></i>
travar <b>batalha</b>	<i>s'engager dans / livrer une <b>bataille</b></i>
(não) vir ao <b>caso</b>	<i>(ne pas) être <b>question de</b></i>

Fontes: *Le Robert* (2007); Kilgarriff; Rychlý (s.d.).

Em relação a estes dados, destacamos que o francês, via de regra, demanda um elemento determinante na ligação do verbo ao substantivo – artigo, preposição, pronome etc. –, como nas colocações *donner l'exemple, prendre au sérieux, joindre ses efforts, s'engager dans / livrer une bataille*. Sobre a inexistência de isomorfia lexical (CARVALHO, 2001), assinala-se que algumas das colocações encontradas em português não mantêm a mesma organização sintagmática em francês, como os itens lexicais *profiter / bénéficier [de quelque chose]*: estes foram os equivalentes mais próximos de **tirar / levar vantagem** encontrados.

Passemos, então, à investigação sobre estas colocações em ferramentas mais usuais que o *corpus*, ainda que tenham importância similar, e onipresentes no processo de ensino-aprendizagem de léxico: as obras lexicográficas monolíngues e bilíngues.

### 3 AS COLOCAÇÕES EM OBRAS LEXICOGRÁFICAS MONOLÍNGUES E BILÍNGUES

Os dicionários são um tipo de obra lexicográfica que oferecem um modelo de descrição lexical mais ou menos padronizado, possibilitando que especialistas e não especialistas de língua tenham acesso a um seleto repertório de itens lexicais e dispondo informações específicas sobre estes itens em verbetes. Há vários tipos de dicionário, com diferentes tamanhos e estruturas organizacionais:

dicionário histórico, dicionário etimológico, minidicionário, dicionário infantil, dicionário fundamental, dicionário escolar, dicionário para aprendizagem, dicionário bilíngue, dicionário semibilíngue, dicionário de uso, dicionário padrão, dicionário ilustrado, dicionário de imagens e dicionário analógico (VILARINHO; MAIA-PIRES, 2016, p. 88).

Dentre as obras lexicográficas tipificadas por Vilarinho e Maia-Pires (2016), importa-nos analisar o quanto as colocações se fazem presentes nas obras de referência para falantes de português e francês. Analisaremos os dicionários monolíngues padrão Houaiss (2009) e Aulete (*on-line*), assim como os bilíngues TheFreeDictionary / K Dictionaries (2013), Michaelis (*on-line*; 2002) e Larousse (1988), nos formatos impresso e informatizado.

#### 3.1 Dicionários monolíngues padrão

Um dicionário monolíngue apresenta o léxico da língua, à exceção de nomes próprios, “do ponto de vista de suas características linguísticas: pronúncia, parte do discurso, sentido”, etimologia etc., e enumera suas “conexões paradigmáticas”, como sinônimos, antônimos etc., e “sintagmáticas”, ou seja, “propriedades de combinatória restrita”, dentre as quais destacamos nosso objeto de estudo, colocações lexicais (POLGUÈRE, 2018, p. 242, 249). Alguns dicionários que apresentam essa gama de informações gramaticais sobre os itens lexicais são considerados dicionários padrão, pois visam “descrever o léxico do idioma na sua totalidade” (BIDERMAN, 1998, p. 165 *apud* VILARINHO, 2013, p. 59), incluindo o registro de “terminologias de áreas do conhecimento, lexemas da língua comum e variantes da norma

padrão”, para “atender a demandas de jovens e adultos das diversas áreas de formação” (VILARINHO; MAIA-PIRES, 2016, p. 99).

Consideramos, assim, para este subcapítulo, as versões informatizadas dos dicionários padrão Houaiss (2009) e Aulete (*on-line*). Ambos possuem cerca de 100.000 verbetes cada, e apresentam características similares de macroestrutura, como o ordenamento alfabético acessado por uma barra de pesquisa e a disposição das palavras-entrada pela forma canônica, e de microestrutura, como informações gramaticais, definições, exemplos lexicográficos etc. (POLGUÈRE, 2018, p. 244-248). Vejam-se, a título de exemplo, os verbetes do substantivo **objetivo**:

**objetivo**

s.m. (1720)

1 aquilo que se pretende alcançar quando se realiza uma ação; propósito

[...]

ETIMOLOGIA

lat.escl. *objectivus*, a, um 'id.'

SINÔNIMOS/VARIANTES

ver sinonímia de propósito

HOMÔNIMOS

**objetivo** (fl. **objetivar**) (HOUAISS, 2009, grifos do autor).

**objetivo**

(ob.je.ti.vo)

[...]

sm.

4. Meta a ser alcançada; OBJETO; ESCOPO; PROPÓSITO: Atingiu seu

objetivo

5. Fil. Diz-se da ideia ou de tudo que se refere aos objetos exteriores ao espírito

6. Mil. Alvo principal das operações ou manobras militares: “A expedição endireitava para o objetivo da luta como se voltasse de uma campanha” (Euclides da Cunha, *Os sertões*)

[F.: Do lat. *objectivus*] (AULETE, grifos do autor).

Em ambos os verbetes, constam informações de ordem gramatical (classe gramatical em ambos, separação silábica em Aulete), semântica (uma definição para se referir a cada acepção da palavra-entrada), sintática (exemplos lexicográficos), etimológica (origem da palavra) e sobre as relações paradigmáticas (sinonímia e antonímia) (POLGUÈRE, 2018).

Como a menção às relações sintagmáticas, ou seja, às possibilidades de combinatória lexical, costuma se restringir às locuções (*ibid.*; CARVALHO, 2017; 2001), logramos verificar a presença de ambos os constituintes das colocações pesquisadas no segundo capítulo deste trabalho nos verbetes dos itens lexicais que as compõem, e se poderiam estar em outras partes da microestrutura, além da seção de locuções – como na definição ou no exemplo lexicográfico.

**Quadro 3** – Ocorrências de colocações verbais em diferentes seções da microestrutura de dicionários monolíngues padrão de PB

<b>Houaiss (2009)</b>	
<b>Verbetes</b>	<b>Ocorrência</b>
<b>'acordo \ô\</b>	<b>Exemplo</b> “1 ajuste formal; pacto, convenção, concerto <os países beligerantes <u>assinaram</u> um a.>”
<b>atingir</b>	<b>Definição</b> “3 t.d. fig. conseguir, obter (p.ex. um <u>objetivo</u> , uma meta)”
<b>batalha</b>	<b>Exemplo</b> “2 troca de golpes; luta, duelo <os gladiadores romanos <u>travavam</u> b. sangrentas>”
<b>caso</b>	<b>Locução e exemplo</b> “ <u>vir</u> ao c.: vir a propósito, ser pertinente <seu comentário não <u>vem</u> ao c.>”
<b>consciência</b>	<b>Locução</b> “ <u>tomar</u> c.: perceber com clareza (a importância, a gravidade, o perigo etc.); compreender, conscientizar-se”
<b>consideração</b>	<b>Locução</b> “ <u>levar</u> ou <u>tomar</u> em c. levar em conta; considerar”
<b>exemplo</b>	<b>Locução</b> “ <u>dar</u> o e.: agir exemplarmente, de modo que os outros se sintam impelidos a adotar o mesmo comportamento”
<b>firmar</b>	<b>Definição</b> “4 t.d. realizar (pacto, <u>acordo</u> ); ajustar, pactuar”
<b>objetivo</b>	<b>Definição</b> “1 aquilo que se pretende <u>alcançar</u> quando se realiza uma ação; propósito”

<b>prestar</b>	<b>Exemplo</b> “9 t.d. realizar (uma ação) por imposição legal; cumprir, fazer <p. juramento>   <p. <u>serviço</u> militar>”
<b>ranger</b>	<b>Definição</b> “1 t.d. atritar (os <u>dentes</u> ) uns contra os outros, por efeito de nervosismo, dor etc.”
<b>sério</b>	<b>Locução</b> “ <u>levar</u> ou <u>tomar</u> a s. : 1 dar importância a; considerar como coisa grave; 2 melindrar-se, magoar-se”
<b>vantagem</b>	<b>Locução</b> “ <u>levar</u> v.: 1 ser ou tornar-se superior a <com esta altura, ele leva v. sobre a turma> 2 tirar proveito de algo ou alguém, ger. de maneira esperta”
<b>Aulete (s.d.)</b>	
<b>Verbetes</b>	<b>Ocorrência</b>
<b>atingir</b>	<b>Definição</b> “3. Fig. Alcançar ( <u>objetivo</u> , finalidade). [...]”
<b>batalha</b>	<b>Definição em outra locução</b> “Batalha naval 1 Mar. Batalha <u>travada</u> entre esquadras inimigas”
<b>caso</b>	<b>Locução</b> “ <u>Vir ao</u> caso 1 Vir à propósito, ter a ver (com algo).”
<b>consideração</b>	<b>Locução</b> “ <u>Levar/tomar em</u> consideração 1 Considerar, levar/ter em conta; lembrar de (coisa, fato etc.) como elemento importante para formar juízo ou decisão; incluir (coisa, fato etc.) entre os demais fatores analisados ou observados”
<b>desculpa</b>	<b>Exemplo</b> “1. Ação ou resultado de desculpar(-se) [+ de, por : <u>pedir desculpas de/por ter chegado tarde.</u> ]”
<b>levar</b>	<b>Exemplo</b> “32. Ter certa compreensão, entendimento ou comportamento em relação a (certo fato, acontecimento, atividade, ação etc.) [td. : <i>Ela leva tudo na brincadeira: <u>Vocês levam tudo a sério.</u> Ele leva a vida na maciota: <u>Levou a conversa a ferro e fogo</u>]</i> ”
<b>objetivo</b>	<b>Definição e exemplo</b> “4. Meta a ser <u>alcançada</u> ; OBJETO; ESCOPO; PROPÓSITO: <u>Atingiu seu objetivo</u> ”

<b>prestar</b>	<b>Exemplo</b> “5. Realizar (algo) por imposição legal; CUMPRIR [td. : prestar <u>serviço</u> militar.]”
<b>ranger</b>	<b>Definição e exemplo</b> “2. Atritar (os <u>dentes</u> ) por dor, medo, raiva etc. [td. : Ficou com medo do avô ao vê-lo <u>ranger</u> os <u>dentes</u> ]”
<b>sério</b>	<b>Locução</b> “ <u>Levar/tomar a sério</u> 1 Dar importância a, considerar com seriedade: <i>Ele levoul/tomou a sério as ameaças e pediu proteção.</i> 2 Considerar seriamente, a ponto de ofender-se, o que foi feito ou dito como brincadeira: <i>Não se ofenda, não precisa me levar/tomar tão a sério.</i> ”
<b>serviço</b>	<b>Definição em outra locução</b> “Serviço de utilidade pública 1 Serviço útil <u>prestado</u> pelo Estado à sociedade mediante pagamento pelos que o utilizam.” <sup>15</sup>

Fontes: Houaiss (2009); Aulete (s.d.).

A partir dos dados acima descritos, constata-se que os itens que constituem as colocações lexicais aparecem, de forma equilibrada, nas seções dedicadas à definição semântica e às locuções dos verbetes, ocorrendo menos na seção de exemplos do dicionário Houaiss (2009). Em contrapartida, os exemplos com colocações aparecem, frequentemente, de forma associada à definição no dicionário Aulete. Neste mesmo dicionário, aliás, há menos ocorrências das combinações pesquisadas na seção de locuções, em comparação ao dicionário Houaiss (2009)<sup>16</sup>.

Verifica-se, também, a ocorrência de remissões a itens lexicais de carga semântica similar, mas que não são colocações. Por exemplo, o verbo “batalhar” possui, em uma de suas acepções, a definição “travar (batalha)” (HOUAISS, 2009), mas conforme discutido por Carvalho (2017, p. 85-86), a forma colocacional pode ser escolhida em detrimento do verbo único, ou o contrário – o uso depende do tipo de interação que motiva sua utilização.

<sup>15</sup> A colocação “prestar serviço” costuma ocorrer na voz passiva, conforme discutido no capítulo 2.1, mas também ocorre na voz ativa quando associada à locução “serviço militar”, como se pode ver nos verbetes “prestar” (HOUAISS, 2009) e “serviço” (AULETE).

<sup>16</sup> Sobre as ocorrências, sinalizamos que, das combinações listadas no Quadro 1, não foi encontrado verbo para “descaradamente” no dicionário Houaiss (2009), e que não foi possível acessar o verbo “acordo” no dicionário Aulete devido a uma mensagem de erro.

Ademais, assinalamos que os itens lexicais de algumas locuções são registrados tanto na variedade brasileira, quanto na variedade europeia do português, como **levar / tomar a sério** (HOUISS, 2009; AULETE). Este lusismo foi confirmado no dicionário bilíngue *Larousse* (1988), direcionado a falantes de ambas as variedades do português, cuja discussão será desenvolvida no próximo subcapítulo desta pesquisa.

Por fim, registre-se que, nestas obras, não foram mencionadas outras possibilidades de combinação de palavras, para além das apontadas nas locuções, e que o público de PB como língua estrangeira, possível consulente deste tipo de dicionário em níveis mais avançados de aprendizagem, pode não captar, de forma transparente, as informações presentes sobre combinatória lexical. Aliás, é importante tomar consciência desse tipo de relação sintagmática, e cremos que o uso de *corpora* em contextos de ensino-aprendizagem, se associados à consulta a verbetes adaptados ao público, podem despertar essa consciência (CARVALHO, 2017; TAGNIN, 2013).

### **3.2 Dicionários bilíngues**

O dicionário bilíngue é, tradicionalmente, um dos recursos mais utilizados no processo de ampliação da rede lexical de falantes de línguas orais. Este tipo de dicionário pode fazer parte do grupo de dicionários elaborados com fins pedagógicos, como os dicionários escolares, fundamentais e infantis (VILARINHO; MAIA-PIRES, 2016, p. 91-92).

Considerando que a categoria bilíngue de dicionários tem público-alvo mais bem definido que a dos dicionários monolíngues padrão, a estrutura desse dicionário tende a ser mais delimitada que a deste último. A microestrutura de um verbete de dicionário bilíngue, nas direções de língua-fonte para língua-alvo e vice-versa, abrange, basicamente, uma palavra-entrada, a informação de sua classe gramatical e equivalência(s). De fato, essa estrutura é bastante tradicional no Brasil: não costumam ser apresentadas definições em

nenhuma das línguas, tampouco são apresentadas propriedades semânticas, sintáticas ou etimológicas das lexias<sup>17</sup> (*ibid.*, p. 95; CARVALHO, 2001, p. 49).

A seguir, apresentaremos a presença das colocações verbais na versão informatizada do *TheFreeDictionary* / K Dictionaries (2013), na versão impressa do *Dictionnaire Français-Portugais Larousse* (1988, coleção “Apollo”) e nas versões impressa e informatizada do dicionário escolar *Michaelis* português/francês (2002; *on-line*).

### 3.2.1 Informatizados

Os dicionários bilíngues informatizados beneficiam da não linearidade, própria ao mundo digital: os verbetes podem ser acessados e conectados uns aos outros a partir de um *hyperlink*, saltando de um ponto a outro da obra lexicográfica de acordo com a necessidade do consulente e enriquecendo sua experiência (SANTAELLA, 2014). Para prosseguir com nossa análise, acessamos a versão para dispositivos móveis do *TheFreeDictionary* / K Dictionaries (2013).

O dicionário disponível em versão *on-line* (gratuita<sup>18</sup>) e eletrônica (disponível para compra em lojas de aplicativos para dispositivos móveis) *TheFreeDictionary*, gerenciado pela empresa norte-americana Farlex, compila várias bases de dados a fim de disponibilizar verbetes distintos para a mesma palavra. Entre bases de dados monolíngues, etimológicas e outras, o que interessa à nossa análise é a base multilíngue Kennerman / K Dictionaries Ltd., gerida pela empresa israelense Lexicala e usada como base para dicionários bilíngues e semibilíngues (LEXICALA API BY K DICTIONARIES). Sobre sua microestrutura, vale ressaltar que há, para além da estrutura tradicional de um dicionário bilíngue, definições e paráfrases curtas que são apresentadas em ambas as línguas, a depender da direção, e que, se o consulente assim desejar, também pode salvar verbetes como favoritos ou criar *flashcards* na

---

<sup>17</sup> Um tipo de dicionário que resolve esta lacuna estrutural, mas que ainda é pouco difundido, é o dicionário semibilíngue, que apresenta, para além da entrada de determinada palavra, a “definição ou paráfrase, o exemplo de uso e a unidade linguística correspondente” na língua de aprendizagem (VILARINHO; MAIA-PIRES, 2016, p. 96). Do par português/francês, encontramos o dicionário impresso *Palavra-chave* (2011), para estudantes brasileiros, e o dicionário informatizado multilíngue Kennerman, para públicos lusófonos. Mais detalhes sobre esta obra lexicográfica serão dados no próximo subcapítulo.

<sup>18</sup> Disponível em: <https://fr.thefreedictionary.com/> (DICTIONNAIRE FRANÇAIS).

versão para dispositivos móveis, o que se mostra profícuo no contexto de criação de redes lexicais (POLGUÈRE, 2018).

Listamos, a seguir, os verbetes em francês e português encontrados nesta base de dados em cujo conteúdo ocorreram colocações. Sinalizamos que as equivalências listadas tendem a considerar a variedade europeia do português.

**Quadro 4** – Colocações verbais no dicionário *TheFreeDictionary* / K Dictionaries (2013)

<b>Direção português-francês</b>	
<b>Verbete</b>	<b>Ocorrência</b>
<b>alcançar</b>	<b>Definição</b> “1. <u>atingir</u> um <u>objectivo</u> . <i>atteindre, arriver à</i> ”
<b>consciência</b>	<b>Definição</b> “2. ter/ <u>tomar</u> consciência - dar-se conta de algo. <i>avoir / prendre conscience</i> ”
<b>objectivo</b>	<b>Definição</b> “1. aquilo que se quer <u>atingir</u> . <i>objectif</i> ”
<b>pedir</b>	<b>Exemplo</b> “1. dizer a alguém o que se quer. pedir <u>desculpa</u> a alguém [...] <i>demander pardon à qqn</i> ”
<b>promover</b>	<b>Definição</b> “2. favorecer o <u>desenvolvimento</u> de algo. <i>promouvoir</i> ”
<b>ranger</b>	<b>Definição</b> “1. roçar os <u>dentes</u> uns nos outros. <i>grincer</i> ”
<b>Direção francês-português</b>	
<b>Verbete</b>	<b>Ocorrência</b>
<b>accord</b>	<b>Exemplo</b> “6. arrangement, traité. <u>conclure</u> / signer un accord. <i>conclure / assinar um acordo</i> ”
<b>atteindre</b>	<b>Exemplo</b> “2. figuré. réussir, parvenir à [...] <i>Atteindre son objectif. Atingir o objetivo [...]</i> ”

<b>conscience</b>	<b>Definição</b>  “1. Avoir / prendre conscience de - savoir se rendre compte de. <i>ter/tomar consciência de</i> ”
<b>demander</b>	<b>Exemplo</b>  “1. dire à qqn ce que l’on veut. [...] demander <u>pardon</u> à qqn. <i>pedir desculpa a alguém</i> ”
<b>excuse</b>	<b>Exemplo</b>  “1. pardon. [...] faire / <u>présenter</u> des excuses à qqn. <i>pedir desculpa a alguém</i> ”
<b>grincer</b>	<b>Definição</b>  “2. grincer des <u>dents</u> - frotter les <u>dents</u> du haut contre celles du bas. <i>Ranger os dentes</i> ”
<b>objectif</b>	<b>Definição</b>  “1. ce que l’on veut <u>atteindre</u> , but. <i>objectivo [...]</i> ”
<b>pardon</b>	<b>Exemplo</b>  “1. action de pardonner. [...] demander <u>pardon</u> à qqn. <i>pedir desculpa a alguém</i> ”
<b>promouvoir</b>	<b>Definição</b>  “2. favoriser le <u>développement</u> de qqch. <i>promouvoir</i> ”
<b>rendre</b>	<b>Definição e exemplo</b>  “3. rendre <u>service</u> à qqn - aller voir qqn.. prestar serviço a alguém”
<b>sérieux</b>	<b>Definição e exemplo</b>  “3. <u>Prendre</u> qqch au sérieux - considérer comme important.. Un problème à <u>prendre</u> au sérieux. <i>Levar um problema a sério</i> ”
<b>unir</b>	<b>Exemplo</b>  “1. mettre ensemble. [...] Il faudrait unir nos <u>efforts</u> . <i>É preciso unir os nossos esforços</i> ”

Fonte: K Dictionaries (s.d.).

A partir dos dados acima, percebe-se como a direção da obra para consulentes de francês como língua estrangeira possui mais dados, tanto pela maior quantidade de verbetes encontrados, quanto pelo maior número de informações neles contidas. Vejam-se, por exemplo, as acepções de “consciência”, na direção português-francês, em relação à “*conscience*”, na direção oposta: a primeira traz a forma colocacional e definição curta, enquanto a segunda traz, além destes dados, uma informação de combinatória gramatical – indica a preposição *de*.

No que concerne às equivalências em diferentes variedades, considerando a mesma direção de francês como língua-alvo, foram constatadas outras equivalências em português europeu, apontadas como menos frequentes no *corpus Portuguese Web 2011*: “atrair atenção” para “*attirer l’attention*”, “tratar-se” para “*être question*”, entre outras. Verificou-se, de igual modo, que o verbo “*profiter*”, apesar de não apresentar estrutura colocacional, como discutido no capítulo anterior, apresenta definição semântica que converge com o seu equivalente binário em português: “obter uma vantagem sobre algo” (K DICTIONARIES, tradução nossa).

### 3.2.2 Impressos

O dicionário bilíngue francês/português *Larousse*, cuja primeira edição data dos anos 1950, é uma obra de referência elaborada para “alunos e viajantes” que utilizam francês e português europeu. Tal informação pode ser inferida desde o prefácio, em que é indicado o espaço dado pela obra a usos mais distantes da norma culta portuguesa, como “os neologismos, a língua familiar e os brasileirismos” (LAROUSSE, 1988).

Sua microestrutura compreende a palavra-entrada, informação de classe gramatical, indicação de acento tônico e pronúncia figurada, seguidas pelas equivalências, que podem contar com notas adicionais em ambas as direções. Nesta edição, há, ainda, uma concepção diferenciada de verbo: agrupam-se itens lexicais, não de forma individual, mas de acordo com os radicais, sejam eles reais ou simulados para fins pedagógicos, como se pode ver em um único verbo “**ren** || **-dre** vt. [...] || **-du, ue** adj. [...]” (*ibid.*, p. 290).

Salvo indicação em contrário, as colocações lexicais encontradas nesta obra são indicadas na seção de locuções, ao final do verbete, como se observa em “sérieux”: “**sérieux, euse** adj. (cêriã, ãz). Sério, ia. || s.m. Seriedade (f.) [...] || Loc. *prendre au sérieux*, tomar a sério” (*ibid.*, p. 310). O quadro a seguir compila as demais colocações – antes, sinalizamos que não foram encontrados verbetes de três palavras, na direção português-francês, e de uma palavra, na direção oposta:

**Quadro 5** – Equivalências de colocações verbais no dicionário *Larousse* (1988)

<b>Direção português-francês</b>	
<b>Verbete</b>	<b>Colocação</b>
batalha	dar batalha [lusismo]
dente	ranger dentes
prestar [indicação de colocação na seção de equivalências]	prestar serviço
serviço	
sério	tomar a sério [lusismo]
<b>Direção francês-português</b>	
<b>Verbete</b>	<b>Colocação</b>
livrer	livrer bataille
pardon	demander pardon
rendre	rendre service
sérieux	prendre au sérieux
tenir	tenir compte

Fonte: Larousse (1988).

Quanto ao dicionário escolar *Michaelis*, assinala-se que o público-alvo é de estudantes brasileiros que aprendem francês, e que seus cerca de 30.000 verbetes apresentam “divisão silábica, transcrição fonética, classe gramatical, área de conhecimento, várias acepções e expressões atuais” (MICHAELIS, 2002).

Todas as equivalências de colocações lexicais identificadas neste dicionário, tanto na direção português-francês, quanto em francês-português, foram encontradas na seção de expressões, logo após as equivalências. À guisa de exemplo, veja-se o verbete “sérieux”:  
**“sérieux, -ieuse** [serjø, -jøz] *adj+nm* **1** sério, razoável, refletido [...] *nm* seriedade. [...] **prendre quelqu’un au sérieux** levar alguém a sério” (*ibid.*, p. 403). Listamos as demais colocações encontradas, e ressaltamos que, pela delimitação do público-alvo e limitação de tamanho da obra, não foram encontrados verbetes para sete palavras constantes do *corpus*, em ambas as direções.

**Quadro 6** – Equivalências de colocações verbais no dicionário *Michaelis* (2002)

<b>Direção português-francês</b>	
<b>Verbete</b>	<b>Colocação</b>
chamar	chamar atenção
desculpa	pedir desculpas
levar	levar a sério
ranger	ranger os dentes
dentes	
tomar	tomar consciência
consciência	
<b>Direção francês-português</b>	
<b>Verbete</b>	<b>Colocação</b>
bataille	livrer bataille
bêtise	dire des bêtises
compte	tenir compte
question	(ne pas) être question de
rendre	rendre service
service	
sérieux	prendre [qqn] au sérieux
suite	donner suite (à)

Fonte: Michaelis (2002).

Algumas equivalências foram mais literais do que as encontradas no *corpus Portuguese Web 2011*, como “*prendre en considération*”, no verbete “consideração”, e “*tirer avantage*”, no verbete “tirar” (MICHAELIS, 2002).

Cabe destacar, também, que os verbetes contidos no Michaelis *on-line* são idênticos aos da versão impressa, mesmo após a revisão de 2016, apontada no *site* (MICHAELIS ON-LINE). Como ferramenta adicional, nota-se uma lista de parônimos logo após o verbete, para facilitar o acesso do consultante a itens lexicais similares.

## CONSIDERAÇÕES FINAIS

Ao longo deste trabalho, definiu-se a relação sintagmática das colocações como a coocorrência restrita de dois (ou mais) itens lexicais e verificou-se a sua grande frequência nas interações em português e francês. Foram analisadas ocorrências deste fenômeno, em sua subdivisão composta por um verbo e um complemento – substantivo, adjetivo ou advérbio –, em *corpora* de ambas as línguas, assim como em dicionários direcionados a seus diversos públicos, considerando as variedades de maior destaque.

Nos *corpora*, foram constatadas as principais características das colocações verbais em português, a partir da lista de 100 verbos e substantivos mais frequentes desta língua (cf. Anexo I), acessadas por três ferramentas de análise. Foram listadas, de igual modo, suas equivalências em francês, que foram mais observadas no capítulo dedicado às obras lexicográficas.

A respeito destas obras, verificou-se a necessidade de adaptação das informações apresentadas para o público de PB como língua estrangeira, sobretudo nas monolíngues. Tal adequação é feita, ligeiramente, pelos dicionários bilíngues que indicam contextos de uso, mas pode melhorar mediante o uso das ferramentas da lexicografia pedagógica e de *corpora* em contextos de ensino-aprendizagem, a fim de se representar e treinar, na medida do possível, a intuição do falante nativo (BINON; VERLINDE, 2000b).

Ainda sobre *corpora*, acredita-se que os resultados apontados nestas compilações apontam o quanto as combinações lexicais são reveladoras de nossa organização sociocultural. Devido à interface entre o nível da frase e o nível textual que há nestas combinações, conforme discutido por Albuquerque e Lôpo Ramos (2019), espera-se que, em pesquisas futuras, seja factível o levantamento e a análise das colocações mais frequentes por gênero textual, a partir das ferramentas das quais já se dispõe, para viabilizar uma reflexão ainda mais apurada sobre a língua em uso.

Assim, além das possibilidades de consulta recorrente a ferramentas de *corpora*, entende-se, também, que as obras lexicográficas possam ser utilizadas para a criação e o aprimoramento de redes lexicais, como algumas versões informatizadas de dicionários já o fazem, aliás (POLGUÈRE, 2018; SANTAELLA, 2014). Todas essas ferramentas podem

contribuir, nos contextos escolares e individuais de ensino-aprendizagem, para usos mais espontâneos e conscientes das implicações culturais dessas relações sintagmáticas convencionadas pelo uso (BARNBROOK et al., 2013; BINON; VERLINDE, 2000a).

## REFERÊNCIAS

- ALBUQUERQUE, R.; RAMOS, A. A. L. Coesão e coerência: da análise frástica à textual. In: DIAS, J. de F. (Org.). *Ler e (re) escrever textos na universidade: da prática teórica e do processo de aprendizagem traço ensino*. 2. ed. Campinas: Pontes, 2019. p. 107-146.
- BARNBROOK, G. et al. *Collocation: applications and implications*. New York: Palgrave Macmillan, 2013.
- BASSO, R. M; GONÇALVES, R. T. *História concisa da língua portuguesa*. Petrópolis, RJ: Vozes, 2014.
- BERBER SARDINHA, T. *Linguística de corpus*. Barueri, São Paulo: Manole, 2004.
- BINON, J.; VERLINDE, S. Como otimizar o ensino e a aprendizagem de vocabulário de uma língua estrangeira ou segunda? In: LEFFA, V. J. *As palavras e sua companhia: o léxico na aprendizagem das línguas*. Pelotas: EDUCAT, 2000a. p. 123-170.
- BINON, J.; VERLINDE, S. A contribuição da lexicografia pedagógica à aprendizagem e ao ensino de uma língua estrangeira ou segunda. In: LEFFA, V. J. *As palavras e sua companhia: o léxico na aprendizagem das línguas*. Pelotas: EDUCAT, 2000b. p. 98-122.
- CARVALHO, O. L. de S. Colocações e português brasileiro como língua estrangeira. In: SILVA, F. C. O.; VILARINHO, M. M. O. (Org.). *Diálogos em português brasileiro como língua adicional*. Brasília: UAB, 2017. p. 79-105. (Série "O que a distância revela"; v. 4).
- CARVALHO, O. L. de S. Dicionários escolares: definição, oracional e texto lexicográfico. In: CARVALHO, O. L. de S; RANGEL; BAGNO, M. (Org.). *Dicionários escolares: políticas, formas e usos*. São Paulo: Parábola, 2011.
- CARVALHO, O. L. de S. *Lexicografia bilíngüe português/alemão: teoria e aplicação à categoria das preposições*. Brasília: Thesaurus, 2001.
- COSERIU, E. *Principios de semántica estructural*. Madrid: Gredos, 1977.
- CRYSTAL, D. *Dicionário de linguística e fonética*. Tradução e adaptação de Maria Carmelita Pádua Dias. Rio de Janeiro: Jorge Zahar, 2008.
- DAVIES, M. *Corpus do Português: 2.5 billion words, Dialects / Genres / Historical*. Provo (Utah), s.d. Disponível em: <http://www.corpusdoportugues.org/web-dial/>. Acesso em: 10 out. 2020.
- DERADRA, S. Vers une typologie des collocations. *Synergies Algérie: revue du groupe Gerflint, Sylvains-lès-Moulins*, v. 3, n. 8, p. 237-252, 2009a.
- DERADRA, S. *Vers une typologie des collocations : l'exemple de « Il était une fois un vieux couple heureux » de Mohammed Khaïr-Eddine*. 2009b. 160 p. Dissertação (Mestrado em

Linguística) – Faculdade de Letras e Ciências Sociais, Universidade Mohammed Boudiaf, M’sila, 2009b. Disponível em: [http://these.univ-msila.dz/pmb/opac\\_css/doc\\_num.php?explnum\\_id=549](http://these.univ-msila.dz/pmb/opac_css/doc_num.php?explnum_id=549). Acesso em: 10 set. 2020.

DICTIONNAIRE FRANÇAIS. *TheFreeDictionary*. Huntingdon Valley, s.d. Disponível em: <https://fr.thefreedictionary.com/>. Acesso em: 10 out. 2020.

GIL, A. C. *Métodos e técnicas de pesquisa social*. São Paulo: Atlas, 2008.

HOUAISS, A. *Dicionário eletrônico Houaiss da língua portuguesa 3.0*. Rio de Janeiro: Objetiva, 2009.

K DICTIONARIES: K Dictionaries Ltd. and partners. In: THEFREEDICTIONARY.COM: FARLEX. *Dicionário +: dictionnaire français-portugais [aplicativo]*. Huntingdon Valley, s.d. Disponível em: <https://play.google.com/store/apps/details?id=thefreedictionary.dictionary>. Acesso em: 10 out. 2020.

KILGARRIFF, A.; RYCHLÝ, P. *Sketch Engine*. East Sussex: Lexical Computing Limited, s.d. Disponível em: <https://app.sketchengine.eu>. Acesso em: 10 set. 2020.

KUHN, T. Z.; FERREIRA, J. P. O Corpus de português escrito em periódicos - CoPEP. *DELTA* [on-line], São Paulo, v. 36, n. 2, 2020. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502020000200408&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502020000200408&lng=en&nrm=iso). Acesso em: 10 set. 2020.

LAROUSSE: DICTIONNAIRE FRANÇAIS-PORTUGAIS. Tradução de Fernando V. Peixoto da Fonseca. Paris: Larousse, 1988. (Coleção “Apollo”).

LEXICALA API BY K DICTIONARIES. *Lexicala API: dictionary API and lexical data*. Tel Aviv, s.d. Disponível em: <https://api.lexicala.com/>. Acesso em: 10 out. 2020.

LINGUATECA. *Acesso a corpos de português: projeto AC/DC*. S.l., s.d. Disponível em: <https://www.linguateca.pt/ACDC/>. Acesso em: 10 out. 2020.

MICHAELIS: DICIONÁRIO ESCOLAR FRANCÊS : FRANCÊS-PORTUGUÊS, PORTUGUÊS-FRANCÊS. Tradução de Jelssa Ciardi Avolio e Mára Lucia Faury. São Paulo: Melhoramentos, 2002.

MICHAELIS ON-LINE: DICIONÁRIO ESCOLAR FRANCÊS. Apresentação. Disponível em: <https://michaelis.uol.com.br/escolar-frances/>. Acesso em: 10 out. 2020.

MIKUS, F. Le syntagme est-il binaire ? *WORD*, Londres, ano 3, n. 1-2, p. 32-38, [1947]. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00437956.1947.11659301>. Acesso em: 20 set. 2020.

NEGRÃO, E. V. A natureza da linguagem humana. In: FIORIN, J. L. et al. (Org.). *Linguística?: Que é isso?*. São Paulo: Contexto, 2013. p. 75-109.

*PALAVRA-CHAVE: DICIONÁRIO SEMIBILINGUE PARA BRASILEIROS: FRANCÊS.* Tradução de Andréa Stahel M. Da Silva. São Paulo: Martins Fontes, 2011.

PERINI, M. A. *Gramática descritiva do português*. 4. ed. São Paulo: Ática, 2005.

POLGUÈRE, A. Lexicologia descritiva. In: POLGUÈRE, A. *Lexicologia e semântica lexical: noções fundamentais*. Tradução de Sabrina Pereira Abreu. São Paulo: Contexto, 2018. p. 237-258.

ROBINS, R. H. *General linguistics: an introductory survey*. 4. ed. Abingdon, Nova York: Routledge, 2013. (Coleção “Longman linguistics library”).

SANTAELLA, L. Gêneros discursivos híbridos na era da hipermídia. *Bakhtiniana*, São Paulo, ano 9, n. 2, p. 206-216, ago./dez. 2014.

SKETCH ENGINE. *TenTen Corpus Family*. Brno, s.d. Disponível em: <https://www.sketchengine.eu/documentation/tenten-corpora/>. Acesso em: 10 set. 2020.

SKETCH ENGINE. *Statistics used in Sketch Engine*. Brno, 8 jul. 2015. Disponível em: <https://www.sketchengine.eu/documentation/statistics-used-in-sketch-engine/#logdice>. Acesso em: 10 out. 2020.

TAGNIN, S. E. O. *O jeito que a gente diz: combinações consagradas em inglês e português*. Barueri: DISAL, 2013.

VILARINHO, M. M. de O. Como usar o dicionário para compreensão e produção textual? In: DIAS, J. de F. (Org.). *Ler e (re) escrever textos na universidade: da prática teórica e do processo de aprendizagem traço ensino*. 2. ed. Campinas: Pontes, 2019. p. 151-171.

VILARINHO, M. M. de O. *Proposta de dicionário informatizado analógico de língua portuguesa*. 2013. 307 f. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2013.

VILARINHO, M. M. de O.; MAIA-PIRES, F. de O. O dicionário e outras obras: tipologias, formatos e política de língua. In: SILVA, F. C. O. da; VILARINHO, M. M. de O. *Apontamentos analíticos*. Logpress: Brasília, 2016. v. 2. p. 82-115. (Série “O que a distância revela”; v. 2).

WILLIAMS, G. Sur les caractéristiques de la collocation. In: CONFÉRENCE SUR LE TRAITEMENT AUTOMATIQUE DES LANGUES NATURELLES, 8., 2001, Tours. *Actes du TALN...* Tours: [s.n.], 2001. Disponível em: <http://talnarchives.atala.org/TALN/TALN-2001/taln-2001-tutoriel-002.pdf>. Acesso em: 10 set. 2020.

**Anexo I – Wordlist: verbos e substantivos mais comuns da língua portuguesa, por Kilgarriff e Rychlý (s.d.)**

#	Noun	Average Reduced Frequency	Verb	Average Reduced Frequency
1	ano	5077821	ser	47466148
2	dia	3861563	ter	14904281
3	vez	2646323	ir	14762624
4	peessoa	2432985	estar	10304610
5	trabalho	2164046	fazer	7529889
6	tempo	2141510	poder	7316396
7	forma	2062651	haver	3889620
8	parte	1962902	dever	3666326
9	estado	1690472	dizer	3612115
10	vida	1684176	dar	3207051
11	país	1571757	ver	2779269
12	projeto	1499026	ficar	2735848
13	mundo	1450943	passar	2679466
14	brasil	1449021	saber	2201658
15	área	1443972	querer	2042009
16	cidade	1429062	vir	1935743
17	caso	1427510	realizar	1750874
18	empresa	1415905	deixar	1533477
19	casa	1308071	apresentar	1515501
20	paulo	1251659	chegar	1483636
21	presidente	1221328	receber	1368713
22	serviço	1205368	levar	1368312
23	informação	1194419	falar	1310673
24	processo	1174977	encontrar	1272857
25	problema	1166222	conhecer	1259030
26	governo	1148255	começar	1243079
27	grupo	1147697	conseguir	1202269
28	hora	1136267	contar	1170054
29	meio	1134777	partir	1168371
30	coisa	1130829	acontecer	1165366

31	ação	1121126	existir	1153350
32	relação	1120416	usar	1115153
33	mês	1085461	criar	1056822
34	ponto	1085081	mostrar	1034276
35	programa	1080861	colocar	1023224
36	sistema	1073652	considerar	993706
37	momento	1068388	tornar	984946
38	exemplo	1032249	continuar	981426
39	atividade	1010498	trabalhar	924985
40	valor	994567	afirmar	916359
41	acordo	975743	seguir	912512
42	tipo	957890	chamar	905468
43	rio	956114	utilizar	900061
44	nome	954239	manter	894733
45	desenvolvimento	940520	precisar	884010
46	direito	935988	permitir	877837
47	região	934823	participar	866289
48	resultado	931406	achar	862382
49	lugar	930831	pensar	846297
50	fim	915436	tratar	837936
51	número	907560	entrar	827707
52	história	902419	desenvolver	816894
53	lado	900085	parecer	803644
54	semana	878655	voltar	792159
55	objetivo	862682	esperar	789427
56	centro	862469	acabar	773359
57	questão	855468	ocorrer	765875
58	saúde	824327	perder	763673
59	situação	817476	gostar	758627
60	mercado	812567	sair	746592
61	qualidade	810731	oferecer	746504
62	lei	799542	tomar	740153
63	final	798193	viver	737229
64	escola	792956	tentar	702119
65	evento	789601	trazer	677396

66	espaço	774521	lembrar	676068
67	filho	769041	ganhar	672595
68	produto	764809	equipar	666814
69	obra	756525	garantir	665911
70	professor	747457	explicar	647394
71	homem	745582	procurar	637736
72	local	733829	ajudar	634798
73	condição	728041	pedir	626913
74	conta	719662	ler	625064
75	recurso	719428	pelar	623841
76	janeiro	706955	possuir	623829
77	educação	701450	incluir	622766
78	período	699688	escrever	616808
79	curso	698983	acreditar	612598
80	sociedade	698480	entender	595366
81	família	694009	promover	591022
82	josé	680904	acompanhar	586814
83	início	672275	representar	586813
84	necessidade	668028	sentir	575371
85	fato	666275	publicar	570603
86	ideia	662930	marcar	561398
87	produção	661795	prever	557865
88	criança	659564	iniciar	552021
89	município	658842	envolver	544440
90	estudo	645088	destacar	540365
91	nível	640967	abrir	536637
92	rede	635400	unir	525716
93	base	634976	servir	522680
94	apoio	634381	aumentar	519578
95	tema	633777	produzir	514868
96	jogo	632447	obter	508471
97	uso	628151	formar	498559
98	site	627774	mudar	491246
99	amigo	622851	atender	491162
100	mulher	619575	determinar	490763