



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Câmpus de São José do Rio Preto

ARIANE DONIZETE DELGADO RIBEIRO CALDAS

**ESTUDO COMPARATIVO DE COLOCAÇÕES ESPECIALIZADAS NO  
ÂMBITO DA INVESTIGAÇÃO CRIMINAL COM VISTAS À  
ELABORAÇÃO DE UM DICIONÁRIO PEDAGÓGICO TRILÍNGUE**

São José do Rio Preto  
2023

**ESTUDO COMPARATIVO DE COLOCAÇÕES ESPECIALIZADAS NO  
ÂMBITO DA INVESTIGAÇÃO CRIMINAL COM VISTAS À ELABORAÇÃO  
DE UM DICIONÁRIO PEDAGÓGICO TRILÍNGUE**

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Orientador: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Adriane Orenha-Ottaiano

C145e	<p>Caldas, Ariane Donizete Delgado Ribeiro</p> <p>Estudo comparativo de colocações especializadas no âmbito da investigação criminal com vistas à elaboração de um dicionário pedagógico trilingue / Ariane Donizete Delgado Ribeiro Caldas. -- São José do Rio Preto, 2023</p> <p>148 f. : il., tabs., fotos</p> <p>Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto</p> <p>Orientadora: Adriane Orenha-Ottaiano</p> <p>1. Linguística de Corpus. 2. Fraseologia Baseada em Corpus. 3. Colocações Especializadas. 4. Dicionários de Colocações. I. Título.</p>
-------	--

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

ARIANE DONIZETE DELGADO RIBEIRO CALDAS

**ESTUDO COMPARATIVO DE COLOCAÇÕES ESPECIALIZADAS NO  
ÂMBITO DA INVESTIGAÇÃO CRIMINAL COM VISTAS À ELABORAÇÃO  
DE UM DICIONÁRIO PEDAGÓGICO TRILÍNGUE**

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Estudos Linguísticos, junto ao Programa de Pós-Graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto.

Financiadora: CAPES

Comissão Examinadora

Prof<sup>a</sup>. Dr<sup>a</sup>. Adriane Orenha-Ottaiano  
Universidade Estadual Paulista  
(UNESP) Orientadora

Prof<sup>a</sup>. Dr<sup>a</sup> Maria João Broa Martins Marçalo  
Universidade de Évora

Prof<sup>a</sup>. Dr<sup>a</sup>. Lídia Amélia De Barros Cardoso  
Universidade Federal do Ceará (UFC)

Prof. Dr. Odair Luiz Nadin Da Silva  
Universidade Estadual Paulista  
(UNESP)

Prof<sup>a</sup>. Dr<sup>a</sup>. Renato Rodrigues Pereira  
Universidade Federal do Mato Grosso  
do Sul (UFMS)

São José do Rio Preto  
02 de fevereiro de 2023

*À minha família, fonte de amor e inspiração.*

## AGRADECIMENTOS

O percurso até aqui foi repleto de linhas sinuosas, altos e baixos, tendo sido percorrido ora só, ora acompanhada, tal como acontece com as palavras, que ora se apresentam sozinhas, ora se juntam a outras para construir novo significado. E esses novos significados permanecerão inscritos na memória ao longo de toda minha vida. A estas pessoas, que se juntaram a mim para compor novo significado, deixo aqui minha gratidão.

Aos meus *pais, Aparecida e Leonildo*, obrigada por serem meus maiores incentivadores em busca dos meus sonhos, por todo apoio durante esta jornada e por não me deixarem dar voz ao desânimo todas as vezes em que ele chamava. À minha *irmã, Ariele*, por sempre acreditar que eu conseguiria realizar meus sonhos. À minha *afilhada e irmã mais nova, Thuany*, por todos os momentos de diversão que tiveram que esperar para que eu pudesse alcançar esse objetivo e por todo seu amor incondicional. Ao meu *namorado, Stéfan*, por insistir em ver apenas o melhor em mim, acreditando sempre em minha capacidade.

À minha *orientadora, Prof<sup>ª</sup>. Dr<sup>ª</sup>. Adriane Orenha-Ottaiano* por ter aberto as portas que permitiram realizar o sonho acadêmico, pela parceria ao longo desta caminhada, pela confiança em meu trabalho e pelas orientações e ensinamentos durante todo esse percurso.

À *Prof<sup>ª</sup> Dr<sup>ª</sup> Maria José Bocorny Finatto*, por ter aceito, mais uma vez, o convite para debater este trabalho no XIII Seminário de Estudos Linguísticos da Unesp, pelas contribuições enriquecedoras e pelos elogios.

À *Prof<sup>ª</sup>. Dr<sup>ª</sup>. Suzi Marques Spatti Cavalari* pelo aceite ao convite e pela orientação do exame de qualificação especial fora de área, pelos ensinamentos que contribuíram sobremaneira para a minha formação enquanto professora e pesquisadora.

À *Prof<sup>ª</sup>. Dr<sup>ª</sup>. Regiani Aparecida Santos Zacarias* e à *Prof<sup>ª</sup>. Dr<sup>ª</sup>. Mariângela de Araújo* por terem composto a banca do exame de qualificação e por contribuírem para o enriquecimento deste trabalho. Aos professores que tão generosamente aceitaram o convite para compor a banca de defesa deste trabalho, *Prof<sup>ª</sup>. Dr<sup>ª</sup> Maria João Broas Martins Marçalo, Prof<sup>ª</sup> Dr<sup>ª</sup> Lídia Amélia de Barros Cardoso, Prof. Dr. Odair Luiz Nadin da Silva* e *Prof. Dr. Renato Rodrigues Pereira*. Obrigada por contribuírem com seu olhar atento à lapidação deste trabalho. Aos professores suplentes, *Prof<sup>ª</sup> Dr<sup>ª</sup> Diva Cardoso de Camargo, Prof. Dr. Ariel Novodvorski* e *Prof<sup>ª</sup> Dr<sup>ª</sup> Rosemeire Selma Monteiro Plantin*, igualmente agradeço pela generosidade e disponibilidade.

Aos professores das disciplinas cursadas durante o Doutorado, os quais foram

responsáveis pela ampliação da minha compreensão sobre Lexicografia, Fraseologia, corpus e legendagem.

Meu muito obrigada aos amigos que fiz durante os anos dedicados à academia e com os quais compartilhei angústias, dúvidas, certezas e vitórias, em especial a *Jean Pimentel*, *Mariele Seco*, *Janaina Camoleze*, *Beatriz Curti-Contessoto* e *Nathalia Soares*. Mais uma vez, vocês fizeram toda a diferença durante esta etapa da minha vida e os levarei sempre comigo.

Por fim, mas não menos importante, agradeço a *Deus* por ser meu refúgio e sustento ao longo de minha vida, em especial, por ter sido a fonte de toda força e perseverança durante a realização desta pesquisa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, à qual agradeço.

## RESUMO

A presente pesquisa, financiada pela CAPES, teve como objetivo geral levantar e analisar colocações especializadas do âmbito investigativo criminal, por meio de corpora compostos pelas legendas das séries *CSI – Crime Scene Investigation*, *NCIS* e *Criminal Minds*, nos idiomas inglês, português e espanhol. Ademais, contou com um corpus comparável composto por artigos científicos, o qual teve como finalidade comprovar o uso das colocações especializadas encontradas no corpus de estudo por profissionais da área investigativa criminal. Como objetivo específico, temos o propósito de elaborar uma amostra de dicionário pedagógico trilingue de colocações especializadas do âmbito investigativo criminal. Dessa forma, a pesquisa contou com o apoio da Fraseologia baseada em corpus, da Linguística de Corpus, da Terminologia, da Fraseografia e da Teoria Funcional da Lexicografia para definir suas bases teóricas e metodológicas. No que tange às colocações especializadas levantadas por meio do corpus de estudo em inglês, estas foram comparadas com os dados obtidos pelo corpus comparável, tendo sido analisadas a partir de um viés quantitativo, a partir do qual buscou-se ressaltar seu *logdice*. Ademais, realizou-se uma análise morfossintática, dando ênfase à classificação taxonômica das colocações obtidas, bem como, uma análise léxico-semântica, ressaltando as semelhanças e diferenças percebidas no comportamento das colocações nos três idiomas sob análise. Entre os resultados, destacam-se um número considerável de colocações do tipo nominal no corpus em língua portuguesa (90 no total), demonstrando, também, que esse corpus apresentou uma maior quantidade total de colocações especializadas (179, em comparação a 130 colocações no corpus inglês e 109 no corpus espanhol). Por sua vez, a análise léxico-semântica empreendida mostrou casos em que diferentes colocações especializadas criminais em língua inglesa apresentam uma mesma tradução em língua espanhola, como acontece com *artight alibi* e *ironclad alibi*, em que ambas foram traduzidas por *coartada irrefutable*, além de enfatizar a análise do termo *unsub*. Considerando as análises e discussões empreendidas, espera-se que esta pesquisa possa contribuir com informações relevantes sobre as colocações especializadas no âmbito criminal, nos três idiomas contemplados, além de lançar luz sobre as informações que um dicionário de colocações pode apresentar aos seus consulentes.

**Palavras-chave:** Linguística de Corpus. Fraseologia Baseada em Corpus. Colocações Especializadas. Dicionário de Colocações.

## **ABSTRACT**

*This research, funded by CAPES, had as general objective the raising and analyzing specialized collocations in the criminal investigative field, through corpora composed of the subtitles of the series CSI – Crime Scene Investigation, NCIS and Criminal Minds, in English, Portuguese and Spanish. In addition, it had a comparable corpus composed of scientific articles, which aimed to prove the use of specialized collocations found in the study corpus by professionals in the criminal investigation area. Considering the collected data, it was proposed, as a specific objective, the elaboration of a sample of a trilingual pedagogical dictionary of specialized collocations of the criminal investigative field. Thus, the research was supported by Corpus-Based Phraseology, Corpus Linguistics, Terminology, Phraseography and the Functional Theory of Lexicography to define its theoretical and methodological bases. Regarding the specialized collocations raised through the study corpus in English, these were compared with the data obtained by the comparable corpus, having been analyzed from a quantitative bias, in which we sought to highlight its logdice. Furthermore, a morphosyntactic analysis was carried out, emphasizing the taxonomic classification of the collocations obtained, as well as a lexical-semantic analysis, emphasizing the perceived similarities and differences in the behavior of the collocations in the three languages under analysis. Among the results, the number of collocations of the nominal type in the corpus in Portuguese (90 in total) is highlighted, making this corpus present a greater total quantity of collocations specialized (179, against 130 collocations in the English corpus and 109 in the Spanish corpus). In turn, the lexical-semantic analysis carried out, showed cases in which different specialized criminal collocations in English present the same translation in Spanish, as happens with airtight alibi and ironclad alibi, in which both were translated by coartada irrefutable, besides to emphasize the analysis of the term unsub. Considering the analyzes and discussions presented here, it is hoped that this research can contribute with relevant information about specialized collocations in the criminal field, in the three languages covered, in addition to shedding light on the information that a collocation dictionary should present to its consultants.*

**Keywords:** *Corpus Linguistics. Corpus-Based Phraseology. Specialized Collocations. Collocations Dictionary.*

## LISTA DE ILUSTRAÇÕES

Figura 1 – Elenco original da série <i>CSI – Crime Scene Investigation</i> .	20
Figura 2 – Elenco da série <i>Criminal Minds</i> .	22
Figura 3 - Elenco da série <i>NCIS – Naval Criminal Investigative Service</i> .	23
Figura 4 - Fraseologia ampla e estreita.	40
Figura 5 - Diagrama representando as colocações especializadas.	49
Figura 6 - Ferramenta <i>Keywords – CCI</i> .	63
Figura 7 - Resultados obtidos com a ferramenta <i>Word Sketch – Unsub</i> .	65
Figura 8 – Informações sobre o Corpus Comparável Criminal II	68
Figura 9 – Resultados obtidos com a ferramenta <i>Word Sketch - Unsub</i>	80
Figura 10 - Print da tela do site <i>Lexico</i> com definição do termo <i>unsub</i> .	92
Figura 11 - Print da tela do site <i>Lexico</i> com definição do termo <i>unsub</i> .	92
Figura 12 - Print da tela do site <i>Merriam-Webster</i> com a definição do termo <i>unsub</i> .	94
Figura 13 - Print da definição de <i>profile</i> do <i>Cambridge Dictionary</i> na direção inglês-português.	99
Figura 14- Print da definição de <i>ironclad</i> do <i>Cambridge Dictionary</i> na direção inglês-português.	104
Figura 15 – Porcentagem de aprendizes de tradução que usam dicionários para buscar colocações	110
Figura 16 – Porcentagem de alunos do curso de Licenciatura em Letras que usam dicionários para buscar colocações	111
Quadro 1 – Análise morfológica das colocações especializadas.	67
Quadro 2 – Informações referentes às entradas	71
Quadro 3 - Informações apresentadas em cada verbete.	72
Quadro 4 - Amostra de microestrutura do dicionário pedagógico de colocações especializadas do âmbito criminal.	73
Quadro 5 - Quantidade de episódios das séries em cada idioma	75
Quadro 6 - Palavras-chave selecionadas	79
Quadro 7 - Colocações Especializadas Adjetivas – <i>Corpus</i> Inglês	84
Quadro 8 - Colocações Especializadas Adjetivas – <i>Corpus</i> Português	84
Quadro 9 - Colocações Especializadas Adjetivas – <i>Corpus</i> Espanhol	85
Quadro 10 - Colocações Especializadas Nominais – <i>Corpus</i> Inglês	86
Quadro 11 - Colocações Especializadas Nominais no CCP - <i>Suspeito</i> como colocado	87

Quadro 12 - Colocações Especializadas Nominais no CCP - <i>Suspeito</i> como base	88
Quadro 13 - Colocações Especializadas Nominais – <i>Corpus</i> Espanhol	88
Quadro 14 - Colocações Especializadas Verbais – <i>Corpus</i> Inglês	89
Quadro 15 – Colocações Especializadas Verbais – <i>Corpus</i> Português	90
Quadro 16 – Colocações Especializadas Verbais – <i>Corpus</i> Português	90
Quadro 17 – Índice em Língua Inglesa	115
Quadro 18 – Informações referentes à palavra de entrada	116
Quadro 19 – Informações apresentadas em cada verbete	117
Quadro 20 – Amostra de microestrutura do dicionário pedagógico de colocações especializadas	118

## LISTA DE TABELAS

Tabela 1 – Índice de chavicidade das Palavras-chave selecionadas no CCI	64
Tabela 2 – Informações obtidas por meio da ferramenta <i>Word Sketch</i>	66
Tabela 3 – Tamanho dos <i>corpora</i>	74
Tabela 4 – Comparativo das listas de palavras-chave	76
Tabela 5 – Palavras-chave selecionadas no CCI.	78
Tabela 6 - Quantidade de colocações obtidas por palavra-chave em cada <i>corpus</i> .	81
Tabela 7 – Tipos de colocações levantadas.	81
Tabela 8 - Exemplo de comparação de dados entre o CCI e o CCCII.	83

## LISTA DE ABREVIATURAS E SIGLAS

<b>adj</b>	Adjetivo / <i>Adjective</i> / <i>Adjetivo</i>
<b>art</b>	Artigo / <i>Article</i> / <i>Artículo</i>
<b>CCE</b>	Corpus Criminal Espanhol
<b>CCI</b>	Corpus Criminal Inglês
<b>CCP</b>	Corpus Criminal Português
<b>det</b>	determinante / <i>determiner</i> / <i>determinante</i>
<b>DP</b>	Dicionário Pedagógico
<b>DPCEC</b>	Dicionário Pedagógico de Colocações Especializadas do Âmbito Criminal
<b>FBC</b>	Fraseologia Baseada em Corpus
<b>l. adj</b>	Locução adjetiva
<b>LC</b>	Linguística de Corpus
<b>LC</b>	língua de chegada
<b>LP</b>	língua de partida
<b>n.</b>	<i>noun</i>
<b>ph. v.</b>	<i>Phrasal verb</i>
<b>prep.</b>	Preposição / <i>Preposition</i> / <i>Preposición</i>
<b>s.</b>	Substantivo / Sustantivo
<b>TFL</b>	Teoria Funcional da Lexicografia
<b>UFs</b>	Unidades Fraseológicas
<b>v.</b>	Verbo / <i>Verb</i> / <i>Verbo</i>
<b>v. f.</b>	Verbo frasal / <i>Verbo Frasal</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	15
<b>2</b>	<b>SÉRIES E LÍNGUA DE ESPECIALIDADE</b>	19
<b>2.1</b>	<b>As séries que compõem o corpus de pesquisa</b>	19
<b>2.2</b>	<b>Língua de especialidade e ficção</b>	24
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	30
<b>3.1</b>	<b>Linguística de Corpus e Terminologia</b>	30
<b>3.2</b>	<b>Fraseologia baseada em <i>corpus</i></b>	37
<b>3.3</b>	<b>Colocações especializadas</b>	44
<b>3.4</b>	<b>Fraseografia e Teoria Funcional da Lexicografia</b>	50
3.4.1	Fraseografia	50
3.4.2	Teoria Funcional da Lexicografia	55
<b>4</b>	<b>PROCEDIMENTOS METODOLÓGICOS</b>	61
<b>4.1</b>	<b>Tratamento dos <i>corpora</i> e levantamento das colocações especializadas</b>	61
4.1.1	O Corpus Comparável Criminal II	67
<b>4.2</b>	<b>Elaboração da amostra de dicionário pedagógico</b>	69
<b>5</b>	<b>ANÁLISE DOS DADOS</b>	74
<b>5.1</b>	<b>Análise quantitativa</b>	74
<b>5.2</b>	<b>Análise morfossintática</b>	83
<b>5.3</b>	<b>Análise léxico-semântica</b>	91
5.3.1	O termo <i>unsub</i>	91
5.3.2	<i>Drive the unsub</i>	96
5.3.3	<i>Profile the unsub</i>	97
5.3.4	<i>Sketch of the unsub</i>	100
5.3.5	<i>Airtight alibi</i>	101
5.3.6	<i>Ironclad alibi</i>	104
5.3.7	<i>Profile of the unsub x Profile the unsub</i>	106
<b>6</b>	<b>DO DICIONÁRIO PEDAGÓGICO DE COLOCAÇÕES ESPECIALIZADAS DO ÂMBITO CRIMINAL</b>	109
<b>6.1</b>	<b>Definindo o Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal</b>	109

<b>6.2 A Macroestrutura do DPCEC</b>	112
<b>6.3 A microestrutura do DPCEC</b>	115
<b>6.4 Amostra do Dicionário Pedagógico de Colocações Especializadas da área Criminal – DPCEC</b>	118
<b>CONSIDERAÇÕES FINAIS</b>	135
<b>REFERÊNCIAS</b>	139

## 1 INTRODUÇÃO

As séries de televisão têm se tornado mais presentes na vida das pessoas, a julgar pela frequência com que novas séries são adicionadas aos *streamings* e como estes têm atingido um público cada vez maior. Nesse sentido, as pesquisas linguísticas também se valem das séries, desde o surgimento de *Friends* (QUAGLIO, 2009), seja utilizando-as como meio de ensino-aprendizagem de idiomas, seja como fonte para análises lexicais e de tradução (CALDAS, ORENHA-OTTAIANO, 2017; FROMM, 2011; ROCHA, ORENHA-OTTAIANO, 2012, 2015; FROMM; BANG, 2013; ESPERANDIO; FINATTO, 2014; ESPERANDIO, 2015).

Entre as séries que mais chamam a atenção do público, destacam-se as séries criminais, que retratam o cotidiano de policiais, investigadores forenses e agentes do FBI, da SWAT e da CIA, por exemplo. Por meio dessas séries, os termos utilizados nas áreas policiais de investigação forense, de criminalística, de análise comportamental, bem como aqueles utilizados por legistas ao realizar as autópsias, deixaram de se restringir a profissionais da área e se tornaram conhecidos por um público muito vasto, uma vez que essas séries aproximaram o público em geral dos trabalhos desenvolvidos por tais profissionais, ainda que por meio da ficção.

Ademais, as pesquisas que abordam aspectos do léxico e da fraseologia que, até então, não tinham tanto destaque, como, por exemplo, as expressões idiomáticas, as colocações e os binômios, aumentaram muito nos últimos anos. Mais especificamente, trabalhos que se dedicam à análise das colocações ganharam força e se destacaram (ZULATEGUI, 2002; ORENHA-OTTAIANO, 2004, 2009, 2012, 2015, 2016, 2020, 2021; ALONSO RAMOS, 2006; FRITZING; HEID, 2009; WALKER, 2011; LAUFER; WALDMAN, 2011; ALMELA, 2011; CALDAS, 2017; ROCHA 2017; 2020).

Nesse sentido, a presente pesquisa buscou analisar os *corpora* constituídos pelas transcrições das legendas em inglês, português e espanhol dos seriados *CSI – Crime Scene Investigation*, *NCIS* e *Criminal Minds*, a fim de identificar colocações especializadas da área criminal, tema central a estas séries, além de compilar um *corpus* comparável constituído por artigos científicos encontrados por meio do Google Acadêmico, tendo em vista a comprovação dos dados obtidos por meio do *corpus* de estudo em língua inglesa. Como objetivo geral, as colocações encontradas nos *corpora* de estudo foram analisadas a fim de verificar a sua taxonomia, de forma a evidenciar a tipologia mais utilizada nas colocações da

área criminal. Outrossim, a pesquisa teve como objetivos específicos: i) verificar o *LogDice*<sup>1</sup> de cada colocação encontrada, o qual indica o quão forte a colocação é; ii) empreender uma análise léxico-semântica comparativa entre as colocações encontradas nos idiomas que compõem nossos *corpora* de pesquisa, buscando enfatizar as semelhanças e diferenças encontradas entre as colocações no idioma de partida e suas respectivas traduções nos idiomas de chegada; iii) verificar a ocorrência das colocações previamente encontradas no *corpus* de estudo em língua inglesa no *corpus* comparável, visando comprovar a utilização destas por profissionais da área investigativa criminal; iv) apresentar informações sobre a composição da microestrutura de dicionários fraseológicos voltados para as colocações, tais como sua composição morfológica, sua base e sua frequência; v) apresentar uma amostra de dicionário construída a partir das colocações especializadas obtidas por meio da análise dos *corpora*.

A justificativa para se empreender uma pesquisa com tais objetivos recai sobre o crescente interesse pelos estudos sobre colocações e pelo destaque que vem sendo conferido a estas por diversos pesquisadores ao elegerem as colocações para comporem seus dicionários e dicionários fraseológicos, tais como Benson (1989), Alonso Ramos (2006), Cowie (2012), Vincze e Alonso Ramos (2013), Buendía e Faber (2014), Orenha-Ottaiano (2017, 2020, 2021), entre outros. Vemos, assim, um crescimento da interdisciplinaridade entre Fraseologia e Lexicografia, resultando em trabalhos fraseográficos, o que serviu de ponto de partida para a elaboração de nossa proposta de dicionário.

Outrossim, a escolha pela âmbito criminal se deu pelo fato de ter sido possível ampliar o estudo e o *corpus* utilizado previamente na pesquisa de mestrado (CALDAS, 2017), uma vez que esta área tem ganhado destaque e relevância entre os aprendizes de idiomas de diferentes níveis e faixas etárias, devido à grande propagação de séries de televisão que têm como tema principal a investigação criminal e seus desdobramentos, aproximando, assim, a linguagem de uma área de especialidade a um público expressivo ao redor de todo o mundo.

Assim, ao considerarmos o tema abordado pelas séries que compõem nosso *corpus* de estudo, o objetivo geral e os objetivos específicos apresentados, partimos da hipótese de que encontraríamos colocações especializadas, comumente presentes na área investigativa/criminal, retratadas nas séries escolhidas.

---

<sup>1</sup> “A pontuação de tipicidade (*LogDice*) indica o quão forte uma colocação é. Quanto maior a pontuação, mais forte é a colocação. Uma pontuação baixa significa que as palavras que na colocação também são combinadas, frequentemente, com muitas outras palavras” (SKETCH ENGINE, acesso em 20 de maio de 2022). “*The typicality score (LogDice) indicates how strong the collocation is; The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words*”.

Consideramos, ainda, a escolha por trabalhar com legendas traduzidas por fãs e disponibilizadas na internet, devido à: i) facilidade de acesso aos dados para constituição do *corpus* de estudo; ii) nossa intenção em promover a utilização de legendas produzidas na informalidade, facilmente consumidas pelos fãs das séries, como fonte de aprendizado e referência para estudos linguísticos, em razão ao tratamento científico apresentado em um dicionário fraseológico, como o proposto ao final desta pesquisa.

Dessa forma, o trabalho segue, no segundo capítulo, com uma apresentação das três séries que compõem o *corpus* de pesquisa, a saber: *CSI – Crime Scene Investigation*, *Criminal Minds* e *NCIS*. Nesse capítulo, ainda é feita uma discussão sobre a utilização de séries como amostra de uma dada língua de especialidade, de forma que justificamos a utilização da linguagem criminal em nosso trabalho com base em autores como Fromm (2011), Esperandio e Finatto (2014) e Castro (2019).

No terceiro capítulo, apresentamos as teorias que servem de base para esse trabalho. Discorreremos sobre a Linguística de Corpus (CRYSTAL, 1995; BIBER, 1998; TOGNINI-BONELLI, 2002; DASH, 2008; McENERY E HARDIE, 2012) e a Terminologia (CABRÉ, 1999; BARROS, 2004, FINATTO et. al. 2010; FINATTO, 2011; KRIEGER, 2013; FADANELLI, 2017). Em seguida, abordamos a Fraseologia baseada em *corpus* (COWIE, 1998; GROOM, 2005; NESSELHAUF, 2005; GRANGER; MEUNIER, 2008; GRANGER; PAQUOT, 2008; BERTRAÑA; BERTRÁN, 2008; GRIES, 2008; TAGNIN, 2013; ORENHA-OTTAIANO, 2004, 2009, 2021).

As colocações especializadas (L’HOMME, 2000, BEVILACQUA, 2004, 2005; ORENHA-OTTAIANO, 2009, 2012, 2020) também são abordadas e discutidas, ainda no segundo capítulo para, em seguida, abordarmos a Fraseografia (OLÍMPIO DE OLIVEIRA, 2007; TRISTÁ PERÉZ, 1998) e a Teoria Funcional da Lexicografia (FUERTES-OLIVERA; BERGENHOLTZ, 2011; FUERTES-OLIVERA; TARP, 2014; TARP, 2014), disciplinas que contribuirão para a amostra do dicionário pedagógico de colocações aqui apresentado.

No quarto capítulo, são apresentados os procedimentos metodológicos utilizados para a compilação e o tratamento dos *corpora* de estudo e de um *corpus* comparável, o qual foi compilado com o objetivo de endossar os dados levantados em nosso *corpus* de estudo em língua inglesa. Para tal, contamos com a plataforma *Sketch Engine* (KILGARIFF, 2004), mais especificamente, com o auxílio da ferramenta *Word Sketch*, presente na referida plataforma, para a obtenção das colocações que são analisadas neste trabalho, além de apresentarmos os procedimentos utilizados na análise das colocações. Ainda no quarto capítulo, são apresentados os procedimentos metodológicos necessários para a proposta de

dicionário pedagógico de colocações especializadas exposto neste trabalho.

O quinto capítulo, por sua vez, traz a análise das colocações especializadas obtidas, abrangendo uma análise quantitativa, passando pela análise morfossintática e, posteriormente, apresentando a análise léxico-semântica. Por fim, no capítulo seis discorreremos sobre a amostra do Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal, o DPCEC, o qual é construído no formato impresso, mas será transferido para o formato online futuramente, uma vez que se insere, também, no projeto de pesquisa intitulado *COMADEC - A Compilação de Materiais Didáticos e Dicionários Especializados de Colocações Baseados em Corpora e A phraseographical methodology and model for an online corpus-based Multilingual Collocations Dictionary Platform*, ambos coordenados pela Profa. Dra. Adriane Orenha-Ottaiano. Nesse sentido, apresentamos os aspectos teóricos que nos levaram a sua definição e proposta, além de apresentarmos a referida amostra compilada a partir das palavras-chave selecionadas. Em seguida, trazemos as considerações finais e as referências.

## 2 SÉRIES E LÍNGUA DE ESPECIALIDADE

Neste capítulo, discorreremos sobre as séries escolhidas para integrarem nosso *corpus* de pesquisa, a saber: *CSI – Crime Scene Investigation*, *Criminal Minds* e *NCIS – Naval Criminal Investigative Service*, além de elucidarmos os motivos que contribuíram para sua escolha e justificarmos a utilização de séries como *corpus* de pesquisa.

O trabalho com *corpus* constituído por legendas de séries de TV teve início no desenvolvimento da pesquisa de mestrado que precede a atual tese de doutoramento. Na referida pesquisa, Caldas (2017) utilizou como ponto de partida para a busca e o levantamento de colocações especializadas do âmbito criminal um *corpus* constituído com as legendas de doze temporadas da série *CSI -Crime Scene Investigation* em língua inglesa, obtendo resultados satisfatórios quanto à riqueza e variedade de colocações especializadas encontradas, as quais tiveram suas frequências comprovadas por meio do emulador de concordanciador *WebCorp*. Tal fato, nos motivou a continuar utilizando as legendas de séries como fonte de busca e levantamento de colocações especializadas.

Sendo assim, para a atual pesquisa mantivemos o *corpus* composto pelas legendas da série *CSI – Crime Scene Investigation*, em língua inglesa, e o ampliamos, contando com as legendas das séries *Criminal Minds* e *NCIS*. Ademais, compilamos mais dois *corpora* paralelos, que contam com as legendas em línguas portuguesa e espanhola das referidas séries.

Na subseção seguinte, discorreremos sobre as três séries para, na subseção posterior, discutirmos sobre a utilização de séries como fonte de pesquisas para estudos do léxico fraseológico.

### 2. 1 As séries que compõem o *corpus* de pesquisa

A série *CSI – Crime Scene Investigation* estreou em outubro de 2000, nos Estados Unidos e, em abril de 2001, no Brasil, tendo ficado conhecida em nosso país como *CSI Las Vegas*. Seu último episódio foi exibido nos Estados Unidos em 27 de setembro de 2015, no formato de um telefilme com duas horas de duração. No Brasil, sua exibição passou por vários canais, variando entre a TV aberta e canais pagos, como a TV Record e os canais *Sony* e *AXN*, respectivamente.

A série busca retratar a rotina de uma equipe de cientistas forenses do Laboratório de Criminalística de Las Vegas, o segundo maior laboratório do país, ficando atrás apenas do

laboratório do FBI. Criada por Anthony E. Zuiker, a série recebeu ao longo de suas temporadas inúmeras indicações aos prêmios importantes na televisão, tendo contabilizado o total de seis *Emmy Awards* (prêmio mais importante da televisão), e sete premiações no Festival de Televisão de Monte Carlo como a série dramática mais assistida no mundo.

Figura 1: Elenco original da série *CSI – Crime Scene Investigation*.



Fonte: <https://gauchazh.clicrbs.com.br/cultura-e-lazer/tv/noticia/2020/10/nos-20-anos-de-csi-investigacao-criminal-confira-sete-curiosidades-sobre-a-serieckfx4hhyl001h016vl8wrujvg.html>. Acesso em 06 de outubro de 2021.

O elenco original da série era composto pelo ator Paul Guilfoyle, que interpretava o capitão James Brass, considerado o líder da equipe. Entretanto, a equipe passa a ser liderada pelo entomologista Gill Grissom, vivido por William Petersen. Grissom fica responsável pelo turno da noite, comandando a equipe composta por Catherine Willows (Marg Helgenberger), Nicholas Stokes (George Eads), Warrick Brown (Gary Dourdan), Gregory Sanders (Eric Szmanda), Sara Sidle (Jorja Fox) e Albert Robbins (Robert David Hall), personagens que compõem o núcleo principal da série até a oitava temporada, quando alguns personagens começam a se desligar, tais como Sara Sidle (Jorja Fox) que se afasta de Las Vegas e da profissão de perita criminal após ter sido sequestrada, e Warrick Brown (Gary Dourdan) que é assassinado no final da oitava temporada por ter descoberto corrupção por parte de outros policiais do departamento.

Devido ao grande sucesso que obteve junto ao público, a série *CSI* se tornou uma

franquia dando origem a três *spin-offs*: *CSI Miami*, *CSI New York* e *CSI Cyber*, tendo sido lançados, respectivamente em 2002, 2013 e 2016 e, assim como a série que deu origem a elas, todas já foram canceladas. Entre os três *spin-offs*, pode-se dizer que *CSI Miami* foi o que obteve maior sucesso junto ao público, tendo ficado no ar de 2002 a 2011 e recebeu premiações importantes, tais como o *Emmy* em 2003 e 2007.

Por sua vez, a segunda série escolhida para compor o *corpus* de estudo desta pesquisa, *Criminal Minds*, estreou nos Estados Unidos em 22 de setembro de 2005 no canal CBS, tendo sido transmitida no Brasil na TV fechada pelo canal AXN. A série foi cancelada em 2019, ano em que foi anunciada que a sua última temporada, a de número 15, iria ao ar em 2020.

Diferentemente do que é retrato na série *CSI*, *Criminal Minds* busca enfatizar o trabalho desenvolvido pela UAC – Unidade de Análise Comportamental (ou *BAU*, *Behavioral Analysis Unit*, em inglês) uma equipe de elite do FBI que analisa o comportamento dos criminosos por meio de diferentes traços e vestígios deixados nas cenas de crimes, bem como, por meio das próprias vítimas. Dessa forma, o foco da série não está no crime em si, mas sim sobre os próprios criminosos, que têm seus perfis elaborados pela equipe.

A sinopse da série, fornecida pelo site do canal AXN reforça a importância do trabalho dos agentes especiais da UAC:

Criminal Minds se passa dentro da Unidade de Análise Comportamental (UAC, ou BAU em inglês) do FBI. Sua equipe é responsável por usar métodos analíticos para descobrir as motivações dos assassinos em série. Dessa forma, eles são capazes de prever seu comportamento e impedir seus próximos passos. Em meio ao clima obscuro que ronda todos os membros da BAU, as tramas pessoais se desenrolam, levando suas próprias mentes ao limite, enquanto trabalham para capturar os criminosos mais perigosos dos EUA. (Retirado de: <https://br.axn.com/series/criminal-minds>. Acesso em 01/12/19).

Além de terem que lidar com os piores criminosos do país, os agentes especiais precisam lidar com situações pessoais complexas, que, por vezes, acabam influenciando no bom andamento das investigações. O elenco da série sofreu algumas mudanças ao longo das temporadas, sendo que muitas delas foram justificadas na própria trama por situações pessoais conflitantes. Durante a maior parte das temporadas, a equipe foi composta por Aaron Hotchner (Thomas Gibson), Derek Morgan (Shemar Moore), Dr. Spence Reid (Matthew Gray Gubler), Jhenifer Jereau (A. J. Cook), Penelope Garcia (Kirsten Vangsness), Emily Prentiss (Paget Brewster) e David Rossi (Joe Mantegna).

Figura 2: Elenco da série *Criminal Minds*.



Fonte: <https://observatoriodocinema.uol.com.br/series-e-tv/2021/02/criminal-minds-vai-voltar-veja-detalhes>. Acesso em 06 de outubro de 2021.

Assim como aconteceu com outras séries, *Criminal Minds* deu origem a dois *spin-offs*: *Criminal Minds Suspect Behavior* e *Criminal Minds: Beyond Borders* e ambos foram cancelados com pouco tempo de duração. Apesar de seus *spin-offs* não terem alcançado sucesso, a série recebeu o *People's Choice Award* em 2017, pela escolha do público como drama criminal favorito, além de ter alcançado uma pontuação de 42% no *Metacritic*<sup>2</sup>, com base em 21 avaliações profissionais, logo em sua primeira temporada. Ademais, em notícia publicada em diversos sites no início de 2021, foi divulgado que novos episódios da série seriam produzidos com a presença do elenco original.

Uma vez que a série se concentra mais no comportamento dos criminosos, buscando traçar seus perfis, o léxico utilizado acompanha o mesmo viés, sendo possível identificar colocações especializadas diferentes daquelas que são mais comumente encontradas no léxico da série *CSI*, por exemplo, o que contribui para o enriquecimento do trabalho e, conseqüentemente, do produto final desta pesquisa, ou seja, a amostra de um dicionário de colocações especializadas.

Seguindo esta mesma linha, sobre o enriquecimento do léxico colocacional que compõe o trabalho, o qual proporciona uma maior abrangência de colocações especializadas do âmbito criminal compreendendo colocações que são usadas em diferentes instâncias de

<sup>2</sup> Site americano responsável por reunir críticas de filmes, programas de tv, jogos e CDs.

investigação, apresentamos a terceira e última série que compõe nosso *corpus* de estudo: *NCIS – Naval Criminal Investigative Service*. Esta série surgiu como um *spin-off* de *JAG – Judge Advocate General*, série americana de aventura e drama que foi ao ar entre 1995 e 2005. *NCIS* foi lançada em setembro de 2003 a partir de dois episódios da série original e se encontra, atualmente, em sua vigésima temporada.

O enredo da série gira em torno da equipe comandada pelo agente especial Leroy Jethro Gibbs, interpretado pelo ator Mark Harmon, responsável por investigar os mais variados tipos de crimes que envolvem membros da Marinha dos Estados Unidos e o Corpo de Fuzileiros Navais, bem como suas famílias. A equipe é composta ainda pelos agentes especiais Anthony DiNozzo (Michael Weatherly), Timothy McGee (Sean Murray), Ziva David (Cote de Pablo), pelo Dr. Donald "Ducky" Mallard (David McCallum) e os diretores(as) do *NCIS*, Jennifer Shepard (Lauren Holly) e Leon Vance (Rocky Carroll). Ao longo das temporadas, a equipe sofre perdas e recebe novos membros, sem que isso interfira nas investigações que empreendem.

Figura 3: Elenco da série *NCIS – Naval Criminal Investigative Service*.



Fonte: <https://www.teleseries.com.br/decima-temporada-de-ncis-estreia-nessa-segunda-feira-no-axn-brasil/>. Acesso em 06 de outubro de 2021.

A série alcançou muito sucesso, chegando a receber o *International Television Audience Award* em 2014 como série de maior audiência mundial na categoria drama. Tamanho sucesso fez com que a série originasse dois *spin-offs* (apesar de ser derivada de outra série já cancelada, conforme mencionamos anteriormente): *NCIS Los Angeles* e *NCIS New Orleans*, sendo que ambos alcançaram tanto sucesso quando a série original.

Nesta série, o foco recai sobre os crimes em si e não nos criminosos, mas, por se tratar de crimes relacionados a uma área muito específica (Marinha e Fuzileiros Navais), ela nos permite uma maior variedade lexical que se reflete no léxico colocacional também. Sendo assim, o *corpus* composto pelas três séries aqui apresentadas objetiva abranger uma maior diversidade lexical recorrente no âmbito criminal em algumas de suas nuances, visando o melhor desenvolvimento e uma maior riqueza colocacional da proposta de dicionário apresentada como produto final desta pesquisa.

## 2. 2 Língua de especialidade e ficção

Apesar do número relevante de pesquisas nas áreas de Terminologia (FROMM, 2011; ESPERANDIO, FINATTO, 2014; CASTRO, 2019), Fraseologia (ROCHA, ORENHA-OTTAIANO, 2012; PEIXOTO, 2014; CALDAS, 2017; CALDAS, ORENHA-OTTAIANO, 2017; YAMAMOTO; LISBOA, 2019) e Linguística de Corpus (CSOMAY, PETROVIC, 2012; FISHEL et al, 2012) que foram desenvolvidas tendo como *corpus* legendas de séries, este é um assunto que ainda merece atenção, haja vista as divergências que podem surgir, principalmente quanto à definição destes *corpora* como amostras de áreas de especialidade.

Sendo assim, convém ressaltar que neste trabalho entendemos como áreas de especialidades aquelas que apresentam uma linguagem própria de seu campo de atuação, ou seja, uma língua de especialidade, a qual entendemos como “o conjunto de todos os recursos linguísticos que são utilizados em um âmbito comunicativo, delimitado por uma especialidade, para garantir a compreensão entre as pessoas que nele atuam” (HOFFMANN, 1988, apud FINATTO; ZILIO, 2015, p. 40 -41).

Segundo Hoffmann (1988) a especificidade de uma língua de especialidade, ou de uma linguagem especializada, se expressa por meio do vocabulário especializado pertencente a tal linguagem, ao qual “pertencem todas as unidades lexicais contidas em textos especializados, já que essas unidades contribuem para a comunicação especializada de maneira direta ou indireta” sendo este vocabulário especializado composto por

a) termos, b) semitermos e c) jargões especializados. Nesse processo, são reconhecidos como termos apenas as palavras cujo conteúdo seja determinado por meio de uma definição normativa; de outro lado, os semitermos não estão definidos em normas, mas são bastante precisos em descrição e denotação. O jargão especializado, por sua vez, não exige precisão” (HOFFMANN, 1988, apud FINATTO; ZILIO, 2015)

O autor define, assim, o vocabulário especializado como sendo um conjunto composto por diferentes unidades lexicais que recebem as classificações de termos, semitermos e jargões especializados, de acordo com o seu grau de definição normativa. De forma semelhante, Barbosa (2006) aborda a questão dos quase-termos técnicos ao discorrer sobre a Etnoterminologia, atribuindo tal classificação ao fato de estas unidades lexicais apresentarem traços da linguagem de especialidade e da linguagem literária.

Considerando, assim, o exposto e defendido pelos dois autores, podemos entender que as séries que foram selecionadas para compor nosso corpus de estudo constituem-se como fontes de vocabulário especializado pois apresentam termos e semitermos - ou quase-termos, nas palavras de Barbosa (2006) – pois, tal como afirma Fromm (2011), desde a década de 1960 é possível perceber um aumento na utilização de termos emprestados de áreas de especialidades por parte de diversas séries norte-americanas.

Assim, Fromm (2011) identificou três padrões nessa utilização, sendo que o último padrão identificado pelo autor vai ao encontro do que é apresentado nas séries que compõem nosso *corpus* de estudo: “c. séries que retratam o cotidiano de médicos, investigadores, cientistas forenses, etc.: neste caso, as terminologias usadas tentam retratar, de forma fiel, aquelas usadas pelos profissionais dessas áreas nos seus afazeres cotidianos” (FROMM, 2011, p. 2).

Conforme apontado pelo autor, vale ressaltar que as séries que abordam as temáticas investigativas, como é o caso das séries selecionadas para comporem nosso *corpus* de estudo - *CSI – Crime Scene Investigation*, *Criminal Minds* e *NCIS* - buscam retratar “de forma fiel” as terminologias que são utilizadas por esses profissionais em suas rotinas diárias.

Nesse sentido, além do proposto por Fromm (2011), ao consultarmos o site do canal de TV *AXN*, responsável por transmitir no Brasil as séries selecionadas como *corpus* de pesquisa, encontramos uma informação sobre uma das séries que vem reforçar a utilização de termos técnicos da área de investigação forense/criminal, a qual afirma que “com *CSI*, foi possível mostrar um processo mais realista que mudou a percepção da investigação criminal, incluindo até mesmo *termos técnicos* que os fãs adotariam mais tarde” (*AXN*, 2021, grifo nosso).

Em outro momento, ainda no referido site, afirma-se que “a discussão técnica detalhada entre os personagens, assim como a representação gráfica de trajetória de balas e métodos de reconstrução das cenas de crimes também foram um diferencial” (*AXN*, 2021). Ademais, o criador da série se inspirou em Daniel Holstein e Yolanda McCrery, aos quais conheceu

enquanto escrevia o roteiro da série e que são “grandes criminalistas do departamento de Las Vegas” (idem) para criar as personagens Gil Grissom e Catherine Willowsn.

A série se tornou responsável, inclusive, por influenciar muitas pessoas a se aprofundarem no mundo da criminologia (AXN, 2021), além de ter influenciado a criação de outras séries que seguem o mesmo estilo, tal como apontado na seção anterior.

Dessa forma, tomando como base a definição apresentada por Fromm (2011), o exposto sobre a série *CSI: Crime Scene Investigation* pela emissora responsável pela sua transmissão e as considerações sobre vocabulário especializado de Hoffmann (1998), bem como a definição de quase-termo proposta por Babosa (2006), entendemos que nosso *corpus* de pesquisa, composto pelas legendas desta e de outras duas séries que retratam o cotidiano de investigadores e cientistas forenses, se encaixa no que é definido no grupo c (FROMM, 2011) apresentado anteriormente, pois retrata as terminologias de suas respectivas áreas de atuação (investigação criminal, ciência forense) de forma fiel.

Nesse sentido, Fromm (comunicação pessoal, 2022) entende, ainda, que os termos são utilizados em um produto ficcional, porém, retratam fielmente uma realidade profissional, ou seja, são termos regulares, com uso diferenciado.

Para o autor, o uso dos termos regulares em ambiente ficcional é expandido para produtos audiovisuais e legitimado pela Etnoterminologia, subárea da Terminologia, a partir da proposta elaborada por Barbosa (2005, 2006, 2009). De acordo com a autora, a Etnoterminologia “busca estudar a norma relativa ao estatuto semântico, sintático e funcional do conjunto das unidades lexicais que caracterizam o universo dos discursos etno-literários, [...]” (BARBOSA, 2006, p. 48).

A autora afirma que esse estudo se justifica pelo fato de essas unidades lexicais possuírem “sememas muito especializados, construídos com semas específicos de discurso em causa, provenientes das narrativas, cristalizados, de modo a tornar-se verdadeiros símbolos dos temas envolvidos” (BARBOSA, 2006, p. 48).

Nesse sentido, a autora explica que

[...] as unidades lexicais do universo de discurso etno-literário têm um estatuto próprio e exclusivo. Nos níveis da norma e da fala, subsumem duas funções, **vocábulo** e **termo**. Associam aspectos referenciais, pragmáticos e simbólicos, em função semiótica, metassemiótica e metametassemiótica, próprias dos vocábulos, mas apresentam, também, características de uma linguagem de especialidade. Tais unidades lexicais têm um significado muito particular, peculiar a esse universo de discurso, e são, ao mesmo tempo, polissemêmicas. Noutras palavras, essas unidades lexicais reúnem qualidades das línguas especializadas e da linguagem literária, de maneira a preservar um valor

semântico social e constituir, simultaneamente, documentos do processo histórico da cultura. Resultam elas do cruzamento de um processo de metaterminologização e de metavocabularização. (BARBOSA, 2006, p. 48).

Com base na definição de Etnoterminologia apresentada por Barbosa (2006), é possível compreender a aplicação dessa subárea da Terminologia por Fromm (2011), ao propor que as séries que retratam o cotidiano de médicos, investigadores e cientistas forenses fazem uso de termos regulares em um ambiente ficcional.

Outrossim, ainda é possível recorrer a Barbosa (2009) para entender, no caso das séries que constituem nosso *corpus* de estudo, que há um processo de banalização dos termos científicos, o qual promove a difusão do conhecimento que antes era restrito aos profissionais das áreas de investigação e de ciência forense, uma vez que a autora enfatiza a importância dos processos de banalização, vulgarização, popularização de linguagens especializadas, como mecanismos de circulação e difusão do conhecimento contando com a “necessária adequação dos discursos a diferentes grupos de destinatários, a comunicação entre especialistas e não especialistas, os distintos níveis de linguagem envolvidos e algumas decorrências transdisciplinares” (BARBOSA, 2009, p. 31).

Sobre o processo de banalização, Barbosa (2009) afirma, ainda, que no

continuum científicidade-banalização, podemos detectar termos/vocábulo que se situam na interface entre o discurso científico e o discurso banal, como por exemplo, o termo/vocábulo câncer. Esse tipo de termos garante a comunicação entre especialistas de uma área, entre leigos, entre os primeiros e os segundo. Por conseguinte, entre o mais alto grau de científicidade e o mais alto grau de banalização existe sempre um subconjunto que tem dupla natureza, a de termo e a de vocábulo. (BARBOSA, 2009, p. 39)

Dessa forma, ao levarmos em consideração o exposto por Fromm (2011) e Barbosa (2006, 2009), entendemos que nosso *corpus* de estudo composto pelas legendas das séries *CSI: Crime Scene Investigation*, *Criminal Minds* e *NCIS*, constitui-se como uma fonte de pesquisa da área criminal/investigativa, uma vez que apresenta termos regulares, ou seja, científicos e técnicos, além de ser composto por semitermos, também, de acordo com a definição de Hoffmann (1988), que se encontram no continuum entre científicidade e banalização, tal como exposto por Barbosa (2009), de forma que, neste trabalho, optamos por denominá-los como termos, buscando uma padronização e melhor descrição dos procedimentos teórico-metodológicos utilizados na descrição e análise das colocações especializadas aqui apresentadas.

Ademais, no que tange a área de Linguística de Corpus, Csomay e Petrovic (2012), ao

empreenderem um estudo de vocabulário técnico baseado nas transcrições de sete filmes e cinco episódios de séries televisivas que apresentam como tema principal a lei, afirmam que a utilização de filmes e séries como fonte de pesquisas na área de aprendizagem incidental já vem sendo debatida e enfatizada há tempos. Os autores mencionam os estudos de Webb e Rogers (2009) e de Webb (2010), para ressaltar os benefícios que alunos de idiomas obtêm ao assistirem séries e filmes relacionados a determinada área. Segundo os autores

[...]. Assistir a filmes com conteúdo específico de uma disciplina fornece uma oportunidade para mais encontros com palavras desconhecidas (Webb, 2010) e, portanto, é propício para a aprendizagem incidental. Por exemplo, o drama de TV que segue a vida da equipe médica em um hospital (ex. *House*, *Grey's Anatomy*) fornecerá dados que compreendem vocabulário relativo a procedimentos e tratamentos médicos. Portanto, presume-se que assistir a filmes e programas de TV relacionados à disciplina pode beneficiar o aprendizado de vocabulário técnico<sup>3</sup> (CSOMAY; PETROVIC, 2012, p. 307).

Embora as pesquisas mencionadas abordem a questão da aprendizagem incidental, fica evidente a importância das legendas de séries e filmes para o aprendizado de uma língua estrangeira ou segunda língua e, até mesmo, como fonte de aprendizado de vocabulário especializado, fato que corrobora a utilização das legendas de séries como *corpus* desta pesquisa.

Cabe ressaltar ainda que, em relação à utilização das legendas obtidas por meio da *web* e compiladas como nosso *corpus* de pesquisa, estas constituem-se como as chamadas legendas não comerciais, conforme aponta Sperandio (2015), uma vez que são produzidas por fãs das séries (ou de filmes) que se organizam em grupos de legendadores, os quais, normalmente, contém um membro que é da área da qual a série trata e que fica responsável por verificar a tradução correta dos termos em questão do idioma de partida para o idioma de chegada, conforme salienta Sperandio (2015).

Nesse sentido, a autora ressalta ainda que

é de senso comum que toda tradução é passível de erro, independentemente da experiência e da qualidade do tradutor. Como prova disso, trazemos um exemplo de problema de tradução encontrado na legendagem comercial do seriado *Dexter*. No primeiro episódio da primeira temporada, a personagem pergunta: “Where the hell do you keep your gun”. A legenda do DVD traduz:

---

<sup>3</sup> *Watching movies with discipline-specific content provides an opportunity for more encounters with unknown words (Webb, 2010) and therefore, is conducive to the opportunity to learn incidentally. For example, a TV drama that follows the lives of medical staff in a hospital (e.g., House, Grey's Anatomy) will provide input comprising vocabulary pertaining to medical procedures and treatments. Hence, the assumption is that watching discipline-related movies and TV shows can benefit the learning of technical vocabulary.* [Tradução nossa].

“Onde guarda seu chiclete?”. A simples troca de fonema de /n/ (gun, ou arma) para /m/ (gum, ou chiclete) implica um erro de tradução que se pode considerar ser de um tradutor inexperiente. Há explicações passíveis, como “o tradutor só teve acesso ao script e ele continha esse erro” ou “o tradutor só teve acesso ao vídeo e não fez a distinção do fonema”. Ainda assim, o erro de tradução é claramente perceptível ao espectador, uma vez que ele ouve a fala do personagem e vê a movimentação da cena. Por sua vez, a legendagem não comercial apresentou: “Onde diabos guarda sua arma?”. Nossa intenção com isso é mostrar que, independentemente de a legenda ser comercial ou não, a tradução é passível de erro. (SPERANDIO, 2015, p. 27).

Conforme salientado pela autora, também acreditamos que toda tradução é passível de erro e, devido a nossa experiência com a pesquisa em legendas não comerciais, podemos dizer que elas não deixam a desejar em comparação às legendas comerciais. Ressaltamos, porém, que, ao identificarmos alguma inconsistência na tradução das legendas que compõem nosso *corpus* de estudo, foram propostas novas possibilidades de tradução.

Dessa forma, tendo apresentado as séries que compõem o *corpus* de estudo, bem como, justificado a utilização destas para a sua compilação, apresentamos no capítulo seguinte os fundamentos que serviram de base para o desenvolvimento desta pesquisa.

### 3 FUNDAMENTAÇÃO TEÓRICA

Apresentamos, neste capítulo, os pressupostos teóricos que serviram de base para o desenvolvimento de nossa pesquisa. Assim, discutimos os aspectos relevantes da Linguística de Corpus e da Terminologia. Em seguida, discorremos sobre a Fraseologia Baseada em *Corpus* e as Colocações Especializadas, para, posteriormente, apresentarmos aspectos relevantes sobre a Fraseografia e a Teoria Funcional da Lexicografia.

#### 3.1 Linguística de Corpus e Terminologia

Desde o surgimento do *Corpus Brown* (1964) a Linguística de Corpus ampliou-se ainda mais, sendo, hoje, uma área da Linguística capaz de auxiliar na promoção de diferentes pesquisas, indo desde a Linguística Histórica, passando pela Linguística Sistêmico-Funcional e a Linguística Cognitiva (NOVODVORSKI; FINATTO, 2014), chegando à Fraseologia, Fraseografia, ao Ensino de Línguas e a Terminologia.

Graças a esse percurso, ao olharmos para trás, vemos que há uma variedade de definições e características atribuídas à Linguística de Corpus (doravante LC). Para Biber et al (1998, p. 1), a LC é vista como uma “perspectiva diferente [...] de enfatizar a língua em uso”<sup>4</sup>, por meio da qual pode-se investigar a exploração dos recursos linguísticos empreendidos por falantes e escritores de dado idioma, ressaltando a forma como esses recursos ocorrem naturalmente na língua sob análise.

Segundo o autor, as características essenciais da LC são: ser empírica, já que analisa a ocorrência de uso em textos naturais; utilizar uma grande compilação de textos, chamada de *corpus* como base de suas análises; utilização de computadores para o desenvolvimento das análises, já que não seria possível empreender tais estudos manualmente; e depender das análises quantitativas e qualitativas, haja vista que a análise baseada em *corpus* deve ir além da contagem de características linguísticas, necessitando, sempre, do olhar qualitativo do linguista responsável por explorar a importância desses padrões em uso no idioma (BIBER, et al, 1998, p. 4 – 5).

Já Tognini-Boneli (2002) define a LC como

---

<sup>4</sup> “A diferente perspective [...] is to emphazise language use”.

uma abordagem empírica, na qual, como todos os tipos de investigação científica, o ponto de partida são dados reais autênticos. O procedimento para descrever os dados que fazem uso de *corpus* é, portanto, indutivo, já que são afirmações de caráter teórico sobre a língua ou a cultura, às quais se chegaram a partir de observações das instâncias reais. A observação dos fatos de linguagem conduz à formulação de uma hipótese para explicar esses fatos; esta, por sua vez, leva a uma generalização baseada na evidência de padrões repetidos utilizados na concordância; a última etapa é a unificação destas observações em uma formulação teórica<sup>5</sup> (TOGNINI-BONELLI, 2002, p. 2).

Ambas as definições apresentam a LC como uma abordagem empírica e enfatizam a utilização de dados autênticos obtidos por meio dos textos naturais que compõem os *corpora* (ou o *corpus*) e que são analisados de maneira quantitativa, observando-se as estatísticas de uso e recorrência, bem como, de maneira qualitativa, na qual a observação dos dados linguísticos e estatísticos passa pelo olhar do linguista que busca, além de descrevê-los, analisá-los a fim de explorar e unificar esses padrões em uma formulação teórica.

No que tange ao *corpus* mencionado por ambos os autores, este é definido por Biber et al (1998, p. 12) como uma “coleção ampla e baseada em princípios de textos naturais”<sup>6</sup>, enquanto Tognini Bonelli (2002) o define como

[...] coleção de textos entendidos como representativos de uma dada língua, reunidos de forma que possam ser usados para análise linguística. Normalmente, pressupõe-se que a língua armazenada em um *corpus* é de ocorrência natural, que é compilada de acordo com critérios, com um propósito específico em mente, e com o intuito de representar os ‘pedaços maiores do idioma’ selecionados de acordo com uma tipologia específica<sup>7</sup>. (TOGNINI-BONELLI, 2002, p.2).

De acordo com as definições de *corpus*, os princípios mencionados por Biber et al (1998) seriam os critérios que devem ser observados ao se compilar o *corpus*, tal como ressalta Tognini-Bonelli (2002). Tais critérios, ou princípios, devem conduzir a compilação do *corpus* de modo que este seja condizente com o objetivo da pesquisa, ou seja, um *corpus*

---

<sup>5</sup> “[...] as an empirical approach in that, like all types of scientific enquiry, the starting point is actual authentic data. The procedure to describe the data that makes use of corpus is therefore inductive in that it is statements of a theoretical nature about the language or the culture which are arrived at from observations of the actual instances. The observation of language facts leads to the formulation of a hypothesis to account for these facts; this in turn leads to a generalization based on the evidence of the repeated patterns in the concordance; the last step is the unification of these observations in a theoretical statement”.

<sup>6</sup> “[...] a large and principled collection of natural texts”.

<sup>7</sup> “[...] a collection of texts assumed to be representative of a given language put together so that it can be used for linguistic analysis. Usually, the assumption is that the language stored in a corpus is naturally-occurring, that is gathered according to explicit design criteria, with a specific purpose in mind, and with a claim to represent larger chunks of language selected according to a specific typology.”

de língua geral, por exemplo, dificilmente permitiria a análise de *chunks* (pedaços maiores do idioma) que pertencem a língua de especialidade.

Já Dash (2008) afirma que a LC

[...] como uma área importante da linguística computacional, desempenha um papel importante. Ela fornece grandes quantidades de base de dados empíricos acumulados de maneira sistemática, de várias áreas de atual língua em uso seguindo alguns métodos estatísticos e técnicas de amostragem de dados. Fornece, também, dispositivos sofisticados para analisar estes *corpora* para extrair dados linguísticos, exemplos e informações necessárias para a linguística aplicada, linguística computacional, e inteligência artificial para entender a linguagem humana da melhor maneira, bem como, para aplicar esses dados e informações em várias áreas do conhecimento humano<sup>8</sup> (DASH, 2008, s/p).

A partir da definição de Dash (2008), evidencia-se a importância da LC para outras áreas da Linguística, enfatizando-se, também, a utilização de dados empíricos que retratem a língua em uso. O autor salienta também a importância do desenvolvimento de ferramentas que permitem a análise dos *corpora*, bem como, ressalta que os resultados das pesquisas obtidas por meio da LC podem contribuir com diferentes áreas do conhecimento humano, não restringindo essas pesquisas à grande área da Linguística, uma vez que a linguagem perpassa todas as áreas do conhecimento.

Por sua vez, McEnery e Hardie (2012) afirmam que a LC “é certamente bem diferente da maioria dos outros tópicos que você deve ter estudado em linguística, e não é diretamente sobre o estudo de qualquer aspecto particular da língua. Ao contrário, é uma área que foca em um conjunto de procedimentos, ou métodos, para estudar a língua”<sup>9</sup> (MCENERY; HARDIE, 2012, p. 1). De acordo com os autores, o diferencial da LC reside no conjunto de ferramentas que a compõem e que auxiliam no desenvolvimento de pesquisas em diversas áreas da linguística.

Nesse sentido, os autores definem a LC nos seguintes termos

A linguística de corpus não é um conjunto monolítico e consensual de métodos

---

<sup>8</sup> “[...] as an important area of computational linguistics, plays an important role. It provides large quantities of empirical databases accumulated in a systematic manner from various fields of actual language use following some statistical methods and techniques devices to analyze these corpora to extract linguistic data, examples, and information necessary in applied linguistics, computational linguistics, and artificial intelligence for understanding human language in a better way as well as for applying this data and information in various fields of human knowledge”.

<sup>9</sup> “It is certainly quite distinct from most other topics you might study in linguistics, as it is not directly about the study of any particular aspect of language. Rather, it is an area which focuses upon a set of procedures, or methods, for studying language [...]”.

e procedimentos para a exploração da linguagem. Embora algumas generalizações possam ser feitas para caracterizar muito do que é chamado de “linguística de corpus”, é muito importante perceber que a linguística de corpus é um campo heterogêneo. [...]. Poderíamos razoavelmente definir a linguística de corpus como lidando com algum conjunto de textos legíveis por máquina que é considerado uma base apropriada para estudar um conjunto específico de questões de pesquisa. O conjunto de textos ou *corpus* tratado é geralmente de um tamanho que desafia a análise à mão e a olho apenas dentro de qualquer prazo razoável. É a grande escala dos dados usados que explica o uso de texto legível por máquina. [...]. A próxima generalização decorre dessa observação: os *corpora* são invariavelmente explorados por meio de ferramentas que permitem aos usuários pesquisá-los de forma rápida e confiável. [...]. As concordâncias e os dados de frequência exemplificam, respectivamente, as duas formas de análise, qualitativa e quantitativa, igualmente importantes para a linguística de corpus<sup>10</sup>. (MCENERY; HARDIE, 2012, p. 1 – 2).

Como se pode observar pelo posicionamento dos autores, embora apontem que a LC possa ser definida como uma metodologia, os princípios presentes no posicionamento de outros autores, como Dash (2008), Tognini-Bonelli (2002) e Biber (1998), também são observados em sua definição. Ou seja, mesmo se posicionando a favor de uma definição mais metodológica, os autores afirmam que a LC é constituída pela análise de *corpora* compostos por textos que são lidos por programas de computadores especialmente desenvolvidos para fins linguísticos, que permitem analisar uma maior quantidade de dados de forma mais rápida e confiável, resultando tanto em uma análise quantitativa, como em uma análise qualitativa por parte do linguista.

Ademais, tal como Dash (2008), McEnery e Hardie (2012) ressaltam a importância da LC para o desenvolvimento de pesquisas em outras áreas da linguística, vindo daí a sua visão de LC enquanto metodologia.

No que diz respeito à definição de *corpus*, Dash utiliza a definição dada por Crystal (1995), a qual diz que este se refere a “uma grande coleção de dados linguísticos, seja de textos escritos ou de transcrições de gravações de fala, o qual pode ser utilizado como um ponto de partida para a descrição linguística ou como um meio de verificar hipóteses sobre

---

<sup>10</sup> “*Corpus linguistics is not a monolithic, consensually agreed set of methods and procedures for the exploration of language. While some generalisations can be made that characterise much of what is called ‘corpus linguistics’, it is very important to realise that corpus linguistics is a heterogeneous field. [...]. We could reasonably define corpus linguistics as dealing with some set of machine-readable texts which is deemed an appropriate basis on which to study a specific set of research questions. The set of texts or corpus dealt with is usually of a size which defies analysis by hand and eye alone within any reasonable timeframe. It is the large scale of the data used that explains the use of machine-readable text. [...]. The next generalisation follows from this observation: corpora are invariably exploited using tools which allow users to search through them rapidly and reliably. [...] Concordances and frequency data exemplify respectively the two forms of analysis, namely qualitative and quantitative, that are equally important to corpus linguistics*”.

a língua”<sup>11</sup> (CRYSTAL, 1995, apud DASH, 2008, s/p).

Dash afirma, ainda, que o *corpus* pode ser definido de maneira acrostica a partir das letras que o compõem, haja vista que em inglês a sua definição se encaixa ao termo em questão. Para o autor, um *corpus* deve ser compatível com homem e máquina, operacional nas aplicações da pesquisa, representativo de uma língua ou variedade, processável tanto pelo homem quanto pela máquina, ser ilimitado na quantidade de dados e amostras e ser sistematizado pela sua formação e pela sua representatividade (DASH, 2008).

A partir do exposto aqui, entendemos que nosso *corpus* de pesquisa, composto pelas legendas das séries *CSI: Crime Scene Investigation*, *Criminal Minds* e *NCIS* apresenta as características necessárias para ser considerado como tal, haja vista sua extensão (mais de 9 milhões de palavras, somente o *corpus* de língua inglesa), sua representatividade da língua de especialidade, dado o fato de ser composto por legendas de séries que retratam o cotidiano de investigadores, agentes e cientistas forenses, contando, assim, conforme ressalta Fromm (2011), com a terminologia da área, além de atenderem nosso propósito de pesquisa, que é o de empreender o levantamento de colocações especializadas da área em questão.

Ademais, uma vez que nosso *corpus* de estudo é composto pelas legendas das séries em três diferentes idiomas, este é classificado como um *corpus* paralelo, o qual é definido por Tognini-Bonelli (2002) como um *corpus* no qual é possível identificar uma “correspondência de sentença por sentença entre os dois textos”<sup>12</sup> (TOGNINI- BONELLI, 2002, p. 6-7). Embora a definição da autora mencione a correspondência entre dois textos, nesta pesquisa, foi possível perceber a correspondência entre três “textos”, uma vez que as legendas se encontram originalmente em língua inglesa e traduzidas para dois idiomas de chegada diferentes, o português e o espanhol.

Ainda de acordo com a autora, nosso *corpus* também pode ser definido como um *corpus* especializado, haja vista que este se ocupa da língua de especialidade, ainda que retratada por meio da ficção, conforme explicado na seção 2.2 do capítulo 2, permitindo-nos olhar mais atentamente para as idiosincrasias do uso autêntico dessa linguagem especializada (TOGNINI-BONELLI, 2002), ou seja, este *corpus* nos permite buscar as colocações especializadas da área investigativa criminal, uma vez que é constituído por textos (legendas de séries) que abordam este tema.

---

<sup>11</sup> “[...] a large collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting point of linguistic description or as a mean of verifying hypotheses about a language”.

<sup>12</sup> “[...] sentence-by-sentence correspondence between the two texts”.

Esta pesquisa conta, ainda, com um *corpus* comparável, composto por artigos científicos que discorrem sobre a área forense. Esse tipo de *corpus* é definido por Tognini-Bonelli (2002) como

[...] *corpora* cujos componentes são escolhidos por serem amostras similares de suas respectivas línguas em termos de critérios externos como língua falada x língua escrita, registro, etc. No entanto, nenhum deles são traduções; nenhum alinhamento é possível, mas as correspondências podem ser estabelecidas entre as principais características linguísticas dos *corpora*<sup>13</sup>. (TOGNINI-BONELLI, 2012, p. 7).

Dessa forma, o *corpus* compilado a partir de artigos científicos que abordam a área investigativa/criminal, bem como aspectos forenses, atende ao requisito proposto pela autora de ser uma amostra similar do conteúdo apresentado em nosso *corpus* de estudo, por meio do qual é possível estabelecer correspondências e comprovar o uso, fora das séries, das colocações especializadas em nossa análise.

Assim, uma vez que esta pesquisa tem como objeto de estudo as colocações especializadas e estas, conforme explicado na seção 3.3, contam com um termo em sua formação, recorreremos aos pressupostos da Terminologia para salientar essa interdisciplinaridade.

A Terminologia é “uma disciplina, filiada à Linguística Aplicada, que se ocupa dos fenômenos da comunicação técnica e científica. Estuda o vocabulário científico e também as práticas textuais, incluindo a comunicação entre leigos e especialistas”(Finatto *et. al.* 2010, p. 227). Segundo Finatto (2011b), a Terminologia ainda tem como objetivo estudar o vocabulário científico e suas práticas textuais e convencionalidades. Esse estudo tem por objeto o termo, também chamado de unidade terminológica, o qual é definido como “uma unidade lexical com um conteúdo específico dentro de um domínio específico” (BARROS, 2004).

Uma vez que tem como objeto de estudo o termo, responsável por evocar conceitos e conteúdo específicos de uma língua de especialidade, a Terminologia difere-se da Lexicografia, já que esta última se ocupa do estudo da palavra. De acordo com Gouadec (1990) o “termo é uma unidade linguística que designa um conceito, um objeto ou um processo. O termo é a unidade de designação de elementos no universo percebido ou

---

<sup>13</sup> “[...] *corpora* whose components are chosen to be similar samples of their respective languages in terms of external criteria such as spoken vs. written language, register, etc. However, none of them are translations; no alignment is possible, but correspondences can be established among the main linguistic features of the *corpora*.”

concebido. Ele, raramente, confunde-se com a palavra ortográfica” (GOUADEC, 1990, apud KRIEGER, 2013, p. 27).

Devido a expansão de seu campo de atuação com o passar do tempo, deixando seu caráter mais normativo, específico de seu surgimento enquanto disciplina, na década de 1930, proposta pela Teoria Geral da Terminologia de Eugen Wüster, passando pela Teoria Comunicativa da Terminologia, proposta por Cabré na década de 1990, a Terminologia passa a abordar o termo de forma poliédrica, considerando seu aspecto linguístico, cognitivo e comunicacional (CABRE, 1999). Essa poliedricidade do termo faz com que esse passe a ser estudado por diferentes subáreas da linguística, como a fraseologia especializada, a definição terminológica e o texto especializado.

No que diz respeito à fraseologia especializada, esta “integra o núcleo de conhecimento de um texto e é equacionada com apoio de verbos e substantivos deverbais, ao que se pode acrescentar que costuma acolher um termo em sua estruturação formal” (KRIEGER, 2013, p. 25). Já o texto especializado é “o habitat natural das terminologias, tornando-se componente essencial para a identificação da existência e do comportamento das unidades de conhecimento especializado” (idem). Ambas as áreas mencionadas contribuem para o desenvolvimento desta pesquisa, sendo que discorreremos mais sobre a fraseologia especializada ao abordarmos as colocações especializadas na seção 3.3.

No que tange a perspectiva textual da Terminologia, Fadanelli (2017) afirma que

a característica mais marcante dos estudos de Terminologia de perspectiva textual é que o enfoque estritamente morfológico – da palavra – e oracional – de seu uso em uma frase - dá lugar ao destaque a estruturas pertinentes ao texto como um todo: recursos de coesão e coerência, fraseologias, macro e microestrutura do texto, sintaxe e semântica do texto, etc, sem desconsiderarem-se, naturalmente, as terminologias. A percepção do termo segue sendo muito importante, mas se dá como um elemento totalmente inserido dentro de um universo textual e vinculado a um todo significativo (FADANELLI, 2017, p. 44).

Amplia-se, dessa forma, a abrangência e a interdisciplinaridade da Terminologia, uma vez que a língua de especialidade perpassa diferentes estruturas linguísticas, como as unidades fraseológicas e os textos. Ainda no que tange à interdisciplinaridade, ressaltamos aquela presente entre LC e Terminologia. De acordo com as definições de termo expostas aqui, os nódulos de buscas extraídos da lista de palavras-chave a partir de nosso *corpus* de estudo, gerada por uma ferramenta criada no âmbito da LC, são classificados como termos, e não como palavras, ou unidades lexicais simples, haja vista que evocam conceito em uma

área de especialidade. Ou seja, os nódulos de busca por nós levantados, são termos, pois, independente do contexto no qual ocorrerem, estarão relacionados diretamente à área da investigação criminal, fazendo com que o leitor/ouvinte faça a relação entre o termo e a área privilegiada nesta pesquisa.

Sendo assim, a seguir apresentamos o aporte teórico da Fraseologia baseada em *corpus*, bem como, discorreremos sobre as colocações especializadas, áreas que apresentam interdisciplinaridade com a LC e com a Terminologia, conforme o exposto nesta seção.

### 3. 2 Fraseologia baseada em *corpus*

De acordo com Granger e Meunier (2008), a Fraseologia

está difundida em todos os campos da linguagem e, ainda assim, - ou talvez precisamente por causa disso - só recentemente se estabeleceu como uma disciplina por direito próprio. Geralmente é apresentada como um subcampo da Lexicologia que lida com o estudo de combinações de palavras, em vez de palavras isoladas<sup>14</sup>. (GRANGER, MEUNIER, 2008, p. XIX).

O surgimento da Fraseologia remonta ao século XVIII, quando Michail Vasilevich Lomonósov (1711-1765) empreendeu uma análise minuciosa, ao incluir parênteses e modismos russos em sua gramática, ressaltando uma semelhança entre a palavra e as frases (MINORESKO, 1997, apud, MONTEIRO-PLANTIN, 2014). Mais tarde, autores como Bréal (1897 [1924]) e Gabelentz (1901) também se dedicaram aos estudos da fraseologia, por meio de fórmulas, locuções, aforismos e sentenças.

Posteriormente, Saussure (1916), em seu Curso de Linguística Geral, afirma que “há, primeiramente, um grande número de expressões que pertencem à língua; são as frases feitas, nas quais o uso proíbe qualquer modificação, mesmo quando seja possível distinguir, pela reflexão, as partes significativas”(SAUSSURE, 1916 [1969], p. 144), de forma que, embora não tenha como objeto de seus estudos as unidades fraseológicas, o linguista genebrino já reconhecia a existência de tais unidades que passaram a constituir o objeto de estudo da Fraseologia.

Os estudos posteriores, empreendidos por autores como Zuluaga (1980), Tristá

---

<sup>14</sup> “*Phraseology is pervasive in all language fields and yet this fact – or perhaps precisely because of it – it has only relatively recently become established as a discipline in its own right. It is usually presented as a subfield of lexicology dealing with the study of word combinations rather than single words.*”

(1988), Corpas Pastor (1996), Iñesta Mena e Pamies Bertrán (2002), buscaram descrever e analisar as unidades fraseológicas, evidenciando as características dessas unidades, bem como, buscando delimitá-las. Esses estudos, assim como outros que seguiram a mesma linha, ficaram conhecidos como sendo pertencentes à Abordagem Tradicional da Fraseologia, à medida em que pesquisas que utilizam *corpora* eletrônicos começaram a surgir na área, dando início à Abordagem Distribucional ou Baseada em Frequência, a qual é caracterizada por se valer da Linguística de Corpus e suas ferramentas como uma forma de busca e identificação das unidades fraseológicas, ficando conhecida, também, como Fraseologia Baseada em *Corpus*.

De acordo com Cowie (1998), a Fraseologia Baseada em *Corpus* teve início com o trabalho dos neo-firthianos, Michael Halliday e John Sinclair, os quais utilizavam *corpora* escritos e falados para empreenderem análises fraseológicas. Já Granger e Paquot (2008) afirmam que John Sinclair foi o responsável pelo surgimento da Fraseologia Baseada em *Corpus* (doravante FBC). Sendo assim, uma vez que o ponto comum para ambos os autores é o trabalho lexicográfico de Sinclair, em 1987, entendemos que este trabalho foi o mais relevante para a área e por isso, nos deteremos sobre ele, em um primeiro momento.

Tal como apontado por Caldas (2017), o trabalho lexicográfico desenvolvido por Sinclair (1991), o *Projeto Cobuild*, ficou conhecido pela utilização da abordagem dirigida por *corpus* (*corpus-driven approach*) e teve como objetivo buscar e identificar as coocorrências lexicais. Além desta busca e identificação das coocorrências lexicais, essa abordagem passou a considerar e analisar sequências como quadros colocacionais e coligações que eram consideradas periféricas na língua.

O autor explica que esse trabalho ganhou destaque por considerar a língua em uso, uma vez que os *corpora* permitiam a busca por unidades léxicas de forma a considerar e analisar, também, o contexto em que tais unidades ocorriam, haja vista que, conforme ressalta o autor, qualquer significado da língua depende dos contextos em que esta está inserida (SINCLAIR, 1991). Por meio da utilização das ferramentas da Linguística de Corpus, o autor passa a considerar o contexto em que as unidades léxicas estão inseridas, observando as *keywords in context* (as palavras-chave em contexto), apresentadas pelas linhas de concordância geradas pelas ferramentas da LC.

Contudo, ainda segundo Caldas (2017), Bertraña e Bertrán (2008) afirmam que os estudos de Sinclair, os quais conciliam Lexicografia e Fraseologia, tiveram início antes do Projeto *Cobuild*, na década de 1960, quando o autor utilizou cálculos estatísticos advindos da LC para o estudo de coocorrência linear de palavras e, posteriormente, em 1974, junto

com Jones, utilizou a frequência de coocorrência como um fator definidor de colocações.

As pesquisas que se seguiram após o trabalho de Sinclair dão continuidade e aprofundam os estudos na área da Fraseologia Baseada em *Corpus*. Autores como Nesselhauf (2005), Groom (2005), Granger e Paquot (2008), Gries (2008), se destacaram internacionalmente ao empreenderem pesquisas que conciliam LC e Fraseologia. No Brasil, temos os trabalhos de Tagnin (2013), Orenha-Ottaiano (2004, 2009, 2021), entre outros.

Nesselhauf (2005), ao empreender um estudo sobre as colocações produzidas por aprendizes de língua inglesa, por meio da utilização de *corpora* de aprendizes, define as colocações a partir da observação do seu *corpus* de pesquisa (definição está que é apresentada na seção 3.3), ressaltando a importância da LC para a Fraseologia.

Por sua vez, Groom (2005) empreende um estudo que busca identificar se e em que medida a fraseologia apresenta variações ou se é consistente em quatro *corpora* compostos por milhões de palavras. Para isso, o autor toma como base os estudos de Huston e Francis (1999), os quais denominam a fraseologia como gramática de padrões, sendo que esta “rejeita a visão tradicional da sintaxe e do léxico como domínios separados e oferece uma alternativa às concepções formalistas da linguagem como uma série de slots gerados sintaticamente que o falante ou o escritor preenche gramaticalmente, itens lexicais semanticamente e pragmaticamente aceitáveis”<sup>15</sup> (GROOM, 2005, p. 258).

Uma vez que a gramática de padrões rejeita a visão tradicional que entende sintaxe e léxico como domínios separados, segundo Groom (2005), essa teoria visualiza “‘frases semi pré-construídas que constituem escolhas únicas’ (SINCLAIR, 1991, p. 110), tanto como a unidade central de significado e estrutura linguística, quanto como o foco principal da análise linguística”<sup>16</sup> (GROOM, 2005, p. 258). Dessa forma, embora receba uma denominação diferente, os estudos empreendidos pela gramática de padrões, desenvolvidos a partir de *corpora*, são, também, estudos fraseológicos, mais especificamente, trabalhos que se valem da FBC.

Granger e Paquot (2008) discutem as fronteiras da Fraseologia com outras disciplinas e afirmam que, quando comparadas, a Abordagem Tradicional e a Abordagem Distribucional ou Baseada em Frequência, a qual chamamos de FBC, a última apresenta fronteiras mais

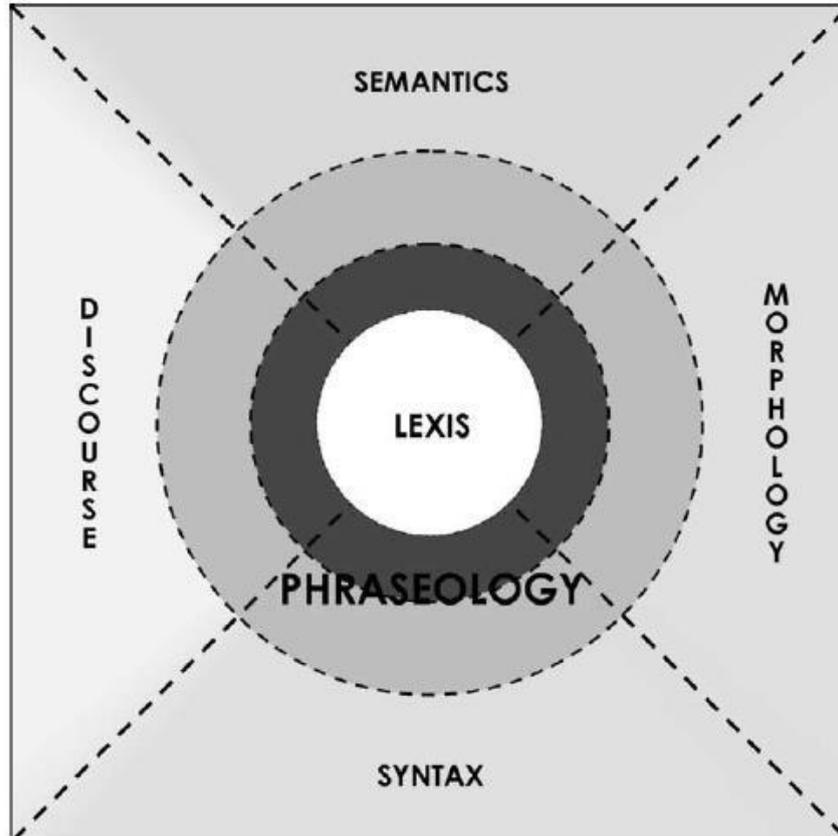
---

<sup>15</sup> “[...] rejects the traditional view of syntax and lexis as separate domains, and offers an alternative to formalist conceptions of language as a series of syntactically generated slots which the speaker or writer fills with grammatically, semantically and pragmatically acceptable lexical items”.

<sup>16</sup> “[...] ‘semi-preconstructed phrases that constitute single choices’ (Sinclair, 1991, p. 110), as both the core unit of linguistic meaning and structure, and as the principal focus of linguistic analysis [...]”.

amplas, conforme pode ser observado na imagem proposta pelas autoras, a qual é reproduzida abaixo:

Figura 4: Fraseologia ampla e estreita<sup>17</sup>.



Fonte: GRANGER; PAQUOT, 2008.

Conforme as autoras explicam, “o território coberto pela abordagem baseada em frequência (representado pelo círculo cinza claro) é muito mais amplo do que a visualização tradicional (em cinza escuro). A Figura 1 destaca a natureza inerentemente multidisciplinar do campo. Como apontado por Mel’cuk (1995: 227), a fraseologia tem que lidar com tudo, o que a torna ‘tão difícil, mas tão atraente!’”<sup>18</sup> (GRANGER; PAQUOT, 2008, p. 29).

Ressalta-se, assim, o caráter multidisciplinar e interdisciplinar da Fraseologia e como essa característica foi expandida ao longo dos estudos empreendidos pela FBC, possibilitando estudos cada vez mais específicos e aprofundados. Nesse sentido, Malá (2020) aponta que,

<sup>17</sup> “*Phraseology wide and narrow*”.

<sup>18</sup> *The territory covered by the frequency-based approach (represented by the light grey circle) is much wider than that of the traditional view (in dark grey). Figure 1 highlights the inherently multidisciplinary nature of the field. As pointed out by Mel’cuk (1995: 227), phraseology has to deal with everything, which makes it ‘so difficult, but so appealing!’*”.

Os métodos de extração e avaliação de dados baseados em *corpus* parecem ser particularmente adequados para a identificação de padrões fraseológicos, uma vez que "são mais indutivos, na medida em que o *corpus* em si são os dados e os padrões de uso da linguagem que ele representa são apontados como formas de expressar regularidades e exceções na linguagem"<sup>19</sup> (Callies 2015: 36 *apud* MALÁ, 2020, p. 77).

Em se tratando da definição da FBC, Malá (2020) retoma as palavras de Gledhill (2000) para enfatizar que a FBC pode ser entendida como “a maneira preferida de dizer as coisas em um discurso particular”<sup>20</sup> (GLEDHILL 2000, *apud* MALÁ, 2020). Esse discurso particular pode ser o discurso empreendido por diferentes comunidades por meio da fala ou da língua escrita e são estudados e analisados por meio de *corpora* escritos e falados.

No Brasil, destacamos o trabalho desenvolvido por Tagnin (2013) que aborda a questão do discurso particular na língua por meio do termo convencionalidade. Segundo a autora, o termo “abrange tudo aquilo que é convencional”, podendo este ser aplicado à língua, “tanto no nível social, isto é, deve-se saber *quando* dizer algo, quanto no nível linguístico, ou seja, saber *como* dizê-lo” (TAGNIN, 2013, p. 21). A autora explica que algumas expressões são convencionais “por estarem intimamente ligadas a um fato social” enquanto em outras “o que é convencional é a sua forma” (*idem*).

Assim, Tagnin (2013) introduz a questão da idiomaticidade, que é percebida quando a “convenção passa para o nível do significado” (TAGNIN, 2013, p. 22). A idiomaticidade, bem como outros aspectos, como a variação, a fixação, e opacidade, são características atribuídas às unidades fraseológicas. Entende-se, assim, que para a autora, a Fraseologia se encontra dentro da convencionalidade linguística, ponto de vista que também é compartilhado por Orenha-Ottaiano (2004).

Uma vez que as unidades fraseológicas estão inseridas na convencionalidade, Tagnin (2013) ressalta que a LC

Oferece uma metodologia que veio facilitar muito a identificação das unidades convencionais da língua. Enquanto, no passado, se dependia de “notar” sua recorrência no dia a dia até nos conscientizarmos de que se tratava de uma unidade fixa, hoje essa conscientização é facilitada pela observação simultânea de uma grande quantidade de dados a partir de um

<sup>19</sup> “*Corpus-driven methods of data extraction and evaluation appear to be particularly well-suited for the identification of phraseological patterns since they “are more inductive, in that the corpus itself is the data and the patterns of language use it represents are noted as ways of expressing regularities and exceptions in language” (Callies 2015: 36)*”.

<sup>20</sup> “[...] *the preferred way of saying things in a particular discourse*”.

*corpus* eletrônico. (TAGNIN, 2013, p. 29).

As unidades fraseológicas (doravante UFs) constituem o objeto de estudo da Fraseologia e, apesar de serem de diferentes tipos, apresentam características comuns que devem ser observadas, em maior ou menor grau, em todos os seus tipos.

Antes de abordar as suas características, convém definir melhor as UFs, haja vista a vasta gama de definições e, até mesmo, de terminologias disponíveis na literatura que se ocupa do objeto de estudos da Fraseologia. Com relação à nomenclatura, Orenha-Ottaiano (2004), aponta que vários autores se ocupam dos fraseologismos (outro nome dado às UFs), fazendo um levantamento dos nomes dados à estas unidades, mencionando, entre eles, os termos: unidades multipalavras (COWIE, 1998), expressões convencionais, unidades frasais (PAWLEY, 2001), blocos pré-fabricados (BOLINGER, 1976), unidades fraseológicas (VINOGRADOV, 1947; AMOSOVA, 1963; CARNEADO E TRISTÁ, s/d, TERMINASOVA, 1992), entre outros (ver ORENHA-OTTAIANO, 2004, p. 15).

Nesta pesquisa, utilizamos o termo unidade fraseológica (UF), uma vez que entendemos, assim como Orenha-Ottaiano (2004), que não há distinção entre as unidades que são abarcadas pelos diferentes termos.

No que tange a definição das UFs, Corpas Pastor (1996) as define como

unidades léxicas formadas por mais de duas palavras gráficas em seu limite inferior, cujo limite superior se situa no nível da oração composta. Estas unidades se caracterizam por sua alta frequência de uso e de coaparição de seus elementos integrantes; por sua institucionalização, entendida em termos de fixação e especialização semântica; por sua idiomaticidade e variação potenciais; assim como pelo grau em que se dão todos estes aspectos nos diferentes tipos<sup>21</sup>. (CORPAS PASTOR, 1996, p.20).

A autora, ao definir as UFs, já elenca as principais características que estas devem conter para serem consideradas como tal. Contudo, desde a publicação desta definição amplamente difundida, outros autores também se empenharam em definir as UFs e como essas se caracterizam, especificamente quando se trata de definir as UFs de acordo com a Fraseologia Baseada em *Corpus*, mencionamos as definições dadas por Greis (2008) e Heid (2008).

---

<sup>21</sup> “[...] unidades léxicas formadas por más de dos palabras gráficas en su límite inferior, cuyo límite superior se sitúa en el nivel de la oración compuesta. Dichas unidades se caracterizan por su alta frecuencia de uso, y de coaparición de sus elementos integrantes; por su institucionalización, entendida en términos de fijación y especialización semántica; por su idiomaticidade y variación potenciales; así como el grado en el cual se dan todos estos aspectos en los distintos tipos.”

Nesse sentido, Gries (2008) acredita que para que uma unidade léxica seja considerada uma UF se deve observar seis critérios, os quais são:

- i. a *natureza* dos elementos envolvidos em um fraseologismo;
- ii. o *número* de elementos envolvidos em um fraseologismo;
- iii. o *número de vezes* que uma expressão deve ser observada antes de contar como fraseologismo;
- iv. a *distância* admissível entre os elementos envolvidos em um fraseologismo;
- v. o grau de *flexibilidade lexical e sintática* dos elementos envolvidos;
- vi. o papel que a *unidade semântica* e a *não composicionalidade* / não previsibilidade semântica desempenham na definição<sup>22</sup>. (GRIES, 2008, p. 4).

Cada um desses critérios contribui para um aspecto da definição de UF proposta pelo autor. Sobre o primeiro, que aborda a questão da natureza dos elementos que compõem a UF, Gries (2008) entende que a UF é uma coocorrência de uma forma ou de um lema lexical ou de um padrão gramatical. Já sobre o segundo critério, o autor afirma que as UFs podem se constituir de apenas dois elementos ou podem incluir um número maior de elementos. Sobre o terceiro critério, aponta que deve se observar uma frequência de ocorrência maior do que a esperada. No que diz respeito à distância admissível entre os elementos constituintes de uma UF, o autor afirma que, apesar de alguns trabalhos considerarem apenas os elementos imediatamente adjacentes, ele reconhece que há UFs que são descontinuas.

No que tange ao quinto critério, Gries (2008) entende que as UFs compreendem padrões lexicais fixos e semifixos, excluindo completamente padrões lexicais inespecíficos e os que são totalmente flexíveis. Sobre o sexto e último critério, o autor entende que, para se considerar uma UF como tal, esta deve conter uma unidade semântica, mas não semântica não-composicional.

Dessa forma, Gries (2008) propõe que os fraseologismos sejam definidos, em suma, como uma “coocorrência de uma forma ou lema de um item lexical e um ou mais elementos linguísticos adicionais de vários tipos que funcionam como uma unidade semântica em uma oração ou frase e cuja frequência de coocorrência é maior do que o esperado com base no acaso<sup>23</sup>” (GRIES, 2008, p. 6).

<sup>22</sup> “i. *the nature of the elements involved in a phraseologism*;  
 ii. *the number of elements involved in a phraseologism*;  
 iii. *the number of times an expression must be observed before it counts as a phraseologism*;  
 iv. *the permissible distance between the elements involved in a phraseologism*;  
 v. *the degree of lexical and syntactic flexibility of the elements involved*;  
 vi. *the role that semantic unity and semantic non-compositionality / non-predictability play in the definition*”

<sup>23</sup> “[...] *co-occurrence of a form or a lemma of a lexical item and one or more additional linguistic elements of*

Heid (2008), por sua vez, usa o termo “expressões multipalavras”<sup>24</sup> (HEID, 2008, p. 339) já que este pode se referir às UFs e, conforme afirma Moon (1998), “não existe um fenômeno unificado para descrever, mas sim um complexo de recursos que interagem de várias maneiras, muitas vezes desordenadas, e representam um amplo continuum entre grupos não composicionais (ou idiomáticos) e composicionais de palavras”<sup>25</sup> (MOON, 1998 apud HEID, 2008, p. 339).

Dito isso, o autor explica que as expressões multipalavras (ou UFs) apresentam relação direta com a noção de palavra que é padrão para a área de processamento natural da língua, a qual entende como palavra “uma linha de caracteres rodeada por espaços”<sup>26</sup> (HEID, 2008, p. 339). Nesse sentido, “qualquer unidade significativa composta por mais de uma linha de caracteres se qualifica como uma expressão multipalavras”<sup>27</sup>. (HEID, 2008, p. 339 – 340).

Tendo discorrido sobre a FBC e UFs, na seção seguinte, apresentamos as colocações especializadas, a UF que constitui o objeto de estudo desta pesquisa.

### 3.3 Colocações especializadas

O termo colocação tem sido amplamente atribuído à Firth (1957), que o utilizou pela primeira vez em seu artigo *Modes of Meaning*, ao explicar os casos de coocorrência lexical. Contudo, conforme ressalta Cowie (1998), Palmer já abordava o termo colocação na década de 1930 e, segundo Orenha-Ottaiano (2020), Jespersen já o fazia em 1917.

De acordo com Cowie (1998), enquanto Palmer estava no Japão trabalhando como professor de língua inglesa, ele desenvolveu um projeto, o qual passou a contar com o auxílio de Hornby posteriormente, visando identificar uma ampla variedade de unidades multipalavras, tendo utilizado o termo colocação em seu Segundo Relatório Provisório de Colocações em Inglês, de 1933, para se referir às combinações de palavras que identificou

---

*various kinds which functions as one semantic unit in a clause or sentence and whose frequency of cooccurrence is larger than expected on the basis of chance”.*

<sup>24</sup> “[...] multi-word expressions”.

<sup>25</sup> “[...] there is no unified phenomenon to describe, but rather a complex of features that interact in various, often untidy ways and represent a broad continuum between non-compositional (or idiomatic) and compositional groups of words”.

<sup>26</sup> “[...] a string of characters surrounded by blankers”.

<sup>27</sup> “[...] any meaningful unit made up of more than one string of characters qualifies as an MWE”.

na fala e na escrita. Nesse relatório, Palmer buscou enfatizar as consequências que a utilização das colocações desempenharia no ensino-aprendizagem de língua inglesa. Para o autor, “uma colocação é uma sucessão de duas ou mais palavras que devem ser aprendidas como um todo integral, e não reunidas a partir de suas partes componentes”<sup>28</sup> (PALMER, 1933, apud COWIE, 1998, p. 211).

Outros autores que também desenvolveram e desenvolvem pesquisas na área das colocações são Hausmann (1984), Heid et. al (1991), Sinclair (1991), sendo que outros são responsáveis por pesquisas que abordem as colocações especializadas como objeto de estudo, tais como L’Homme (2000), Bevilacqua (2004, 2005), Orenha-Ottaiano (2009, 2012a, 2020).

Para Hausmann (1984 apud ORENHA-OTTAIANO, 2009), as colocações são produtos semiacabados de uma língua, fato que faz com que estas sejam objeto de interesse de pesquisadores, de professores e alunos de idiomas, haja vista que, conforme ressalta o autor, os usuários da língua apenas reutilizam-nas.

Ainda segundo o autor, os dois elementos que compõem as colocações recebem o nome de base e colocado, sendo esta classificação determinada por aquilo que o autor chama de hierarquia, já que um dos itens é responsável por determinar (a base), sendo totalmente autônomo, e o outro é determinado (o colocado) pelo primeiro item (ORENHA-OTTAIANO, 2009, p. 39).

Por sua vez, Sinclair (1991) utiliza as denominações *node* e *collocate* para se referir ao que Hausmann (1984) chama de base e colocado, sendo que a base ora corresponderá ao *node*, ora corresponderá ao *collocate*, o que dependerá da palavra que está em busca (ORENHA-OTTAIANO, 2009). Por exemplo, se a palavra de busca for *gunshot* (tiro) e a colocação em análise for *gunshot residue* (resíduo de tiro), ela (a palavra de busca) corresponderá ao *node* de Sinclair e a base de Hausmann; entretanto, como em *gunshot wound* (ferida de tiro) ela seria o *node* de Sinclair, enquanto seria o colocado para Hausmann.

Considerando, assim, o fato de que as colocações são formadas por uma base e um colocado, Hausmann (1985) propõe uma classificação taxonômica para as colocações, que também foi utilizada por Orenha-Ottaiano (2004, 2009), de forma que seja possível identificar colocações nominais, verbais, adjetivas e adverbiais, a qual é apresentada abaixo, com exemplos extraídos do *corpus* de estudo.

---

<sup>28</sup> “[...] A collocation is a succession of two or more words that must be learned as an integral whole, and not pieced together from its component parts”.

### Verbais

- Substantivo **base** + Verbo **colocado** = *unsub targets*
- Verbo **base** + Determinante + Substantivo **colocado** = *verify an alibi*
- Verbo **colocado** + Determinante + Substantivo **base** = *change the victimology*

### Nominais

- Substantivo **base**+ Substantivo **colocado** = *gunshot wound*
- Substantivo **colocado**+ Preposição + Substantivo **base** = *evidence in autopsy*

### Adjetiva

- Adjetivo **colocado**+ Substantivo **base** = *similar victimology*

### Adverbial

- Verbo **base**+ Advérbio **colocado** = *shot directly*
- Advérbio **colocado**+ Verbo **base** = *easy kill*

Corpas Pastor (1996), por sua vez, afirma que a noção de colocação é “esboçada na teoria semântica de Coseriu (1966, 1978, 1986 [1977]), concretamente dentro das denominadas *solidariedades léxicas*”<sup>29</sup> (CORPAS PASTOR, 1996, p. 63) e as denomina como

[...] as combinações assim resultantes, ou seja, as unidades fraseológicas formadas por duas unidades léxicas em relação sintática, que não constituem, por si mesmas, atos de fala nem enunciados; e que, devido a sua fixação na norma, apresentam restrições de combinação estabelecidas pelo uso, geralmente de base semântica: o colocado autônomo semanticamente (a base) não só determina a eleição do colocativo, como também, seleciona neste uma aceção especial, frequentemente de caráter abstrato e figurativo.<sup>30</sup> (CORPAS PASTOR, 1996, p. 66)

Ademais, para a autora, uma colocação é entendida como “aquela propriedade pela qual os falantes das línguas tendem a produzir certas combinações de palavras entre

<sup>29</sup> “[...] aparece ya esbozada en la teoría semántica de Coseriu (1966, 1978, 1986 [1977]), concretamente dentro de las denominadas *solidariedades léxicas*”.

<sup>30</sup> “[...] a las combinaciones así resultantes, es decir, a las unidades fraseológicas formadas por dos unidades léxicas en relación sintática, que no constituyen, por sí mismas, actos de habla ni enunciados; y que, debido a su fijación en la norma, presentan restricciones de combinación establecidas por el uso, generalmente de base semántica: el colocado autónomo semanticamente (la base) no sólo determina la elección del colocativo, sino que, además, selecciona en éste una acepción especial, frecuentemente de carácter abstracto o figurativo.”

uma grande quantidade de combinações teoricamente possíveis”<sup>31</sup> (CORPAS PASTOR, 1996, p, 66).

Para Orenha-Ottaiano (2020), as colocações são

combinações ubíquas, recorrentes, arbitrárias e convencionalizadas, que são lexicamente e / ou sintaticamente fixas em um certo grau e podem ter um alcance colocacional (mais ou menos) restrito. São combinações específicas de uma língua e de uma cultura e, como tal, a colocabilidade de seus elementos pode variar significativamente de um idioma para outro, sendo, assim, constituídas por sua própria rede de colocações<sup>32</sup> (ORENHA-OTTAIANO, 2020, p. 62).

Assim como a autora, entendemos que as colocações são combinações de palavras que apresentam como características principais a arbitrariedade, a recorrência, certo grau de fixação (uma vez que umas podem apresentar uma fixação maior que outras) e que são convencionalizadas nos mais diferentes idiomas e culturas, bem como nas mais diferentes áreas de especialidades de um idioma, recebendo a denominação de colocações especializadas quando utilizadas nestes contextos.

Nesta pesquisa, nos deteremos mais sobre os aspectos concernentes às colocações especializadas, uma vez que estas configuram nosso objeto de estudo. Nesse sentido, Bevilacqua (2004, 2005) afirma que há uma variedade de denominações atribuídas às colocações especializadas, haja vista que alguns autores as tratam como unidades fraseológicas, outros como colocações. Contudo, a autora opta por denominá-las Unidades Fraseológicas Especializadas, as quais são definidas como:

Unidades formadas por um núcleo eventivo, considerado como tal por ser de base verbal ou derivada de verbo (nominalização ou particípio), e por um núcleo terminológico (termo). Entre estes dois núcleos se estabelecem relações sintáticas, mas especialmente semânticas, determinadas pelas propriedades do texto em que são utilizadas. Portanto, são unidades que se conformam no e pelo texto em que são utilizadas. Cumprem, tal como os termos, a função de representar e transmitir conhecimento especializado. (BEVILACQUA, 2005, p. 83).

---

<sup>31</sup> “[...] aquella propiedad de las lenguas por la que los hablantes tienden a producir ciertas combinaciones de palabras entre una gran cantidad de combinaciones posibles”.

<sup>32</sup> “[...] pervasive, recurrent, arbitrary and conventionalized combinations, which are lexically and/or syntactically fixed to a certain degree and may have a (more or less) restricted collocational range. They are a language and a culture’s specific combinations and, as such, the collocability of their elements may vary significantly from a language to another, being thus made up of their own collocations network”.

Levando em consideração o destaque dado ao termo presente na colocação especializada por Bevilacqua (2005), retomamos Orenha-Ottaiano (2009, 2012a), quem nos diz que a principal diferença entre as colocações especializadas e as colocações da língua geral consiste no fato de a base das colocações da língua geral ser constituída por uma unidade lexical pertencente à própria língua geral, enquanto que a base das colocações especializadas é composta por uma unidade terminológica de uma língua de especialidade.

Dessa forma, Orenha-Ottaiano (2012a), baseada em L’Homme (2000), afirma que, assim como as colocações da língua geral são definidas por meio da convenção estabelecida pelos falantes de uma dada língua, o mesmo se observa no caso das colocações especializadas, as quais são determinadas pela convenção acordada por um determinado grupo de falantes de uma dada especialidade, fato que corrobora a presença de um termo de uma área específica, na formação da colocação especializada.

Ademais, Orenha-Ottaiano (2012a), ao retomar L’Homme (2000), afirma que

as colocações especializadas podem ser melhor descritas em termos de coocorrência lexical livre (L’HOMME, 2000, p. 90), e isso pode ser verificado por meio de diferentes propriedades: 1) pelo fato de a não composicionalidade não ser um critério predominante para a identificação de uma combinação lexical especializada; e 2) em razão de ser altamente produtiva, a definição dos grupos de termos semanticamente relacionados está associada a coocorrentes. (ORENHA-OTTAIANO, 2012a, p. 161).

Evidencia-se, assim, que para Orenha-Ottaiano (2012a) as colocações especializadas apresentam mais semelhanças que diferenças com as colocações da língua geral, como são comumente denominadas, haja vista a sua não composicionalidade e sua característica de coocorrência lexical livre.

Patiño (2014), ao empreender um estudo buscando definir as colocações especializadas, afirma que estas podem ser descritas como um

tipo de expressão com várias palavras compostas de pelo menos um termo que serve como o nódulo. Seus colocados podem ser substantivos, verbos, adjetivos ou advérbios em uma relação sintática direta com o nódulo. Esses constituintes formam uma combinação lexical que pode ser imprevisível e semicomposicional e têm tendência de preferência interna e estatística. Esta definição implica que os constituintes de uma colocação especializada não são necessariamente adjacentes um ao outro.

[...] Qualquer termo, seja ele composto de um ou mais lexemas, pode entrar em uma colocação especializada com um conjunto restrito de outros substantivos, verbos, adjetivos e advérbios<sup>33</sup>. (PATIÑO, 2014, p.125-126).

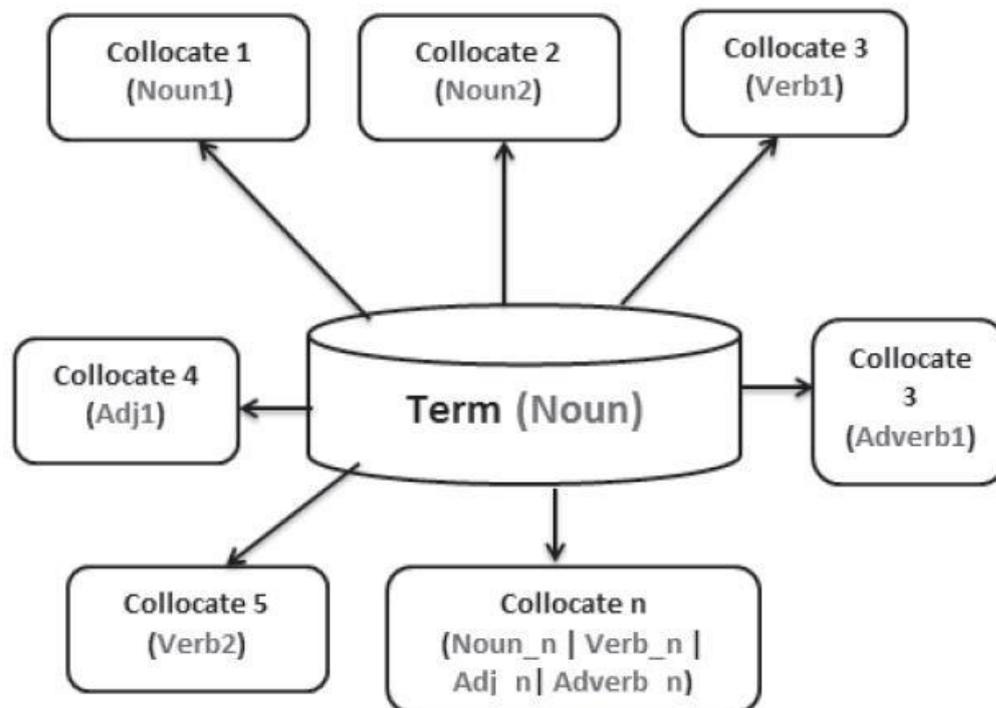
---

<sup>33</sup> “[...] type of multiword expression composed of at least a term that serves as the node. Its collocates can be

De acordo com essa definição do autor, o termo que atua como base da colocação (o qual ele denomina como nódulo) pode ser tanto um substantivo, quanto um verbo, ou um adjetivo ou, ainda, um advérbio, podendo, da mesma forma, ter como colocado uma lexia que se enquadre em qualquer uma destas classes gramaticais, não havendo, assim, restrições a bases verbais ou derivadas de verbos, conforme aponta Bevilacqua (2004).

O autor apresenta, ainda, uma imagem que tem como objetivo representar a noção de colocação especializada por ele defendida, a qual reproduzimos a seguir.

Figura 5: Diagrama representando as colocações especializadas.



Fonte: PATIÑO, 2014.

Dessa forma, a partir do exposto aqui, entendemos que as colocações especializadas são combinações de palavras recorrentes, arbitrárias e convencionalizadas em uma dada área de especialidade, compostas por duas ou mais palavras, denominadas de base e colocado, sendo a base um termo, podendo este ser um substantivo, um verbo, um adjetivo ou um

---

*nouns, verbs, adjectives or adverbs in a direct syntactic relation with the node. These constituents make a lexical combination that can be unpredictable and semicompositional and have an internal and statistical tendency of preference. This definition implies that the constituents of a specialized collocation are not necessarily adjacent to each other.*

*[...] Any term, whether it is composed of one or more lexemes, may enter into a specialized collocation with a restricted set of other nouns, verbs, adjectives and adverbs”.*

advérbio, que está sob análise e que determina os seus colocados de acordo com a área de especialidade em questão.

Na seção seguinte, apresentamos os pressupostos teóricos da Fraseografia e da Teoria Funcional da Lexicografia, disciplinas que contribuem para o desenvolvimento da proposta de dicionário de colocações apresentada como produto final deste trabalho.

### **3. 4 Fraseografia e Teoria Funcional da Lexicografia**

Nesta seção, discorreremos sobre os pressupostos teóricos de duas disciplinas que apresentam uma interdisciplinaridade entre si e que são fundamentais para o desenvolvimento de nossa proposta de dicionário pedagógico de colocações especializadas: a Fraseografia e a Teoria Funcional da Lexicografia. Ao abordarmos a Fraseografia, também discorreremos sobre a Teoria Funcional da Lexicografia, uma vez que esta teoria é apontada por Olímpio de Oliveira (2016) como fundamental para o desenvolvimento das pesquisas na área de Fraseografia.

#### 3.4.1 Fraseografia

O termo Fraseografia surgiu há mais de duas décadas na linguística soviética, com o intuito de designar a subárea da lexicografia responsável pelos aspectos teóricos e práticos envolvidos na produção de dicionários fraseológicos, tendo sido introduzido na linguística espanhola por Carneado Moré e Tristán Pérez. Contudo, o termo se firmou na linguística espanhola a partir do final dos anos 1990 e início dos anos 2000 (OLÍMPIO DE OLIVEIRA, 2007).

A Fraseografia, hoje, é considerada uma disciplina independente, embora tenha surgido e se mantido, durante muito tempo, entre a Fraseologia e a Lexicografia, sendo considerada um fruto da relação indiscutível existente entre ambas as áreas, conforme aponta Olímpio de Oliveira (2007). Ainda segundo a autora, foi Carneado Moré (1989) a responsável por empreender um estudo global no qual buscava estabelecer as bases históricas da Fraseografia e que culminou no apontamento de três etapas principais sobre o desenvolvimento da área.

A primeira etapa dos estudos fraseográficos, que compreende desde a Antiguidade até a primeira metade do século XIX, se caracteriza pela “elaboração de obras que recolhem

unidades de cunho popular, como provérbios e refrões surgidos em épocas muito antigas”<sup>34</sup> (OLÍMPIO DE OLIVEIRA, 2007, p. 29). Já na segunda etapa do estudo, que se estende desde a segunda metade do século XIX até a primeira metade do século XX, “além de compilar provérbios e refrões, registram-se frases proverbiais e expressões idiomáticas”<sup>35</sup> (OLÍMPIO DE OLIVEIRA, 2007, p. 30), havendo, também, nesta etapa, a utilização de critérios linguísticos na explicação de características de significado, etimológicas e estilísticas das unidades selecionadas.

Já a terceira etapa da Fraseografia, que tem início dos anos 1950 até a atualidade, a qual decorre do grande desenvolvimento da Lexicografia e da Fraseologia, traz à luz obras baseadas em “estudos práticos e teóricos que estimulam a utilização de critérios científicos na hora de selecionar, ordenar e apresentar material fraseológico nos dicionários”<sup>36</sup> (OLÍMPIO DE OLIVEIRA, 2007, p. 30-31). Esta etapa ocorre após a Segunda Guerra Mundial, período em que se observou um aumento significativo sobre o ensino-aprendizagem de idiomas.

Segundo Carneado Moré (1989), as obras publicadas nesta etapa da Fraseografia são responsáveis por marcar o nascimento dos dicionários fraseológicos, uma vez que

incluem prólogos onde se analisam tanto os problemas práticos relacionados com a compilação de dicionários plurilíngues, como os problemas teóricos da fraseologia e da fraseografia. Os autores destes dicionários fundamentam sua concepção de fraseologia, especificam a composição do Dicionário, elaboram princípios de ordenação do material fraseológico e da apresentação do material fático no texto. Além disso, incluem anotações estilísticas, exemplos ilustrativos, etc”<sup>37</sup> (CARNEADO MORÉ, 1989, apud OLÍMPIO DE OLIVEIRA, 2007, p. 31).

Como explanado pela pesquisa de Carneado Moré (1989) e apresentado por Olímpio de Oliveira (2007), o surgimento da Fraseografia está atrelado ao desenvolvimento das

---

<sup>34</sup> “[...] elaboración de obras que recogen unidades de cuño popular, como proverbios y refranes surgidos en épocas muy antiguas”.

<sup>35</sup> “[...] además de compilarse proverbios y refranes, se registran frases proverbiales y expresiones idiomáticas”.

<sup>36</sup> “[...] estudios prácticos y teóricos que estimulan la utilización de criterios científicos a la hora de seleccionar, ordenar y presentar el material fraseológico en los diccionarios”.

<sup>37</sup> “incluyen prólogos donde se analizan tanto los problemas prácticos relacionados con la compilación de diccionarios plurilingües, como los problemas teóricos de la fraseología y la fraseografía. Los autores de estos diccionarios fundamentan su concepción de la fraseología, precisan la composición del Diccionario, elaboran los principios de ordenamiento del material fraseológico y la presentación del material fático en el texto. Además, incluyen acotaciones estilísticas, ejemplos ilustrativos, etc”.

pesquisas lexicográficas e fraseológicas, principalmente, a partir da Segunda Guerra Mundial, quando se deu, também, o aprimoramento de pesquisas na área de ensino de idiomas.

No que tange à Fraseografia em língua portuguesa, mais especificamente no Brasil, considera-se que a primeira obra fraseográfica publicada em nosso país tenha sido a *Colleção de provérbios, adágios, rifaos. Anexins, sentenças Moraes e idiotismos da língua portuguesa*, datada de 1848, elaborada por Paulo Perestello da Câmara (OLÍMPIO DE OLIVEIRA, 2007).

Contudo, conforme aponta Xatara (2012 *apud* MIRANDA, 2014), entre 1900 a 1970, apenas três obras fraseográficas foram publicadas no país, uma abordando o latim e outras duas as variantes ibérica e americana da língua portuguesa, sendo que as obras bilíngues deste tipo surgem apenas na década de 1980.

Neste sentido, algumas definições foram atribuídas à Fraseografia. Segundo Carneado Moré (1985) “a fraseografia se ocupa da elaboração dos métodos teóricos e práticos e dos princípios para confeccionar dicionários fraseológicos, assim como de analisar e classificar o caudal fraseológico nos dicionários”<sup>38</sup> (CARNEADO MORÉ, 1985 *apud* OLÍMPIO DE OLIVEIRA, 2007, p. 22).

Tristá Pérez (1998) afirma que “como consequência natural do desenvolvimento alcançado na esfera da fraseologia, surgiu na União Soviética um novo ramo da lexicografia que investiga os problemas teóricos e práticos que dizem respeito à elaboração de dicionários fraseológicos: a fraseografia”<sup>39</sup> (TRISTÁ PÉREZ, 1998, p. 297).

No trecho apresentado, a autora, além de retomar o surgimento da Fraseografia e de enfatizar que esse ocorreu na União Soviética, define a referida disciplina como sendo responsável por investigar os problemas teóricos e práticos relacionados à elaboração de dicionários fraseológicos, deixando de lado a análise do tratamento dado às UFs nos dicionários gerais.

Nesse sentido, Olímpio de Oliveira (2007) propõe uma definição mais abrangente para a disciplina em questão. Segundo a autora,

---

<sup>38</sup> “[...] *La fraseografía se ocupa de la elaboración de los métodos teóricos y prácticos y de los principios para confeccionar diccionarios fraseológicos, así como de analizar y clasificar el caudal fraseológico en los diccionarios*”.

<sup>39</sup> “[...] *como consecuencia natural del desarrollo alcanzados en la esfera de la fraseología, surgió en la Unión Soviética una nueva rama de la lexicografía que investiga los problemas teóricos y prácticos que plantea la confección de los diccionarios fraseológicos: la fraseografía*”.

a fraseografia é uma disciplina linguística que se ocupa, por uma parte, dos princípios teóricos e práticos que regem a inclusão da fraseologia em compilações léxicas (dicionários, léxicos, vocabulários, glossários, concordâncias, etc), tanto restringidas como gerais e, por outra, do estudo crítico e descritivo destas compilações, no que se refere ao tratamento da fraseologia, o que significa dizer que o âmbito de interesse da fraseografia compreende desde a apresentação tipográfica seguida na obra até a sua adequação aos usuários”<sup>40</sup> (OLÍMPIO DE OLIVEIRA, 2007, p. 27).

A partir dessa definição, podemos perceber que a Fraseografia, além de ocupar-se da análise e do tratamento dado às unidades fraseológicas, também dá conta da compilação das obras lexicográficas que abarcam essas unidades, ou seja, se pode falar de uma Fraseografia Teórica ou Metafraseografia, responsável por estudos que abarcam a história, a crítica, a investigação e a teoria fraseográficas, e de uma Fraseografia Prática, responsável pela elaboração de dicionários fraseológicos e por estabelecer a técnica e a metodologia que devem se observar para tal (OLÍMPIO DE OLIVEIRA, 2007, p. 27-28).

Dessa forma, segundo Olímpio de Oliveira (2007), cabe ao fraseólogo, além do que foi mencionado anteriormente em relação aos dicionários fraseológicos, empreender estudos sobre a utilização das UFs nos dicionários gerais, analisar o processamento dessas informações, bem como determinar a metodologia que será utilizada para introduzir tais unidades nesses dicionários.

De acordo com essa disciplina, quando se considera o trabalho do fraseólogo voltado para um dicionário geral, existem duas vertentes que interferem na determinação das UFs que integrarão a obra em questão: uma que diz respeito à seleção das UFs antes das demais unidades léxicas e outra que diz respeito à seleção de tipos de UFs dentro da própria classe de fraseologismos (OLÍMPIO DE OLIVEIRA, 2007). Contudo, quando se trata de uma obra fraseográfica propriamente dita, não existem problemas teóricos que devam ser enfrentados, pois “uma vez que se decide o tipo de unidade que se irá pesquisar, deve-se apenas atentar-se estritamente às determinações iniciais e proceder à coleta do material”<sup>41</sup> (OLÍMPIO DE OLIVEIRA, 2007, p. 59).

---

<sup>40</sup> “[...] la fraseografía es una disciplina lingüística que se ocupa, por una parte, de los principios teóricos y prácticos que rigen la inclusión de la fraseología en compilaciones léxicas (diccionarios, léxicos, vocabularios, glosarios, concordancias, etc.), tanto restringidas como generales y, por otra, del estudio crítico y descriptivo de estas compilaciones, en lo que al tratamiento de la fraseología se refiere, lo que significa decir que el ámbito de interés de la fraseografía comprende desde la presentación tipográfica seguida en la obra hasta la adecuación a los usuarios”.

<sup>41</sup> “Una vez que se decide el tipo de unidad que se va a recoger, sólo queda atenderse estrictamente a las determinaciones iniciales y proceder a la recolección del material”.

No que tange à seleção de UFs para um dicionário fraseológico, Olímpio de Oliveira (2007) explica que Hanks (2001), tomando como base um dicionário de expressões idiomáticas, apontou alguns fatores que podem interferir nesse processo, tais como a política editorial e os usuários a quem se destina a obra em questão. Para o autor, estas questões devem ser consideradas devido à baixa frequência das UFS que seriam inseridas no dicionário de expressões idiomáticas que tomou como exemplo.

Para a autora, fatores como o conhecimento fraseológico do autor ou pesquisador e o estado em que se encontra o estudo devem ser levados em consideração no momento da elaboração de uma obra fraseográfica (OLÍMPIO DE OLIVEIRA, 2007).

Outro aspecto importante abordado por Olímpio de Oliveira (2016) é o caráter bilíngue da Fraseografia, o qual, ao ser considerado a partir de um ponto de vista tradicional, é entendido como “uma subárea da fraseografia, a qual inclui todos os trabalhos relacionados aos dicionários bilíngues e dicionários bilíngues fraseológicos”<sup>42</sup> (p. 577).

Sobre essa subárea da Fraseografia, a autora afirma que

como um ramo específico da fraseografia, ela deve incluir, em primeiro lugar, uma perspectiva metafraseográfica (de historiografia, teoria, crítica e pesquisa). Ademais, deve conter um aspecto prático com foco na compilação de dicionários fraseológicos bilíngues (DFB) (incluindo aspectos técnicos e metodológicos), bem como tudo que pertença ao tratamento de elementos fraseológicos em qualquer tipo de repertório léxico<sup>43</sup> (OLÍMPIO DE OLIVEIRA, 2016, p. 577).

A partir do exposto, e com base na definição apresentada acima, entendemos que este trabalho se filia a essa subárea da Fraseografia, uma vez que tem como objetivo final propor um dicionário pedagógico de colocações especializadas em inglês, português e espanhol. Embora esse seja, de fato, um dicionário trilíngue ou multilíngue, e não bilíngue, a própria autora enfatiza a existência e importância da fraseografia multilíngue, uma vez que este “campo despertou interesse considerável e tem sido assunto de uma série de artigos (BERTHEMET, 2009, 2012; DAWES, 2014), motivo pelo qual não deve ser deixado de fora desta área da fraseografia”<sup>44</sup> (OLÍMPIO DE OLIVEIRA, 2016, p. 578).

---

<sup>42</sup> “[...] a sub-area of phraseography which includes all the phraseographical work regarding general bilingual dictionaries and bilingual phraseological dictionaries”.

<sup>43</sup> “As a specific branch of phraseography, it should include, in the first place, a metafraseographical perspective (of historiography, theory, criticism and research). Furthermore, it should contain a practical aspect focusing on the compilation of bilingual phraseological dictionaries (BPD) (including technical and methodological aspects), as well as everything pertaining to the treatment of phraseological elements in any type of bilingual lexical repertoire.”

Ao analisarmos os trabalhos de Berthemet (2009, 2012), por exemplo, percebemos um interesse pelo caráter multilíngue, uma vez que a autora propôs um dicionário fraseológico que contemplava as UFs em quatro diferentes idiomas: inglês, francês, alemão e russo. Em seu trabalho de 2012, a autora propôs a construção de uma base de dados fraseológica multilíngue, contemplando os mesmos idiomas com os quais trabalhou em 2009. Essa base de dados foi desenvolvida para ser de domínio público, no formato de um aplicativo disponibilizado na web, possibilitando a pesquisa multidirecional, tanto semasiológica quanto onomasiológica das UFs.

Ademais, Olímpio de Oliveira (2016) defende que para o desenvolvimento dos estudos na área da Fraseografia Bilíngue, a Teoria Funcional da Lexicografia deve ser levada em consideração. Sendo assim, na seção seguinte, abordamos melhor essa teoria.

### 3.4.2 Teoria Funcional da Lexicografia

A Teoria Funcional da Lexicografia surgiu no ano de 1992, com a tese de doutorado de Tarp e se tornou mais conhecida por meio dos trabalhos de Bergenholtz e Tarp (1995). Contudo, conforme o próprio autor afirma, essa teoria não teve destaque até 2001, quando Wiegand fez uma crítica a ela, fato que levou Bergenholtz e Tarp a revisitarem e reformularem seus escritos (TARP; BERGENHOLTZ, 2002, 2003, 2004). Posteriormente, com o avanço dos recursos tecnológicos e digitais, a Teoria Funcional da Lexicografia passou por novo desenvolvimento, que levou às publicações de Fuertes-Olivera e Tarp (2014) e Tarp (2014).

De acordo com Tarp (2014, p.65), para a Teoria Funcional da Lexicografia (doravante TFL)

os dicionários e demais obras lexicográficas são ferramentas de informação que podem ter usuários potenciais em diversos contextos e situações, as quais, em última instância, determinam as características de um projeto de dicionário. A teoria funcional, além disso, considera que a lexicografia é uma disciplina independente com seu próprio sistema de teorias, métodos, etc., disciplina que, ao mesmo tempo, tem uma grande vocação interdisciplinar que, ao longo da história, tem se relacionado com quase todas as atividades e áreas do pensamento do ser humano<sup>44</sup> (TARP 2014, p. 65)

<sup>44</sup> “[...] field has aroused considerable interest and has been the subject of study in a series of recent articles (Berthemet 2009; 2012; Dawes 2014), for which reason it must not be left out of the field of phraseography”.

<sup>45</sup> “[...] los diccionarios y demás obras lexicográficas son herramientas de información, por lo que deben ser las necesidades de información que puedan tener usuarios potenciales en diversos contextos y situaciones, las

Dessa forma, o cerne desta teoria reside no fato de se ter em mente o usuário potencial ao qual o dicionário que se pretende elaborar é destinado, a fim de que as necessidades desse usuário sejam atendidas pela obra lexicográfica. Nesse sentido, a TFL considera fundamental que

[...] qualquer tipo de dicionário é uma ferramenta de uso concebida para ser consultada de forma rápida e fácil, com o fim de satisfazer necessidades pontuais de informação relacionadas a qualquer ciência, disciplina e área de atividade humana, e que tem tipos específicos de usuários que se encontram em tipos específicos de situações extralxicográficas.

As necessidades de informação dos usuários se resolvem dando acesso a dados lexicográficos explicitamente preparados, dos quais os usuários podem extrair a informação requerida que logo pode empregar para um sem fim de propósitos relacionados com as situações extralxicográficas nas quais as necessidades originalmente se produziram<sup>46</sup> (TARP, 2015, p. 34).

A partir do exposto, entende-se que as necessidades lexicográficas dos usuários são atendidas pelos dicionários e que estas, por sua vez, colaboram para que as necessidades extralxicográficas possam ser sanadas, independente de qual seja a situação extralxicográfica na qual o usuário / consulente se encontre.

Para que seja possível atender as necessidades lexicográficas dos usuários é preciso, primeiramente, definir quem é esse usuário. De acordo com Fuertes-Olivera e Tarp (2014), o primeiro passo para a delimitação do usuário de uma obra lexicográfica é a divisão feita entre as necessidades desse usuário e outros tipos de necessidades que as pessoas possam ter.

Nesse sentido, os autores recorrem à Hausmann (1977), quem distingue entre necessidades globais e pontuais dos usuários, para afirmarem que as obras lexicográficas devem se concentrar nas necessidades pontuais dos usuários, uma vez que essa dualidade deve ser entendida como uma relação entre o todo e a parte. Em outras palavras, um

---

*que, en última instancia, determinen las características de un proyecto de diccionario. La teoría funcional, además, considera que la lexicografía es una disciplina independiente con su propio sistema de teorías, métodos, etc., disciplina que, al mismo tiempo, tiene una gran vocación interdisciplinaria que, a lo largo de la historia, la ha relacionado con casi todas las actividades y áreas de pensamiento del ser humano; cf. Tarp (2008) y Fuertes Olivera & Tarp (2014)*”.

<sup>46</sup> “[...] Cualquier tipo de diccionario es una herramienta de uso concebida para ser consultada de forma rápida y fácil con el fin de satisfacer necesidades puntuales de información relacionadas con cualquier ciencia, disciplina y área de actividad humana, y que tienen tipos específicos de usuarios que se encuentran en tipos específicos de situaciones extra-lexicográficas.

*Las necesidades de información de los usuarios se resuelven dando acceso a datos lexicográficos explicitamente preparados, de los cuales los usuarios pueden extraer la información requerida que luego pueden emplear para un sinfín de propósitos relacionados con las situaciones extra-lexicográficas en las que las necesidades originalmente se produjeron*”.

dicionário especializado da área do direito em economia, por exemplo, contempla apenas uma parte do todo, ou seja, uma parte da grande área do direito.

As necessidades pontuais do usuário dizem respeito, assim, às informações que devem ser incluídas na obra em desenvolvimento. Além dessas informações, existem dois fatores fundamentais sobre o estabelecimento das necessidades do usuário, segundo Fuertes-Olivera e Tarp (2014): “as características da pessoa concreta que experiencia essas necessidades”<sup>47</sup> e a “situação social ou o contexto no qual essas necessidades acontecem”<sup>48</sup> (FUERTES-OLIVERA; TARP, 2014, p. 30).

No que tange a esses fatores, os autores explicam que saber as características físicas de uma ou mais pessoas que consultaram o dicionário não influenciará nas decisões lexicográficas, mas, saber as características de desenvolvimento social e econômico que afetam essas pessoas poderá influenciar, por exemplo, na escolha do veículo dessa obra lexicográfica. Para ilustrar melhor esses fatores, os autores recorrem ao exemplo de dois estudantes de inglês como língua estrangeira que apresentam o mesmo nível de proficiência, mas que, devido à suas condições sociais e econômicas – um é dinamarquês e outro é namibiano – o veículo da mesma obra lexicográfica deverá ser diferente, já que o primeiro poderá acessá-la por meio da internet, enquanto o segundo, devido ao fato de não ter acesso sequer a energia elétrica, deverá consultá-la de forma impressa (FUERTES-OLIVERA; TARP, 2014).

Assim, os autores propõem uma série de perguntas que devem ser respondidas para que se possa determinar as características dos usuários:

- Características do usuário relevantes para a função:
- Qual é a língua materna ou primeira do usuário?
- Qual é o nível de proficiência em língua materna do usuário?
- Com qual método o usuário aprendeu a língua materna ou primeira? Qual é o nível de proficiência do usuário em segunda, terceira, etc. língua?
- Com qual método o usuário aprendeu a segunda, terceira, etc. língua? Qual é o nível do usuário em cultura geral e enciclopédica?
- Qual é a experiência do usuário em tradução entre um conjunto específico de línguas?
- Qual é conhecimento do usuário em uma disciplina específica (leigo, semiespecialista, especialista)?
- Qual é o conhecimento comparado do usuário de uma disciplina cultural específica, conforme expresso em duas áreas específicas do idioma?
- Qual é o nível de proficiência do usuário em uma determinada língua de especialidade?
- Qual é a experiência do usuário na tradução entre um conjunto

<sup>47</sup> “[...] characteristics of the concrete person [...]”.

<sup>48</sup> “[...] social situation or context where the needs occur.”

específico de línguas de especialidade?  
Etc.... etc....<sup>49</sup> (FUERTES-OLIVERA; TARP, 2014, p. 30)

Após apresentar as perguntas que servem como guia para a definição das características dos usuários, os autores apresentam as perguntas que devem ser feitas ao se definir as características que são relevantes para a consulta da obra lexicográfica por parte do usuário. São elas:

- Características do usuário relevantes para a consulta:

Qual é a experiência do usuário em consultas lexicográficas?  
O usuário é cego, surdo ou sofre de qualquer outra deficiência que possa limitar o uso de tipos específicos de ferramentas lexicográficas? O usuário tem eletricidade e luz elétrica?  
O usuário possui dispositivo com acesso à Internet? O usuário sabe distinguir entre direita e esquerda?  
Etc.... etc....<sup>50</sup>(FUERTES-OLIVERA; TARP, 2014, p. 30)

Como se pode observar, são várias as perguntas e os aspectos que devem ser considerados ao iniciar a elaboração de uma obra lexicográfica e, embora essas perguntas sirvam como norte para o estabelecimento do usuário e de suas necessidades, os autores ressaltam que pode haver dúvidas sobre como respondê-las. Nesse sentido, Fuertes-Olivera e Tarp (2014) propõem três maneiras possíveis para se obter as respostas: i) pedir para que um grupo heterogêneo de possíveis usuários respondam questionários e entrevistas; ii) escolher um grupo de possíveis usuários específico e consultar alguém que tenha

---

<sup>49</sup>“ – *Function-relevant user characteristics:*

*Which language is the user's mother tongue or first language? What is the user's proficiency level in the mother tongue?*

*With which method is the user learning the mother tongue or first language? What is the user's proficiency level in a second, third, etc., language?*

*With which method is the user learning a second, third, etc., language? What is the user's general cultural and encyclopaedic level?*

*What is the user's experience in translation between a specific set of languages? What is the user's knowledge of a specific discipline (layman, semi-expert or expert)?*

*What is the user's comparative knowledge of a specific culture-bound discipline as expressed in two specific language areas?*

*What is the user's proficiency level in a specific specialised language?*

*What is the user's experience in translation between a specific set of specialised languages? Etc., etc...”*

<sup>50</sup>“ – *Consultation-relevant user characteristics:*

*What is the user's experience of lexicographical consultations?*

*Is the user blind, deaf, or suffers from any other handicap which may limit the use of specific types of lexicographical tools?*

*Does the user have electricity and electric light?*

*Does the user possess a device with access to the Internet? Does the user know to distinguish between right and left? Etc., etc...”*

conhecimento íntimo do grupo; iii) definir o grupo de usuários com base no conhecimento que o lexicógrafo apresenta da área em questão, listando as perguntas apresentadas anteriormente.

Os autores argumentam que a primeira forma de se conseguir essas respostas não se configura como sendo a mais indicada devido ao tempo demandado e às questões estatísticas de representatividade e confiabilidade que podem influenciar na leitura dos resultados, bem como pelo investimento em moeda que deve ser feito para a realização de tais questionários. Já com relação aos dois últimos métodos de responder as perguntas propostas, os autores afirmam que

Quando, portanto, um perfil de usuário está sendo projetado para um dicionário especializado específico, pode ser considerado recomendável o uso de um dos dois últimos métodos, ou uma combinação deles; por exemplo, a primeira versão do conceito é preparada de acordo com o terceiro método, e então é refinado por meio do segundo. Em comparação ao primeiro dos três métodos mencionados, os dois últimos são relativamente mais rápidos, fáceis e baratos de aplicar, embora sejam eficientes como uma base sólida para garantir a correspondência entre as necessidades previstas e reais do grupo de usuários pretendido<sup>51</sup>. (FUERTES-OLIVERA, 2014, p. 31).

Nesse sentido, em nossa pesquisa, definimos como usuário específico da proposta de dicionário pedagógico de colocações especializadas os alunos dos cursos de Bacharelado em Letras e de Licenciatura em Letras. Entretanto, assim como acontece com a maioria das obras lexicográficas, nossa proposta de dicionário de colocações especializadas do âmbito criminal poderá ser consultada, também, pelos demais aprendizes dos idiomas contemplados nesta pesquisa (inglês, português e espanhol), assim como por tradutores.

Ressaltamos que a definição deste grupo de usuários específicos se deu por meio da junção entre o terceiro método para a obtenção de respostas para as perguntas propostas, conforme enfatizado pelos próprios autores, ou seja, pela experiência do lexicógrafo, aqui obtida pelos conhecimentos compartilhados pela coordenadora do grupo de estudo ao qual essa pesquisa se filia, uma vez que desenvolve atividades direcionadas ao público-alvo definido há mais de vinte anos, com o primeiro método, que prevê a utilização de formulários

---

<sup>51</sup> “When, therefore, a user profile is being designed for a specific specialised dictionary, it may be considered recommendable to use one of the two last methods, or a combination of them; for instance, the first version of the concept is prepared according to the third method, and then it is refined by means of the second. In comparison with the first of the three methods mentioned, the latter two are relatively quicker, easier and cheaper to apply, whilst being efficient as a solid basis for ensuring correspondence between the foreseen and real needs of the intended user group”.

respondidos por um grupo heterogêneo de participantes. Tais formulários foram elaborados como parte da pesquisa desenvolvida por Alves (em andamento), junto ao grupo de estudos ao qual esta pesquisa se filia.

Ainda de acordo com a TFL, após a definição do público-alvo, deve-se levar em conta as funções que este possa necessitar diante das situações em que se encontrará. Nesse sentido, a TFL propõe quatro situações: i) situações comunicativas; ii) situações cognitivas; iii) situações operativas; e iv) situações interpretativas (FUERTES- OLIVERA; TARP, 2014).

Os autores afirmam que, antes da TFL, a primeira situação não era considerada. Contudo, “como um grande número de dicionários são tradicionalmente concebidos e consultados com o objetivo de solucionar problemas de comunicação textual, a teoria funcional operou, por razões de conveniência, desde o início, com uma quarta situação fundamental”<sup>52</sup>(FUERTES-OLIVERA; TARP, 2014, p. 31).

Assim, cada uma das quatro situações diz respeito às informações que os usuários precisam, respectivamente, ao:

- Produzir, receber (por meio da leitura ou da audição), traduzir e revisar textos orais ou escritos;
- Adquirir conhecimento ou acessar conhecimento prévio para a realização de uma tarefa específica;
- Compreender ou interpretar dado fenômeno, sinal, símbolo, texto, etc.;
- Realizar uma ação de caráter físico, mental ou linguístico. (FUERTES-OLIVERA; TARP, 2014).

A partir do exposto, a amostra de obra lexicográfica que propomos aqui visa sanar as necessidades de seus usuários específicos em situações comunicativas, auxiliando, ainda, na aquisição de novas colocações especializadas por meio da memorização, uma vez que esta se caracteriza como um dos processos de aquisição de novo vocabulário.

No próximo capítulo, ao abordarmos os aspectos metodológicos utilizados no desenvolvimento desta pesquisa, discorreremos, também, sobre a metodologia empregada na elaboração de nossa amostra de dicionário, tendo como base a TFL aqui exposta.

---

<sup>52</sup> “[...] as a large number of dictionaries have traditionally been conceived and consulted with a view to solving problems of text communication, the function theory has, for convenience reasons, operated from the very outset with a fourth fundamental situation”.

## 4 PROCEDIMENTOS METODOLÓGICOS

Nesta seção, apresentamos os procedimentos metodológicos adotados nesta pesquisa, tanto no que diz respeito à busca, ao levantamento e à análise das colocações, quanto ao desenvolvimento da amostra do dicionário que será apresentado como produto final deste trabalho. Dessa forma, o capítulo é dividido em duas seções: na primeira, explicamos os procedimentos utilizados para a análise dos *corpora* de estudo e do *corpus* comparável, além de abordarmos o levantamento e a análise das colocações; na segunda, discorremos sobre os procedimentos adotados para a elaboração da amostra de um dicionário pedagógico de colocações especializadas.

### 4.1 Compilação dos *corpora* e levantamento das colocações especializadas

Conforme mencionamos anteriormente, nossos *corpora* de estudo são constituídos pelas legendas das séries de TV norte-americanas *Criminal Minds*, *CSI – Crime Scene Investigation* e *NCIS* em línguas inglesa, portuguesa e espanhola. Ao compilá-los, levamos em consideração alguns aspectos importantes, tais como: representatividade da língua, propósito de pesquisa linguística e ocorrência natural. Dessa forma, a compilação do *corpus* se deu com base na definição de *corpus* proposta por Tognini-Bonelli (2002, p.2), já mencionada anteriormente.

As séries já apresentadas foram escolhidas para compor os *corpora* de estudo por corroborarem o aspecto de ocorrência natural dos textos, uma vez que retratam a rotina de trabalho das pessoas envolvidas em investigações criminais, de forma que o vocabulário específico do âmbito criminal seja evidenciado. Dessa forma, os *corpora* de estudo atendem a outro aspecto importante proposto por Tognini-Bonelli (2002), o propósito de pesquisa linguística, ou seja, os *corpora* nos fornecem amostras reais de uso da língua especializada, o que possibilitou a busca e análise das colocações especializadas da referida área.

Quanto à compilação dos *corpora*, os procedimentos adotados são descritos a partir do *Corpus Criminal Inglês*, nome dado ao *corpus* composto pelas legendas em língua inglesa das três séries selecionadas e já mencionadas. Contudo, faz-se importante ressaltar que os mesmos procedimentos foram adotados para a compilação dos outros dois *corpora* de estudo, a saber: o *Corpus Criminal Português* e o *Corpus Criminal Espanhol*, respectivamente.

Dessa forma, o início da coleta das legendas se deu ao acessar o site

<http://www.tvsubtitles.net/> e verificar a disponibilidade das legendas das séries definidas nos idiomas selecionados. Verificamos, assim, que as legendas de todos os episódios de todas as temporadas da série *CSI – Crime Scene Investigation* estavam disponíveis em inglês, assim como as legendas em inglês de 12 temporadas da série *Criminal Minds* e de 15 temporadas da série *NCIS*, uma vez que na data em que a coleta dos dados foi realizada, no início desta pesquisa (2018), as duas últimas séries ainda estavam sendo veiculadas e lançaram mais temporadas (a série *Criminal Minds* foi cancelada em 2019, totalizando 15 temporadas; e *NCIS* ainda está sendo veiculada, contando com um total de 20 temporadas).

No que diz respeito ao *Corpus Criminal Português*, foi possível obter as legendas de 12 temporadas da série *CSI – Crime Scene Investigation*; de 9 temporadas da série *Criminal Minds*; e de 10 temporadas da série *NCIS*. Já em relação ao *Corpus Criminal Espanhol*, obtivemos as legendas de 5 temporadas da série *CSI – Crime Scene Investigation*; de 8 temporadas da série *Criminal Minds*; e de 8 temporadas da série *NCIS*.

Assim, tendo realizado o *download* das legendas em formato *.txt* e organizado os *corpora* de estudo, passamos a analisá-los com o auxílio das ferramentas da plataforma online *Sketch Engine* (KILGARIFF, 2004). Após fazer o *upload* dos *corpora* na plataforma, e levando em conta o nível de automatização da plataforma, buscamos analisar a lista de palavras-chave obtida por meio da ferramenta *Keywords*, a fim de que fosse possível selecionarmos as palavras-chave que atuariam como termos de busca para as colocações. Esta ferramenta utiliza o *corpus* de estudo e cruza os dados deste com um *corpus* de referência, tendo sido utilizado para essa finalidade o *corpus English Web 2020 (enTenTen20)*, próprio da plataforma *Sketch Engine* (KILGARIFF, 2004), e que contém um total de 43.125.207.462 tokens e 36.501.273.153 palavras. A seguir, trazemos o print da tela obtido a partir do resultado apresentado ao selecionarmos o *Corpus Criminal Inglês* (doravante CCI).

Figura 6: Ferramenta *Keywords* – CCI.

(Items: 77,542, total frequency (focus corpus): 9,643,089, total frequency (reference corpus): 24,728,263,034)

reference corpus: English Web 2018 (enTenTen18)

Word	Word	Word	Word	Word
1 unsub	11 mm-hmm	21 whoa	31 got	41 hmm
2 dinozzo	12 grissom	22 hotchner	32 okay	42 ooh
3 gibbs	13 elderman	23 garcia	33 bolo	43 p0
4 mcgee	14 i-i	24 prentiss	34 csi	44 secnav
5 ziva	15 jethro	25 huh	35 rossi	45 uh-huh
6 ncis	16 yeah	26 hey	36 fornell	46 reid
7 uh	17 probie	27 abbs	37 uhh	47 hodge
8 hotch	18 gonna	28 ohh	38 jareau	48 ecklie
9 ducky	19 aah	29 um	39 mallard	49 autopsy
10 abby	20 warrick	30 tox	40 shh	50 vance

Rows per page: 50 1-50 of 1,000 1 / 20

Fonte: Plataforma *Sketch Engine* com a utilização de dados próprios da pesquisa.

Como pode ser observado na imagem acima, a primeira palavra-chave obtida pela análise feita entre o CCI e o *enTenTen20* foi *unsub*, fato que nos levou a dar mais atenção a esta palavra e à sua escolha como termo de busca para as colocações especializadas em nosso *corpus*.

Ressaltamos ainda que, embora tenha sido realizado o mesmo procedimento como *Corpus Criminal Português* (doravante CCP) e o *Corpus Criminal Espanhol* (doravante CCE) para a obtenção de uma lista de palavras-chave referente a cada *corpus*, optamos por utilizar como ponto de partida para a seleção das palavras-chave que atuariam como termo de busca apenas a lista obtida como resultado da análise empreendida no CCI, uma vez que esta pesquisa tem como ponto de partida a língua inglesa. Dessa forma, selecionamos as 5 palavras-chave que apresentaram maior índice de chavicidade<sup>53</sup> e que se relacionam, de alguma forma, com o âmbito criminal na lista apresentada acima e utilizamos palavras correspondentes a elas, em língua portuguesa e em língua espanhola, para o levantamento das colocações especializadas nos *corpora* CCP e CCE.

A seguir, apresentamos as cinco palavras que foram selecionadas como termos de busca e seus respectivos índices de chavicidade, os quais recebem a denominação de *score* pela ferramenta utilizada.

<sup>53</sup> O índice de chavicidade corresponde a pontuação obtida como resultado dos cálculos realizados de forma automática pela plataforma *Sketch Engine*, o qual diz que quanto maior o seu número, maior é também a frequência dessa palavra no *corpus* de estudo em relação ao *corpus* de referência. Do mesmo modo, quanto menor seu número, menor a frequência dessa palavra no *corpus* de estudo quando comparada ao *corpus* de referência.

Tabela 1: Índice de chavidade das Palavras-chave selecionadas no CCI

Palavras-chave selecionadas no CCI	
Palavra-chave	Índice de Chavidade
<i>Unsub</i> (suspeito)	259.500
<i>Autopsy</i> (autópsia)	18.000
<i>Gunshot</i> (tiro)	17.700
<i>Alibi</i> (álibi)	17.700
<i>Victimology</i> (vitimologia)	17.100

Fonte: Próprio da pesquisa.

Como mencionamos anteriormente, a palavra-chave *unsub*, que aparece logo na primeira posição da lista, apresenta um índice de chavidade bem alto, o que justifica a sua escolha. Já as demais palavras apresentam um equilíbrio entre seus índices de chavidade.

Após a seleção das palavras-chave, passamos a utilizar a ferramenta *Word Sketch*, também da referida plataforma. Essa ferramenta nos apresenta, de maneira automatizada, as colocações que são formadas a partir da palavra que estamos buscando. A seguir, apresentamos uma imagem da ferramenta mencionada com os resultados obtidos a partir do levantamento das colocações formadas com o termo de busca *unsub*:

Figura 7: Resultados obtidos com a ferramenta *Word Sketch – Unsub*.

modifiers of "unsub"	verbs with "unsub" as object	verbs with "unsub" as subject	"unsub" and/or ...	prepositional phrases
<b>female</b> The female unsub 9 9.4	<b>believe</b> We believe the unsub 37 10.1	<b>know</b> the unsub knew 52 9.0	<b>subject</b> unsub, or unknown subject 4 10.4	... of "unsub" 1.9%
<b>male</b> I think the male unsubs might even be 6 9.1	<b>think</b> think the unsub 43 8.9	<b>kill</b> unsub killed 33 8.8	<b>case</b> in this case, the unsub 8 9.7	... to "unsub" 1.4%
<b>multiple</b> multiple unsubs 8 8.9	<b>mean</b> Which means the unsub 21 8.7	<b>use</b> unsub is using 30 8.7	<b>victim</b> unsub and his victims 10 9.5	... for "unsub" 1.1%
<b>single</b> a single unsub 6 8.3	<b>profil</b> We profiled the unsub 7 8.4	<b>want</b> unsub wants to 34 8.6		... with "unsub" 0.9%
<b>same</b> this is the same unsub 29 8.2	<b>suggest</b> suggests the unsub 8 8.3	<b>have</b> the unsub has 128 8.5		... like "unsub" 0.8%
<b>The</b> The unsubs are 16 8.0	<b>organize</b> unsub 's organized 6 8.2	<b>choose</b> unsub chose 16 8.3		"unsub" in ... 0.7%
<b>second</b> second unsub 4 7.1	<b>meet</b> met the unsub 10 7.9	<b>do</b> unsub did n't 87 8.3		... about "unsub" 0.5%
<b>real</b> the real unsub 4 6.7	<b>drive</b> unsub is driven by 7 7.7	<b>be</b> unsub is 495 8.1		... at "unsub" 0.4%
<b>old</b> the older unsub is 5 6.5	<b>assume</b> assume the unsub 5 7.7	<b>target</b> the unsub is targeting 13 8.1		"unsub" with ... 0.4%
	<b>say</b> said the unsub 20 7.7	<b>feel</b> the unsub feels 14 8.0		... from "unsub" 0.4%
	<b>connect</b> unsub is n't connected 5 7.6	<b>take</b> the unsub took 21 7.9		... as "unsub" 0.4%
	<b>appear</b> appear the unsub 4 7.6	<b>get</b> unsub got 32 7.7		... on "unsub" 0.3%

Fonte: Plataforma *Sketch Engine* com a utilização de dados próprios da pesquisa.

Nesta tela, é possível identificar, no topo de cada coluna aparente, uma classificação de acordo com a função desempenhada pelas palavras que acompanham o termo de busca. Por exemplo, na primeira coluna temos combinações de palavras que são candidatas a colocações, que atuam como modificadores do termo de busca em questão, *unsub*; enquanto na segunda coluna as palavras listadas são verbos os quais têm *unsub* como objeto; etc. Ademais, ao clicar sobre o sinal de reticências ao lado de cada palavra, a ferramenta disponibiliza uma tela com informações de frequência de ocorrência e o *LogDice*<sup>54</sup> de cada colocação, bem como as linhas de concordância com os exemplos da combinação de palavras em análise. Essas informações foram salvas em tabelas, para que fosse possível utilizá-las no momento da análise. As informações relacionadas ao *LogDice* foram utilizadas para a análise quantitativa, enquanto que a informação sobre a classificação morfológica e os exemplos serviram para o desenvolvimento da análise qualitativa. Estas tabelas seguiram o modelo apresentado a seguir:

<sup>54</sup> “A pontuação de tipicidade (*LogDice*) indica o quão forte uma colocação é. Quanto maior a pontuação, mais forte é a colocação. Uma pontuação baixa significa que as palavras que na colocação também são combinadas, frequentemente, com muitas outras palavras” (SKETCH ENGINE, acesso em 20 de maio de 2022). “*The typicality score (LogDice) indicates how strong the collocation is; The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words*”.

Tabela 2: Informações obtidas por meio da ferramenta *Word Sketch*.

Informações obtidas por meio da <i>Word Sketch</i>			
Colocação	<i>LogDice</i>	Categoria <i>Word Sketch</i>	Exemplo
<i>Multiple unsubs</i> (múltiplos suspeitos)	8.9	Modificadores de <i>unsub</i>	<i>The fact that entire families are being target suggests multiple unsubs.</i>
<i>Single unsub</i> (único suspeito)	8.3	Modificadores de <i>unsub</i>	<i>What if we're looking for a single unsub?</i>
<i>Real unsub</i> (verdadeiro suspeito)	6.7	Modificadores de <i>unsub</i>	<i>Bringing in another suspect draws attention away from the real unsub.</i>
<i>Catch the unsub</i> (pegar o suspeito)	7.5	Verbos com <i>unsub</i> como objeto	<i>Then do whatever you can to catch the unsub.</i>
<i>Profile the unsub</i> (perfilar o suspeito)	8.4	Verbos com <i>unsub</i> como objeto	<i>Guys, something doesn't add up. We profiled this unsub would have an intimate knowledge of his kill zone.</i>

Fonte: Próprio da pesquisa.

Com os resultados desse levantamento automatizado empreendido pela própria ferramenta e demonstrado acima, passamos a analisar cada sugestão de combinação de palavras, a fim de levantarmos as colocações que são compostas a partir da base *unsub*. No que tange à análise das colocações especializadas, nos pautamos nos pressupostos teóricos de Hausmann (1984, 1985) e de Orenha-Ottaiano (2009, 2017, 2020), a fim de analisar sintaticamente as colocações e, assim, promovermos sua classificação de acordo com sua taxonomia. Para tal, utilizamos quadros, como o apresentado a seguir, nos quais buscamos identificar a base e o colocado de cada colocação, bem como sua respectiva taxonomia.

Quadro 1: Análise morfológica das colocações especializadas.

Análise Sintática		
Colocação	Estrutura sintática	Taxonomia
<i>Infectious autopsy</i> (autópsia infecciosa)	Adjetivo colocado + subst base	Adjetiva
<i>Psychological autopsy</i> (autópsia psicológica)	Adjetivo colocado + subst base	Adjetiva
<i>Autopsy report</i> (relatório de autópsia)	Subst colocado + subst base	Nominal
<i>Evidence in autopsy</i> (evidência na autópsia)	Subst colocado + prep + subst base	Nominal
<i>Assisting in an autopsy</i> (auxiliar em uma autópsia)	Verbo colocado + prep + det + subst base	Verbal
<i>Observe an autopsy</i> (observar uma autópsia)	Verbo colocado + det + subst base	Verbal

Fonte: Próprio da pesquisa.

Conforme ressaltamos anteriormente, nesta seção utilizamos o CCI para descrever os procedimentos metodológicos concernentes à compilação dos *corpora*, seu tratamento, busca e levantamento das colocações especializadas, porém os procedimentos aqui descritos foram empreendidos nos demais *corpora*, a saber: o CCP e o CCE. Esses dados são apresentados e discutidos no capítulo 5.

No que tange a análise léxico-semântica, apresentada na seção 5.3, contamos com o auxílio dos dicionários on-line *Cambridge Dictionary* e *Léxico Dictionary*. Essas consultas se fizeram necessárias para elucidarmos os processos de busca que podem ser utilizados por estudantes dos curso de Letras – Licenciatura e Letras - Tradutor, considerados público-alvo desta pesquisa, ao terem dificuldade de compreensão das colocações especializadas apresentadas e discutidas neste trabalho, uma vez que tais estudantes, movidos pela ausência de uma ferramenta fraseográfica que os auxiliem, buscam compreender as colocações olhando de forma isolada para as partes que as compõem (base e colocado) para só então tentarem depreender o seu significado composicional.

Na subseção seguinte, abordamos os procedimentos utilizados para a compilação e análise do Corpus Comparável Criminal, além de justificarmos a sua inclusão em nossa análise.

#### 4.1.1 O Corpus Comparável Criminal II

Após os resultados obtidos por meio da análise de nossos *corpora* de estudo,

passamos para a compilação de um *corpus* comparável, ao qual chamamos de Corpus Comparável Criminal II, a fim de que pudéssemos ter uma fonte de comprovação de ocorrência e a frequência das colocações aqui levantadas, tal como realizado durante a pesquisa de mestrado (CALDAS, 2017) que antecedeu e originou essa tese de doutoramento. Dessa vez, no entanto, adotamos um procedimento diferente do que foi utilizado anteriormente.

Para a compilação de nosso Corpus Comparável Criminal II (doravante CCCII), recorreremos ao site de busca *Google Acadêmico* e pesquisamos por artigos que abordassem os temas investigativos em língua inglesa, por meio das palavras de busca *investigation*, *forensic investigation* e *criminology*. Como resultado dessa busca, foram selecionados 32 artigos que contemplam a área definida pelas palavras-chave. De posse desses artigos, foi feito o *upload* do *corpus* na plataforma *Sketch Engine* (KILGARIFF, 2004) para que o CCCII pudesse ser analisado seguindo os mesmos passos dos *corpora* de estudo compostos pelas legendas das séries selecionadas para esta pesquisa (*CSI: Crime Scene Investigation*, *Criminal Minds* e *NCIS*).

A seguir, apresentamos a tela contendo as informações sobre o CCCII:

Figura 8: Informações sobre o Corpus Comparável Criminal II – CCCII

The screenshot displays the Sketch Engine interface for the Corpus Comparável Criminal II (CCCII). The interface includes a navigation bar with buttons for 'MANAGE CORPUS', 'MANAGE SUBCORPORA', 'COMPARE CORPORA', and 'TEXT TYPE ANALYSIS'. The main content area is divided into three sections: 'GENERAL INFO', 'COUNTS', and 'TEXT TYPES'. 'GENERAL INFO' shows the language as English and provides links for 'CORPUS DESCRIPTION & BIBLIOGRAPHY', 'TAGSET', 'WORD SKETCH GRAMMAR', and 'TERM GRAMMAR'. 'COUNTS' displays the following statistics: Tokens (379,583), Words (283,782), Sentences (13,291), and Documents (1). 'TEXT TYPES' shows a list of text types with their respective counts: <doc> (2), File ID, doc.id (1), File name, doc.filename (1), <archive\_file> (1), <g> (0), <s> (0).

Text Type	Count
<doc>	2
File ID, doc.id	1
File name, doc.filename	1
<archive_file>	1
<g>	0
<s>	0

Fonte: Plataforma *Sketch Engine* com a utilização de dados próprios da pesquisa.

Como pode ser observado na imagem acima, nosso *corpus* comparável, o CCCII conta com 32 arquivos, já que os arquivos dos artigos foram adicionados à plataforma no formato PDF, mesmo formato em que são disponibilizados na internet, uma vez que a plataforma *Sketch Engine* (KILGARRIF, 2004) faz a leitura de arquivos nesse formato, não necessitando, assim, que os textos fossem salvos em outro formato. Ademais, esses 32 artigos nos proporcionaram um *corpus* contendo o total de 283.782 palavras, ou seja, um *corpus* suficiente para o propósito para o qual foi compilado: atestar o uso das colocações especializadas encontradas em nosso CCI, a fim de comprovar que as séries fazem uso da

linguagem especializada utilizada por profissionais da área investigativa/criminal.

Convém ressaltar que compilamos o CCCII apenas em língua inglesa, devido ao fato de esta ser a língua de partida de nossa pesquisa, bem como a língua na qual as séries são produzidas originalmente. Dessa forma, os estudantes de tradução e os tradutores de séries que desejarem consultar nossa amostra de dicionário, apresentada ao final deste trabalho, partirão da língua inglesa para as línguas portuguesa e/ou espanhola.

Tendo esclarecido, dessa forma, os processos utilizados para o levantamento e análise das colocações especializadas, bem como, os processos de compilação dos *corpora* utilizados nesta pesquisa, a seguir, passamos aos aspectos metodológicos concernentes à elaboração da amostra de dicionário pedagógico de colocações especializadas do âmbito criminal.

#### 4. 2 Elaboração da amostra de dicionário pedagógico

No que diz respeito à elaboração de nossa amostra de dicionário pedagógico de colocações especializadas, tomamos como ponto de partida a afirmação de Fuertes-Olivera e Tarp (2008) segundo a qual

para determinar as necessidades dos usuários, estas devem se relacionar não somente a um determinado tipo de usuário (já que os diferentes tipos de usuário têm diferentes tipos de necessidades), mas também a um determinado tipo de situação social que ainda influencia com maior peso do que aquele na caracterização do tipo de necessidades<sup>55</sup> (FUERTES OLIVERA; TARP, 2008, p. 78).

Nesse sentido, começamos por definir os usuários a quem se destina nossa amostra de dicionário, levando em conta não apenas o usuário em si, mas considerando, também, uma situação social em que este usuário se beneficiaria do conteúdo expresso em nosso dicionário. A partir dessas afirmações, e levando em consideração a experiência da coordenadora do grupo de pesquisa ao qual nos filiamos e do projeto de pesquisa sobre dicionários de colocações no qual nos baseamos<sup>56</sup>, de mais de vinte anos de trabalho com os

---

<sup>55</sup> “[...] para determinar las necesidades de los usuarios, estas deben relacionarse no sólo a un determinado tipo de usuario (ya que los diferentes tipos de usuarios tienen diferentes necesidades), sino también a un determinado tipo de situación social que incluso influye con mayor peso que aquel en la caracterización del tipo de necesidades.”

<sup>56</sup> Grupo de Pesquisa “Fraseologia e Colocações a partir de *Corpora* (FRASCORP), devidamente cadastrado no Diretório de Pesquisas do CNPq e do Projeto de Pesquisa “A phraseographical methodology and model for an online *corpus*-based Multilingual Collocations Dictionary Platform”, ambos coordenados pela Profa. Dra. Adriane

alunos dos cursos de graduação delimitados como público-alvo, bem como as respostas que os mesmos forneceram aos questionários aplicados por Alves (tese de doutorado em andamento), definimos como usuários em potencial estudantes dos cursos de Licenciatura em Letras e de Bacharelado em Letras com Habilitação para Tradutor, os quais poderão fazer uso das informações do dicionário que compilamos em seus estudos, a fim de colaborarem com sua formação. Acreditamos, ainda, que tradutores, principalmente aqueles que se dedicam a tradução de séries, seja para legendagem, seja para dublagem, também se beneficiarão das informações apresentadas em nosso dicionário. Outrossim, estudantes brasileiros das línguas estrangeiras contempladas nesta pesquisa (inglês, português e espanhol) podem se beneficiar igualmente das informações que são apresentadas nesta amostra.

Além de determinar o usuário ao qual se destina a amostra de obra lexicográfica que propomos aqui, esta deve apresentar, ainda, uma função lexicográfica bem definida. Tal como abordado na seção 3.4.2, sobre a TFL (Teoria Funcional da Lexicografia) proposta por Bergenhotlz e Tarp (2002, 2003) e explorada por Fuertes-Olivera e Tarp (2014), a referida teoria define quatro situações para as quais há quatro funções definidas. São elas: comunicativa, cognitiva, operativa e interpretativa.

Nesta pesquisa, buscamos contemplar a função comunicativa dos usuários aos quais nossa amostra de dicionário pedagógico se dirige, uma vez que esta contempla “tudo o que diz respeito à produção, recepção, tradução e revisão de textos orais ou escritos”<sup>57</sup> (FUERTES-OLIVERA; TARP, 2014, p. 31). Ressaltamos que esta função foi escolhida justamente pela necessidade que os estudantes dos cursos de graduação em questão (Licenciatura em Letras e Bacharelado em Letras) deparam-se com várias situações de aprendizagem nas quais devem resolver atividades didáticas, fazer traduções e versões de diferentes textos, também.

Dessa forma, após a delimitação dos usuários a que se destina o produto final desta pesquisa e da função lexicográfica a ele atribuída, passamos a considerar as informações que seriam integradas à macro e microestrutura do dicionário. No que diz respeito à macroestrutura, esta será organizada de maneira a facilitar a busca pelas colocações organizando a lista de verbetes de forma alfabética e vertical, para que seja possível uma

---

Orenha-Ottaiano.

<sup>57</sup> “[...] everything that concerns the production, reception, translation, and revision of oral or written texts [...]”.

rápida visualização dos conteúdos dos verbetes.

Ademais, pautamo-nos na metodologia lexicográfica desenvolvida pelo Projeto de Pesquisa *PLATCOL - Plataforma On-line de Dicionários Bilíngues de Colocações* (2017-2019), conforme Orenha-Ottaiano (2017), e nas novas propostas do projeto *A phraseographical methodology and model for an online corpus-based Multilingual Collocations Dictionary Platform* (ORENHA-OTTAIANO, 2020), aos quais este trabalho se insere.

Assim, tomamos como base o proposto por Orenha-Ottaiano et. al (2021) para elaborar a microestrutura dos verbetes apresentados a partir de entradas, as quais são constituídas pelas palavras-chave que serviram como termo de busca para o levantamento das colocações especializadas, conforme descrito na seção anterior e na seção 3.3. Dessa forma, cada entrada conta com as informações descritas no quadro abaixo.

Quadro 2: Informações referentes às entradas.

<b>Palavra de entrada na LP:</b> corresponde à base das colocações e pode ser um substantivo, um verbo ou um adjetivo.
<b>Classe gramatical da palavra na LP:</b> a classe gramatical da palavra de entrada é posicionada logo após a palavra em questão. Em nossa amostra de dicionário de colocações, as palavras de entrada podem ser um substantivo (s), um verbo (v) ou um adjetivo (adj). Caso a palavra de entrada apresente mais de uma classe gramatical, essas aparecerão em mais de uma entrada, por exemplo <i>murder</i> (s) e <i>murder</i> (v).
<b>Definição:</b> uma breve definição para cada palavra de entrada e para cada indicação de classe gramatical desta, indicando o sentido a ela atribuído em cada situação
<b>Palavra de entrada na LC:</b> corresponde à tradução da palavra de entrada apresentada na língua de chegada.
<b>Classe gramatical da palavra de entrada na LC:</b> corresponde à classe gramatical atribuída a essa palavra na língua de chegada
<b>Definição:</b> uma breve definição para a palavra de entrada na língua de chegada
<b>Frequência:</b> ao lado cada palavra de entrada haverá um número indicando a frequência dessa palavra no <i>corpus</i> de estudo, tanto na LP quanto na LC.

Fonte: Baseado em ORENHA-OTTAIANO et al (2021).

Quanto às definições das entradas, estas foram elaboradas ao tomarmos como base as definições das referidas palavras-chave encontradas no dicionário online *Cambridge*

*Dictionary*. Abaixo de cada entrada, seguem as colocações que apresentam como base a referida entrada, listadas como verbetes individuais e apresentadas de acordo com sua classificação taxonômica, ou seja, todas as colocações adjetivas, seguidas das colocações nominais e, por fim, as colocações verbais. Cada verbete contém as informações descritas no quadro abaixo:

Quadro 3: Informações apresentadas em cada verbete.

<b>Colocação na LP:</b> exemplo de colocação na língua de partida.
<b>Estrutura sintática da colocação na LP:</b> as colocações são analisadas e organizadas de acordo com o proposto por Hausmann (1985, 1989) e Orenha-Ottaiano (2004, 2009), conforme apresentado na seção anterior e na seção 3.3.
<b>Exemplo de uso na LP:</b> é apresentado um exemplo de uso na língua de chegada.
<b>Frequência na LP:</b> ao lado de cada colocação há um número indicando a sua frequência no <i>corpus</i> de estudo, na língua de partida.
<b>Colocação na LC:</b> exemplo de colocação na língua de chegada.
<b>Estrutura sintática da colocação na LC:</b> as colocações são analisadas e organizadas de acordo com o proposto por Hausmann (1985, 1989) e Orenha-Ottaiano (2004, 2009), conforme apresentado na seção anterior e na seção 3.3.
<b>Exemplo de uso na LC:</b> é apresentado um exemplo de uso na língua de chegada.
<b>Frequência na LC:</b> ao lado cada colocação há um número indicando a sua frequência no <i>corpus</i> de estudo, na língua de chegada.

Fonte: Baseado em ORENHA-OTTAIANO et al (2021).

No que tange aos exemplos dados na língua de chegada, sempre que possível estes foram retirados dos *corpora* correspondentes, a saber: CCP para a língua portuguesa e CCE para a língua espanhola. Contudo, uma vez que estes *corpora* são menores que o CCI, quando não houver a possibilidade de encontrar o exemplo selecionado no CCI nos demais *corpora*, estes exemplos foram traduzidos pela pesquisadora, a fim de que fosse mantido a correspondência entre os idiomas. Assim, os verbetes são apresentados da seguinte forma:

Quadro 4: Amostra de microestrutura do Dicionário Pedagógico de Colocações Especializadas do âmbito criminal.

Autopsy (s). Freq. 595.	Autópsia (s). Freq. 447.	Autopsia(s). Freq. 186.
The examination of a body in order to discover the cause of death.	Exame realizado em um corpo o qual visa atestar a causa da morte.	El examen de un cuerpo para descubrir la causade la muerte.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<b>Psychological autopsy: (6) (adj + n)</b> - We're gonna build what we call psychological autopsies to determine whether the victims killed themselves. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia psicológica: (7) (s + adj)</b> - Nós vamos fazer o que chamamos de autópsia psicológica para determinar se as vítimas se mataram. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia psicológica: (6) (s + adj)</b> - Construiremos lo que llamamos <b>autopsias psicológicas</b> para determinar si las víctimas sesuicidaron. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).

Fonte: Próprio da pesquisa.

A amostra de dicionário que propomos ao final desta pesquisa seguirá, então, as características definidas e apresentadas nesta seção. Outrossim, no capítulo no qual trazemos a amostra do referido dicionário, discutimos melhor cada uma dessas características.

Tendo apresentado os aspectos metodológicos seguidos para a realização deste trabalho, no capítulo seguinte, apresentamos a análise das colocações obtidas por meio do levantamento realizado em nossos *corpora* de estudo.

## 5 ANÁLISE DOS DADOS

Neste capítulo, além de apresentar as colocações especializadas levantadas a partir da busca descrita na metodologia, trazemos as análises que foram empreendidas e que contribuíram para a seleção das colocações especializadas do âmbito criminal que compõem a proposta de dicionário pedagógico deste trabalho. Dessa forma, expomos a análise quantitativa e comparativa feita a partir dos dados obtidos, a análise morfossintática, que buscou identificar a tipologia das colocações especializadas levantadas, bem como a análise léxico-semântica da palavra-chave *unsub* e de algumas colocações que podem causar dificuldade de compreensão por parte dos alunos dos cursos de Licenciatura em Letras e de Bacharelado em Letras.

### 5.1 Análise quantitativa

Nesta seção, apresentamos e discutimos os dados quantitativos referentes aos *corpora* de estudo e às colocações que foram levantadas e analisadas, de forma comparativa entre os três idiomas que integram esta pesquisa. Além disso, explicamos, por meio de aspectos quantitativos, como se deu a seleção das palavras-chave que foram utilizadas como base para a busca das colocações que compõem a amostra do dicionário proposto nesta pesquisa.

Dessa forma, iniciamos a análise com a apresentação dos dados quantitativos gerais, referentes aos três *corpora* de estudo e ao *corpus* comparável, a saber: o Corpus Criminal de Língua Inglesa, o Corpus Criminal de Língua Portuguesa, o Corpus Criminal de Língua Espanhola e o Corpus Comparável Criminal II (doravante CCI, CCP, CCE, CCCII respectivamente).

Tabela 3: Tamanho dos *corpora*

Tamanho dos <i>Corpora</i>				
Tipo	CCI	CCP	CCE	CCCII
<i>Tokens</i>	9.643.089	6.709.702	3.843.577	379.583
<b>Words</b>	4.741.879	2.952.368	1.910.225	283.782

Fonte: Próprio da pesquisa

Como se pode observar, o CCCII é menor que os demais *corpora* que compõem esta pesquisa, o que se deve ao fato de ele ter sido compilado para servir como uma fonte de comprovação da utilização das colocações especializadas obtidas por meio da análise do CCI,

*corpus* que serviu como ponto de partida para a busca pelas colocações especializadas encontradas e analisadas nesta pesquisa.

Já com relação aos *corpora* de estudo, tal como se observa na tabela anterior, o CCI é o maior *corpus*, o que se deve ao fato de ter sido possível o levantamento das legendas de quase todos os episódios das séries que o compõem, o que não foi possível para os demais idiomas (português e espanhol), já que estas legendas não foram encontradas na *web*. A isso, também, se deve ao fato de o CCE ser o menor entre os três *corpora*: foram encontradas as legendas de poucos episódios das séries selecionadas na *web*, conforme apresentado no quadro abaixo:

Quadro 5: Quantidade de episódios das séries em cada idioma

Episódios das séries que compõem os <i>corpora</i>			
Idioma	<i>CSI</i>	<i>Criminal Minds</i>	<i>NCIS</i>
Inglês	337	272	350
Português	279	174	222
Espanhol	97	147	154

Fonte: Próprio da pesquisa

A diferença no número de legendas que foram obtidas para cada idioma se explica por essas terem sido coletadas por meio de um site no qual fãs das séries disponibilizam as legendas e suas respectivas traduções de forma gratuita para todos os que venham a precisar delas, não havendo, assim, traduções para todos os episódios que foram lançados.

Convém ressaltar, ainda, que a escolha por trabalhar com essas legendas se deu pelo fato de serem facilmente acessadas por qualquer pessoa que necessite ou queira utilizar as legendas, além de serem feitas por tradutores que são fãs das séries em questão, os quais contam com um profissional da área criminal/investigativa em sua equipe de tradução, conferindo, assim, fidedignidade às traduções e promovendo o alcance dessas a um grande público.

Ademais, conforme ressalta Fromm (2011) “o que nos importa é o material linguístico disponibilizado para análise” (FROM, 2011, p. 2-3). Tomando como ponto de partida o fato de o CCI ser o maior entre os três *corpora*, além de ser o inglês a língua de origem dos textos (legendas) que os compõem, toda a análise e busca empreendida nos *corpora* teve como ponto de partida os dados em língua inglesa. Assim, a seleção de

palavras-chave foi feita com base nos dados encontrados no CCI e, posteriormente, buscou-se a palavra correspondente a cada uma das palavras-chave levantadas nos outros dois *corpora*. A seguir, apresentamos a tabela com as palavras-chave selecionadas em cada um dos *corpora*.

Tabela 4: Comparativo de listas de palavras-chave.

Lista de Palavras-chave					
Corpus Criminal Inglês (CCI)		Corpus Criminal Português (CCP)		Corpus Criminal Espanhol (CCE)	
Termo	Score	Termo	Score	Termo	Score
<i>Unsub</i>	259.500	<i>suboficial</i>	63.200	<i>contramaestre</i>	32.400
<i>Autopsy</i>	18.000	<i>unsub</i>	49.900	<i>Victimología</i>	21.700
<i>Gunshot</i>	17.700	<i>fuzileiro</i>	33.900	<i>Asesino</i>	18.600
<i>Alibi</i>	17.700	<i>autópsia</i>	31.800	<i>Unsub</i>	18.200
<i>Victimology</i>	17.100	<i>legista</i>	22.300	<i>Marine</i>	16.200
<i>Killer</i>	16.800	<i>xerife</i>	22.300	<i>Sospechoso</i>	14.900
<i>Lieutenant</i>	11.700	<i>Assassino</i>	20.600	<i>Coartada</i>	13.900
<i>Ballistics</i>	11.600	<i>Álibi</i>	19.200	<i>Apuñalar</i>	13.700
<i>Postmortem</i>	11.200	<i>Detetive</i>	17.700	<i>Sádico</i>	13.400
<i>Sadist</i>	10.800	<i>Vitimologia</i>	17.100	<i>Sargento</i>	13.200
<i>Sargent</i>	10.200	<i>Suspeito</i>	16.700	<i>Detective</i>	13.000
<i>Cop</i>	10.000	<i>Esfaquear</i>	16.700	<i>Autopsia</i>	12.700
<i>Copycat</i>	9.700	<i>Atirador</i>	15.400	<i>Bala</i>	11.800
<i>Victim</i>	9.000	<i>DNA</i>	13.600	<i>Postmortem</i>	11.600
<i>Fingerprint</i>	8.900	<i>Bala</i>	12.700	<i>Rastrear</i>	11.000
<i>Murder</i>	8.700	<i>Necrotério</i>	12.400	<i>Balística</i>	10.900
<i>Stab</i>	8.700	<i>Balística</i>	12.100	<i>Polígrafo</i>	10.400
<i>Kidnapper</i>	8.600	<i>Tenente</i>	11.600	<i>Imitador</i>	10.200
<i>Homicide</i>	8.500	<i>Serial</i>	11.200	<i>Teniente</i>	10.200
<i>Abduction</i>	8.200	<i>Assassinato</i>	11.200	<i>Matar</i>	10.200

Fonte: Própria da Pesquisa

Na tabela acima, são apresentadas as 20 (vinte) primeiras palavras-chave de cada *corpus*, sendo que cada palavra é acompanhada pelo seu *score*. Como pode ser observado, existe uma grande diferença entre o *score* apresentado pelas palavras-chave do CCI em relação ao apresentado pelas palavras-chave dos dois outros *corpora*, além de haver diferença em relação à ordem em que as palavras-chave aparecem em cada uma das listas apresentadas.

Com relação as palavras-chave, essas foram obtidas por meio de cálculos estatísticos realizados de forma automática pela plataforma *Sketch Engine*, para os quais são

consideradas as frequências das palavras que compõem o *corpus* de estudo e as palavras que compõem um *corpus* de referência. A partir desse cálculo, feito com base na fórmula abaixo, é obtida “a pontuação de chavicidade”<sup>58</sup> (KILGARRIFF et al, 2014,p. 3):

$$\frac{fpm_{focus} + n}{fpm_{ref} + n}$$

na qual,  $fpm_{focus}$  é a frequência normalizada (por milhão) da palavra no *corpus* de estudo,  $fpm_{ref}$  é a frequência normalizada (por milhão) da palavra no *corpus* de referência,  $n$  é o parâmetro de Matemática simples ( $n = 1$  é o valor padrão) (KILGARRIFF et al, 2014, p.3).

O resultado desses cálculos, feitos para cada uma das palavras que compõem os *corpora* de estudo, é apresentado na coluna *score* das tabelas presentes nesta pesquisa, tal como é apresentado pela própria plataforma *Sketch Engine* ao mostrar os referidos resultados.

Dessa forma, ao considerarmos que os *corpora* de português e espanhol não contêm as legendas de todos os episódios das séries, tal como acontece com o *corpus* de língua inglesa, é possível entender que o *score* das palavras-chave obtidas poderia sofrer variação, o que influenciaria diretamente no resultado obtido. Assim, optamos por desenvolver a busca e o levantamento de colocações a partir da lista de palavras-chave do CCI, uma vez que esse *corpus* contém as legendas de quase todos os episódios das séries que compõem os *corpora* (a série NCIS, por exemplo, ainda se encontra em exibição) e, conseqüentemente, apresenta uma lista de palavras-chave que “refletem muito bem o domínio do *corpus* em foco”<sup>59</sup> (KILGARRIFF et al, 2015, p. 3), ou seja, o âmbito criminal.

Nesse sentido, para emprendermos a busca e posterior análise das colocações, selecionamos as 5 primeiras palavras de conteúdo<sup>60</sup> relacionadas ao âmbito criminal entre a lista de palavras-chave obtida a partir do CCI, as quais serviram de busca para as colocações nos três idiomas contemplados pelos *corpora* de estudo, tendo sido utilizadas as traduções correspondentes de cada palavra-chave para a busca de colocações no CCP e no CCE, respectivamente.

As palavras-chave selecionadas no CCI são apresentadas na tabela abaixo seguidas

---

<sup>58</sup> “The keyness score [...]”

<sup>59</sup> “[...] reflect the domain of the focus corpus very well [...]”.

<sup>60</sup> Palavras de conteúdo é como nos referimos às palavras que apresentam conteúdo semântico, a saber: substantivos, verbos, adjetivos e advérbios; ocorrem em oposição às palavras gramaticais, as quais são identificadas como posições, artigos e conjunções, por exemplo, e não apresentam conteúdo semântico.

pelas suas frequências no *corpus* de estudos (CCI) e no *corpus* de referência (*enTenTen20*), bem como, pelo índice de chavicidade.

Tabela 5: Palavras-chave selecionadas no CCI

Palavras-chave selecionadas no CCI			
Palavra-chave	Freq. CCI	Freq. REF	Score
<i>Unsub</i> (suspeito)	2699	2124	259.500
<i>Autopsy</i> (autópsia)	595	64276	18.000
<i>Gunshot</i> (tiro)	454	44323	17.700
<i>Alibi</i> (álibi)	297	20696	17.700
<i>Victimology</i> (vitimologia)	173	2756	17.100

Fonte: Próprio da pesquisa

Assim, ao observarmos a lista de palavras-chave apresentadas na tabela anterior, percebemos que a primeira palavra-chave retirada do CCI (*corpus* em língua inglesa) é *unsub* e que esta apresenta um *score* de 259.500, ou seja, essa palavra é altamente frequente em nosso *corpus* de estudo de língua inglesa, muito mais frequente que a segunda palavra-chave apresentada, *autopsy*, que apresenta um *score* de 18.000 apenas. Isso nos mostra que, além de ser extremamente frequente no *corpus* de estudo, a palavra *unsub* configura-se como uma palavra importante para as séries que compõem o *corpus*. Sobre essa palavra e sua importância, bem como seu significado, discutiremos melhor na seção 5.3.1.

Ressaltamos, ainda, que foram selecionadas apenas estas palavras-chave para o levantamento de colocações devido ao fato de que o objetivo principal desta pesquisa é o de propor um modelo de dicionário pedagógico de colocações especializadas, para o qual as colocações especializadas obtidas a partir do levantamento feito por meio destas palavras são suficientes para promover uma discussão e demonstrar os procedimentos adotados, bem como, para o esclarecimento do comportamento destas nos três idiomas que compõem os *corpora* desta pesquisa.

Dessa forma, antes de apresentarmos os números que descrevem a quantidade de colocações obtidas por meio da busca e análise empreendidas a partir das palavras-chave selecionadas, tanto em inglês, quanto nos demais idiomas (português e espanhol), trazemos um quadro com as palavras-chave selecionadas no CCI e suas respectivas correspondentes no CCP e no CCE, já que, conforme dissemos anteriormente, utilizamos apenas a lista de

palavras-chave do CCI como ponto de partida para a busca e levantamento das colocações nos três *corpora* de pesquisa, a saber: CCI, CCP e CCE.

Quadro 6: Palavras-chave selecionadas

Palavras-chave selecionadas em cada <i>corpus</i>		
CCI	CCP	CCE
<i>Unsub</i>	Suspeito	<i>Sospechoso</i>
<i>Autopsy</i>	Autópsia	<i>Autopsia</i>
<i>Gunshot</i>	Tiro	<i>Disparo</i>
<i>Alibi</i>	Álibi	<i>Coartada</i>
<i>Victimology</i>	Vitimologia	<i>Victimología</i>

Fonte: Próprio da pesquisa

Convém ressaltar, ainda, que a delimitação das colocações especializadas que foram incluídas nesta pesquisa se deu pela observação do *LogDice* de cada colocação. O *LogDice*, como apontado anteriormente no capítulo 4, é um índice utilizado pela plataforma *Sketch Engine* (KILGARIFF, 2004) para a identificação de colocações. Segundo a definição da própria plataforma, o *LogDice* é

uma medida estatística para identificar a coocorrência (dois itens que aparecem juntos). O *Sketch Engine* o usa para identificar colocações. Expressa a tipicidade (ou força) da colocação. É usado no recurso de esboço de palavras e também ao calcular colocações de uma concordância. Baseia-se apenas na frequência do nóculo e do colocado e na frequência de toda a colocação (coocorrência do nóculo e do colocado). O *LogDice* não é afetado pelo tamanho do *corpus* e, portanto, pode ser usado para comparar pontuações entre diferentes *corpora*. O *LogDice* é a medida estatística preferida para grandes *corpora*.<sup>61</sup> (KILGARIFF, 2004. Disponível em: <https://www.sketchengine.eu/guide/glossary/?letter%20=L&cat=statistics>)

Como se observa, essa medida estatística serve para identificar a força de uma colocação por meio dos seus componentes e não sofre alteração ainda que o tamanho do *corpus* seja alterado. Ao considerarmos o *LogDice* como índice estatístico, conferimos ao nosso trabalho maior credibilidade, no que tange as colocações especializadas da área criminal, já que este índice mostra a força das colocações encontradas, corroborando o uso de séries como fonte de estudos fraseológicos de áreas de especialidade, tal como apontamos no capítulo 2.

<sup>61</sup> “a statistic measure for identifying co-occurrence (=two items appearing together). Sketch Engine uses it to identify collocations. It expresses the typicality (or strength) of the collocation. It is used in the word sketch feature and also when computing collocations from a concordance. It is only based on the frequency of the node and the collocate and the frequency of the whole collocation (co-occurrence of the node and collocate). logDice is not affected by the size of the corpus and, therefore, can be used to compare scores between different corpora. logDice is the preferred statistic measure for large corpora”.

Ademais, conforme salienta Orenha-Ottaiano (2019), baseada em Glabasova, Brezina e McEnergy (2017), é importante delimitar o valor de  $LogDice \geq 3$ . Sendo assim, as colocações devem apresentar, no mínimo, esse valor para serem incluídas nessa pesquisa. A seguir, apresentamos a tela da ferramenta *Word Sketch* da plataforma *Sketch Engine* (KILGARIFF, 2004), no qual é possível identificarmos o valor de  $LogDice$  para cada colocação.

Figura 9: Resultados obtidos com a ferramenta *Word Sketch – Unsub*.

Category	Collocate	Count	LogDice
modifiers of "unsub"	female	9	9.4
	male	6	9.1
	multiple	8	8.9
	single	6	8.3
	same	29	8.2
	The	16	8.0
	second	4	7.1
	real	4	6.7
	old	5	6.5
	verbs with "unsub" as object	believe	37
think		43	8.9
mean		21	8.7
profiteer		7	8.4
suggest		8	8.3
organize		6	8.2
meet		10	7.9
drive		7	7.7
assume		5	7.7
say		20	7.7
verbs with "unsub" as subject	know	52	9.0
	kill	33	8.8
	use	30	8.7
	want	34	8.6
	have	128	8.5
	choose	16	8.3
	do	87	8.3
	be	495	8.1
	target	13	8.1
	feel	14	8.0
"unsub" and/or ...	subject	4	10.4
	victim	10	9.5
prepositional phrases	... of "unsub"		1.9%
	... to "unsub"		1.4%
	... for "unsub"		1.1%
	... with "unsub"		0.9%
	... like "unsub"		0.8%
	"unsub" in ...		0.7%
	... about "unsub"		0.5%

Fonte: Plataforma *Sketch Engine* com a utilização de dados próprios da pesquisa.

Tal como pode ser observado, as colocações especializadas encontradas a partir da palavra de busca *unsub* atendem ao requisito de valor mínimo de  $LogDice$  igual ou superior a 3, justificando assim, sua inclusão nesta pesquisa para análise e utilização na amostra de dicionário pedagógico proposto ao final do trabalho.

Quanto ao número total de colocações encontradas nos *corpora* de pesquisa a partir das 5 palavras-chave selecionadas para essa análise, foi possível identificar 133 colocações em língua inglesa, 179 em língua portuguesa, e 111 em língua espanhola, resultado este que é melhor demonstrado na tabela abaixo.

Tabela 6: Quantidade de colocações obtidas por palavra-chave em cada *corpus*.

Quantidade de colocações obtidas			
Palavras-chave	CCI	CCP	CCE
<i>Unsub</i>	24	69	16
<i>Autopsy</i>	45	35	25
<i>Gunshot</i>	13	40	31
<i>Alibi</i>	20	21	16
<i>Victimology</i>	28	13	20
<b>TOTAL</b>	132	178	110

Fonte: Próprio da pesquisa.

Como se pode observar na tabela anterior, existe um número maior de colocações especializadas levantadas no CCP, sendo que dentro deste *corpus* foi possível identificar 69 colocações especializadas composta pelo termo *suspeito*, o equivalente para *unsub*, tal como apresentado no quadro 5. Essa diferença se deve ao fato de, no momento da tradução, ter sido utilizado o termo *suspeito* como correspondente tanto para *unsub*, palavra-chave selecionada e apresentada nesta análise, quanto para *suspect*, palavra-chave que aparece na 419<sup>a</sup> posição na lista de palavras-chave obtida por meio da ferramenta *Keywords* da plataforma *Sketch Engine* (KILGARIFF, 2004), fato que colaborou para que esta palavra (*suspect*) não tenha sido selecionada para integrar essa análise.

Ainda assim, ao realizar a busca por colocações especializadas que contenham *suspect* em sua formação, foi possível identificar um total de 50 colocações. Sobre a escolha tradutória das colocações formadas por *unsub* e por *suspect*, bem como, sobre a própria palavra de busca *unsub*, discutiremos melhor na seção 5.3.

Dessa forma, ao concluir o levantamento das colocações a partir da palavra-chave *unsub*, em inglês, e seus equivalentes nos *corpora* em português (*suspeito*) e em espanhol (*sospechoso*), estas foram analisadas, a fim de identificar sua classificação taxonômica, a qual será melhor explicada na seção a seguir. Aqui, apresentamos os resultados quantitativos obtidos por meio desta análise, na tabela abaixo:

Tabela 7: Tipos de colocações levantadas.

Tipos de colocações obtidas				
Palavras-chave	CCI	CCP	CCE	TOTAL
Adjetivas	39	34	33	106
Nominiais	42	90	31	163
Verbais	51	53	45	151

Fonte: Próprio da pesquisa

Pode-se perceber que há certo equilíbrio entre as colocações nominais e verbais encontradas no CCI e no CCE, havendo uma menor ocorrência de colocações adjetivas nos três *corpora*. Ademais, percebe-se um destaque maior no número de colocações nominais encontradas no CCP (90 no total), sendo que esse número se deve ao fato de o termo *suspeito* ter sido utilizado como tradução para *unsub* e *suspect*, conferindo, assim, um maior número de colocações especializadas nominais formadas por *suspeito* em língua portuguesa.

Quanto ao CCCII, este serviu para comprovarmos a utilização das colocações especializadas levantadas em nosso *corpus* de estudo em língua inglesa, de forma que consideramos, também, o *LogDice* das colocações.

Contudo, devido ao fato deste *corpus* ser composto por artigos científicos que, conseqüentemente, são escritos em uma linguagem mais formal do que a linguagem apresentada em nosso *corpus* de estudo composto por séries que retratam o cotidiano de profissionais das áreas investigativa e forense, não foi possível identificarmos colocações formadas pela palavra-chave *unsub*, selecionada em nosso *corpus* de estudo, uma vez que esta surgiu a partir de uma abreviação e é classificada como pertencente a linguagem informal, de acordo com o que é explicitado na seção 5.3.

Ainda assim, optamos por manter a sua seleção para esta pesquisa, bem como a utilização das colocações especializadas por ela compostas, devido ao seu alto índice de chavidade dentro do *corpus* de estudo em língua inglesa (259.500) e aos altos índices de *LogDice* apresentados pelas colocações levantadas, conforme pôde ser visto na figura 9, anteriormente.

Outrossim, sua inclusão na pesquisa se justifica, ainda, pelo fato de tradutores e aprendizes de tradutores que compõem o público-alvo de nossa amostra de dicionário, terem acesso a informações que serão úteis no processo de tradução de novas séries que abordem a mesma temática.

Dito isso, apresentamos abaixo, uma tabela com exemplos em que são comparados os valores de *LogDice* de colocações especializadas entre o *corpus* de estudo em língua inglesa, o CCI, e o *corpus* comparável, o CCCII, para exemplificar os dados obtidos em ambos, bem como a comprovação da utilização das colocações especializadas tanto nas séries que compõem o *corpus* de estudo, quanto por profissionais das áreas investigativa e forense.

Tabela 8: Exemplo de comparação de dados entre o CCI e o CCCII

Comparação entre o CCI e o CCCII		
Colocação	LogDice CCI	LogDice CCCII
<i>Psychological autopsy</i>	10.3	10.8
<i>Full autopsy</i>	8.4	10.5
<i>Forensic autopsy</i>	7.7	4.8
<i>Perform autopsy</i>	11.4	10.3
<i>Gunshot residue</i>	12.1	12.2
<i>Gunshot wound</i>	11.1	12.3

Fonte: Próprio da pesquisa

Como se pode observar pelos exemplos apresentados na tabela acima, o CCCII permitiu validar a utilização das colocações especializadas encontradas em nosso *corpus* de estudo por parte de profissionais da área criminal, corroborando, assim, nossa amostra de dicionário pedagógico apresentada ao final deste trabalho, como fonte de consulta para o público-alvo previamente definido, a saber: alunos dos cursos de Bacharelado e de Licenciatura em Letras, bem como, para tradutores profissionais que precisem traduzir áudios e legendas de séries que abordem a mesma temática.

A seguir, apresentamos a análise morfossintática empreendida para chegarmos aos resultados demonstrados aqui.

## 5.2 Análise morfossintática

Apresentamos, aqui, a análise morfossintática das colocações especializadas obtidas por meio do levantamento realizado com o auxílio da ferramenta *Word Sketch* da plataforma *Sketch Engine* (KILGARIFF, 2004), para a palavra de busca *unsub*, e ressaltamos que os mesmos procedimentos de análise demonstrados por meio delas foram adotados para a análise das colocações obtidas a partir das demais palavras de busca selecionadas.

Esta análise tem por objetivo contrastar a formação das colocações nos três diferentes idiomas que compõem nossos *corpora* e promover a sua classificação taxonômica, uma vez que esta informação se constitui relevante para a elaboração da proposta de dicionário que apresentaremos ao final de nossa pesquisa. Para tal, tomamos por base os estudos de Orenha-Ottaiano (2009) que, por sua vez, se fundamenta em Hausmann (1984, 1985), tal como apresentado na seção 3.3.

Isto posto, ao analisar as colocações obtidas a partir da palavra de busca *unsub*,

obtivemos um total de 35 colocações encontradas no CCI, tendo sido possível identificarmos, a partir de seus correspondentes *suspeito* e *sospechoso*, respectivamente, 81 colocações no CCP e 19 no CCE, conforme apresentado na seção anterior. Sobre a diferença de colocações encontradas nos três *corpora*, reforçamos que o CCE é o menor dos *corpora* que compõem esta pesquisa e que, por isso, esperávamos uma quantidade menor de colocações para cada palavra de busca. Já com relação às 81 colocações encontradas no CCP, partimos da hipótese de que esse número reflete as escolhas tradutórias para colocações formadas com diferentes termos em língua inglesa, a saber: *unsub* e *suspect*, as quais tiveram suas versões em língua portuguesa elaboradas com o mesmo termo, *suspeito*. Tal fato será melhor explicado na seção 5.3.1, onde esperamos comprovar nossa hipótese.

A partir da análise taxonômica, foi possível identificarmos colocações especializadas adjetivas, nominais e verbais nos três *corpora* de estudo. A seguir, apresentamos exemplos colocações adjetivas que têm como base o termo *unsub* e seus correspondentes, a saber: *suspeito* e *sospechoso*.

Quadro 7: Colocações Especializadas Adjetivas – *Corpus* Inglês.

<b>Colocações Especializadas Adjetivas CCI</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Multiple unsubs</i> (múltiplos suspeitos)	Adjetivo colocado + subst. base	Adjetiva
<i>Single unsub</i> (único suspeito)	Adjetivo colocado + subst. base	Adjetiva
<i>Real unsub</i> (suspeito verdadeiro)	Adjetivo colocado + subst. base	Adjetiva

Fonte: Próprio da pesquisa.

Quadro 8: Colocações Especializadas Adjetivas – *Corpus* Português.

<b>Colocações Especializadas Adjetivas CCP</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Possível suspeito</i>	Adjetivo colocado + subst. base	Adjetiva
<i>Principal suspeito</i>	Adjetivo colocado + subst. base	Adjetiva
<i>Novo suspeito</i>	Adjetivo colocado + subst. base	Adjetiva
<i>Múltiplos suspeitos</i>	Adjetivo colocado + subst. base	Adjetiva

Fonte: Próprio da pesquisa

Quadro 9: Colocações Especializadas Adjetivas – *Corpus* Espanhol

Colocações Especializadas Adjetivas CCE		
Colocação	Estrutura Sintática	Taxonomia
<i>Sustancia sospechosa</i> (substância suspeita)	Subst. base + adjetivo colocado	Adjetiva
<i>Actividad sospechosa</i> (atividade suspeita)	Subst. base + adjetivo colocado	Adjetiva
<i>Vehículo sospechoso</i> (veículo suspeito)	Subst. base + adjetivo colocado	Adjetiva
<i>Ahogamiento sospechoso</i> (afogamento suspeito)	Subst. base + adjetivo colocado	Adjetiva
<i>Elemento sospechoso</i> (elemento suspeito)	Subst. base + adjetivo colocado	Adjetiva

Fonte: Próprio da pesquisa

Como pode ser observado, ao compararmos os resultados obtidos nos *corpora* de estudo, as colocações apresentam a mesma estrutura sintática nos três idiomas, indicando, assim, um traço comum entre os *corpora*. A estrutura apresentada pelas colocações adjetivas vai ao encontro da proposta de Orenha-Ottaiano (2009) e do exposto em Caldas (2017), uma vez que o adjetivo atua como colocado, enquanto que o substantivo, a palavra de busca, nesses casos, atua como base.

Contudo, como pode se observar pelos exemplos encontrados no CCE, embora as colocações apresentem a mesma estrutura proposta por Orenha-Ottaiano (2009), a palavra *sospechoso* não atua como base da colocação, uma vez que, nos exemplos apresentados no quadro 9 (como *elemento sospechoso*, *ahogamiento sospechoso*) *sospechoso* é classificado como um adjetivo e não como um substantivo. Dessa forma, a palavra de busca, em língua espanhola, tende a atuar ora como base, ora como colocado nas colocações especializadas encontradas, conforme poderá ser observado pela alternância de sua função dentro da colocação pelos exemplos apresentados anteriormente.

Tal alternância de função do termo de busca dentro das colocações, aparecendo ora como base, ora como colocado, já havia sido apontada nas colocações especializadas da área criminal em língua inglesa por Caldas (2017). Sendo assim, os dados obtidos pela pesquisa atual, além de corroborar o exposto em 2017, amplia a ocorrência dessa característica para a língua espanhola e para a língua portuguesa, conforme pôde ser observado pelos exemplos já mencionados e pelos exemplos que são apresentados na sequência.

Quanto às colocações especializadas classificadas como nominais, conforme pode ser observado nos exemplos que apresentamos abaixo, temos uma prevalência de sintagmas preposicionados.

Quadro 10: Colocações Especializadas Nominais – *Corpus* Inglês

<b>Colocações Especializadas Nominais CCI</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Type of unsub</i> (tipo de suspeito)	Subst base + prep + subst colocado	Nominal
<i>Projection of the unsub</i> (projeção do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Sketch of the unsub</i> (retrato-falado do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Intelligence of the unsub</i> (inteligência do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Kind of the unsub</i> (tipo de suspeito)	Subst base + prep + subst colocado	Nominal
<i>Description of the unsub</i> (descrição do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Perspective of the unsub</i> (perspectiva do suspeito)	Subst base + prep + subst. colocado	Nominal
<i>Preference of the unsub</i> (preferência do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Desire of the unsub</i> (desejo do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Profile of the unsub</i> (perfil do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Team of unsub</i> (equipe do suspeito)	Subst base + prep + subst colocado	Nominal
<i>Connection to the unsub</i> (conexão com o suspeito)	Subst base + prep + subst. colocado	Nominal
<i>Group of unsubs</i> (grupo de suspeitos)	Subst base + prep + subst colocado	Nominal

Fonte: Próprio da pesquisa

Quadro 11: Colocações Especializadas Nominais no CCP - *Suspeito* como colocado

<b>Colocações Especializadas Nominais CCP – Suspeito como colocado</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Lista de suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Perspectiva do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Maturação do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Artifício do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Revólver do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Psique do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Psicose do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Prosecução do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Impressão do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Carro do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Arma do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Razão do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Grupo de suspeitos</i>	Subst base + prep + subst colocado	Nominal
<i>Mente do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Confissão do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Preferência do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Veículo do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Perfil do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Stress do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Tortura do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Descrição do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Troféu do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Perseguição do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Munição do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Mentalidade do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Eliminação de suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Atos do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Suspeito da morte</i>	Subst base + prep + subst colocado	Nominal
<i>Motivos do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Comportamento do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Habilidade do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Objetivo do suspeito</i>	Subst base + prep + subst colocado	Nominal
<i>Arsenal do suspeito</i>	Subst base + prep + subst colocado	Nominal

Fonte: Prório da pesquisa

Quadro 12: Colocações Especializadas Nominais no CCP - *Suspeito* como base

<b>Colocações Especializadas Nominais CCP</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Suspeito de assassinato</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de homicídio</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de crime</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito da policia</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito do caso</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito da investigação</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de estupro</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de abuso</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de roubo</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de assalto</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de rapto</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito de tráfico</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito na cena</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito no caso</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito no assassinato</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito em movimento</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito em ação</i>	Subst base + prep. + subst. colocado	Nominal
<i>Suspeito em potencial</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito em custódia</i>	Subst base + prep. + subst colocado	Nominal
<i>Suspeito em uma investigação</i>	Subst base + prep. + subst colocado	Nominal

Fonte: Próprio da pesquisa

Quadro 13: Colocações Especializadas Nominais – *Corpus Espanhol*

<b>Colocações Especializadas Nominais CCE</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Sospechoso de asesinato</i> (suspeito de assassinato)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso de robos</i> (suspeito de roubos)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso del crimen</i> (suspeito do crime)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso de homicidio</i> (suspeito de homicídio)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso de los atracos</i> (suspeito de roubos)	Subst base + prep + subst colocado	Nominal
<i>Sospecho del atropellamiento</i> (suspeito de atropelamento)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso en una investigación</i> (suspeito em uma investigação)	Subst base + prep + subst colocado	Nominal
<i>Sospechoso en custodia</i> (suspeito sob custódia)	Subst base + prep + subst colocado	Nominal

Fonte: Próprio da pesquisa

Conforme se observa nos quadros anteriores, as colocações especializadas nominais levantadas são formadas por sintagmas nominais preposicionados, contando com a utilização das palavras *unsub* e *sospechoso* como base em todas as colocações levantadas. Já, nas colocações encontradas no CCP, o termo *suspeito* aparece ora como base (*suspeito de assassinato, suspeito de homicídio*), ora como colocado das colocações (*lista de suspeitos, descrição do suspeito*). A presença do sintagma preposicionado nos três *corpora* de estudo indica uma semelhança na estrutura dos idiomas nos casos analisados.

Já, no que tange às colocações especializadas verbais, foi possível também identificar colocações em que a palavra de busca exerce a função de objeto, conforme pode ser observado nos quadros abaixo.

Quadro 14: Colocações Especializadas Verbais – *Corpus* Inglês

Colocações Especializadas Verbais CCI		
Colocação	Estrutura Sintática	Taxonomia
<i>Believe the unsub</i> (acreditar no suspeito)	Verbo base + det + subst colocado	Verbal
<i>Profile the unsub</i> (perfilar o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Drive the unsub</i> (dirigir o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Catch the unsub</i> (pegar o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Meet the unsub</i> (conhecer o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Know the unsub</i> (conhecer o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Find the unsub</i> (encontrar o suspeito)	Verbo base + det + subst colocado	Verbal
<i>Deal with an unsub</i> (lidar com o suspeito)	Verbo frasal base + det + subst colocado	Verbal

Fonte: Próprio da pesquisa

Quadro 15: Colocações Especializadas Verbais – *Corpus* Português

<b>Colocações Especializadas Verbais CCP</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Descrever o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Analisar o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Identificar o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Interrogar o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Perseguir o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Prender o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Procurar um suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Encontrar o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Achar o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Levar ao suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Rastrear o suspeito</i>	Verbo base + prep. + susbt colocado	Verbal
<i>Lidar com suspeito</i>	Verbo base + prep. + susbt colocado	Verbal

Fonte: Próprio da pesquisa

Quadro 16: Colocações Especializadas Verbais – *Corpus* Espanhol

<b>Colocações Especializadas Verbais CCE</b>		
<b>Colocação</b>	<b>Estrutura Sintática</b>	<b>Taxonomia</b>
<i>Haber sospechoso</i> (haver suspeito)	Verbo base + prep. + susbt colocado	Verbal
<i>Parecer sospechoso</i> (parecer suspeito)	Verbo base + prep. + susbt colocado	Verbal
<i>Tener sospechoso</i> (ter suspeito)	Verbo base + prep. + susbt colocado	Verbal

Fonte: Próprio da pesquisa

Ao olhar para os quadros expostos, percebemos que há a predominância dos termos *unsub*, *suspeito* e *sospechoso* como colocados das colocações especializadas. Com relação às colocações especializadas verbais encontradas no CCE, apresentamos apenas 3 exemplos, uma vez que não foi possível identificarmos mais nenhuma colocação desse tipo no *corpus* de estudo em espanhol, no que diz respeito às colocações compostas pela palavra-chave *sospechoso*, traduzidas a partir de *unsub*. Contudo, ao considerarmos as outras quatro palavras-chave que foram selecionadas para o levantamento e a análise das colocações, foi possível identificar 46 colocações verbais no CCE, número bem próximo do encontrado nos demais *corpora*.

Todavia, no que tange às colocações verbais formadas por *sospechoso*, apesar de ocorrer em menor número, apresentam a mesma formação, tendo um verbo como colocado e o substantivo *sospechoso* – a base da colocação – atuando como objeto do verbo em questão, sendo que essa formação também pode ser observada nas colocações obtidas no CCI e no CCP. Há, ainda, uma única colocação, apenas, composta por um verbo frasal (*phrasal*

*verb*) em língua inglesa (*deal with*).

Ademais, foi possível perceber que as colocações especializadas do âmbito criminal tendem a apresentar estrutura semelhante nos três idiomas que compõem os *corpora* de estudo, havendo, assim, uma proximidade morfossintática entre elas.

No que tange ao comportamento do termo sob análise, atuando ora como base, ora como colocado nas colocações obtidas, já identificado por Caldas (2017), ressaltamos que tal fato não foi levado em consideração ao propormos a amostra de dicionário apresentada no capítulo 6, tendo as colocações sido organizadas sob entradas delimitadas a partir das palavras-chave sob análise, não considerando assim seu papel em cada uma das colocações (base e colocado), a fim de facilitar a sua localização por parte dos consulentes.

Na seção seguinte, apresentamos a análise léxico-semântica desenvolvida a partir das colocações especializadas levantadas, além de discutirmos a utilização do termo *suspeito* para a tradução dos termos *unsub* e *suspect*.

### 5.3 Análise léxico-semântica

Ao analisarmos as colocações levantadas, a fim de que estas possam ser incluídas na amostra de dicionário que apresentamos ao final desta pesquisa, buscamos evidenciar informações relevantes a respeito delas, tais como informações referentes a morfossintaxe e a frequência de ocorrência. Nesta seção, abordamos as questões léxico-semânticas dos termos que foram selecionados como palavras de busca e das colocações especializadas obtidas.

#### 5.3.1 O termo *unsub*

Começamos esta seção pelo termo *unsub*, que figura na lista de palavras-chave com um índice de chavidade bem alto (259.500), além de ocorrer 2699 vezes no CCI, *corpus* utilizado como ponto de partida para o levantamento de todas as colocações nesta pesquisa.

Uma vez que esse termo é utilizado para se referir aos suspeitos nas três séries que compõem os *corpora* de estudo, buscamos informações sobre sua definição em dicionários on-line a fim de comprovar esse uso. A partir da busca realizada, não foi possível identificar uma entrada para o referido termo no dicionário on-line *Cambridge Dictionary*. Por sua vez, o site *Lexico* apresenta as seguintes definições para o termo *unsub*:

Figura 10: Print da tela do site *Lexico* com definição do termo *unsub*

Main meanings of **unsub** in English: [unsub<sup>1</sup>](#) [unsub<sup>2</sup>](#)

# unsub<sup>1</sup>

Pronunciation  /'ʌnsʌb/ 

---

**VERB** (unsubs, unsubbing, unsubbed)

[NO OBJECT]

*informal*  
Unsubscribe.

[+ More example sentences](#)

---

**Origin**  
1990s abbreviation.

Fonte: <https://www.lexico.com/definition/unsub> acesso em 17 de maio de 2021.

Figura 11: Print da tela do site *Lexico* com definição do termo *unsub*

Main meanings of **unsub** in English: [unsub<sup>1</sup>](#) [unsub<sup>2</sup>](#)

# unsub<sup>2</sup>



---

**NOUN**

*informal US*  
(in police use) a person of unknown identity who is the subject of a criminal investigation.

[+ More example sentences](#)

---

**Origin**  
1960s abbreviation of unknown subject or unidentified subject.

---

**Pronunciation** 

**unsub** /'ʌnsʌb/ 

Fonte: <https://www.lexico.com/definition/unsub> acesso em 17 de maio de 2021.

Como se pode observar, essa página traz duas definições para o termo *unsub*, sendo que a primeira não se encaixa no contexto em que as colocações são utilizadas. A segunda definição, contudo, se encaixa perfeitamente no contexto em que as colocações são utilizadas nas séries de TV que compõem os *corpora* de pesquisa, haja vista que, ainda que receba uma marcação de uso informal, se refere a “uma pessoa de identidade desconhecida que é suspeita em uma investigação criminal”<sup>62</sup> (LEXICO, 2021).

Outro dado que convém ressaltar é a explicação que a página traz para a origem do termo, que surgiu como uma abreviação de *unknown subject* ou *unidentified subject* (ambos podendo ser traduzidos para o português como *suspeito desconhecido* ou *suspeito não identificado*, respectivamente) e passou a ser convencionalmente utilizado no meio criminal, tanto que foi inserido entre os termos do âmbito criminal utilizados para comporem os textos de séries que buscam retratar o trabalho desenvolvido nessa área.

Por sua vez, o site *Merriam-Webster* apresenta a seguinte definição para o termo *unsub*:

---

<sup>62</sup> “[...] a person of unknow identity who is the subject of a criminal investigation”.

Figura 12: Print da tela do site *Merriam-Webster* com a definição do termo *unsub*.

The screenshot shows the Merriam-Webster website interface. At the top, there is a navigation bar with links for 'GAMES & QUIZZES', 'THESAURUS', 'WORD OF THE DAY', 'FEATURES', and 'SHOP'. The Merriam-Webster logo and 'SINCE 1828' are on the left. A search bar contains the word 'unsub', with 'Dictionary' and 'Thesaurus' tabs below it. The main content area displays the word 'unsub' as a noun, with a 'Save Word' button, phonetic transcription 'un-sub | \ 'ən-, səb', and plural 'unsubs'. The definition is 'US law enforcement, informal', followed by the example 'the unknown perpetrator of a crime'. Two usage examples are provided: one from a TV guide about 'Criminal Minds' and another from Jeffery Deaver's work. The 'First Known Use of unsub' is noted as 1965, in the meaning defined above.

Fonte: <https://www.merriam-webster.com/dictionary/unsub>. Acesso em: 17 de maio 2021.

Para o dicionário *Merriam-Webster*, o termo *unsub*, além de ser classificado como um vocábulo informal, é visto como pertencente ao vocabulário da “aplicação da lei dos EUA”<sup>63</sup> (MERRIAM-WEBSTER, 2021), enfatizando, assim, o seu pertencimento ao âmbito criminal, tal como denominada nesta pesquisa. No que tange à sua definição, o dicionário menciona que este termo se refere ao “o autor desconhecido de um crime”<sup>64</sup> (idem) e, a título de exemplificação, traz uma sentença que faz menção à série *Criminal Minds* (uma das séries que compõem nosso *corpus* de estudo) e que foi retirada de um guia de TV, além de uma sentença escrita por Jeffery Deaver, jornalista norte-americano que, após cursar a faculdade de Direito, se dedicou à escrita de romances policiais.

Outro dado apresentado na imagem e que chama a atenção é a data em que foi registrado o primeiro uso do termo com o significado apresentado: 1965. Ou seja, o

<sup>63</sup> “*US law enforcement*”

<sup>64</sup> “*the unknown perpetrator of a crime*”

termo, embora não figure como uma entrada no dicionário on-line *Cambridge Dictionary*, de fato vem sendo utilizado, há quase 60 anos, para se referir a um suspeito, tal como identificamos em nossos *corpora* de estudo, e além de ser uma marca de informalidade, como ressalta o site *Lexico*, é também utilizado no vocabulário técnico das forças da lei dos Estados Unidos, fatos que corroboram a inserção deste termo em nossa pesquisa.

Quanto à diferença entre o número de colocações especializadas formadas por *unsub* e o de colocações especializadas formadas por *suspeito*, respectivamente em línguas inglesa e portuguesa, constatamos que essa se deu pela utilização do mesmo termo em língua portuguesa – *suspeito* – para traduzir colocações formadas por *unsub* e por *suspect*, tal como pode ser observado nos exemplos a seguir.

- *Perspectiva do suspeito:*

Neste caso, a colocação especializada em língua portuguesa foi utilizada para traduzir uma colocação em língua inglesa formada pelo termo *unsub: perspective of the unsub*:

- *All right, so we need to start over, go back over both Gail and Molly's cases and look at everything from the perspective of a female unsub.* (Exemplo extraído do CCI).

- “Tudo bem, por isso precisamos recomeçar, voltar a tanto aos casos de Gail e Molly olhar para tudo a partir da perspectiva de uma suspeita feminina”. (Exemplo extraído do CCP).

- *Fazer de alguém suspeito*

Essa colocação, extraída do CCP, embora contenha o termo *suspeito* como base, foi utilizada para traduzir a colocação *make someone a suspect* encontrada no CCI, conforme mostram os exemplos abaixo:

- *Your video is very similar to the one the killer left behind.*

- *So that makes me a suspect?* (Exemplo extraído do CCI).

- “O seu vídeo é muito similar ao que o assassino deixou.

- Isso faz de mim um suspeito?” (Exemplo extraído do CCP).

Uma vez que as colocações em língua portuguesa foram traduzidas a partir de colocações em língua inglesa formadas por diferentes termos, em nossa amostra de dicionário pedagógico proposta ao final desta pesquisa, procuramos trazer exemplos de colocações formadas tanto por *unsub* quanto por *suspect*, a fim de enriquecer o repertório lexical apresentado.

A seguir, apresentamos algumas colocações como exemplo da análise léxico-semântica empreendida. As colocações que são apresentadas a seguir foram selecionadas tomando como base seu nível de transparência, ou seja, aquelas com significado menos transparente, ou que podem acarretar dúvidas aos estudantes das línguas sob análise.

### 5.3.2 *Drive the unsub*

A colocação especializada *drive the unsub* teve sete ocorrências encontradas no CCI, nosso Corpus Criminal Inglês, e o valor de 7,7 para seu *LogDice*, foi selecionada para integrar essa análise pois, embora possa parecer simples e completamente transparente, seu significado pode não ser totalmente claro para estudantes dos cursos de Licenciatura em Letras e de Bacharelado em Letras que, por ventura, ainda se encontrem nos anos iniciais dos referidos cursos.

Esses alunos podem entender a colocação como *dirigir o suspeito*, fazendo uma tradução literal que pode não ser adequada ao contexto em que a colocação é utilizada. Vejamos o exemplo a seguir:

*That responsibility is probably what **drove the unsub** into action.*  
(Fonte: <http://www.tvsubtitles.net/>).

Como se pode observar pelo exemplo acima, ao traduzir a colocação por *dirigiu o suspeito* o aluno pode interpretar que a responsabilidade dirigiu o suspeito até aquela ação. Entretanto, essa tradução literal não é a forma mais utilizada por falantes de português, que dificilmente afirmam que *algo dirigiu alguém a desenvolver uma dada ação*. Nesse caso, a tradução mais adequada seria *levou o suspeito à*, tal como apresentado no exemplo a seguir, extraído do CCP.

*Essa responsabilidade é, provavelmente, o que levou o unsub em ação.* (Fonte: <http://www.tvsubtitles.net/>).

Tal como se observa, o exemplo apresenta a tradução encontrada no *corpus* de língua portuguesa e apresenta a construção *levou o unsub*, a qual nos mostra dois fatores interessantes: primeiro, a tradução de *drive* para *levar* (levou, no caso), já que esse verbo mostra de forma mais clara o significado do verbo *drive* originalmente utilizado em

língua inglesa; segundo, a presença do termo *unsub* na própria língua inglesa, sem que tenha sido traduzido para suspeito. Esse último fator pode ser explicado pela noção de empréstimo linguístico, de forma que se mantém o termo em língua de partida, no caso, em língua inglesa, para a língua de chegada, a língua portuguesa.

De toda forma, o que mais nos interessa é a tradução de *drive* por *levou*, tradução essa que nos parece mais próxima da realidade dos falantes brasileiros ao se considerar o significado e o contexto de uso da expressão.

### 5.3.3 *Profile the unsub*

A colocação *profile the unsub*, que ocorre 7 vezes no *corpus* de língua inglesa e apresenta o valor de *LogDice* de 8,4 e, apesar de ter certo nível de transparência<sup>65</sup>, pode causar dúvidas para alunos e demais pessoas que se interessem pela língua inglesa, mas que não estejam familiarizados com a área de investigação criminal. Como apresentado na seção anterior, a colocação é formada por um verbo colocado, um determinante e o substantivo base.

Assim, ao buscarmos pelo verbo *profile* no dicionário on-line *Cambridge Dictionary*, encontramos a palavra *profile* definida, apenas, como um substantivo, o qual se refere ao *perfil*, tanto enquanto uma descrição de alguém ou de alguma empresa, quanto a visão lateral do rosto de alguém, conforme pode ser observado na imagem abaixo:

---

<sup>65</sup> Nível de transparência aqui é entendido como significado transparente e utilizado para se referir a colocações que, normalmente, são de fácil entendimento a partir da compreensão das palavras que a compõem, não causando, assim, dúvidas quanto ao seu significado, tampouco apresentando significado mais composicional.

Figura 13: Print da definição de *profile* do *Cambridge Dictionary* na direção inglês-português.



Fonte: Cambridge Dictionary. Disponível em: <https://dictionary.cambridge.org/pt/dicionario/ingles-portugues/profile>. Acesso em 12 dez 2021.

Contudo, quando comparamos as definições apresentadas pelo referido dicionário com um exemplo extraído do CCI, percebemos que estas definições não são suficientes para ajudar na compreensão do aluno. Veja o exemplo abaixo:

- *Guys, something doesn't add up. We **profiled this unsub** would have an intimate knowledge of his kill zone. Right?* (Fonte: <http://www.tvsubtitles.net/>).

De acordo com o exemplo apresentado, a palavra *profile* é utilizada como um verbo, tal como mencionado na seção de análise sintática. Nesse caso, o aluno que não tem muita experiência com o idioma, ou com a área em questão, chegaria a uma tradução literal incoerente com o apresentado no exemplo anterior, obtendo algo como ‘*Pessoal, algo não se encaixa. Nós perfil deste suspeito teria um conhecimento intimo desta zona de morte*’. O aluno acabaria percebendo o estranhamento e correria o risco de se frustrar com a aprendizagem do idioma, ou com a aprendizagem dos processos de tradução, prejudicando assim o seu progresso.

Na verdade, a tradução adequada para a colocação *profile the unsub* seria *perfilar o suspeito*, ou seja, fazer, construir o perfil do suspeito com base nas informações que se

conhecem dele, como pode ser observado nos exemplos a seguir:

- *We did profile the unsub would be someone close to the investigation.*
- *But only Ruiz and I heard him say it.* (Fonte: <http://www.tvsubtitles.net/> ).
- *We profiled that the unsub may have had his daughter taken away from him.* (Fonte: <http://www.tvsubtitles.net/> ).
- *We profiled the unsub. And we were able to locate her before he harmed her.* (Fonte: <http://www.tvsubtitles.net/> ).
- *We profiled the unsub felt limited in his job.*
- *What's more powerless than being a bit player on the Washington stage?* (Fonte: <http://www.tvsubtitles.net/> ).

Essa tradução se mostra a mais adequada ao observamos os exemplos apresentados, embora esta não tenha sido utilizada pelos tradutores de maneira homogênea nas legendas em língua portuguesa e em língua espanhola, tal como pode ser observado abaixo:

- *Nós fizemos o perfil do unsub como alguém que poderia ser próximo à investigação.*
- *Mas só Ruiz e eu ouvimos ele dizer isso.* (Fonte: <http://www.tvsubtitles.net/> ).
- *Fizemos o perfil do suspeito limitado no trabalho dele.*
- *O que é mais impotente do que ser uma pequena parte do palco de Washington?* (Fonte: <http://www.tvsubtitles.net/> ).
- *Perfilamos que el sospechoso pude haber perdido a su hija.* (Fonte: <http://www.tvsubtitles.net/> ).
- *Hicimos el perfil del sospechoso y la localizamos antes de que le*

*hiciera daño*. (Fonte: <http://www.tvsubtitles.net/> ).

Vemos, assim, uma inconsistência nas traduções, sendo que estas não afetam a compreensão dos expectadores/leitores das legendas, uma vez que a opção utilizada nas traduções transmite o significado da colocação original em língua inglesa. Dessa forma, para a elaboração da proposta de dicionário apresentada no capítulo seguinte, utilizaremos como tradução *perflar o suspeito*, para língua portuguesa, e *perflar el sospechoso*, para língua espanhola.

#### 5.3.4 *Sketch of the unsub*

A colocação especializada *sketch of the unsub* ocorre três vezes em nosso *corpus* de língua inglesa, o CCI, e tem um valor de *LogDice* de 10,6, indicando assim que é uma colocação forte, já que as palavras que a compõem tendem a se combinar com mais frequência, do que com outras palavras, motivo pelo qual foi selecionada para esta análise.

Ademais, ela pode causar dificuldade de compreensão para os alunos de língua inglesa, pois o seu significado, o conceito que ela evoca, pode não ser de fácil compreensão, principalmente se, ao fazer uma tradução literal, os alunos forem levados a pensar em um *esboço do suspeito*, por exemplo.

Devemos levar em consideração que, apesar da lexia *sketch* ser traduzida para a língua portuguesa como *esboço*, ou *esboçar* (de acordo com o *Cambridge Dictionary*), e para a língua espanhola como *boceto* ou *dibujar* (também de acordo com o dicionário *Cambridge Dictionary*), a compreensão da colocação não deve levar em conta apenas o significado isolado desta lexia, mas sim o contexto em que ela é utilizada.

Nesse sentido, apresentamos os três exemplos retirados do CCI em que ocorrem a colocação sob análise

- *You need to get a **sketch of the unsub** out to every camera shop in buffalo*. (Fonte: <http://www.tvsubtitles.net/>).

- *Composite **sketch of the unsub***.

- *Garcia's running facial recognition against her list of blue van owners*. (Fonte: <http://www.tvsubtitles.net/> ).

- *Composite sketch of the unsub. I already went wide to the press.*

(Fonte: <http://www.tvsubtitles.net/> ).

Ao considerarmos o contexto específico em que a colocação é usada, percebemos que a versão mais adequada para a mesma em língua portuguesa seria *retrato falado do suspeito*, a qual ocorre apenas uma vez no CCP; já, na língua espanhola, teríamos a colocação *boceto del sospechoso*, que ocorre apenas uma vez no CCE.

- *Disponibilizamos uma linha de emergência e um **retrato-falado do suspeito** será divulgado. Quem tenha qualquer informação sobre o atirador entre em contato imediatamente.* (Fonte: <http://www.tvsubtitles.net/> ).

- *Todos recibieron copias del **boceto del sospechoso**. Deben distribuirlos a todos los oficiales de la ley en el área de los tres estados.* (Fonte: <http://www.tvsubtitles.net/> ).

Como podemos ver, além da diferença de ocorrências entre os *corpora*, não foi possível identificarmos as traduções dadas para todas as sentenças em que a colocação foi utilizada no CCI, o que se explica por não ter sido possível encontrarmos todas as legendas das séries que compõem os *corpora* em língua portuguesa e em língua espanhola, conforme mencionado anteriormente. Contudo, as colocações encontradas no CCP e no CCE nos mostram o real significado da colocação em questão.

Ademais, ao compararmos os exemplos extraídos dos três *corpora*, percebemos uma semelhança enquanto a estrutura morfossintática, uma vez que todos os exemplos são constituídos por um sintagma nominal preposicionado.

### 5.3.5 *Airtight alibi*

A colocação *airtight alibi*, que ocorre dez vezes no CCI e tem *LogDice* de 12,4, indicando que é uma combinação forte, sendo que as palavras que a compõem não tendem a combinar com outras palavras, é apresentada aqui devido ao seu colocado, *airtight*, uma

palavra que pode causar dúvidas no momento da tradução para as línguas portuguesa e espanhola, interferindo na compreensão do significado da colocação em questão.

*Airtight* é traduzida, tanto para o português, quanto para o espanhol, como *hermético* pelo dicionário online *Cambridge Dictionary*, sendo que o exemplo utilizado pelo dicionário alude a um contêiner fechado hermeticamente (CAMBRIDGE DICIONARY, 2022). Ora, ao consultar esse dicionário, o estudante de tradução pode ser levado a traduzir a colocação *airtight alibi* por *álibi hermético*. Contudo, como se verificou em nosso *corpus* de estudo, não há registros dessa colocação em língua portuguesa ou em língua espanhola, como se pode observar pelos exemplos apresentados a seguir.

- *Garcia, has the boyfriend been questioned by local P.D.?*  
 - *He has by phone. He's got an **airtight alibi**. Him and his band have been on tour to Asia for the last month.* (Fonte: <http://www.tvsubtitles.net/>).

- *Garcia, o namorado foi interrogado pela polícia local?*  
 - *Foi, por telefone. Tem um **álibi muito forte**. Ele e sua banda estão em turnê pela Ásia há um mês.* (Fonte: <http://www.tvsubtitles.net/>).

Os exemplos acima mostram que a escolha feita pelos tradutores, nesse caso, foi em traduzir a colocação especializada *airtight alibi* como um *álibi muito forte*, para a língua portuguesa. Já para a língua espanhola, não foi possível encontrar a correspondência exata para a tradução desse trecho. Apesar de ser uma tradução muito boa, parece que o *muito forte* não dá conta de toda a carga semântica de *airtight*, considerando-se a tradução apresentada pelo dicionário *Cambridge Dictionary*.

Nesse sentido, apresentamos mais dois exemplos de traduções que foram empregadas para essa colocação e que foram identificados no *corpus* de estudo.

- *Very public arguments. Kaiser even makes threats.*  
 - *However, **airtight alibi**. In New York at the time of the murder. Just got back this morning.* (Fonte: <http://www.tvsubtitles.net/>).

- *Discussões públicas. Kaiser até fez ameaças.*  
 - *Entretanto, **álibi consistente**. Em Nova Iorque na hora do*

*homicídio. Voltou hoje pela manhã.* (Fonte: <http://www.tvsubtitles.net/>).

- *Debates públicos. Kaiser incluso hizo amenazas.*

- *Sin embargo, una **fuerte coartada**. En Nueva York a la hora del crimen. Volvió esta mañana.* (Fonte: <http://www.tvsubtitles.net/>).

Nesse caso, ao identificarmos a correspondência de traduções para o mesmo trecho nos três *corpora*, constatamos uma nova tradução em língua portuguesa para a colocação em questão (*álibi consistente*) e uma versão em língua espanhola que se equipara a primeira versão apresentada em língua portuguesa, no exemplo anterior, *fuerte coartada*.

Já nos trechos apresentados abaixo, temos uma correspondência de tradução entre as línguas portuguesa e espanhola para a colocação *airtight alibi*. Observe:

- *GPS records show that Sikes was 20 miles away from the crime scene at the time of the murder.*

- *He's got an **airtight alibi**, the US government.* (Fonte: <http://www.tvsubtitles.net/>).

- *Os registros do GPS mostram que Sikes estava a 30 quilômetros da cena do crime na hora do assassinato.*

- *Ele tem um **álibi irrefutável**, o governo dos Estados Unidos.* (Fonte: <http://www.tvsubtitles.net/>).

- *Las grabaciones del GPS muestran que Sikes estaba 20 millas de distancia del lugar del crimen en el momento del asesinato.*

- *Tiene una **coartada irrefutable**, el gobierno de los EE.UU.* (Fonte: <http://www.tvsubtitles.net/>).

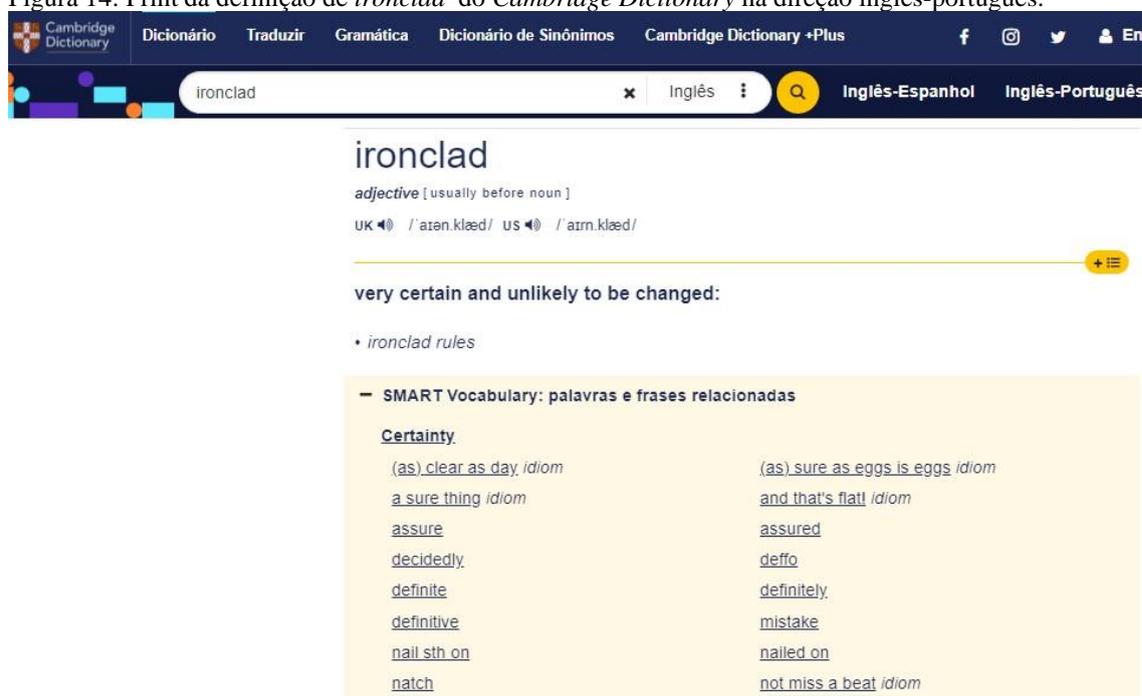
Como se pode observar por meio dos exemplos apresentados, existe uma variedade de traduções possíveis para a colocação *airtight alibi*. Entretanto, ao analisarmos o exposto aqui, nos parece que as colocações *álibi irrefutável*, em língua portuguesa, e *coartada irrefutable*, em língua espanhola, transmitem melhor a carga semântica contida na palavra originalmente utilizada em língua inglesa, a saber: *hermético*. Dessa forma, ao apresentarmos

a proposta de dicionário no capítulo seguinte, traremos essa tradução (*álibi irrefutável* e *coartada irrefutável*) e apresentamos as demais como versões possíveis.

### 5.3.6 *Ironclad alibi*

A colocação *ironclad alibi*, que ocorre três vezes no CCI e tem *LogDice* de 10,9, é formada por um colocado que não apresenta tradução para o português no *Cambridge Dictionary*, tal como pode ser observado na imagem abaixo:

Figura 14: Print da definição de *ironclad* do *Cambridge Dictionary* na direção inglês-português.



Fonte: Cambridge Dictionary. Disponível em: <https://dictionary.cambridge.org/pt/dicionario/ingles-portugues/profile>. Acesso em 12 dez 2021.

Uma vez que não existe a entrada em língua portuguesa, o aluno consulente deste dicionário precisa depreender o significado desse colocado a partir do exposto na entrada em língua inglesa, apresentada acima. De acordo com a entrada, *ironclad* se refere a algo que é “muito certo ou pouco provável de ser alterado”<sup>66</sup> (CAMBRIDGE DICTIONARY, 2021). Entretanto, empregar esse significado à colocação pode causar confusão aos alunos dos cursos de licenciatura e de bacharelado em Letras que se encontram nos anos iniciais dos cursos e àqueles que não estão familiarizados com o âmbito criminal.

Assim, apresentamos os seguintes exemplos encontrados no CCI e suas respectivas

<sup>66</sup> “very certain and unlikely to be changed”.

correspondências, encontradas no CCP e no CCE:

*Well, Brass' gut told him that the brother was involved, but the guy had an **ironclad alibi**.* (Fonte: <http://www.tvsubtitles.net/>).

- *Hey! The DNA matched. Laura Seeger has an **ironclad alibi**.*

- *I know.* (Fonte: <http://www.tvsubtitles.net/>).

- *Laura Seeger who'd get diddly if her divorce went through. Green was in D.C. with witnesses.*

- *Seeger was in the bank on video.*

*How can our two great suspects both have **ironclad alibis**?* (Fonte: <http://www.tvsubtitles.net/>).

- *O instinto de Brass dizia que o irmão estava envolvido, mas o cara tinha um **álibi incontestável**.*

(Fonte: <http://www.tvsubtitles.net/>).

- *Ei! O DNA combina. Laura Seeger tem **álibi incontestável**.*

- *Eu sei.* (Fonte: <http://www.tvsubtitles.net/>).

- *El instinto de Brass le dictaba que el hermano estaba involucrado pero el tipo tenía una **coartada irrefutable**.*

(Fonte: <http://www.tvsubtitles.net/>).

Como se pode observar, foram encontradas as traduções correspondentes de dois trechos em língua portuguesa, o primeiro e o segundo, e apenas uma correspondência em língua espanhola. No que tange as versões em língua portuguesa, podemos observar a predominância de uma escolha tradutória para a colocação *ironclad alibi*, a saber: *álibi incontestável*, enquanto que na versão em língua espanhola identificada no CCE, temos o emprego de *coartada irrefutable* como tradução de *ironclad alibi*.

Essa colocação em língua espanhola é a mesma empregada para a tradução da colocação *airtight alibi*, o que nos leva a considerar duas hipóteses: i) para os tradutores, não existe uma diferença clara entre a *ironclad* e *airtight*, assumindo que ambas as palavras se referem ao *álibi* que, de fato, é irrefutável, não pode ser questionado; ii) não encontrando

uma palavra melhor, no momento da tradução, ou devido à falta de repertório lexical, os tradutores optaram por traduzi-las da mesma maneira. Ao considerar que ambas as palavras fazem referência à um álibi incontestável, irrefutável, existe uma equiparação de ambas as colocações especializadas na língua de partida a uma única colocação na língua de chegada, ou seja, ambas são vistas como quase sinônimas, fato que pode facilitar a compreensão e utilização das colocações por parte de aprendizes.

Já com relação à escolha tradutória para a língua portuguesa, essa nos parece mais fiel ao sentido que a própria palavra *ironclad* traz à colocação, sendo que *incontestável* abarca o significado expresso pelo *Cambridge Dictionary*, o de ser um álibi muito certo, que pouco provavelmente será alterado ou questionado. Em se tratando da diferença de significado presente entre *airtight alibi* e *ironclad alibi*, a julgar pelos exemplos encontrados nos *corpora* de estudo e suas respectivas versões em língua portuguesa e língua espanhola, acreditamos que elas diferenciam-se na medida em que a primeira, *airtight alibi*, refere-se a um álibi extremamente perfeito, do qual não se pode encontrar nenhuma abertura ou falha, enquanto a segunda, *ironclad alibi*, diz respeito a um álibi que, apesar de ser incontestável, pode ser que sofra, em algum momento ou por alguma razão, algum questionamento.

### 5.3.7 *Profile of the unsub* x *Profile the unsub*

Apresentamos aqui uma comparação entre as colocações *profile of the unsub* e *profile the unsub*, apesar de a colocação *profile the unsub* (que ocorre sete vezes no CCI e tem *LogDice* de 8,4) ter sido analisada no item 5.3.3.

A colocação *profile of the unsub* ocorre apenas duas vezes no CCI e tem *LogDice* de 9,4 diferencia-se da colocação *profile the unsub* presença da preposição *of*, enquanto é observada a ausência dessa preposição na segunda colocação. O fato de a primeira ser um sintagma preposicionado indica que o colocado, *profile*, é um substantivo, enquanto na segunda colocação, como não há um sintagma preposicionado, temos *profile* como verbo, conforme explicado no item 5.3.2.

Essa diferenciação se faz necessária, principalmente, para estudantes dos anos iniciais dos cursos de Licenciatura e de Bacharelado em Letras, bem como para aqueles que não estão familiarizados com o âmbito criminal. A seguir, apresentamos alguns exemplos, extraídos do CCI, que ilustram a diferença entre ambas as colocações, bem como, suas respectivas versões encontradas no CCP e no CCI.

- *Lynn Dempsey was an executive assistant. She has no expertise with chemicals. She doesn't fit the **profile of the unsub**.* (Fonte: <http://www.tvsubtitles.net/> ).

- *Lynn Dempsey era uma auxiliar executiva. Ela não tinha experiência com químicos. E não se encaixa no **perfil de nosso suspeito**.* (Fonte: <http://www.tvsubtitles.net/> ).

- *Creo que nada en el **perfil del sospechoso** sugiere motivación racial.*

De acordo com os exemplos apresentados acima, a colocação *profile of the unsub*, formada por um sintagma nominal, corresponde ao *perfil do suspeito* em língua portuguesa e ao *perfil del sospechoso* em língua espanhola. Não apresentamos uma correspondência em língua espanhola desse mesmo trecho devido ao fato de esse *corpus* ser o menor e não ter sido possível encontrar a correspondência perfeita, bem como a diferença de uso entre os termos *unsub* e *suspect*, explicada no item 5.3.1. A seguir, apresentamos um exemplo, no qual temos o sintagma verbal *profile the unsub*.

- *We **profiled the unsub** felt limited in his job.*

- *What's more powerless than being a bit player on the Washington stage?* (Fonte: <http://www.tvsubtitles.net/> ).

- *Fizemos o **perfil do suspeito** limitado no trabalho dele.*

- *O que é mais impotente do que ser uma pequena parte do palco de Washington?* (Fonte: <http://www.tvsubtitles.net/> ).

- ***Perfilamos que el sospechoso** pude haber perdido a su hija.*

(Fonte: <http://www.tvsubtitles.net/> ).

Conforme ressaltamos no item 5.3.3, a colocação *profile the unsub* deveria ter sido traduzida por *perfilar o suspeito* em língua portuguesa, tal como aconteceu com a tradução em língua espanhola. Entretanto, apesar da utilização do verbo fazer para se referir ao perfil do suspeito, é evidente a diferença existente entre as colocações *profile of the unsub* e *profile*

*the unsub.*

Assim, finalizamos a análise léxico-semântica das colocações selecionadas e, no capítulo seguinte, ao abordarmos essas colocações na amostra de dicionário pedagógico que propomos, trazemos as duas colocações com suas respectivas informações.

## 6 Do Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal

Neste capítulo, empreendemos uma discussão acerca da macro e da microestrutura do dicionário, assim como discorremos acerca da classificação de dicionário pedagógico adotada para nos referirmos à amostra de dicionário aqui proposto. Dessa forma, este capítulo se divide em três partes: na primeira, discutimos sobre nossa proposta de dicionário pedagógico; na segunda, abordamos a macroestrutura do dicionário; e, na terceira, discorremos sobre a microestrutura adotada no Dicionário Pedagógico de Colocações Especializadas (DPCE).

### 6.1 Definindo o Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal

Ao longo deste trabalho, abordamos as questões teóricas e metodológicas a respeito da Teoria Funcional da Lexicografia, respectivamente nos capítulos 3 (Fundamentação Teórica) e 4 (Metodologia da Pesquisa). Neste subcapítulo, apresentamos as características lexicográficas que nos levam a classificar o dicionário ora proposto como sendo um Dicionário Pedagógico (doravante DP).

Ao considerarmos o proposto pela Teoria Funcional da Lexicografia, é possível dizer que a definição do público-alvo do dicionário é fundamental, bem como a definição de sua função social, tal como apontam Fuertes Olivera e Tarp (2008) ao afirmarem que para

determinar as necessidades dos usuários, estas devem relacionar-se não apenas a um determinado tipo de usuário (já que os diferentes tipos de usuários têm diferentes necessidades), mas também a um determinado tipo de situação social que inclusive influencia com maior peso que aquele da caracterização do tipo de necessidades<sup>67</sup> (FUERTES OLIVERA; TARP, 2008, p. 78)

Dessa forma, ao pensarmos na proposta do Dicionário Pedagógico de Colocações Especializadas da área Criminal (doravante DPCEC), definimos como público-alvo alunos dos cursos de Bacharelado em Letras e de Licenciatura em Letras, bem como tradutores profissionais, mais especificamente aqueles que se ocupam da tradução de séries

---

<sup>67</sup> [...] las necesidades de los usuarios, estas deben relacionarse no sólo a un determinado tipo de usuario (ya que los diferentes tipos de usuarios tienen diferentes necesidades), sino también a un determinado tipo de situación social que incluso influye con mayor peso que aquel en la caracterización del tipo de necesidades.

(legendagem e dublagem) já que estes mais se beneficiariam com as informações aqui apresentadas, uma vez que estas auxiliarão na compreensão das colocações especializadas criminais e, conseqüentemente, no processo de tradução dos textos nesta área, contemplando assim, a definição de um público-alvo e a função social que este exerce e que será atendida pelo DPCEC.

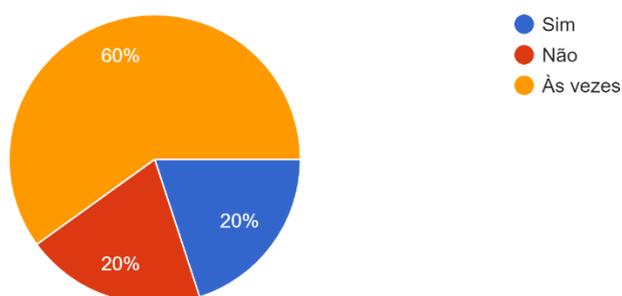
Outrossim, tal como mencionado anteriormente, para a definição desde público-alvo, levamos em consideração a experiência do lexicógrafo, aqui granjeada pelos conhecimentos coparticipados pela coordenadora do grupo de estudos ao qual essa pesquisa se filia, por meio das atividades desenvolvidas junto ao público-alvo há mais de vinte anos, além de considerarmos as respostas obtidas por meio dos formulários elaborados por Alves (em desenvolvimento), os quais também estão inseridos no grupo de estudos ao qual nos filiamos.

A partir desses formulários, que foram enviados a alunos dos cursos contemplados nessa pesquisa, apresentamos o gráfico abaixo, quanto ao uso de dicionários.

Figura 15: Porcentagem de aprendizes de tradução que usam dicionários para buscar colocações.

Você costuma usar o dicionário para buscar colocações?

25 respostas



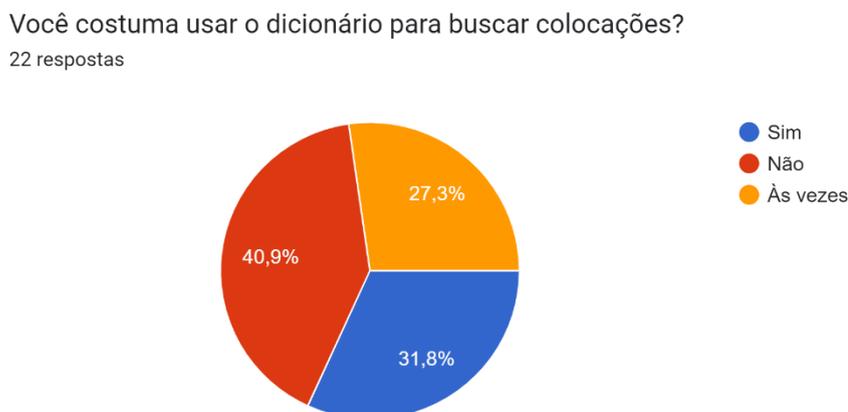
Fonte: Grupo de pesquisa “Fraseologia e Colocações a partir de *Corpora* (FRASCORP)”

Como se pode observar pelos resultados parciais obtidos até o momento pela pesquisa que está sendo desenvolvida por Alves, 60% dos aprendizes do curso de Bacharel em Letras às vezes usam os dicionários como fonte de informação sobre colocações, havendo um empate de 20% entre aqueles que usam e aqueles que não usam os dicionários. Esse dado vem ao encontro da escolha desses alunos como público-alvo de nosso DPCEC, uma vez que um dicionário específico de colocações pode contribuir para o aumento desse número entre os aprendizes de tradução.

Já no que tange os alunos do curso de Licenciatura em Letras, temos o seguinte resultado

parcial obtido por meio das respostas dos formulários propostos por Alves (em andamento):

Figura 16: Porcentagem de alunos do curso de Licenciatura em Letras que usam dicionários para buscar colocações.



Fonte: Grupo de pesquisa “Fraseologia e Colocações a partir de *Corpora* (FRASCORP)”

Diferentemente do resultado obtido entre os alunos do curso de Bacharel em Letras, entre os alunos do curso de Licenciatura em Letras que responderam à pesquisa, 40,9% não têm o hábito de utilizar dicionários como fonte de informação sobre colocações, enquanto que 31,8% o fazem e apenas 27,3% utilizam os dicionários às vezes. Esses dados, tal como acontece com os dados apresentados anteriormente, corroboram a criação de dicionários de colocações voltados para esse público-alvo, a fim de que estes possam facilitar o aprendizado desses alunos, bem como, para que possam servir de auxílio para a criação do hábito de estudar, aprender e, futuramente, ensinar colocações, sejam elas especializadas ou da língua geral.

Nesse sentido, qualquer aprendiz de língua inglesa ou língua espanhola poderá se beneficiar, também, do conteúdo de nosso DPCEC, uma vez que as informações foram pensadas e organizadas para aprendizes, afinal, de acordo com o que diz Molina Garcia ao tomarmos como princípio que o

(...) dicionário é um veículo para o ensino, o dicionário fraseológico é um veículo para o ensino das UF. Tal tipo de obra deve ser concebida a partir de um ponto de vista pedagógico para que tenha um fim pedagógico. Ademais, como qualquer outro tipo de dicionário, pode e deve ser explorado para fazer chegar ao aprendiz de línguas a informação relativa à fraseologia que o faça assimilar-las de forma efetiva<sup>68</sup>. (MOLINA GARCÍA, 2006, p.5).

<sup>68</sup> “[...] diccionario es un vehículo para la enseñanza. el diccionario fraseológico es un vehículo para la enseñanza de las UF. Tal tipo de obra debe ser concebida desde un punto de vista pedagógico para que tenga un fin pedagógico. Además, como cualquier otro tipo de diccionario, puede y debe ser explotado para hacer llegar al aprendiz de lenguas la información relativa a la fraseología que le haga asimilarlas de forma efectiva”.

Outro aspecto do DPCEC que merece destaque é que ao utilizarmos o termo aprendiz para se referir ao nosso consulente, seja ele o aprendiz de tradução ou o aprendiz de uma língua estrangeira, temos como base o conceito de aprendizagem autônoma, proposto por Molina García (2006, p.23), o qual afirma que a partir da utilização do termo aprendiz é possível abranger tanto o aluno que precisa ser guiado por um professor em seu processo de aquisição de conhecimentos linguísticos, quanto aquele que apresenta uma competência didática autônoma. Dessa forma, as informações apresentadas na macro e na microestrutura devem permitir essa autonomia dos consulentes.

Assim, entendemos que o DPCEC pode ser considerado um dicionário pedagógico pelas informações e a forma como estas são dispostas em sua estrutura, além de ser um dicionário que permite o desenvolvimento da autonomia de seus consulentes, os aprendizes dos cursos de Bacharel e de Licenciatura em Letras, no que diz respeito à aquisição e desenvolvimento dos conhecimentos linguísticos.

A seguir, apresentamos a organização da macro e da microestrutura do DPCEC.

## **6.2 A Macroestrutura do DPCEC**

Conforme ressaltamos no capítulo 4, no qual discorremos sobre a metodologia adotada para a elaboração de nosso DPCEC, utilizamos como base os pressupostos metodológicos propostos pelo Projeto de Pesquisa *PLATCOL - Plataforma On-line de Dicionários Bilíngues de Colocações* (2017-2019), conforme Orenha-Ottaiano (2017), e na nova proposta do projeto *A phraseographical methodology and model for an online corpus-based Multilingual Collocations Dictionary Platform* (ORENHA-OTTAIANO, 2020), aos quais este trabalho se insere.

Dessa forma, ao considerarmos que os referidos projetos desenvolveram plataformas on-line, ressaltamos que nossa proposta de macroestrutura leva em consideração aspectos que são necessários a um dicionário impresso, formato no qual apresentamos nossa proposta e para tal, contamos com o suporte teórico-metodológico proposto por Orenha-Ottaiano (2004, 2009, 2017) e Rocha (2017), baseados em Barbosa (1990, 1999). Entretanto, futuramente, os dados aqui apresentados serão inseridos em um dicionário on-line, tal como previsto pelo projeto *COMADEC - A Compilação de Materiais Didáticos e Dicionários Especializados de Colocações Baseados em Corpora* (ORENHA-OTTAIANO, 2021).

Assim, para Barbosa (1990, p. 154), “[...] a macroestrutura é constituída pelo conjunto

das entradas, ordenadas, submetidas à leitura vertical, sendo que os conceitos de polissemia e homonímia podem alterar substancialmente o número de entradas e a própria organização dos artigos ou verbetes”. Nesse sentido, ressaltamos que o conceito de polissemia apresentado pela autora não é utilizado neste dicionário, uma vez que de acordo com os dados levantados por meio da análise não foi possível a identificação de casos em que os termos, utilizados como palavras-chave e entradas no dicionário, ocorram com significados diferentes a depender da situação em que são utilizados.

Ademais, no que tange as questões relacionadas a polissemia e homonímia, devido à especificidade conferida ao trabalho por seu âmbito criminal e conforme ressalta Orenha-Ottaiano (2004), quando se trata de colocações de uma área de especialidade, é possível falarmos em parassinonímia, “única relação possível entre colocações e colocados de uma dada base, não havendo, portanto, relação de sinonímia” (ORENHA- OTTAIANO, 2004, p. 146). A autora se apoia em Barbosa (1997) e em Greimas (1979) para definir essa relação de parassinonímia.

Segundo Barbosa (1997), a sinonímia acontece quando “a dois ou mais elementos do significante corresponde um e somente um elemento do significado” (BARBOSA, 1997, p. 21 apud ORENHA-OTTAIANO, 2004, p. 146). Já para Greimas(1979)

sinonímia é a relação de identidade que duas ou mais grandezas (nesse caso chamadas de sinônimos) do plano do conteúdo seriam suscetíveis de contrair entre si. Tal relação, entre dois lexemas, por exemplo, seria verificável pelo teste de substituição: nesse caso, os dois lexemas seriam substituíveis em todos os contextos, demonstrando, assim, que os semas contextuais – que entram na composição de seus sememas – são idênticos. (GREIMAS, 1979, p. 477 apud ORENHA-OTTAIANO, 2004, p. 146).

Assim, tal como a consideração feita por Orenha-Ottaiano (2004), entendemos que não é possível dizer que uma colocação especializada, composta por um termo técnico, possa ser sinônimo de outra colocação, ou de um lexema, da linguagem comum, haja vista que para que isso ocorra é preciso que “os lexemas envolvidos tenham sempre os mesmos sememas em todos os contextos” (ORENHA-OTTAIANO, 2004, p. 147).

Dessa maneira, segundo os dados analisados neste trabalho, a única relação possível entre as colocações especializadas aqui levantadas é a de parassinonímia, a qual é definida por Greimas (1979) como “a identidade parcial de dois ou mais lexemas, que se reconhecem pela possibilidade que têm de se substituírem apenas em certos contextos” (GREIMAS, 1979, p. 326 apud ORENHA-OTTAIANO, 2004, p.148).

Nesse sentido, é possível falar em relação de parassinonímia entre colocações e entre

colocados, como é o caso das colocações *airtight alibi* e *ironclad alibi*, em língua inglesa, que compartilham um colocado quando traduzidas para a língua portuguesa, a saber: *incontestável*, conforme apresentado no capítulo 5, e são consideradas parassinônimas por compartilharem a característica de incontestabilidade em língua inglesa. Discutiremos mais sobre essa relação de parassinonímia na seção 6.3, ao abordarmos a microestrutura de nossa amostra de dicionário, uma vez que esta relação se dá no âmbito da microestrutura.

Desse modo, além das considerações feitas sobre a definição de macroestrutura proposta por Barbosa (1990), entendemos a macroestrutura como a

organização interna de uma obra lexicográfica ou terminográfica. Esse tipo de organização está relacionado às características gerais do repertório, ou seja, à estruturação das informações em verbetes (que podem se suceder verticalmente e/ou horizontalmente), a presença ou não de anexos, índices remissivos, ilustrações, setores temáticos, mapa conceptual e outros. (BARROS, 2004, p. 151).

Assim, organizamos a macroestrutura de nosso DPCE seguindo os princípios propostos por Orenha-Ottaiano (2004) e por Rocha (2017), os quais são explicados a continuação:

- i) Apresentação da obra: texto introdutório no qual apresentamos a obra fraseográfica em questão, discorrendo sobre a área de especialidade contemplada, a saber: o âmbito criminal, seu público-alvo, bem como sobre os procedimentos teórico-metodológicos utilizados na elaboração da mesma, além de apresentarmos a própria organização interna da obra;
- ii) Sistema de abreviações: seção na qual são apresentadas e explicadas as abreviações com as quais os consulentes se deparam ao utilizar a obra fraseográfica, de acordo com os idiomas que compõe o DPCE, conforme os exemplos apresentados abaixo:

<i>adj.</i> → <i>adjetivo / adjective / adjetivo</i>
<i>det</i> → <i>determiner / determinante / determinante</i>
<i>l. adj</i> → <i>locução adjetiva</i>
<i>n.</i> → <i>noun</i>
<i>ph. v.</i> → <i>phrasal verb</i>
<i>prep.</i> → <i>preposição / preposition / preposición</i>
<i>s.</i> → <i>substantivo / substantivo</i>
<i>v.</i> → <i>verbo / verb / verbo</i>

- iii) Índice: nesta seção da macroestrutura, trazemos as palavras-chave selecionadas como base para o levantamento das colocações especializadas e apresentadas no capítulo anterior, dispostas em ordem alfabética nas línguas alvo, de forma a facilitar a utilização do consulente. A seguir, mostramos o índice, em ordem alfabética em língua inglesa, tal como é exibido na obra lexicográfica.

Quadro 17: Índice em Língua Inglesa

Índice em Língua Inglesa
<i>Alibi</i>
<i>Autopsy</i>
<i>Gunshot</i>
<i>Unsub</i>
<i>Victimology</i>

Fonte: Próprio da pesquisa.

Tendo recorrido sobre a macroestrutura de nosso DPCEC, trataremos sobre a organização microestrutural dos verbetes no item a seguir, para posteriormente passarmos aos verbetes que integram os DPCEC.

### 6.3 A microestrutura do DPCEC

De acordo com o abordado no item 4.2, no capítulo de metodologia, a microestrutura utilizada para o desenvolvimento de nosso DPCEC toma por base a proposta teórico-metodológica empregada no recente projeto *A phraseographical methodology and model for an online corpus-based Multilingual Collocations Dictionary Platform* (ORENHA-

OTTAIANO, 2021) e no Projeto de Pesquisa *PLATCOL - Plataforma On-line de Dicionários Bílingues de Colocações* (ORENHA- OTTAIANO, 2017-2019), aos quais esta pesquisa se insere.

Dessa forma, a microestrutura dos verbetes é agrupada a partir de entradas, as quais são constituídas pelas palavras-chave que serviram como termo de busca para o levantamento das colocações especializadas, bem como constituem-se como a base das colocações em questão, conforme descrito nos itens 3.3 e 4.2, respectivamente.

Nesse sentido, a estrutura dos verbetes foi definida tomando como base o modelo de verbete proposto por Orenha-Ottaiano *et. al* (2021), adequando, assim, esta pesquisa, aos projetos aos quais se insere. Ademais, uma vez que a amostra de dicionário apresentada nesta pesquisa se caracteriza pelo aspecto multilíngue (inglês, português e espanhol), baseado no modelo de verbete proposto por Orenha-Ottaiano *et. al* (2021), apresentamos as informações constantes nos idiomas e direções abordadas, ou seja, na direção inglês → português / → espanhol. Dessa forma, cada entrada conta com as informações descritas no quadro abaixo.

Quadro 18: Informações referentes às palavras de entrada.

<b>Palavra de entrada na LP:</b> corresponde à base das colocações e pode ser um substantivo, um verbo ou um adjetivo.
<b>Classe gramatical da palavra na LP:</b> a classe gramatical da palavra de entrada é posicionada logo após a palavra em questão. Em nossa amostra de dicionário de colocações, as palavras de entrada podem ser um substantivo (s), um verbo (v) ou um adjetivo (adj). Caso a palavra de entrada apresente mais de uma classe gramatical, essas aparecerão em mais de uma entrada, por exemplo <i>murder</i> (s) e <i>murder</i> (v).
<b>Definição:</b> uma breve definição para cada palavra de entrada e para cada indicação de classe gramatical desta, indicando o sentido a ela atribuído em cada situação
<b>Palavra de entrada na LC :</b> corresponde à tradução da palavra de entrada apresentada na língua de chegada.
<b>Classe gramatical da palavra de entrada na LC:</b> corresponde à classe gramatical atribuída a essa palavra na língua de chegada
<b>Definição:</b> uma breve definição para a palavra de entrada na língua de chegada
<b>Frequência:</b> ao lado cada palavra de entrada haverá um número indicando a frequência dessa palavra no <i>corpus</i> de estudo, tanto na LP quanto na LC.

Fonte: Baseado em ORENHA-OTTAIANO et al (2021).

Abaixo de cada palavra de entrada, a qual recebeu a sua definição com base no dicionário online *Cambridge Dictionary*, são listadas as colocações que a apresentam como base ou como colocado, no formato de verbetes individuais e de acordo com sua classificação taxonômica, ou seja, são apresentadas todas as colocações adjetivas para aquela base, seguidas das colocações nominais e, por fim, as colocações verbais. Cada verbete contém as informações descritas no quadro abaixo.

Quadro 19: Informações apresentadas em cada verbete.

<b>Colocação na LP:</b> exemplo de colocação na língua de partida.
<b>Estrutura sintática da colocação na LP:</b> as colocações são analisadas e organizadas de acordo com o proposto por Hausmann (1985, 1989) e Orenha-Ottaiano (2004, 2009), conforme apresentado na seção anterior e na seção 3.3.
<b>Exemplo de uso na LP:</b> é apresentado um exemplo de uso na língua de chegada.
<b>Frequência na LP:</b> ao lado de cada colocação há um número indicando a sua frequência no <i>corpus</i> de estudo, na língua de partida.
<b>Colocação na LC:</b> exemplo de colocação na língua de chegada.
<b>Estrutura sintática da colocação na LC:</b> as colocações são analisadas e organizadas de acordo com o proposto por Hausmann (1985, 1989) e Orenha-Ottaiano (2004, 2009), conforme apresentado na seção anterior e na seção 3.3.
<b>Exemplo de uso na LC:</b> é apresentado um exemplo de uso na língua de chegada.
<b>Frequência na LC:</b> ao lado cada colocação há um número indicando a sua frequência no <i>corpus</i> de estudo, na língua de chegada.

Fonte: Baseado em ORENHA-OTTAIANO et al (2021).

No que tange à escolha das colocações especializadas que integram a amostra do DPCEC apresentada na seção seguinte, inserimos as colocações especializadas que apresentaram o *LogDice* com valor igual ou maior que 3, conforme descrito no capítulo de análise. Ressaltamos, ainda, que conforme mencionado no capítulo 4, os exemplos nas línguas de chegada, a saber: em língua portuguesa e em língua espanhola, foram traduzidos pela pesquisadora quando não foram encontrados nos *corpora* de estudo, CCP e CCE, recebendo, nesses casos, o símbolo  $\diamond$  para sinalizar essa tradução.

Ainda falando sobre a escolha das colocações especializadas que integram a amostra apresentada na seção seguinte, ressaltamos que, devido ao fato de termos encontrado casos de relação de parassinonímia entre as colocações (ORENHA-OTTAIANO, 2004, 2009), como o caso abordado na seção anterior (*airtigh alibi* e *ironclad alibi*), e entre as colocações

*meet the unsub* e *know the unsub*, as quais, ao serem traduzidas para as línguas portuguesa e espanhola compartilham um mesmo significado, *conhecer* e *conocer*, respectivamente, ao serem utilizadas nos exemplos apresentados no DPCEC, acrescentamos a informação *see also, ver también, ver también*, logo abaixo das informações apresentadas em cada verbete aqui mencionado.

Assim, os verbetes são apresentados da seguinte forma:

Quadro 20: Amostra de microestrutura do dicionário pedagógico de colocações especializadas.

<b>Autopsy (s). Freq. 595.</b>	<b>Autópsia(s). Freq. 447.</b>	<b>Autopsia(s). Freq. 186.</b>
The examination of a body in order to discover the cause of death.	Exame realizado em um corpo o qual visa atestar a causa da morte.	El examen de un cuerpo para descubrir la causa de la muerte.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<b>Psychological autopsy: (6) (adj + n)</b> - We're gonna build what we call psychological autopsies to determine whether the victims killed themselves. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia psicológica: (7) (s + adj)</b> - Nós vamos fazer o que chamamos de autópsia psicológica para determinar se as vítimas se mataram. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia psicológica: (6) (s + adj)</b> - Construiremos lo que llamamos <b>autopsias psicológicas</b> para determinar si las víctimas sesuicidaron. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).

Fonte: Próprio da pesquisa.

Assim, na seção seguinte, apresentamos o DPCEC.

#### **6.4 Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal – DPCEC**

Antes de apresentarmos o DPCEC, nossa amostra de dicionário pedagógico de colocações especializadas, convém ressaltar que os verbetes que contém as colocações especializadas identificadas em nossa análise foram agrupados todos sob a mesma entrada, de forma que é possível identificarmos colocações em que o termo sob análise aparece ora como base, ora como colocado, conforme mencionado na seção 5.2, a fim de facilitar a busca pelas colocações compostas pela mesma palavra-chave, por parte do consulente.

<b>Alibi (s).</b> Freq. 297	<b>Álibi(s).</b> Freq. 198	<b>Coartada(s).</b> Freq. 105
An excuse for something bad or for a failure, especially a crime.	Uma justificativa para algo ruim ou para uma falha, especialmente um crime.	Una justificación para algo malo o un fracaso, especialmente un crimen
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<b>Airtight alibi: (10) (adj + n)</b> - GPS records show that Sikes was 20 miles away from the crime scene at the time of the murder. - He's got an airtight alibi, the US government. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Álibi irrefutável: (6) (s + adj)</b> - Os registros do GPS mostram que Stikes estava a 30 quilômetros da cena do crime na hora do assassinato. - Ele tem um <b>álibi irrefutável</b> , o governo dos Estados Unidos. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Coartada irrefutable<sup>1</sup>: (3) (s + adj)</b> - Las grabaciones del GPS muestran que Stikes estaba millas de distancia del lugar del crimen en el momento del asesinato. - Tiene una <b>coartada irrefutable</b> , el gobierno de los EE.UU. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>See also: ironclad alibi</b>	<b>Ver também: álibi incontestável</b>	<b>Ver también: coartada irrefutable<sup>2</sup></b>
<b>Good alibi: (4) (adj + n)</b> - Boss, Hitch has got a very good alibi for last night. He was in New York promoting his club. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Bom álibi: (1) (s + adj)</b> - Chefe, Hitch tem um <b>álibi bom</b> para a noite passada. Estava em Nova York promovendo sua boate. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Buena coartada: (1) (s + adj)</b> - Jefe, Hitch tiene una muy <b>buena coartada</b> para lo de anoche. Estaba en Nueva York promocionando su club. ◇
<b>Ironclad alibi: (3) (adj + n)</b> - Well, Brass' gut told him that the brother was involved, but the guy had an ironclad alibi. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Álibi incontestável: (3) (s + adj)</b> - O instinto de Brass dizia que o irmão estava envolvido, mas o cara tinha um <b>álibi incontestável</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Coartada irrefutable<sup>2</sup>: (3) (s + adj)</b> El instinto de Brass le dictaba que el hermano estaba involucrado, pero el tipo tenía una <b>coartada irrefutable</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>See also: airtight alibi</b>	<b>Ver também: álibi irrefutável</b>	<b>Ver también: coartada irrefutable<sup>1</sup></b>
<b>Reasonable alibi: (2) (adj + n)</b> And Mr. Higgins knows a lot about reasonable alibis. He's a big fan of true crime. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Álibi razoável: (1) (s + adj)</b> E o Sr. Higgins sabe muito sobre <b>álibis razoáveis</b> . Ele é um grande fã do verdadeiro crime. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Coartada razonable: (2) (s + adj)</b> Y el Sr. Higgins sabe mucho de <b>coartadas razonables</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Solid alibi: (7) (adj + n)</b> - But she seems to have a solid alibi. - One of her errands was to the bank. Charlie could get us the surveillance tapes. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Álibi sólido: (2) (s + adj)</b> - Mas ela parece ter um <b>álibi sólido</b> . - Uma de suas tarefas foi ir ao banco. Charlie poderia conseguir fitas de segurança. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Coartada sólida: (4) (s + adj)</b> - Pero ella parece tener una <b>coartada sólida</b> . Uno de sus mandados fue al banco. Charlie podría conseguirnos las cintas de vigilancia. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).

Verbal collocations	Colocações verbais	Colocaciones verbales
<p><b>Alibi check out: (31) (n + ph v)</b> - And we know Drew Rich didn't do it... his alibi checks out. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Álibi conferir: (14) (s + v)</b> - Sabemos que não foi Drew Rich, o <b>álibi confere</b>.(Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Comprobar una coartada<sup>1</sup>: (4) (v + det + s)</b> - Y sabemos que Drew Rich no lo hizo... su <b>coartada se comprueba</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Alibi hold up: (2) (n + ph v)</b> - Well, let's see if Laura Seeger's alibi holds up. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Álibi sustentar-se: (1) (s + v)</b> - Bem, vejamos se o <b>álibi</b> de Laura Seeger <b>se sustenta</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Sostener una coartada: (1) (v + det + s)</b> - Bueno, veamos si la <b>coartada</b> de Laura Seeger <b>se sostiene</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Check an alibi: (14) (v + det + n)</b> -She's staying here until we check her alibi. -Alibi? For what? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Verificar um álibi: (10) (v + det + s)</b> - Ela ficará aqui até <b>verificarmos seu álibi</b>. - Álibi? Para o quê? ◇</p>	<p><b>Comprobar una coartada<sup>2</sup>: (4) (v + det + s)</b> -Se quedará aquí hasta que <b>comprobemos su coartada</b>. -¿Coartada? ¿Para qué? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Confirm an alibi: (4) (v + det + n)</b> - I confirmed his alibi with two of the other onboard cooks. They say he was working all day on the mess decks. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Confirmar um álibi: (7) (v + det + v)</b> - <b>Confirmei o álibi</b> dele com dois cozinheiros a bordo. Dizem que trabalhou o dia todo no convésaté o exercício. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Confirmar una coartada: (3) (v + det + s)</b> - <b>Confirmé su coartada</b> con dos de los otros cocineros a bordo. Dicen que estuvo trabajandotodo el día en las cubiertas del comedor. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Get an alibi: (19) (v + det + n)</b> Well, you got an alibi or just your sunshinyattitude? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Conseguir um álibi: (1) (v + det + v)</b> - Então, <b>conseguiu um álibi</b> ou apenas sua atitude brilhante? ◇</p>	<p><b>Tener una coartada: (10) (v + det + s)</b> - Bueno, ¿<b>tienes una coartada</b> o solo tu actitud alegre? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p> <p><b>Ver también: tener coartada</b></p>
<p><b>Give an alibi: (2) (v + det + n)</b> Danforth gave an alibi for the day of the explosion. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Dar um álibi: (5) (v + det + v)</b> Danforth <b>deu-lhe um álibi</b> para o dia da explosão. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Dar una coartada: (1) (v + det + s)</b> - Danforth dio una coartada para el día de la explosión. ◇</p>
<p><b>Have an alibi: (67) (v + det + n)</b> -They've all had polygraphs. Everybody's beenvetted. -And they all have alibis for the night of the kidnapping? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Ter um álibi: (65) (v + det + v)</b> - Todos passaram pelo polígrafo, todos foram testados. Todos <b>têm álibis</b> para a noite do sequestro?(Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Tener coartada: (10) (v + det + s)</b> -Todos han tenido polígrafos. Todo el mundo hasido investigado. - ¿Todos <b>tienen coartadas</b> para la noche del secuestro? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p> <p><b>Ver también: tener coartada</b></p>

<p><b>Verify an alibi: (4) (v + det + n)</b>          - Jim? I need you to verify the alibi for Daniel Moore.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Verificar um álibi: (10) (v + det + v)</b>          - Jim? Eu preciso de você para <b>verificar o álibi</b> de Daniel Moore.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Verificar una coartada: (6) (v + det + s)</b>          - ¿Jim? Necesito que <b>verifiques la coartada</b> de Daniel Moore.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
--	---	---

Autopsy (s). Freq. 595.	Autópsia(s). Freq. 447.	Autopsia(s). Freq. 186.
The examination of a body in order to discover the cause of death.	Exame realizado em um corpo o qual visa atestar a causa da morte.	El examen de un cuerpo para descubrir la causa de la muerte.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<p><b>Abbreviated autopsy: (3) (adj + n)</b>          - Besides, Homicide reported that nothing at the scene indicated foul play, so, there was no reason to do more than an abbreviated autopsy.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia resumida: (1) (s + adj)</b>          - Além disso, a Homicídios informou que nada na cena era suspeito, por isso, não havia porque não fazer uma <b>autópsia resumida</b>.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autopsia abreviada: (1) (s + adj)</b>          - Además, Homicidios informó que nada en la escena era sospechoso, por lo que no había razón para hacer más que una <b>autopsia abreviada</b>. ◊</p>
<p><b>Forensic autopsy: (2) (adj + n)</b>          - I can just see those FBI forensic weasels sitting around, laughing, picturing what I'm going to look like performing a forensic autopsy on this poor creature.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia forense: (1) (s + adj)</b>          - Eu posso ver aquelas doninhas forenses do FBI sentados, rindo, imaginando como vou ficar realizando uma <b>autópsia forense</b> nesta pobre criatura. ◊</p>	<p><b>Autopsia forense: (1) (s + adj)</b>          - Puedo ver esas comadreas forenses del FBI sentadas, riendo, imaginando cómo me veré realizando una <b>autopsia forense</b> a esta pobre criatura. ◊</p>
<p><b>Formal autopsy: (2) (adj + n)</b>          - I would prefer to wait until I've begun my formal autopsy.          (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia formal: (1) (s + adj)</b>          - Eu preferiria esperar até começar minha <b>autópsia formal</b>. ◊</p>	<p><b>Autopsia formal: (1) (s + adj)</b>          - Preferiría esperar hasta que haya comenzado mi <b>autopsia formal</b>. ◊</p>
<p><b>Full autopsy: (5) (adj + n)</b>          - I'll know more when I conduct a full autopsy. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia completa: (1) (s + adj)</b>          - Eu saberei mais quando conduzir uma <b>autópsia completa</b>. ◊</p>	<p><b>Autopsia completa: (1) (s + adj)</b>          - Sabré más cuando realice una <b>autopsia completa</b>. ◊</p>
<p><b>Infectious autopsy: (4) (adj + n)</b>          - How do you alert visitors when conducting an infectious autopsy?          - We hang a decomposing body in the corridor. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia infecciosa: (1) (s + adj)</b>          - Como você alerta os visitantes quando está conduzindo uma <b>autópsia infecciosa</b>?          - Penduramos um corpo em</p>	<p><b>Autopsia infecciosa: (1) (s + adj)</b>          - ¿Cómo alerta a los visitantes cuando se realiza una <b>autopsia infecciosa</b>?          - Colgamos un cuerpo en descomposición en el</p>

<a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	decomposição nocorredor. ◇	pasillo.◇
<b>Preliminary autopsy: (2) (adj + n)</b> - Well, the <b>preliminary autopsy</b> came back. The victim's name is Becky Williams. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia preliminar: (3) (s + adj)</b> - A <b>autópsia preliminar</b> chegou. O nome da vítima era Becky Williams. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia preliminar: (1) (s + adj)</b> - Bueno, llegó la <b>autopsia preliminar</b> . El nombre dela víctima es Becky Williams. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Psychological autopsy: (6) (adj + n)</b> - We're gonna build what we call psychological autopsies to determine whether the victims killedthemselves. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia psicológica: (5) (s + adj)</b> - Vamos montar o que chamamos de <b>autópsia psicológica</b> para determinar se as vítimas se mataram. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia psicológica: (6) (s + adj)</b> - Construiremos lo que llamamos <b>autopsias psicológicas</b> para determinar si las víctimas sesuicidaron. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Remote autopsy: (2) (adj + n)</b> - Ever done this before? -All the time. It's called a remote autopsy. So, you ask total novices to just chop up dead (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia remota: (2) (s + adj)</b> - Você já fez isso antes? - Sempre. Chama-se <b>autópsia remota</b> . - Então você pede para novatos cortarem defuntos sempre? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia remota: (1) (s + adj)</b> - ¿Alguna vez has hecho esto antes? -Todo el tiempo. Se llama <b>autopsia remota</b> . - Así que les pides a los novatos que simplemente corten muertos. ◇
<b>Routine autopsy: (2) (adj + n)</b> - <b>This morning, he was on</b> his way to Bethesda for a routine autopsy when fate intervened. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia de rotina: (2) (s + prep + s)</b> - Nesta manhã, ele estava indo para Bethesda, para uma <b>autópsia de rotina</b> quando o destino interviu. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia de rutina: (2) (s + adj)</b> - Esta mañana, se dirigía a Bethesda para una <b>autopsia de rutina</b> cuando intervino el destino. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Solo autopsy: (2) (adj + n)</b> - Hello, this is... - Don't say "hello". - Okay. This is the first solo autopsy of AssistantMedical Examiner David Gregory Phillips. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia solo: (2) (s + adj)</b> Olá, isto é... - Não diga "olá". - Ok. Esta é a primeira <b>autópsia solo</b> do assistente David Gregory Phillips. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autopsia solo: (1) (s + adj)</b> - Hola este es... - No digas "hola". - Okey. Esta es la primera <b>autopsia solo</b> del médico forense asistente David Gregory Phillips. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Nominal collocations</b>	<b>Colocações nominais</b>	<b>Colocaciones nominales</b>
<b>Autopsy garage: (2) (n + det + n)</b> - Cameras showed Cobb forced EJ and Palmer to take him to the autopsy garage. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Garagem da autópsia: (2) (s + prep + det + s)</b> - As câmeras mostram Cobb forçando EJ e Palmer até a <b>garagem da autópsia</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Garaje de autopsia: (1) (s + prep + s)</b> - Las cámaras mostraron que Cobb obligó a EJ yPalmer a llevarlo al <b>garaje de autopsias</b> . ◇
<b>Autopsy photo: (5) (n + det + n)</b>	<b>Foto da autópsia: (4) (s + prep + det + s)</b>	<b>Foto de la autopsia: (1) (s + prep + det + s)</b>

<p>- He's doing time for assault, and he ID'd the victim's autopsy photo as his son Eduardo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- Este é Carlos Gomez, está preso por agressão, e identificou a <b>foto da autópsia</b> da vítima como seu filho, Eduardo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- Está cumpliendo condena por agresión e identificó la <b>foto de la autopsia</b> de la víctima como su hijo Eduardo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy protocol: (2) (n + det + n)</b>  - Okay, then how about the fact that on one hand, we have paranoid psychosis...  - But the autopsy protocol says what? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Protocolo da autópsia: (2) (s + prep + det + s)</b>  - Ok, então como de um lado temos um psicopata paranoico...  - Mas o <b>protocolo da autópsia</b> diz o quê? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Protocolo de la autopsia: (1) (s + prep + det + s)</b>  - Analizando los hechos desde un punto de vista, padecería psicosis paranoica...  - Pero el <b>protocolo de la autopsia</b> dice ¿qué? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy records: (2) (n + det + n)</b>  - Duck, all three of these women's deaths were ruled accidental.  - Pull their autopsy records, along with the guys who died with them. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Registro da autópsia: (2) (s + prep + det + s)</b>  - Ducky, as mortes dessas três mulheres foram dadas como acidentais.  - Recupere os <b>registros da autópsia</b> desses dois sujeitos que morreram com elas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Registro de autopsia: (1) (s + prep + s)</b>  - Duck, las tres muertes de estas mujeres fueron declaradas accidentales.  - Obtenga sus <b>registros de autopsia</b>, junto con los tipos que murieron con ellos. ◇</p>
<p><b>Autopsy report: (64) (v + det + n)</b>  - Ah, colonel, your C.O. just called. Requested copies of my autopsy report. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Relatório da autópsia: (31) (s + prep + det + s)</b>  - Coronel, seu chefe acabou de ligar. Requisitou cópias do meu <b>relatório da autópsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Informe de autopsia: (27) (s + prep + s)</b>  - Ah, coronel, su C.O. acabó de llamar. Solicité copias del <b>informe de la autopsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy result: (11) (v + det + n)</b>  - No. The police told us they died in the accident.  - Yes, sir. I understand that.  - We had to wait for the autopsy results. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Resultado da autópsia: (5) (s + prep + det + s)</b>  - Não. A polícia disse que eles morreram no acidente!  - Sim, senhor. Sei disso.  - Tínhamos que esperar pelo <b>resultado da autópsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Resultado de la autopsia: (6) (s + prep + det + s)</b>  No. La policía nos dijo murieron en el accidente.  - Sí señor. Entiendo que.  - Tuvimos que esperar los <b>resultados de la autopsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy room: (5) (v + det + n)</b>  - The man who supposedly kidnapped Commander Tanner is lying dead in our autopsy room. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Sala de autópsia: (5) (s + prep + s)</b>  - O homem que supostamente raptou o Comandante Tanner está morto na nossa <b>sala de autópsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Sala de autopsia: (1) (s + prep + s)</b>  El hombre que supuestamente secuestró el comandante Tanner yace muerto en nuestra <b>sala de autopsias</b>. ◇</p>
<p><b>Autopsy table: (9) (v + det + n)</b></p>	<p><b>Mesa da autópsia: (6) (s + prep + det + s)</b></p>	<p><b>Mesa de autopsia: (5) (s + prep + s)</b></p>

<p>- You know, that... that image of me lying dead on...on the autopsy table and... and how that makes me feel. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- Aquela imagem de mim morta na <b>mesa de autópsia</b> e como isso me faz sentir. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- Ya sabes, esa... esa imagen de mí yaciendo muerto en... en la <b>mesa de autopsias</b> y... y cómo me hace sentir eso. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Body in autopsy: (8) (n + det + n)</b> - DNA confirms the body in autopsy is NCIS Agent Abog Galib. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Corpo na autópsia: (3) (s + prep + det + s)</b> - O <b>corpo na autópsia</b> é do agente da NCIS, AbogGalib. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Cuerpo en la autopsia: (1) (s + prep + det + s)</b> - El ADN confirma que el <b>cuerpo en la autopsia</b> es el agente del NCIS Abog Ghalib. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Evidence in autopsy: (1) (n + prep + n)</b> - Labels on the pack say as recently as two weeks ago. There was no evidence in autopsy. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Evidência na autópsia: (1) (s + prep + det + s)</b> - Etiquetas na embalagem dizem tão recente como há duas semanas. Não houve <b>evidência na autópsia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Evidencia en la autopsia: (1) (s + prep + det + s)</b> Las etiquetas en el paquete dicen tan reciente como hace dos semanas. No <b>hubo evidencia en la autopsia</b>. ◇</p>
<p><b>Incongruity in autopsy: (1) (n + prep + n)</b> - You came down here hoping that I had found some subtle incongruity in my autopsy that would shed some light on this dastardly mystery. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Incongruência na autópsia: (1) (s + prep + det + s)</b> - Você veio aqui esperando que eu tenha achado alguma <b>incongruência na minha autópsia</b>, que levante alguma explicação sobre esse maliciosomistério. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Incongruencia en la autopsia: (1) (s + prep + det + s)</b> - Viniste aquí con la esperanza de que hubiera encontrado alguna sutil <b>incongruencia en mi autopsia</b> que arrojara algo de luz sobre este cobarde misterio. ◇</p>
<p><b>Verbal collocations</b></p>	<p><b>Colocações verbais</b></p>	<p><b>Colocaciones verbales</b></p>
<p><b>Autopsy confirm: (3) (n + v)</b> - They were dead before the accident. The autopsy confirmed that and the cause of death. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia confirmar: (3) (s + v)</b> - Estavam mortos antes do acidente. A <b>autópsia confirmou</b> isso e a causa da morte. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autopsia confirmar: (2) (s + v)</b> - Ellos estaban muertos antes del accidente. La <b>autopsia confirmó</b> eso y la causa de la muerte. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy reveal: (4) (n + v)</b> - Sex lube, found at the crime scene. Autopsy revealed that Becky was sexually active. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autópsia revelar: (3) (s + v)</b> - Lubrificante sexual, encontrado na cena do crime. A <b>autópsia revelou</b> que Becky era sexualmente ativa. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autopsia revelar: (2) (s + v)</b> - Lubrificante sexual, encontrado en la escena del crimen. La <b>autopsia reveló</b> que Becky era sexualmente activa. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Autopsy say: (4) (n + v)</b> - Ok, the autopsy says Michelle Watson's time of death was 4:30 in the afternoon. So, he edited</p>	<p><b>Autópsia dizer: (3) (s + v)</b> - A <b>autópsia diz</b> que a morte de Michelle Watson foi às 16h30. Então, ele tirou 7 horas na edição. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Autopsia decir: (2) (s + v)</b> - Ok, la <b>autopsia dice</b> que la hora de la muerte de Michelle Watson fue a las 4.30 de la tarde. Así que, él editó sobre</p>

out 7 hours. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	las 7 horas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Autopsy show: (3) (n + v)</b> - He did more than silence her, autopsy shows that Veronica Gilmore sustained more than 50 blows to her head and body. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Autópsia mostrar: (1) (s + v)</b> - Ele fez mais do que silenciar, sua <b>autópsia mostra</b> que Veronica Gilmore sofreu mais de 50 golpes na cabeça e no corpo.◊	<b>Autopsia mostrar: (2) (s + v)</b> - Hizo más que silenciar, su <b>autopsia muestra</b> que Veronica Gilmore recibió más de 50 golpes en la cabeza y el cuerpo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Begin the autopsy: (7) (v + det + n)</b> - Then we can get this bag off and begin the autopsy. There are ligature marks on her wrists. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Começar a autópsia: (4) (s + v)</b> - Então podemos tirar este saco e <b>começar a autópsia</b> . Há marcas de ligaduras nos pulsos, sugerindo que foi amarrada. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Empezar la autopsia: (5) (v + det + s)</b> - Entonces podemos sacar esta bolsa y <b>empezar la autopsia</b> . Hay marcas de ligadura en sus muñecas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Conduct the autopsy: (7) (v + det + n)</b> - Yes, I once conducted an autopsy on a man who drowned in his kitchen sink. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Conduzir uma autópsia: (1) (s + det + v)</b> - Sim, uma vez <b>conduzi uma autópsia</b> de um homem que se afogou na pia.◊	<b>Realizar la autopsia<sup>1</sup>: (7) (v + det + s)</b> - Sí, una vez <b>realicé la autopsia</b> de un hombre que se ahogó en el fregadero de su cocina. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
-	-	<b>Ver también: realizar la autopsia<sup>2</sup></b>
<b>Complete the autopsy: (3) (v + det + n)</b> - From the angle of entry, I would say that the shooter was above the body. - Distance? - I'll need to complete the autopsy first. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Completar a autópsia: (1) (s + det + v)</b> - Pelo ângulo de entrada, eu diria que o atirador estava acima do corpo. - Distância? Eu precisarei <b>completar a autópsia</b> primeiro.◊	<b>Completar la autopsia: (1) (v + det + s)</b> - Desde el ángulo de entrada, yo diría que el tirador estaba arriba del cuerpo - ¿Distancia? Necesito <b>completar la autopsia</b> primero.◊
<b>Do an autopsy: (46) (v + det + n)</b> - Morgan took those samples back to the lab. What about the body, did Dr. Kinney do the autopsy? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Fazer uma autópsia<sup>1</sup>: (57) (s + det + v)</b> - Morgan mandou as amostras ao laboratório. E o corpo... O Dr. Kinney <b>fez a autópsia</b> ? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Hacer la autopsia: (29) (v + det + s)</b> - Morgan llevó esas muestras al laboratorio. ¿Qué pasa con el cuerpo, el Dr. Kinney <b>hizo la autopsia</b> ? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>See also: perform an autopsy</b>	<b>Ver também: fazer uma autópsia<sup>2</sup></b>	-
<b>Finish the autopsy: (12) (v + det + n)</b> - He just finished the	<b>Terminar a autópsia: (8) (s + det + v)</b> - Ele apenas terminou a	<b>Terminar la autopsia: (1) (v + det + s)</b> - Acaba de <b>terminar la</b>

autopsy, told me to put her to bed and then he said he was going to Norfolk. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	autópsia, mandou colocá-la em uma maca...e depois disse que ia para Norfolk. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>autopsia</b> , me dijo que la acostara y luego dijo que se iba a Norfolk. ◇
<b>Handle the autopsy: (2) (v + det + n)</b> - He's all yours, Ducky. NCIS will handle the autopsy. -Okay, Aldridge. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Lidar com a autópsia: (1) (s + prep + det + v)</b> - Ele é todo seu, Ducky. NCIS <b>lidará com a autópsia</b> . - Ok, Aldridge. ◇	<b>Encargar de la autopsia: (1) (v + prep + det + s)</b> - Es todo tuyo, Ducky. NCIS se <b>encargará de la autopsia</b> . -Está bien, Aldridge (Exemplo dado pela pesquisadora)
<b>Perform the autopsy: (24) (v + det + n)</b> -Only a licensed pathologist can perform an autopsy. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Fazer uma autópsia<sup>2</sup>: (57) (s + det + v)</b> - Apenas um patologista licenciado pode <b>fazer uma autópsia</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Realizar la autopsia<sup>2</sup>: (7) (v + det + s)</b> - Solo un patólogo con licencia puede <b>realizar una autopsia</b> . (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>See also: do an autopsy</b>	<b>Ver também: fazer uma autópsia<sup>1</sup></b>	<b>Ver también: realizar la autopsia<sup>1</sup></b>

<b>Gunshot (s). Freq. 454.</b>	<b>Tiro(s). Freq. 1024.</b>	<b>Disparo (s). Freq. 412</b>
The shooting of a gun; the sound it makes.	O disparo de uma arma; o som que ela faz.	El disparo de un arma; el sonido que hace.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<b>Medium-caliber gunshot: (2) (adj + n)</b> - Shotgun guy. One distant-range, medium-caliber gunshot wound to the forehead. The bullet penetrated the brain and was recovered in the occipital lobe. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Tiro de médio calibre: (3) (s + prep + l. adj)</b> - Espingarda, cara. Um passo, o <b>tiro de médio-calibre</b> acerta a testa. A bala penetra o cérebro. E foi recuperado no lóbulo occipital.(Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Disparo de mediano calibre: (1) (adj + s)</b> - Tipo de la escopeta. Una herida de <b>disparo de mediano calibre</b> a distancia en la frente. La bala penetró en el cerebro y fue recuperada en el lóbulo occipital. ◇
<b>Multiple gunshots: (5) (adj + n)</b> - Victim was a well-dressed woman, young, black. She's at Desert Palm Hospital: multiple gunshots to the head, possible sexual assault. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Múltiplos tiros: (4) (adj + s)</b> - A vítima era uma mulher bem vestida, jovem, negra. Ela está no Hospital Desert Palm: <b>múltiplos tiros</b> na cabeça e possível agressão sexual. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Múltiplos disparos: (1) (adj + s)</b> - La víctima era una mujer bien vestida, joven, negro. Ella está en el Hospital Desert Palm: <b>múltiplos disparos</b> en la cabeza posible agresión sexual. ◇
<b>Self-inflicted gunshot: (3)</b>	<b>Tiro auto infligido: (1) (s +</b>	<b>Disparo auto-infligido: (1)</b>

<p><b>(adj + n)</b> - Here on his front lobe, but the exit wound is not consistent with a self-inflicted gunshot. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>l. adj)</b> - Aqui no lobo frontal, mas a ferida de saída não é consistente com um <b> tiro auto infligido</b>. ◇</p>	<p><b>(s + adj)</b> - Aquí en su lóbulo frontal, pero la herida desalida no es consistente con un <b>disparo autoinfligido</b>. ◇</p>
<p><b>Single gunshot: (21) (adj + n)</b> - Single gunshot to the head with a .22. Surveillance cameras have captured video of 3 of the murders. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Único tiro: (27) (adj + s)</b> - Um <b>único tiro</b> na cabeça com uma .22. Ascâmeras de segurança gravaram três dos assassinatos. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Único disparo: (6) (adj + s)</b> - <b>Único disparo</b> en la cabeza con calibre .22. Las câmaras de vigilancia han capturado videos de 3 de los asesinatos. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Small-caliber gunshot: (2) (adj + n)</b> - Check this out. Three small-caliber gunshots to the chest. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Tiro de pequeno calibre: (1) (s + prep + l. adj)</b> - Veja isso. Três <b> tiros de pequeno calibre</b> no peito. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Disparo de pequeño calibre: (1) (s + adj)</b> - Mira esto. <b>Tres disparos de pequeño calibre</b> en el pecho. ◇</p>
<p><b>Nominal collocations</b></p>	<p><b>Colocações nominais</b></p>	<p><b>Colocaciones nominales</b></p>
<p><b>Evidence of a gunshot: (1) (n + prep + det + n)</b> - The record states that he sustained an injury during a work-related accident on base in Korea, but Dr. Mallard found evidence of a gunshot wound. Not something I'd expect to see in a work-related injury. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Evidência de tiro: (1) (s + prep + s)</b> - O registro afirma que ele sofreu uma lesão durante um acidente relacionado ao trabalho na base na Coreia, mas o Dr. Mallard encontrou <b>evidências de um tiro</b>. Não é algo que eu esperaria ver em uma lesão relacionada ao trabalho. ◇</p>	<p><b>Evidencia de un disparo: (1) (s + prep + det + s)</b> - El registro establece que sufrió una lesión durante un accidente relacionado con el trabajo en la base en Corea, pero el Dr. Mallard encontró <b>evidencia de una herida de disparo</b>. No es algo que esperaría ver en una lesión relacionada con el trabajo. ◇</p>
<p><b>Gunshot residue: (36) (n + n)</b> - When exactly did war become a party? - Gunshot residue, close range .32 auto. Belly got it. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Resíduos de tiro: (4) (s + prep + s)</b> - Quando exatamente a guerra virou festa? <b>Resíduos de tiro</b> a curta distância .32 automática. Em cheio. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Residuos de disparo: (13) (s + prep + s)</b> - ¿Cuándo exactamente la guerra se convirtió en una fiesta? - <b>Residuos de disparo</b>, corto alcance .32 automático. El vientre lo entendió. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Gunshot victim: (12) (n + n)</b> - Never seen a gunshot victim with so much blood left in his body. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vítima de tiro: (7) (s + prep + s)</b> - Nunca vi uma <b>vítima de tiro</b> com tanto sangue no seu corpo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vítima de disparo: (3) (s + prep + s)</b> - Nunca había visto una <b>vítima de un disparo</b> con tanta sangre en el cuerpo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Gunshot wound: (60) (n + n)</b> - Just a single gunshot wound</p>	<p><b>Ferida de tiro: (3) (s + prep + s)</b> - Apenas uma única <b>ferida</b></p>	<p><b>Herida de disparo: (7) (s + prep + s)</b> - Sólo una única <b>herida de</b></p>

that nicked his carotid artery. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>de tiro</b> que rompeu sua artéria carótida. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>disparo</b> que le cortó laarteria carótida. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Noise of a gunshot: (1) (n + prep + det + n)</b> - As I would have expected, and if you jam a gun into a pillow to suppress the noise of a gunshot, there would be more powder burns. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Barulho de tiro: (1) (s + prep + s)</b> - Não há tanto quanto eu esperava, e se você enfiar uma arma em um travesseiro para suprimir o <b>barulho de um tiro</b> , haveria mais queimaduras de pólvora.◇	<b>Ruido de disparo: (1) (s + prep + s)</b> - Como esperaba, y si metes un arma en una almohada para suprimir el <b>ruido de un disparo</b> , habría más quemaduras de pólvora. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).
<b>Sound of a gunshot: (1) (n + prep + det + n)</b> - Know many 60-year-olds with a boom box? Trying to mask the sound of the gunshots? Could be. Or he was torturing them with sound. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Som de tiro: (1) (s + prep + s)</b> - Conhece muitas pessoas de 60 anos com umacaixa de som? Tentando mascarar o <b>som dos tiros</b> ? - Poderia ser. Ou ele os estava torturando com som. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Sonido de disparo: (1) (s + prep + s)</b> - ¿Conoce a muchas personas de 60 años con unestéreo portátil? ¿Tratando de enmascarar el <b>sonido de los disparos</b> ? Podría ser. O los estaba torturando con sonido.(Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).

Unsub (s). Freq. 2.489	Suspeito (s). Freq. 1971	Sospechoso (s). Freq. 756
<b>Someone unknow who is suspect in a crime.</b>	Alguém desconhecido que é suspeito em um crime.	Alguien desconocido que es sospechoso de un crimen.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<b>Multiple unsubs: (8) (adj + n)</b> - We think we might have a serious problem. We have multiple unsubs. They're disciplined, they're using counter surveillance, they know the FBI movements, there's a hierarchy. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Múltiplos suspeitos: (6) (adj + s)</b> - Acho que temos um problema sério. Temos <b>múltiplos suspeitos</b> , são organizados. Estão usando policiais e vigilância. Sabem os movimentos do FBI. Têm uma hierarquia. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Múltiplos sospechosos: (1) (adj + s)</b> - Creemos que podríamos tener un problema grave. Tenemos <b>múltiplos sospechosos</b> . Son disciplinados, están usando contra vigilancia, conocen los movimientos del FBI, hay una jerarquía. ◇
<b>Real unsub: (4) (adj + n)</b> - Bringing in another suspect draws attention away from the real unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Suspeito verdadeiro: (1) (s + adj)</b> - Trazer outro suspeito desvia a atenção do <b>verdadeiro suspeito</b> .◇	<b>Verdadero sospechoso: (1) (adj + s)</b> - Traer a otro sospechoso desvía la atención del <b>verdadero sospechoso</b> . ◇
<b>Single unsub: (6) (adj + n)</b> - What if we're looking for a single unsub? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Único suspeito: (4) (adj + s)</b> - E se estivermos procurando por um <b>único suspeito</b> ? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a> ).	<b>Único sospechoso: (1) (adj + s)</b> - ¿Y sol o que buscamos es un <b>único sospechoso</b> ?◇

Nominal collocations	Colocações nominais	Colocaciones nominales
<p><b>Description of the unsub: (5) (n + prep + det +n)</b></p> <p>- The man who asked which way the purse- snatcher went matches the description of the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Descrição do suspeito: (5) (s + prep + s)</b></p> <p>- O homem que perguntou para que lado o ladrão de bolsas foi corresponde à <b>descrição do suspeito</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Descripción del sospechoso: (1) (s + prep + det + s)</b></p> <p>- El hombre que preguntó en qué dirección se fue el ladrón de bolsos coincide con la <b>descripción del sospechoso</b>. ◇</p>
<p><b>Desire of the unsub: (1) (n + prep + det +n)</b></p> <p>- This is also satisfying some internal desire of the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Desejo do suspeito: (1) (s + prep + s)</b></p> <p>- Isso também está satisfazendo algum desejo interno do suspeito. ◇</p>	<p><b>Deseo del sospechoso: (1) (s + prep + det + s)</b></p> <p>- Esto también satisface algún <b>deseo</b> interno <b>del sospechoso</b>. ◇</p>
<p><b>Intelligence of the unsub: (2) (n + prep + det +n)</b></p> <p>- And the lower the risk of the victim, the higher the intelligence of the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Inteligência do suspeito: (1) (s + prep + s)</b></p> <p>- E quanto menor o risco da vítima, maior a <b>inteligência do suspeito</b>. ◇</p>	<p><b>Inteligencia del sospechoso: (1) (s + prep + det + s)</b></p> <p>- Y cuanto menor sea el riesgo de la víctima, mayor será la <b>inteligencia del sospechoso</b>. ◇</p>
<p><b>Perspective of the unsub: (1) (n + prep + det +n)</b></p> <p>- All right, so we need to start over, go back over both Gail and Molly's cases and look at everything from the perspective of a female unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perspectiva do suspeito: (1) (s + prep + s)</b></p> <p>- Tudo bem, por isso precisamos recomeçar, voltar a tanto de Gail e Molly casos e olhar paratudo a partir da <b>perspectiva de um suspeito</b> feminino. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perspectiva del sospechoso: (1) (s + prep + det + s)</b></p> <p>- Muy bien, tenemos que empezar de nuevo, repasar los casos de Gail y Molly y verlo tododesde la <b>perspectiva de una sospechosa</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Profile of the unsub: (2) (n + prep + det +n)</b></p> <p>- Lynn Dempsey was an executive assistant. She has no expertise with chemicals. She doesn't fit the profile of the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perfil do suspeito: (6) (s + prep + s)</b></p> <p>- Lynn Dempsey era uma auxiliar executiva. Ela não tinha experiência com químicos. E não se encaixa no <b>perfil de nosso suspeito</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perfil del sospechoso: (1) (s + prep + det + s)</b></p> <p>- Lynn Dempsey fue asistente ejecutiva. Ella no tiene experiencia con productos químicos. Ella no encaja en el <b>perfil del sospechoso</b>. ◇</p>
<p><b>Projection of the unsub: (2) (n + prep + det +n)</b></p> <p>- Absent child. The abducted baby could be a projection of the unsub himself. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Projeção do suspeito: (1) (s + prep + s)</b></p> <p>- Criança ausente. O bebê raptado pode ser uma <b>projeção do próprio suspeito</b>. ◇</p>	<p><b>Proyección del sospechoso: (1) (s + prep + det + s)</b></p> <p>- Niño ausente. El bebé secuestrado podría ser una <b>proyección del sospechoso</b>. ◇</p>
<p><b>Sketch of the unsub: (3) (n + prep + det +n)</b></p>	<p><b>Retrato-falado do suspeito: (1) (s + prep + s)</b></p>	<p><b>Boceto del sospechoso: (2) (s + prep + det + s)</b></p>

<p>- You need to get a sketch of the unsub out to every camera shop in Buffalo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- Você precisa levar um <b>retrato falado do suspeito</b> para todas as lojas de câmeras em Buffalo.◇</p>	<p>- Tienes que llevar un <b>boceto del sospechoso</b> a todas las tiendas de cámaras de Buffalo. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Team of unsub: (1) (n + prep + n)</b>          - That's why the girls kept saying, "they're" after us. There wasn't a team of unsubs, but one sadist copying the behavior of another. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Equipe de suspeitos: (1) (s + prep + s)</b>          - É por isso que as garotas ficavam dizendo, "eles estão" atrás de nós. Não havia uma <b>equipe de suspeitos</b>, mas um sádico copiando o comportamento de outro. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Equipo del sospechoso: (1) (s + prep + det + s)</b>          - Por eso las chicas seguían diciendo, "están" detrás de nosotros. No había un <b>equipo de sospechosos</b>, sino un sádico que copiaba el comportamiento de otro. ◇</p>
<p><b>Type of unsub: (11) (n + prep + n)</b>          - Why go from low- to high-risk victims?          Yeah, these types of unsubs typically target victims like runaways, transients, and prostitutes. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Tipo de suspeito: (9) (s + prep + s)</b>          - Por que passar de vítimas de baixo para alto risco?          - Sim, esses <b>tipos de suspeitos</b> normalmente têm como alvo vítimas como fugitivos, transeuntes e prostitutas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Tipo de sospechoso: (2) (s + prep + s)</b>          - ¿Por qué pasar de víctimas de bajo riesgo a víctimas de alto riesgo?          - Sí, estos <b>tipos de sospechosos</b> normalmente se dirigen a víctimas como fugitivos, transeuntes y prostitutas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Verbal collocations</b></p>	<p><b>Colocações verbais</b></p>	<p><b>Colocaciones verbales</b></p>
<p><b>Assume the unsub: (5) (v + det + n)</b>          - Factoring in Monica's abduction site and her apartment, I would assume the unsub lives somewhere within this area. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Presumir que o suspeito: (1) (v + prep + det + s)</b>          Analisando o local do rapto e a casa de Monica, <b>presumo que o suspeito</b> mora por aqui. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Asumir que el sospechoso: (1) (v + prep + det + s)</b>          - Tomando en cuenta el sitio del secuestro de Mónica y su apartamento, <b>asumo que el sospechoso</b> vive en algún lugar dentro de esta área.◇</p>
<p><b>Catch the unsub: (6) (v + det + n)</b>          - You know, they're right. Odds are we'll catch the unsub when he dumps the body. Or when he tries to snatch another kid. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Pegar o suspeito: (7) (v + det + s)</b>          - Sabe, eles estão certos. Devemos <b>pegar o suspeito</b> quando ele despejar o corpo ou tentar raptar outra criança. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Atrapar el sospechoso: (1) (v + det + s)</b>          - Sabes, tienen razón. Lo más probable es que <b>atrapemos al sospechoso</b> cuando tire el cuerpo. Cuando intenta arrebatar a otro niño. ◇</p>
<p><b>Drive the unsub: (7) (v + det + n)</b>          - That responsibility is probably what drove the unsub into action. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Levar o suspeito: (3) (v + det + s)</b>          - Essa responsabilidade é, provavelmente, o que <b>levou o suspeito</b> em ação. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Llevar al sospechoso: (1) (v + prep + det + s)</b>          - Esa responsabilidad es probablemente lo que <b>llevó al sospechoso</b> a actuar. ◇</p>
<p><b>Find the unsub: (15) (v +</b></p>	<p><b>Achar o suspeito: (5) (v +</b></p>	<p><b>Encontrar al sospechoso:</b></p>

<p><b>det + n</b> - He's following the patterns. We track the storms; we find the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>det + s</b> - Está seguindo os padrões. Rastreamos as tempestades, achamos o suspeito. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>(1) (v + prep + det + s)</b> Está siguiendo los patrones. Seguimos las tormentas; <b>encontramos al sospechoso</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Know the unsub: (27) (v + det + n)</b> - You also posited early on that Vanessa Holden might know the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Conhecer o suspeito<sup>1</sup>: (7) (v + det + s)</b> - Também previu que Vanessa Holden poderia <b>conhecer o suspeito</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Conocer al sospechoso<sup>1</sup>: (3) (v + prep + det + s)</b> - También postuló desde el principio que Vanessa Holden podría <b>conocer al sospechoso</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>See also: meet the unsub</b></p>	<p><b>Ver também: conhecer o suspeito<sup>2</sup></b></p>	<p><b>Ver también: conocer al sospechoso<sup>2</sup></b></p>
<p><b>Meet the unsub: (10) (v + det + n)</b> - What we need is more information on the movements of our victims before they met the unsub. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Conhecer o suspeito<sup>2</sup>: (7) (v + det + s)</b> - O que precisamos é de mais informações sobre os movimentos das vítimas antes de <b>conhecerem o suspeito</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Conocer al sospechoso<sup>2</sup>: (3) (v + prep + det + s)</b> Lo que necesitamos es más información sobre los movimientos de nuestras víctimas antes de <b>conocer al sospechoso</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>See also: know the unsub</b></p>	<p><b>Ver também: conhecer o suspeito<sup>1</sup></b></p>	<p><b>Ver también: conocer al sospechoso<sup>1</sup></b></p>
<p><b>Profile the unsub: (12) (v + det + n)</b> - We did profile the unsub would be someone close to the investigation. - But only Ruiz and I heard him say it. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perfilar o suspeito: (1) (v + det + s)</b> - Nós <b>perfilamos o suspeito</b> como alguém que poderia ser próximo à investigação. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Perfilar el sospechoso: (1) (v + det + s)</b> - Sí <b>perfilamos que el sospechoso</b> sería alguien cercano a la investigación. Pero solo Ruiz y yo lo escuchamos decirlo. ◊</p>

Victimology (s). Freq. 173	Vitimologia (s). Freq. 116	Victimología (s). Freq. 88
The study of the effects experienced by victims of crimes; the study about the victims of crimes.	O estudo dos efeitos experienciados pelas vítimas de crimes; o estudo sobre as vítimas de crimes.	El estudio de los efectos experimentados por las víctimas de los delitos; el estudio sobre las víctimas de los delitos.
<b>Adjectival collocations</b>	<b>Colocações adjetivas</b>	<b>Colocaciones adjetivas</b>
<p><b>Common victimology: (2) (adj + n)</b> - It doesn't make any sense. There's no common victimology, no sexual component, no robbery, no geographical connection. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia comum: (1) (s + adj)</b> - Não faz sentido. Não há <b>vitimologia comum</b>, nenhum componente sexual, nenhum roubo, nenhuma conexão geográfica. ◊</p>	<p><b>Victimología común: (1) (s + adj)</b> - No tiene ningún sentido. No hay <b>victimología común</b>, no hay componente sexual, no hay robo, no hay conexión geográfica. ◊</p>

<p><a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p> <p><b>Different victimology: (5) (adj + n)</b>  - Different locations, different victimology. It's possible we're looking for someone who'll hunt indiscriminately.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia diferente: (1) (s + adj)</b>  - Locais diferentes, <b>vitimologia diferente</b>. Podemos estar procurando alguém que caça indiscriminadamente.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología diferente: (1) (s + adj)</b>  - Distintas localizaciones, <b>diferente victimología</b>. Es posible que estemos buscando a alguien quecazará indiscriminadamente. ◇</p>
<p><b>High-risk victimology: (1) (adj + n)</b>  - This guy started with prostitutes, a high-risk victimology. Took a year off, came back, killed asocialite.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia de alto risco: (1) (s + adj)</b>  - Este cara começou com prostitutas, uma <b>vitimologia de alto risco</b>. Parou por um ano, voltou, matou uma socialite.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología de alto riesgo: (1) (s + prep + adj)</b>  - Este tipo empezó con prostitutas, una <b>vitimología de alto riesgo</b>. Se tomó un año sabático, volvió, mató a un miembro de la alta sociedad. ◇</p>
<p><b>Ideal victimology: (1) (adj + n)</b>  - He's narrowing in on his ideal victimology. Babysitting somehow fits into the scenario he's trying to recreate.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia ideal: (1) (s + adj)</b>  - Ele está se estreitando em sua <b>vitimologia ideal</b>. O serviço de babá de alguma forma se encaixa no cenário que ele está tentando recriar. ◇</p>	<p><b>Victimología ideal: (2) (s + adj)</b>  - Se está acercando a su <b>vitimología ideal</b>. El cuidado de los niños de alguna manera encaja en el escenario que está tratando de recrear.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Limited victimology: (1) (adj + n)</b>  - We've been compiling preliminary profile thoughts based on the limited victimology that we have.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia limitada: (1) (s + adj)</b>  - Temos feito um perfil preliminar baseado na <b>vitimologia limitada</b> que temos.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología limitada: (1) (s + adj)</b>  - Hemos estado compilando pensamientos de perfil preliminares basados en la <b>vitimología limitada</b> que tenemos. ◇</p>
<p><b>Narrow victimology: (2) (adj + n)</b>  - JJ, we need you on the news. Narrow victimology. Let people know that he's out there right now.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia estreita: (1) (s + adj)</b>  - JJ, precisamos de você no noticiário. <b>Vitimologia estreita</b>. Deixe as pessoas saberem que ele está lá fora agora. ◇</p>	<p><b>Victimología estrecha: (1) (s + adj)</b>  - JJ, te necesitamos en las noticias. <b>Victimología estrecha</b>. Que la gente sepa que él está ahí fuera ahora mismo.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Organized victimology: (1) (adj + n)</b>  - No, it's organized victimology. The father figures are likely surrogates for his own.  (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia organizada: (1) (s + adj)</b>  - Não, é <b>vitimologia organizada</b>. As figuras paternas são provavelmente substitutas das suas. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología organizada: (1) (s + adj)</b>  - No, es <b>vitimología organizada</b>. Las figuras paternas son probablemente sustitutos de los suyos. ◇</p>
<p><b>Similar victimology: (1) (adj)</b></p>	<p><b>Vitimologia similar: (1) (s +</b></p>	<p><b>Victimología similar: (1) (s</b></p>

<p><b>+ n)</b> - Sounds like a musical. Similar victimology: young professionals killed 3 miles apart. Is there a gang situation in Portland? (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>adj)</b> - Parece um musical. <b>Vitimologia similar:</b> jovens profissionais mortos a 3 milhas de distância. Existe uma situação de gangue em Portland?◇</p>	<p><b>+ adj)</b> - Suena como el título de un musical. <b>Victimología similar</b> ... jóvenes profesionales asesinados a 5 kilómetros. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Specific victimology: (2) (adj + n)</b> - And for the sexual aspects, we need to look at the specific victimology. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia específica: (1) (s + adj)</b> - E para os aspectos sexuais, precisamos olhar para a <b>vitimologia específica</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología específica: (1) (s + adj)</b> - Y por el aspecto sexual, necesitamos mirar la <b>victimología específica</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Nominal collocations</b></p>	<p><b>Colocações nominais</b></p>	<p><b>Colocaciones nominales</b></p>
<p><b>Shift in victimology: (3) (n + prep + n)</b> - If this is our unsub, taking a pregnant woman would be a huge shift in victimology. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Mudança na vitimologia: (3) (s + prep + det + s)</b> - Se esse é o nosso suspeito, pegar uma mulher grávida seria uma grande <b>mudança na vitimologia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Cambio en la victimología: (4) (v + prep + det + s)</b> - Si este es nuestro sospechoso, llevándose a una mujer embarazada. Sería un gran <b>cambio en la victimología</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Victimology pool: (1) (n + n)</b> - Both of our victims shared the same diagnosis, and Brian has it, too. That's a small victimology pool. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Grupo de vitimologia: (1) (s + prep + s)</b> - Ambas as nossas vítimas compartilharam o mesmo diagnóstico, e Brian também o tem. Isso é um pequeno <b>grupo de vitimologia</b>.◇</p>	<p><b>Grupo de victimología: (1) (s + prep + s)</b> - Nuestras dos víctimas compartieron el mismo diagnóstico, y Brian también lo tiene. Es un pequeño <b>grupo de victimología</b>. ◇</p>
<p><b>Verbal collocations</b></p>	<p><b>Colocações verbais</b></p>	<p><b>Colocaciones verbales</b></p>
<p><b>Change the victimology: (9) (v + det + n)</b> - Something must have happened between the last prostitute and Vanessa Holden making him change his victimology. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Mudar a vitimologia: (8) (v + det + s)</b> - Algo deve ter acontecido entre a última prostituta e Vanessa Holden, para fazê-lo <b>mudar sua vitimologia</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Cambiar la victimología: (5) (v + det + s)</b> - Algo debe haber pasado entre la última prostituta y Vanessa Holden que le <b>hizo cambiar su victimología</b>. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Cover victimology: (3) (v + n)</b> - We need to cover victimology. Garcia, find out everything you can about the officers killed. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Cobrir a vitimologia: (3) (v + det + s)</b> - Nós precisamos <b>cobrir a vitimologia</b>. Garcia, encontre tudo que puder sobre os oficiais mortos. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Cubrir la victimología: (1) (v + det + s)</b> - Tenemos que <b>cubrir la victimología</b>. García, averigua todo lo que puedas sobre los agentes asesinados. ◇</p>
<p><b>Fit the victimology: (8) (v + det + n)</b> - Dad walked out and the unsub had to punish him for</p>	<p><b>Encaixar na vitimologia: (6) (v + prep + det + s)</b> - O papai saiu e o suspeito teve que puni-lo por isso.</p>	<p><b>Encajar en la victimología: (3) (v + prep + det + s)</b> - Papá se fue y el sospechoso tuvo que castigarlo por eso.</p>

<p>that. - James doesn't fit the victimology. It doesn't make sense the unsub would take him. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- James não se <b>encaixa na vitimologia</b>. Não faz sentido que o suspeito o leve. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p>- James no <b>encaja en la victimología</b>. No tiene sentido que el sospechoso se lo lleve. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Invert the victimology: (1) (v + det + n)</b> - He wanted them to feel the pain of losing their families. - I think we've inverted our victimology. We've been focused on the murdered family members. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Inverter a vitimologia: (1) (v + det + s)</b> - Ele queria que eles sentissem a dor de perder suas famílias. - Acho que <b>invertemos nossa vitimologia</b>. Estamos focados nos membros da família assassinados. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Invertir la victimología: (1) (v + det + s)</b> - Quería que sintieran el dolor de perder a sus familias. - Creo que <b>hemos invertido nuestra victimología</b>. Nos hemos centrado en los miembros de la familia asesinados. ◇</p>
<p><b>Isolate victimology: (1) (v + n)</b> - His first two victims may be from the same social class, but if we isolate victimology this soon, it may cause him to go outside of his own comfort zone. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Isolar a vitimologia: (1) (v + det + s)</b> - As primeiras 2 vítimas podem ser da mesma classe, mas se <b>isolarmos a vitimologia</b> agora, ele pode sair de sua zona de conforto. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Aislar la victimología: (1) (v + det + s)</b> - Sus primeras dos víctimas podrán haber sido del mismo estrato social, pero si <b>aislamos su victimología</b> tan pronto, podría hacer que vaya por fuera de su zona de seguridad, y luego será aún más difícil de rastrear. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>
<p><b>Match victimology: (2) (v + n)</b> - JJ, have Garcia match victimology and abduction sites to HSK findings for the last year. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Combinar a vitimologia: (1) (v + det + s)</b> - JJ, faça com que Garcia <b>combine a vitimologia</b> e os locais de abdução com as descobertas do HSK no ano passado. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Relacionar la victimología: (1) (v + det + s)</b> - JJ, haz que García <b>relacione la victimología</b> y los sitios de abducción con los hallazgos del HSK del último año. ◇</p>
<p><b>Victimology vary: (1) (v + n)</b> - Besides gender, the victimology varies. He doesn't care about age or race. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Vitimologia variar: (1) (s + v)</b> - Além do gênero, a <b>vitimologia varia</b>. Ele não se importa com idade ou raça. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Victimología variar: (1) (s + v)</b> - Además del género, la <b>victimología varía</b>. No le importa la edad ni la raza. ◇</p>
<p><b>Work victimology: (5) (v + n)</b> - Dave, you and Morgan work the dumpsite, and I'll work victimology at the local Sheriff's station. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Trabalhar na vitimologia: (4) (v + prep + det + s)</b> - Dave, você e Morgan o trabalho do lixão, e eu vou <b>trabalhar na vitimologia</b> na estação do xerife local. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>	<p><b>Trabajar en victimología: (5) (v + prep + s)</b> - Dave, tú y Morgan trabajáis en el vertedero, y yo <b>trabajaré en victimología</b> en la comisaría del sheriff local. (Fonte: <a href="http://www.tvsubtitles.net/">http://www.tvsubtitles.net/</a>).</p>

## Considerações finais

O trabalho que aqui se encerra teve como objetivo principal analisar, de forma contrastiva entre os idiomas inglês, português e espanhol, do ponto de vista morfossintático e léxico-semântico as colocações especializadas criminais obtidas a partir do *corpus* de estudo composto pelas legendas das séries *CSI – Crime Scene Investigation*, *Criminal Minds* e *NCIS*, bem como o de propor um modelo de Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal.

Esta tese iniciou-se com a apresentação das séries que compõem o *corpus* de estudo, bem como, com a justificativa para a utilização destas para abordar o léxico pertencente ao âmbito criminal, já que esta se configura como uma área de especialidade. Tal justificativa foi feita buscou evidenciar a presença de vocabulário especializado composto por termos e semitermos (HOFFMANN, 1998) nas séries selecionadas, uma vez que elas retratam de maneira fiel o trabalho desenvolvido por investigadores e cientistas forenses, tal como afirma Fromm (2011). Além disso, por meio da Etnoterminologia proposta por Barbosa (2006, 2009) entendemos que existe um continuum entre a cientificidade e banalização dos termos, fazendo com que estes ocorram em discursos mais científicos entre especialistas, além de ocorrerem variações a medida em que são banalizados, ou seja, em que ocorrem em discursos entre especialistas e leigos e entre apenas entre leigos.

Ademais, discorremos sobre como as legendas de séries são utilizadas para desenvolver a aprendizagem incidental de idiomas, abordando as pesquisas de Csomay e Petrovic (2012), de Webb e Rogers (2009) e de Webb (2010) que enfatizam os benefícios do uso de séries como *corpus* em estudos sobre o tema.

Outro aspecto que considerado, ainda neste capítulo, foi a utilização de legendas não comerciais, produzidas por fãs, uma vez que estas são disponibilizadas na internet e, de fácil acesso por parte de todos que tenham interesse em utilizá-las nos próprios vídeos das séries ou estudar, seja um idioma, sejam os processos de tradução utilizados pelas equipes de fãs, já que estas, como afirma Sperandio (2015), têm sempre um membro que é especialista na área em que a série se enquadra.

Em seguida, apresentamos os pressupostos teóricos que servem de base para esta pesquisa, discorrendo sobre a Linguística de Corpus (BIBER, 1998; TOGNINI- BONELLI, 2002) e a Terminologia (CABRÉ, 1999; BARROS, 2004, FINATTO et. al. 2010; FINATTO, 2011; KRIEGER, 2013; FADANELLI, 2017), sobre como a interdisciplinaridade entre ambas as disciplinas contribuem para a identificação das palavras-

chave que servem como nódulos de busca para as colocações especializadas em nossa pesquisa.

Depois, abordamos a Fraseologia baseada em *corpus* (SINCLAIR, 1991; COWIE, 1998; GROOM, 2005; NESSELHAUF, 2005; GRANGER E MEUNIER, 2008; GRANGER EPAQUOT, 2008; BERTRAÑA E BERTRÁN, 2008; GRIES, 2008; TAGNIN, 2013; ORENHA-OTTAIANO, 2004, 2009, 2021), outra disciplina que se utiliza do arcabouço teórico-metodológico da LC para empreender suas análises, enfatizando como ela se relaciona com este trabalho.

As colocações especializadas (L'HOMME, 2000, BEVILACQUA, 2004, 2005; ORENHA-OTTAIANO, 2009, 2012a, 2020) foram abordadas na sequência, de forma que pudéssemos evidenciar os estudos teóricos que contribuem para a sua identificação e classificação taxonômica, enfatizando como nos baseamos nos estudos de Orenha-Ottaiano (2004, 2009, 2011) para analisarmos a estrutura morfossintática das colocações visando, além de identificar sua taxionomia, obter informações que são utilizadas, posteriormente, na elaboração da amostra de dicionário pedagógico que trazemos.

Por fim, abordamos a Fraseografia (OLÍMPIO DE OLIVEIRA, 2007; Tristá Pérez, 1998) e a Teoria Funcional da Lexicografia (FUERTES-OLIVERA; BERGENHOLTZ, 2011; FUERTES-OLIVERA; TARP, 2014; TARP, 2014), disciplinas que contribuíram para as bases teórico-metodológicas de nossa amostra de dicionário de colocações especializadas.

No quarto capítulo, são expostos os procedimentos metodológicos empregados para a compilação do *corpus* e sua posterior análise, a qual foi realizada por meio das ferramentas *Keywords* e *Word Sketch* da plataforma on-line *Sketch Engine* (KILGARIFF, 2004), além de termos demonstrado como se deu o processo de análise das colocações especializadas obtidas. Nesse capítulo, foi retratado, ainda, o procedimento para a elaboração do Dicionário Pedagógico de Colocações Especializadas do âmbito Criminal, o DPCEC.

No quinto capítulo, apresentamos os resultados obtidos por meio das análises quantitativas, a qual revelou que o CCI (Corpus Criminal Inglês) tem um total de 9.643.089 *tokens*, enquanto o CCP (Corpus Criminal Português) apresenta um total de 6.709.702 e o CCE (Corpus Criminal Espanhol) conta com um total de 3.843.577 de *tokens*. Ainda por meio da análise quantitativa foram selecionadas as 5 palavras-chave que serviram de ponto de partida para a busca e o levantamento das colocações especializadas apresentadas nesta pesquisa, a saber: *unsub*, *autopsy*, *gunshot*, *alibi* e *victimology*, sendo *unsub* a palavra-chave que apresenta o maior *score* para o índice de chavidade (259.500) no CCI.

A partir da análise morfossintática, também apresentada no capítulo cinco, as colocações especializadas obtidas por meio das palavras-chave selecionadas foram classificadas de acordo com a sua estrutura morfossintática, uma vez que esta classificação, além de ser necessária para a identificação da taxonomia das colocações especializadas, se faz necessária para a obtenção de informações sintáticas que foram inseridas na amostra do DPCEC apresentada no capítulo seis. Por meio dessa análise, identificamos que as colocações especializadas que ocorrem no âmbito criminal são formadas por termos que ocorrem ora como base, ora como colocado.

Outrossim, foi possível identificar colocações dos tipos nominal, adjetiva e verbal, não havendo nenhuma ocorrência de colocações do tipo adverbial. Tal fato nos leva a crer que este tipo de colocação (adverbial) não seja comum na área investigativa/criminal. Entretanto, reforçamos que mais dados devem ser levantados e análises devem ser realizadas, a fim de comprovar, ou não, a baixa frequência de colocações adverbiais na área em questão.

Ainda no capítulo cinco, foi realizada uma análise léxico-semântica, na qual se buscou contrastar a escolha léxica em cada um dos idiomas que compõem os *corpora* de estudo, bem como, discorrer sobre aspectos semânticos das colocações expostas, apontando e discutindo aspectos relevantes sobre o termo *unsub*, amplamente utilizado em nosso *corpus* de estudo, além de apontar semelhanças entre as colocações *airtgh alibi* e *ironclad alibi* e entre *know the unsub* e *meet the unsub*, sendo esses pares de colocações traduzidos da mesma forma para as línguas de chegada contempladas nessa pesquisa (português e espanhol).

Ademais, apresentamos uma análise do termo *unsub*, que teve origem a partir da abreviação de *unknown subject* ou *unidentified subject*, encontrando-se já dicionarizado nos dicionários *Lexico* e *Merriam-Webster*, conforme apresentado na seção 5.3.1, incluindo uma marcação de uso informal por membros da polícia, fato que nos levou a mantê-lo em nossa análise.

No último capítulo trazemos o protótipo do DPCEC, sua macro e microestrutura, além de realizarmos uma discussão sobre a definição do público-alvo ao qual se destina. São apresentadas como justificativa para a escolha das informações inseridas na microestrutura do dicionário pedagógico o fato de que este se insere em projetos de pesquisa maiores intitulados *COMADEC - A Compilação de Materiais Didáticos e Dicionários Especializados de Colocações Baseados em Corpora* e *A phraseographical methodology and model for an online corpus-based Multilingual Collocations Dictionary Platform*, ambos coordenados pela Profa. Dra. Adriane Orenha-Ottaiano. Em seguida, apresentamos uma amostra do DPCEC contendo colocações especializadas levantadas a partir das cinco

palavras-chave selecionadas no capítulo cinco.

Assim, ao concluir este trabalho, esperamos ter conseguido ressaltar a importância da Linguística de Corpus para a Fraseologia, bem como a importância da Teoria Funcional da Lexicografia para a Fraseografia, enfatizando a importância dessas áreas para o desenvolvimento de pesquisas fraseológicas e fraseográficas, especialmente no que concerne a produção de materiais fraseográficos voltados para as colocações especializadas.

Outrossim, espera-se que a discussão aqui apresentada sobre a utilização de séries de TV como corpus de pesquisas linguísticas de áreas de especialidade seja aprofundada e levada adiante, a fim de que mais pesquisas que utilizem corpora como esses sejam desenvolvidas, haja vista os benefícios que tais pesquisas trazem tanto para as áreas envolvidas, quanto para estudantes das áreas de tradução e para tradutores que possam vir a trabalhar com esse tipo de texto.

Por fim, esperamos que, futuramente, seja possível aprimorar e ampliar os estudos aqui apresentados, a fim de desenvolver uma proposta de Dicionário Pedagógico de Colocações Especializadas *on-line*, visando facilitar e amplificar a divulgação das colocações, não apenas do âmbito criminal, mas também de outras áreas de especialidade que possam beneficiar os estudantes dos cursos de Licenciatura em Letras Bacharelado em Letras com Habilitação para Tradução.

## REFERÊNCIAS

- ALVES, E. H. **Plataforma Multilíngue de Dicionários de Colocações como ferramenta pedagógica para a formação de tradutores aprendizes**: um estudo experimental. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto. Pesquisa em andamento. Set. 2022.
- AIRTIGHT. *In*: **Cambridge Dictionary**. Disponível em: <https://dictionary.cambridge.org/pt/dicionario/ingles-portugues/airtight>. Acesso em: 10 jan. 2022.
- ALMELA, M. The case of verb-adjective collocations: corpus-based analysis and lexicographical treatment. **Revista de Lingüística y Lenguas Aplicadas**, Valência, v. 6, 2011. DOI: <https://doi.org/10.4995/rlyla.2011.892>. Disponível em: <https://polipapers.upv.es/index.php/rdlyla/article/view/892>. Acesso em: 15 jun. 2019.
- ALONSO RAMOS, M.: Entón, ¿é unha colocación ou non? análise contrastiva das colocacións. **Cadernos de Fraseloxía Galega**, Santiago de Compostela, n. 8, p. 29-43. 2006. Disponível em: [http://cirp.es/pub/docs/cfg/cfg08\\_02.pdf](http://cirp.es/pub/docs/cfg/cfg08_02.pdf). Acesso em: 15 jun. 2019.
- AXN. Disponível em: <https://br.axn.com/>. Acesso em: 15 set. 2021.
- BALLY, C. **Précis de stylistique**. Genève: Eggiman, 1905.
- BALLY, C. **Traité de stylistique française**. Paris: Klincksieck, 1909. v. 2
- BANG, M.; FROMM, G. Terminologia em série: House M. D. **Revista Entreletras** (Araguaína), v. 4, n. 2, p. 114-133, 2013. Disponível em: <https://sistemas.uft.edu.br/periodicos/index.php/entreletras/article/view/995>. Acesso em: 20 mar. 2018.
- BARBOSA, M. A. Terminologia e Lexicologia: plurissignificação e tratamento transdisciplinar das unidades lexicais nos discursos etno-literários. **Revista de Letras**, Fortaleza, n. 27, v.1/2, p. 103-107, 2005. Disponível em: <https://repositorio.ufc.br/handle/riufc/17417>. Acesso em: 24 jun. 2022.
- BARBOSA, M. A. Para uma etno-terminologia: recortes epistemológicos. **Ciência e Cultura**, São Paulo, v. 58, n. 2, Apr/June, 2006. Disponível em: [http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200018&script=sci\\_arttext&tlng=pt](http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200018&script=sci_arttext&tlng=pt). Acesso em: 24 jun. 2022.
- BARBOSA, M. A. Terminologia aplicada: percursos interdisciplinares. **Polifonia**. Cuiabá, n.17, p.29-44, 2009. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/polifonia/article/view/1006>. Acesso em 24 jun. 2022.
- BARROS, L. A. **Curso básico de Terminologia**. São Paulo: EDUSP, 2004.

BARTSCH, S. Structural and functional properties of collocations in English. **A corpus study of lexical and pragmatic constraints on lexical co-occurrence**. Tübingen: Gunther NarrVerlag, 2004.

BENSON, M. The Structure of collocational dictionary. **International Journal of Lexicography**, v. 2, n. 1, p. 1-14, 1989. DOI: <https://doi.org/10.1093/ijl/2.1.1>. Disponível em: <https://academic.oup.com/ijl/article-abstract/2/1/1/962135>. Acesso em: 20 mar. 2018.

BERGENHOLTZ, H.; TARP. S. The concept of dictionary usage. **Nordic Journal of English Studies**, n. 3, 2004, p. 23-36. Disponível em: <https://ojs.uib.no/index.php/njes/article/viewFile/231/228>. Acesso em: 20 mar. 2018.

BERGENHOLTZ, H.; TARP. S. Two opposing theories: On H.E. Wiegand's recent discovery of lexicographic functions. **Journal of Linguistics**, Hermes, n. 31, p. 171-196, 2003 DOI: <https://doi.org/10.7146/hjlc.v16i31.25743>. Disponível em: <https://tidsskrift.dk/her/article/view/25743>. Acesso em 20 mar. 2018.

BERGENHOLTZ, H.; TARP. S. Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen. **Lexicographica**, n. 18, p. 253-263, 2002. DOI: <https://doi.org/10.1515/9783484604476.253>. Disponível em: <https://www.degruyter.com/document/doi/10.1515/9783484604476.253/html?lang=de>. Acesso em: 20 mar. 2018.

BERTHEMET, E. Colidioms: An Online Software for Phraseography and Paremiography. **Proceedings of the 15<sup>th</sup> EURALEX International Congress**, 2012. Disponível em: [https://www.euralex.org/elx\\_proceedings/Euralex2012/pp883-888%20Berthemet.pdf](https://www.euralex.org/elx_proceedings/Euralex2012/pp883-888%20Berthemet.pdf). Acesso em: 20 mar. 2018.

BERTHEMET, E. Compiling Phraseological Plurilingual Dictionary. KARPOVA, O.; KARTASHKOVA, F. (eds): **Lexicography and Terminology: a worldwide outlook**. Newcastle upon Tyne: Cambridge Scholars, 2009, p. 2-13.

BEVILACQUA, C. R. Fraseologia: perspectiva da língua comum e da língua especializada. **Revista Língua e Literatura**, v. 6 e 7, n. 10/11, p. 73-86, 2004/2005. Disponível em: <http://www.revistas.fw.uri.br/index.php/revistalinguaeliteratura/article/view/40>. Acesso em: 05 ago. 2016.

BEVILACQUA, C. R. Unidades fraseológicas especializadas: estado da questão em relação a sua definição, denominação e critérios de seleção. **TradTerm**, São Paulo, v.11, p. 237-253, 2005. <https://www.lume.ufrgs.br/handle/10183/183098>. Acesso em: 05 ago. 2016.

BIBER, D.; CONRAD, S.; REPPEN, R. **Corpus Linguistics: investigating language structure and use**. Cambridge: Cambridge University Press, 1998.

BRETAÑA, J. M. P.; BERTRÁN, A. P. Combined statistical and grammatical criteria for the retrieval of phraseological units in an electronic corpus. GRANGER, S.; MEUNIER, F. **Phraseology: an interdisciplinary perspective**. Amsterdam /Philadelphia: John Benjamins, 2008.

BUENDÍA, M.C.; FABER, P. Collocation Dictionaries: a comparative analysis. **MonTi: Monografías de Traducción e Interpretación**, Universidad de Valencia, Alicante, n. 6, p. 203-235, 2014. DOI: <http://dx.doi.org/10.6035/MonTI.2014.6.7>. Disponível em: <https://www.redalyc.org/pdf/2651/265134675007.pdf>. Acesso em: Acesso em 20 mar. 2018.

CABRÉ, M. T. **La terminología**: representación y comunicación. Barcelona, Intitut Universitari de Lingüística Aplicada. 1999.

CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Barcelona, Antártida. 1993.

CALDAS, A. D. D. R.: **A identificação de colocações especializadas extraídas do Corpus CSI e do Corpus Comparável Criminal para a elaboração de atividades didáticas**. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas São José do Rio Preto, 2017. 137 f.

CALDAS, A. D. D. R.; ORENHA-OTTAIANO, A. A compilação de atividades didáticas a partir de colocações especializadas da área criminal extraídas do corpus *CSI* – *Crime Scene Investigation*. **Revista Caminhos em Linguística Aplicada**, Taubaté, v.17, n. 3, p. 373-196, 2017. Disponível em: <http://periodicos.unitau.br/ojs/index.php/caminhoslinguistica/article/view/2351>. Acesso em: 22 set. 2017.

CAMBRIDGE DICTIONARY. Disponível em: <https://dictionary.cambridge.org/pt/>. Acesso em: 10 jan. 2022.

CASTRO, I. S. **A terminologia da série Grimm: uma análise com base em corpus paralelo inglês/português**. Monografia (Trabalho de Conclusão do Curso de Graduação em Bacharel em Línguas Estrangeiras Aplicadas ao Multilinguismo e à Sociedade da Informação) – Instituto de Letas, Universidade de Brasília, Brasília. 2019.

CARNEADO MORÉ, Z. V.: Tipología de las obras fraseográficas. **Anuario L/L**. n. 20. 1989. p. 66-82.

CORPAS PASTOR, G. **Manual de fraseología española**. Madrid: Editorial Gredos, 1996.

COSERIU, E. Structure lexicale et enseignement du vocabulaire. Actes du Premier Colloque International de Linguistique Appliquée, Nancy, 26-31 octobre, 1964. Faculté des Lettres et des Sciences Humaines de l'Université de Nancy. **Annales de l'Est**, Memorie, v. 31, p. 175- 217, 1966.

COWIE, A.P.: IJL: Dictionaries, language learning and phraseology. **International Journal of Lexicography**, v. 25, n. 4, p. 386-392, 2012. DOI: <https://doi.org/10.1093/ijl/ecs023>. Disponível em: <https://academic.oup.com/ijl/article-abstract/25/4/386/923782>. Acesso em: 20 mar. 2018.

COWIE, A.P. **Phraseology: Theory, Analysis, and Applications**. Clarendon Press.Oxford. 1998.

CRYSTAL, D. **The Cambridge Encyclopedia of the English Language**. Cambridge University Press. 1995.

**CSI: Crime Scene Investigation**. Disponível em <https://br.axn.com/series/csi>. Acesso em: 20 out. 2021.

**CSI: Crime Scene Investigation**. Disponível em: <http://www.tvsubtitles.net>. Acesso em: 01 jul. 2014.

CSOMAY, E.; PETROVIC, M. “Yes, your honor!”: A corpus-based study of technical vocabulary in discipline-related movies and TV shows. **System**, v. 40, n. 2, p. 305-315, 2012. DOI: <https://doi.org/10.1016/j.system.2012.05.004>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0346251X12000449>. Acesso em 20 mar. 2018.

DASH, N.S.: Corpus Linguistics: an introduction. **Encyclopedia of Life Support Systems**. 2008.

DAWES, E. ¿Lo small sigue siendo beautiful? La representación lexicográfica de usos no convencionales de frases hechas. GARCÉS GÓMEZ, M. P. (ed): **Lexicografía teórica y práctica**, A Coruña: Universidade da Coruña, p. 104-119, 2014.

ESPERANDIO, I. B. **Legendas de seriados de tema sobrenatural: uma abordagem terminológica para tradutores**. 2015. 229 f. Dissertação (Mestrado em Teorias Linguísticas do Léxico) – Universidade Federal do Rio Grande do Sul, Porto Alegre. 2015.

ESPERANDIO, I. B.; FINATTO, M. J. B. A definição terminológica na legendagem de seriados. **Cadernos de Letras**, Pelotas, n. 22, p. 17-38, 2014.

DOI: [HTTPS://DOI.ORG/10.15210/CDL.V0I22.4533](https://doi.org/10.15210/CDL.V0I22.4533). Disponível em: <https://periodicos.ufpel.edu.br/ojs2/index.php/cadernodeletras/article/view/4533>. Acesso em: 15 set. 2016.

FADANELLI, S. B. **Terminografia didático-pedagógica: metodologia para elaboração de recursos voltados para o ensino de inglês para fins específicos**. 2017. 198 f. Tese (Doutorado em Teorias Linguísticas do Léxico) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

FINATTO, M. J. B.; ZILIO, L. **Textos e termos por Lothar Hoffman**. Porto Alegre: Palotti, 2015.

FINATTO, M. J. B.; EVERS, A.; OLIVEIRA, C. M.; SILVA, M. C. A: Das terminologias às construções recorrentes: um percurso de estudos sobre linguagens especializadas. **Íkala**, Revista de Lenguaje y Cultura. v. 15. n. 25. 2010. Disponível em: <http://www.scielo.org.co/pdf/ikala/v15n25/v15n25a10.pdf>. Acesso em: 20 mar. 2018.

FISHEL, M. *et al.*: From subtitles to parallel corpora. **Proceedings of the 16th EAMT Conference**, Trento, Italy, 2012.

FIRTH, J. R. **Papers in linguistics 1934-1951**. Londres, Oxford University Press, 1957.

FRITZINGER, F.; HEID, U. Automatic grouping of morphologically related collocations. **Proceedings of the Corpus Linguistics 2009 Conference**, Liverpool/UK, 2009.

FROMM, G.: Ficção, tradução, terminografia e linguística de *corpus*: confluências. **Anais do SILEL**, Uberlândia, EDUFU, v. 2, n. 2, 2011. Disponível em: <http://www.ileel.ufu.br/anaisdosilel/pt/arquivos/silel2011/318.pdf>. Acesso em: 15 set. 2016.

FUERTES-OLIVERA, P. A.; BERGENHOLTZ, H. eds.: e-Lexicography: The Internet, Digital Initiatives and Lexicography. Londres e Nova York, **Continuum**, 2011.

FUERTES-OLIVERA, P. A.; TARP, S.: Theory and practice of specialised online dictionaries: Lexicography versus terminography. Berlín y Boston, **De Gruyter**, 2014.

GLEDHILL, C. **Collocations in Science Writing**. Tübingen: Gunter Narr, 2000.

GRANGER, S. Prefabricated patterns in advanced EFL writing: collocations and formulae. COWIE A. P. (ed.). **Phraseology: theory, analysis and applications**, Oxford: Oxford University Press, p. 145-160, 1998.

GRANGER, S.; MEUNIER, F. **Phraseology: an interdisciplinary perspective**. Amsterdam / Philadelphia: John Benjamins, 2008.

GRANGER, S.; PAQUOT, M. Disentangling the phraseological web. In: GRANGER, S.; MEUNIER, F. **Phraseology: an interdisciplinary perspective**. Amsterdam / Philadelphia: John Benjamins, 2008.

GOUADEC, D. **Terminologie: constitution des données**. Paris. AFNOR, 1990. GRIES, S. T. Phraseology and linguistic theory: a brief survey. GRANGER, S.; MEUNIER, F. **Phraseology: an interdisciplinary perspective**. Amsterdam / Philadelphia: John Benjamins, 2008.

GROOM, N. Pattern and meaning across genres and disciplines: An exploratory study. **Journal of English for Academic Purposes**. Elsevier, v. 4, p. 257-277, 2005. DOI: <https://doi.org/10.1016/j.jeap.2005.03.002>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1475158505000160>. Acesso em 15 jun. 2019.

HAUSMANN, F. J. Kollokationen im deutschen wörterbuch: ein beitrag zur theorie des lexikographischen biespiels. BERGENHOLTZ, H; MUGDAN, J. (eds). **Lexikographie und Grammatik**. Tübingen: Niemeyer, 1985.

HAUSMANN, F. J. Wortschatzlernen ist kollokationslernen: zum lheren und lernen französischerwortverbindungen. **Praxis des neusprachlichen Unterrichts**, 31, p. 305-406, 1984.

HEID, U.; MARTIN, W.; POSCH, I. An overview of approaches towards the description of collocations. **Feasibility of standards for collocational description of lexical items**. In: **Eurotra 7 – Report**, Stuttgart/Amsterdam, 1991.

HEID, U. Computational phraseology: an overview. GRANGER, S.; MEUNIER, F. **Phraseology: an interdisciplinary perspective**. Amsterdam / Philadelphia: JohnBenjamins, 2008.

HOFFMANN, L. Grundbegriffe der Fachsprachenlinguistik. **Germanistisches Jahrbuch für Nordeuropa**, n. 7. Helsinki, Estocolmo, 1988, p. 9-16. Deutsche Fachsprachen in Forschung und Lehre. FINATTO, M. J. B.; ZILIO, L. Textos e termos por Lothar Hoffmann. Porto Alegre: Palotti, 2015.

IÑESTA MENA, E. V.; PAMIES BERTRÁN, A. **Fraseología y metáfora: aspectos tipológicos y cognitivos**. Granada: Granada Lingüística, 2002.

KILGARRIFF, A. *et al.* **The Sketch Engine: ten years on**. *Lexicography*, 1: 7-36, 2014.

KRIEGER, M. G. A heterogeneidade do léxico especializado e perfis terminológicos. MURAKAWA, C. A. A.; NADIN, O. L. **Terminologia: uma ciência interdisciplinar**. São Paulo: Cultura Acadêmica, 2013.

LAUFER, B.; WALDMAN, T. Verb-noun collocations in second language writing: a corpus analysis of learners' English. **Language Learning**. Ann Arbor, 61:2, p. 647-672, 2011. DOI: <https://doi.org/10.1111/j.1467-9922.2010.00621.x>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9922.2010.00621.x>. Acesso em: 25 fev. 2015.

L'HOMME, M. C. Understanding specialized lexical combinations. **Terminology**, Amsterdam, v. 6, n. 1, p. 89-110, 2000.

L'HOMME, M. C.; BERTRAND, C. Specialized lexical combinations: should they be described as collocations or in terms of selectional restrictions? **Proceedings**. NinthEuralex International Congress, p. 497-506, 2000. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d4fe3922fef5ee930a20e9dde7586ad7ffc9a54>. Acesso em 25 fev. 2015.

MALÁ, M. Phraseology in learner academic English: Corpus-driven approaches. **Discourse and Interaction**, n. 2, p.75-88, 2020. Disponível em: <https://www.ceeol.com/search/article-detail?id=975553>. Acesso em 30 fev. 2021.

MEL'ČUK, I. Phrasemes in language and phraseology in linguistics. EVERAERT, M., VAN DER LINDEN, E. J.; SCHENK, A. (eds.) **Idioms: Structural and Psychological Perspectives**, Hillsdale, NJ: Lawrence Erlbaum Associates, p. 167– 232, 1995.

MIRANDA, A. K. P. Os avanços na pesquisa fraseográfica no brasil. DORNELES, M. R. H.; FONSECA, J. Z. B. (org.). **Simpósio Nacional de Línguas e Literaturas. Encontro Nacional de Literatura e Filosofia**. Anais eletrônicos, Aquidauana, p.102-114, 2015. Disponível em: <https://periodicos.ufms.br/index.php/ansimp/article/view/1068>. Acesso em: 30 out. 2020.

MIRONESKO, E. **Fraseología rusa: teoría y práctica**. Barcelona: Librería Universitaria, 1997.

MONTEIRO-PLANTIN, R. S. **Fraseologia**: era uma vez um Patinho Feio no ensino de língua materna. Fortaleza: Imprensa Universitária, 2014.

NESSELHAULF, N. **Collocations in a learner corpus**. Amsterdam: John Benjamins Publishing Company, 2005.

NOVODVORSKI, A.; FINATTO, M. J. B: Linguística de Corpus no Brasil: uma aventura mais do que adequada. **Letras e Letras**, Uberlândia, v. 30, n. 2, p. 7-16, 2014. DOI: <https://doi.org/10.14393/LL60-v30n2a2014-1>. Disponível em: <https://seer.ufu.br/index.php/letraseletras/article/view/28516/>. Acesso em: 30 out. 2020.

OLÍMPIO DE OLIVEIRA, M. E. Issues and Perspectives in Bilingual Phraseography. **Open Linguistics**, n.2, p. 577–591, 2016. DOI: <https://doi.org/10.1515/opli-2016-0031>. Disponível em: <https://www.degruyter.com/document/doi/10.1515/opli-2016-0031/html>. Acesso em: 15 jan. 2019.

OLÍMPIO DE OLIVEIRA, M. E. **Fraseografía teórica y práctica**. Frankfurt and Main: Peter Lang, 2007.

ORENHA-OTTAIANO, A.; GARCIA, M.; OLÍMPIO DE OLIVEIRA, M. E.; L'HOMME, M.; ALONSO RAMOS, M.; VALÊNCIO, C. R.; TENÓRIO, W.: Corpus- based methodology for an Online Multilingual Collocations Dictionary: First Steps. **Electronic Lexicography in the 21st century - Proceedings of eLex**, Lexical Computing CZ s.r.o., Brno, Czech Republic, p. 1-28, 2021. Disponível em: <https://www.grupolys.org/biblioteca/OreGarSilLHoAloValTen2021a.pdf>. Acesso em: 10 nov. 2021.

ORENHA-OTTAIANO, A The creation of an online English collocations platform to help develop collocational competence. **Phrasis: Revista di studi fraseologici e paremiologic**. Associazione Italiana di Fraseologia e Paremiologia, v. 1, p. 59-81, 2020. Disponível em: <http://www.phrasis.it/rivista/index.php/rp/article/view/57>. Acesso em 10 jan. 2021.

ORENHA-OTTAIANO, A The compilation of an Online Corpus-Based Bilingual Collocations Dictionary: motivations, obstacles and achievements. **E-Lex 2017**, 2017, Leiden, Holanda. **Electronic Lexicography in the 21st Century - Proceedings of E-Lex Conference 2017**. Leiden, Holanda, p. 458-473, 2017. Disponível em: <https://elex.link/elex2017/wp-content/uploads/2017/09/paper27.pdf>. Acesso em: 10 jan. 2018.

ORENHA-OTTAIANO, A. The compilation of an online corpus-based bilingual collocations dictionary. **CORPAS PASTOR**, G. (Org). **Computerized and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives**. Genebra: Editions Tradulex, v. 1, p. 486-493, 2016.

ORENHA-OTTAIANO, A. Ensino de inglês como LE e contribuições pedagógicas de um glossáriobílingue de colocações. **Signótica**, Goiania, v. 27, n. 2, p. 485-510, 2015. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=6323023>. Acesso em: 15 fev. 2016.

ORENHA-OTTAIANO, A. An English collocations e-workbook designed to Brazilian Portuguese speakers. **Learner Corpus Research Abstract Book**, Bergen, p. 116-117, 2013.

ORENHA-OTTAIANO, A. Collocations and the design of teaching materials for second language learners. **TALC10 Proceedings of the 10th International Conference on Teaching and Language Corpora**, Varsovia, Polonia. Warsaw: Institute of Applied Linguistics, p. 93-103, 2013a.

ORENHA-OTTAIANO, A. English collocations extracted from a corpus of university learners and its contribution to a language teaching pedagogy. **Acta Scientiarum: language and culture**. v.34, p. 241-251, 2012. DOI: Doi: 10.4025/actascilangcult.v34i2.17130. Disponível em: <https://repositorio.unesp.br/handle/11449/74079>. Acesso em: 15 fev. 2016.

ORENHA-OTTAIANO, A. Semelhanças e diferenças entre colocações e colocações especializadas. ORTIZALVAREZ, M. L. (Org.). **Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia**. Campinas: Editora Pontes, v.2, p.147- 163, 2012a.

ORENHA-OTTAIANO, A. **Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não- juramentado**. 2009. 282 f. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociência, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2009.

ORENHA-OTTAIANO, A. **A compilação de um glossário bilíngue de colocações, na área de jornalismo de negócios, baseado em corpus comparável**. 2004. 246 f. Dissertação (Mestrado em Estudos Linguísticos e Literários) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2004.

ORENHA-OTTAINO, A.; ROCHA, J. M. P. Colocações especializadas na área médica extraídas a partir do corpus House M.D. **Cadernos do IL**, Porto Alegre, n.44, p. 295-318, 2012. Disponível em: <https://repositorio.unesp.br/handle/11449/122632>. Acesso em: 15 fev. 2015.

PATIÑO, P. (2014). Towards a definition of specialized collocation. QUIROZ, G. & PATIÑO, P. (eds). **LSP in Colombia: advances and challenges**. Bern: Peter Lang, 2014.

PEIXOTO, L. M. Identificação de unidades fraseológicas no vocabulário de Star Trek: abordagens corpus-driven e corpus-based. **Domínios de Linguagem**, Uberlândia, v. 8, n. 2, 2014. DOI : <https://doi.org/10.14393/DL16-v8n2a2014-8>. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/27630>. Acesso em: 15 fev. 2015.

POLIVÁNOV, E. D. **Travaux du Cercle Linguistique de Prague**, 1931.

QUAQLIO, P. **Television dialogue: the sitcom Friends vs. natural conversation**. Amsterdam / Philadelphia: John Benjamins, 2009

ROCHA, J. M. P. **Fraseologia jurídico-comercial e proposta de um glossário de colocações especializadas trilingue baseado em corpus**. 2017. 292 f. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2017.

SAUSSURE, F. **Curso de linguística geral**. São Paulo: Cultrix, 1969.

SINCLAIR, J. **Corpus, concordance, collocation**. Oxford: Oxford University Press, 1991.

SINCLAIR, J. **Looking up: an account of the COBUILD project in lexical computing**. London: Collins Cobuild, 1987.

TAGNIN, S. E. O. Collecting data for a bilingual dictionary of verbal collocations: from scraps of paper to corpora research. LEWANDOSKA-TOMASZCZYK, B.; MELIA, P. J. (eds). *Palc' 99: Practical Applications in Language Corpora*. **Papers from the International Conference at the University of Lodz**, Frankfurt and Main: Peter Lang GmbH, p. 399-407, 1999.

TAGNIN, S. E. O. **Expressões idiomáticas e convencionais**. São Paulo: Ática, 1989.

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

TARP, S. Reflexiones sobre el papel y el diseño de los diccionarios de traducción especializada. **MonTi: Monografías de Traducción e Interpretación**, Universidad de Valencia, Alicante, n. 6, p. 63-89, 2014. DOI: <https://doi.org/10.6035/MonTI.2014.6.2>. Disponível em: <http://www.e-revistas.uji.es/index.php/monti/article/view/1668>. Acesso em: 13 jun. 2020.

TARP, S. La teoría funcional en pocas palabras. **Estudios de Lexicografía**. Revista Mensual del Grupo de Las Dos Vidas de Las Palabras. n. 4, p. 31-42, 2015. Disponível em: [https://cc.au.dk/fileadmin/04\\_La\\_teori\\_a\\_funcional\\_en\\_pocas\\_palabras.pdf](https://cc.au.dk/fileadmin/04_La_teori_a_funcional_en_pocas_palabras.pdf). Acesso em: 13 jun. 2020.

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

TRISTÁ, A. M. **Fraseología y contexto**. Cuba: Editorial de Ciencias Sociales LaHabana, 1988.

TRISTÁ, A. M.; CARNEADO MOREÉ, Z. V. **Estudios de fraseología**. La Habana: Academia de Ciencias de Cuba, 1985.

UNSUB. In: **Lexico**. Disponível em: <https://www.dictionary.com/browse/unsub>. Acesso em: 10 ago. 2021.

UNSUB. In: **Merriam-Webster**. Disponível em: <https://www.merriam-webster.com/dictionary/unsub>. Acesso em: 10 ago. 2021.

VINCZE, O.; ALONSO RAMOS, M.: Incorporating frequency information in a collocation dictionary: establishing a methodology. **Procedia - Social and Behavioral Sciences**, v. 95, p. 241-248, 2013. DOI: <https://doi.org/10.1016/j.sbspro.2013.10.644>  
Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877042813041645>.  
Acesso em 13 jun. 2020.

VINOGRADOV, V. V. Ob osnovnuikh tipakh frazeologicheskikh edinit v russkom yazuike. SHAKHMATOV, A. A. **Sbornik statey i materialov**. Moscow: Nauka, 1947, p. 339-64.

WALKER, C. P. A Corpus-based study of the linguistic features and processes which influence the way collocations are formed: some implications for the learning of collocations. **Tesol Quarterly**, v. 45, n. 2, p. 291 – 312, 2011. DOI: <https://doi.org/10.5054/tq.2011.247710>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.5054/tq.2011.247710>. Acesso em: 15 fev. 2015.

WEBB, S.; ROGERS, H. The lexical coverage of movies. **Applied Linguistics**, v. 30, p. 407-427, 2009. DOI: <https://doi.org/10.1093/applin/amp010>. Disponível em: <https://academic.oup.com/applij/article-abstract/30/3/407/229264>. Acesso em 13 jun. 2020.

WEBB, S.: A corpus driven study of the potential for vocabulary learning through watching movies. **International Journal of Corpus Linguistics**, v. 15, p. 497- 519, 2010. DOI: <https://doi.org/10.1075/ijcl.15.4.03web>. Disponível em: <https://www.jbe-platform.com/content/journals/10.1075/ijcl.15.4.03web>. Acesso em: 20 fev. 2019.

WEBCORP. Disponível em: <http://www.webcorp.org.uk/live/>. Acesso em: 20 abr. 2016.

YAMAMOTO, M. I.; LISBOA, J. L. R.: Linguística de Corpus e fraseologia em séries televisivas: uma experiência que deu certo. **Itinerarius Reflectionis**, Goiânia, v. 15, n. 2, p. 01–19, 2019. DOI: 10.5216/rir.v15i2.58713. Disponível em: <https://revistas.ufg.br/rir/article/view/58713>. Acesso em 20 mar. 2020.

ZULATEGUI, J. P.L: En torno a la semántica de las colocaciones fraseológicas. **Estudios de Lingüística**, Universidad de Alicante, n, 16, p. 121-138, 2002. Disponível em: <https://rua.ua.es/dspace/handle/10045/6204>. Acesso em: 20 fev. 2019.

ZULUAGA, A. **Introducción al estudio de las expresiones fijas**. Frankfurt: Peter Lang, 1980.