

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Programa de pós-graduação em estudos linguísticos

DIEGO MENDES DE CARVALHO

ESTUDO BASEADO EM *CORPUS*: colocações em livros adaptados

Belo Horizonte

2024

Diego Mendes de Carvalho

ESTUDO BASEADO EM *CORPUS*: colocações em livros adaptados

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Linguística Aplicada.

Orientador: Profa. Dra. Bárbara Orfano

Belo Horizonte

2024

C331e Carvalho, Diego Mendes de.
Estudo baseado em *corpus* [manuscrito] : colocações em livros adaptados / Diego Mendes de Carvalho. – 2024.
1 recurso online (209 f.: il., tab., color.) : pdf.
Orientadora: Bárbara Orfano.
Área de concentração: Linguística Aplicada-Ensino e Aprendizado de Línguas.
Linha de pesquisa: Linguística de Corpus.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.
Bibliografia: f. 165 - 170.
Anexos: f. 171 – 209.
Exigências do sistema: Adobe Acrobat Reader.

1. Linguística de corpus – Teses. 2. Linguagem e línguas – Estudo e ensino – Teses. 3. Linguística aplicada – Teses. I. Orfano, Bárbara Malveira. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 410



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESTUDO BASEADO EM CORPUS: COLOCAÇÕES EM LIVROS ADAPTADOS

DIEGO MENDES DE CARVALHO

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA APLICADA, linha de pesquisa Ensino/Aprendizagem de Línguas Estrangeiras.

Aprovada em 16 de maio de 2024, pela banca constituída pelos membros:

Prof(a). Bárbara Malveira Orfanó - Orientadora
UFMG

Prof(a). Gustavo Leal Teixeira
UFMG

Prof(a). Miriam Piedade Mansur Andrade
UFMG

Belo Horizonte, 16 de maio de 2024.



Documento assinado eletronicamente por **Barbara Malveira Orfano, Professora do Magistério Superior**, em 23/05/2024, às 10:25, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Miriam Piedade Mansur Andrade, Professora do Magistério Superior**, em 23/05/2024, às 18:30, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gustavo Leal Teixeira, Professor do Magistério Superior**, em 04/07/2024, às 10:50, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site

[https://sei.ufmg.br/sei/controlador_externo.php?](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0)

[acao=documento_conferir&id_orgao_acesso_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3237276** e o código CRC **85821286**.

Dedico este trabalho aos meus amigos que sempre acreditaram em mim.

AGRADECIMENTOS

Agradeço sinceramente à Taylor Swift. Suas inspiradoras canções forneceram a trilha sonora motivadora durante o processo de escrita e desenvolvimento desta dissertação de mestrado. Suas composições serviram como uma fonte constante de estímulo criativo, impulsionando-me a alcançar novos patamares em minha pesquisa. Expresso minha profunda gratidão aos meus amigos, cujo contínuo apoio e crença em meu potencial foram fundamentais para enfrentar os desafios deste percurso acadêmico. Suas palavras de encorajamento e presença constante foram verdadeiramente revigorantes.

À minha orientadora, dedico uma gratidão especial. Sua gentileza, paciência e atenção dedicada à minha pesquisa foram elementos cruciais para o sucesso deste trabalho. Seu comprometimento e orientação esclarecedores foram fundamentais para o desenvolvimento não apenas da pesquisa, mas também do meu crescimento acadêmico e pessoal. Agradeço pela confiança depositada em mim, pelos *insights* valiosos e pela postura sempre acolhedora.

À minha amiga, Isabel, que me ajudou nos cálculos estatísticos, expresso a minha mais profunda gratidão.

A todos que, de alguma maneira, contribuíram para este percurso, expresso meu mais profundo reconhecimento.

[...] *We are all literary chameleons. It's a continuation of the idea that we are so many things all the time.* (Taylor Swift)

RESUMO

O presente estudo teve como objetivo principal analisar as colocações mais relevantes dos verbos *make* e *take* em *Graded Readers*. Para alcançar o referido objetivo, o estudo teve seu suporte teórico na Linguística de *Corpus*, se embasando em autores como Johns (1991), Halliday (1966), Sinclair (1991), Sardinha (2000), Biber (2015), Brezina (2015), entre outros. O estudo trabalhou com um *corpus* especializado, composto por obras literárias de linguagem adaptada conforme o Quadro Comum Europeu de Referência para Línguas (CEFR). O *corpus* foi dividido em *subcorpora*, classificados em A1, A2, B1, B2, C1 e C2, segundo a sua etiquetagem linguística, referente ao CEFR. Foi utilizado o software *LancsBox X* para o processamento e análise do *corpus*. Primeiramente, foi utilizada a ferramenta *GraphColl* para identificação das colocações mais relevantes, com base nos resultados do teste estatístico LogDice. Segundamente, foi utilizada a ferramenta *KWIC* para a análise das estruturas e dos significados dos colocados selecionados. Por fim, utilizando a abordagem *Data-driven Learning* (DDL), foi elaborada uma sequência didática com os resultados encontrados da análise. Os resultados da pesquisa revelaram que os *subcorpora* A1 e A2 apresentam colocados com sentenças diretas e simples, com poucas classes gramaticais entre o nó e o colocado, geralmente possuindo apenas um significado; os *subcorpora* B1 e B2 apresentam colocados com sentenças com diversidade em conjugação verbal, com mais classes gramaticais entre o nó e o colocado, geralmente apresentando mais de um significado; os *subcorpora* C1 e C2 apresentam colocados com sentenças mais complexas, como condicionais com *IF* e sentenças no *present perfect*. Em suma, este estudo contribuiu para a compreensão das colocações mais relevantes de *make* e *take* em um *corpus* especializado e para as pesquisas que envolvem a abordagem DDL, ao propor uma atividade que integra os dados ao ensino de línguas por meio dos resultados gerados na análise das colocações.

Palavras-chave: Linguística Aplicada, Linguística de Corpus, Corpus Pedagógico, Ensino Movido a Dados, Ensino de Línguas

ABSTRACT

The present study aimed to analyze the most relevant collocations of the verbs *make* and *take* in graded readers. To achieve this objective, the study relied on Corpus Linguistics, drawing on authors such as Johns (1991), Halliday (1966), Sinclair (1991), Sardinha (2000), Biber (2015), and Brezina (2015), among others. The study worked with a specialized corpus composed of literary works with language adapted according to the Common European Framework of Reference for Languages (CEFR). The corpus was divided into subcorpora, classified as A1, A2, B1, B2, C1, and C2, according to their linguistic labeling, referring to the CEFR. LancsBox X software was used for corpus processing and analysis. Firstly, the GraphColl tool was used to identify the most relevant collocations based on the results of the LogDice statistical test. Secondly, the KWIC tool was used for the analysis of the structures and meanings of the selected collocations. Finally, using the Data-driven Learning (DDL) approach, a didactic sequence was developed based on the analysis results. The research results revealed that subcorpora A1 and A2 present collocations with direct and simple sentences, with few grammatical classes between the node and the collocate, usually having only one meaning; subcorpora B1 and B2 present collocations with sentences showing diversity in verb conjugation, with more grammatical classes between the node and the collocate, generally presenting more than one meaning; subcorpora C1 and C2 present collocations with more complex sentences, such as conditionals with 'IF' and sentences in the present perfect tense. In summary, this study contributed to understanding the most relevant collocations of *make* and *take* in a specialized corpus and researches involving the DDL approach by proposing an activity that integrates data into language teaching through the results generated in the analysis of collocations.

Keywords: Applied Linguistics, Corpus Linguistics, Pedagogical Corpus, Data-driven Teaching, Language Teaching

LISTA DE FIGURAS

Figura 1 — Aplicações da Linguística de <i>Corpus</i>	23
Figura 2 — Tipos de <i>Corpus</i>	25
Figura 3 — Uso dos <i>corpora</i> de aprendizes	28
Figura 4 — Linha do Tempo - DDL	32
Figura 5 — Total de estudos empíricos de DDL ao longo dos anos.....	34
Figura 6 — Conceitos de Sinclair (1991)	42
Figura 7 — Gráfico da Extensão do <i>Corpus</i>	52
Figura 8 — Visão geral do <i>Graphcoll</i>	55
Figura 9 — Gráfico do <i>Graphcoll</i>	56
Figura 10 — <i>Graphcoll</i> - <i>make</i>	59
Figura 11 — Comparação dos resultados dos métodos estatísticos.....	61
Figura 12 — Colocados de <i>make</i> no <i>subcorpus</i> A2.....	67
Figura 13 — <i>Make</i> - A1	70
Figura 14 — Linhas de Concordância de <i>make noise</i> - A1.....	73
Figura 15 — <i>Make</i> - A2	75
Figura 16 — Linhas de Concordância de <i>make sure</i> - A2	77
Figura 17 — Linhas de Concordância de <i>make rich</i> - A2	82
Figura 18 — Linhas de Concordância de <i>make tea</i> - A2	84
Figura 19 — Linhas de Concordância de <i>make friends</i> - A2	86
Figura 20 — Linhas de Concordância de <i>make noise</i> - A2.....	88
Figura 21 — Linhas Concordância de <i>make money</i> - A2	91
Figura 22 — <i>Make</i> - B1	93
Figura 23 — Linhas de Concordância de <i>make sure</i> - B1	95
Figura 24 — <i>Make</i> - B2	97
Figura 25 — Linhas de Concordância de <i>make sure</i> - B2	99
Figura 26 — Linhas de Concordância de <i>make sense</i> - B2.....	101
Figura 27 — Linhas de Concordância de <i>make feel</i> - B2	103
Figura 28 — Linhas de Concordância de <i>make feel</i> - A2	104
Figura 29 — Linhas de Concordância de <i>make feel</i> - B2	104
Figura 30 — Linhas de Concordância de <i>make money</i> - B2	106
Figura 31 — Linhas de Concordância de <i>make happy</i> - B2	108
Figura 32 — <i>Make</i> - C1	111
Figura 33 — Linhas de Concordância de <i>make sure</i> - C1	112
Figura 34 — <i>make</i> - C2.....	114
Figura 35 — Linhas de Concordância de <i>make sense</i> - C2	115
Figura 36 — Linha de Concordância de <i>make difference</i> - C2.....	117
Figura 37 — <i>Take</i> - A1	119
Figura 38 — Linhas de Concordância de <i>take off</i> - A1	121
Figura 39 — <i>Take</i> - A2.....	123

Figura 40 — Linhas de Concordância de <i>take care</i> - A2	125
Figura 41 — Linhas de Concordância de <i>take off</i> - A2	127
Figura 42 — Linhas de Concordância com POS de <i>take off</i> - A2.....	127
Figura 43 — Linhas de Concordância com POS de <i>take off</i> - A2.....	127
Figura 44 — Linhas de Concordância de <i>take away</i> - A2.....	130
Figura 45 — <i>Take</i> - B1	132
Figura 46 — Linhas de Concordância com POS de <i>take care</i> - B1.....	134
Figura 47 — Linhas de Concordância de <i>take off</i> - B1	136
Figura 48 — <i>Take</i> - B2.....	137
Figura 49 — Linhas de Concordância de <i>take care</i> - B2	139
Figura 50 — Linhas de Concordância de <i>take off</i> - B2	141
Figura 51 — Linhas de Concordância de <i>take away</i> - B2.....	143
Figura 52 — <i>Take</i> - C1.....	145
Figura 53 — Linhas de Concordância de <i>Take care</i> - C1	147
Figura 54 — <i>Take</i> - C2.....	148
Figura 55 — Linha de Concordância de <i>take easy</i> - C2	150
Figura 56 — Linhas de Concordância de <i>take look</i> - C2	152
Figura 57 — Resultados de <i>Make</i> obtidos com o <i>LancsBox</i>	156
Figura 58 — Resultados de <i>take</i> obtidos com o <i>LancsBox</i>	157

LISTA DE QUADROS

Quadro 1 — Situação exemplo com o verbo <i>make</i>	58
Quadro 2 — Relação de ocorrências no corpus	63
Quadro 3 — Níveis de exclusividade	64
Quadro 4 — Quadro de colocados mais exclusivos	71
Quadro 5 — Estruturas de <i>Make Noise</i> - A1	72
Quadro 6 — Colocados mais exclusivos de <i>Make</i> - A2	76
Quadro 7 — Estruturas de <i>make sure</i> - A2	76
Quadro 8 — Estruturas de <i>make feel</i> - A2	79
Quadro 9 — Estruturas de <i>make rich</i> - A2	81
Quadro 10 — Estruturas de <i>make tea</i> - A2	83
Quadro 11 — Estruturas <i>make friends</i> - A2	85
Quadro 12 — Estruturas <i>make noise</i> - A2	87
Quadro 13 — Estruturas de <i>make money</i> - A2	90
Quadro 14 — Frequência da colocação	94
Quadro 15 — Estruturas <i>make sure</i> - B1	94
Quadro 16 — Frequência da colocação	98
Quadro 17 — Estruturas <i>make sure</i> - B2	98
Quadro 18 — <i>make sense</i>	100
Quadro 19 — Estruturas com <i>make feel</i> - B2	102
Quadro 20 — Estruturas de <i>make money</i> - B2	105
Quadro 21 — Estruturas de <i>make happy</i> - B2	107
Quadro 22 — Frequência da colocação	111
Quadro 23 — Estruturas de <i>make sure</i> - C1	112
Quadro 24 — Frequência da colocação	114
Quadro 25 — Estruturas <i>make sense</i> - C2	115
Quadro 26 — Estruturas <i>make difference</i> - C2	117
Quadro 27 — Frequência da colocação	120
Quadro 28 — Estruturas <i>take off</i> - A1	120
Quadro 29 — Frequência da colocação	124
Quadro 30 — Estruturas <i>take care</i> - A2	124
Quadro 31 — Estruturas <i>take off</i> - A2	125
Quadro 32 — Estruturas <i>take away</i> - A2	128
Quadro 33 — Frequência da colocação	133
Quadro 34 — Estruturas <i>take care</i> - B1	133
Quadro 35 — Estruturas <i>take off</i> - B1	134
Quadro 36 — Frequência da colocação	138
Quadro 37 — Estruturas de <i>take care</i> - B2	138
Quadro 38 — Estruturas de <i>take off</i> - B2	140

Quadro 39 — Estruturas de <i>take away</i> - B2	142
Quadro 40 — Frequência da colocação	146
Quadro 41 — Estruturas de <i>take care</i> - C1.....	146
Quadro 42 — Frequência da colocação	149
Quadro 43 — Estruturas de <i>take care</i> - C2.....	149
Quadro 44 — Estruturas de <i>take easy</i> - C2	150
Quadro 45 — Estruturas de <i>take look</i> - C2	151

LISTA DE ABREVIATURAS E SIGLAS

CC	Conjunção Coordenativa
CEFR	Common European Framework of Reference for Languages
CPN	Collocation Parameters Notation
DDL	Data-driven learning
DT	Determinante
EAM	Experiência de Aprendizagem Mediada
EAP	Inglês Para Fins Acadêmicos
EFL	Inglês como Língua Estrangeira
FLA	Aquisição da Primeira Língua
IN	Preposição
JJ	Adjetivo
JNR	Adjetivo Comparativo
LC	Linguística de Corpus
LI	Língua Inglesa
NN	Substantivo, singular
NNP	Nomes próprios
NNS	Substantivo Plural
PRP	Pronome Pessoal
PRP\$	Pronome Possessivo
RB	Advérbio
SLA	Aquisição de Segunda Língua
UB	Baseada no Uso
VB	Verbo base
VBP	Verbo na primeira pessoa do singular ou plural
VBZ	Verbo na Terceira Pessoa do singular
ZDP	Zona de Desenvolvimento Proximal

SUMÁRIO

CONSIDERAÇÕES INICIAIS	17
2 FUNDAMENTAÇÃO TEÓRICA	22
2.1 Linguística de <i>corpus</i>	22
2.1.1 Tipos de corpus e extensão	24
2.2 <i>Corpus</i> e Ensino	26
2.2.2 Abordagens	29
2.2.3 Data-driven Learning (DDL)	31
2.2.4 Teorias de ensino e DDL	36
2.3 Colocações	39
2.3.1 Abordagem baseada na frequência	40
2.3.2 Estudos envolvendo as colocações	43
2.3.3 Graded readers	45
3 ASPECTOS METODOLÓGICOS DA PESQUISA	50
3.1 <i>Corpus</i> de estudo	50
3.2 <i>Software</i>	53
3.2.1 Graphcoll	54
3.3 Medidas estatísticas	57
3.3.1 Log Dice	57
4 CRITÉRIOS DE IDENTIFICAÇÃO DE COLOCAÇÕES RELEVANTES	62
5 DESCRIÇÃO E DISCUSSÃO DA ANÁLISE DOS DADOS	69
5.1 Colocações com o verbo <i>make</i> no <i>subcorpus</i> a1	69
5.1.1 Noise	71
5.2 Colocações com o verbo <i>make</i> no <i>subcorpus</i> a2	74
5.2.1 Sure 76	
5.2.2 Feel 79	
5.2.3 Rich 80	
5.2.4 Tea 83	
5.2.5 Friends	84
5.2.6 Noise	87
5.2.7 Money	90
5.3 Colocações com o verbo <i>make</i> no <i>subcorpus</i> b1	92
5.3.1 Sure 94	
5.4 Colocações com o verbo <i>make</i> no <i>subcorpus</i> b2	96
5.4.1 Sure 98	
5.4.2 Sense	100
5.4.3 Feel 102	
5.4.4 Money	105
5.4.5 Happy	107
5.5 Colocações com o verbo <i>make</i> no <i>subcorpus</i> c1	110

5.5.1 Sure	112
5.6 Colocações com o verbo <i>make</i> no <i>subcorpus c2</i>	113
5.6.1 Sense	115
5.6.2 Difference	116
5.7 Colocações com o verbo <i>take</i> no <i>subcorpus a1</i>	118
5.7.1 Off	120
5.8 Colocações com o verbo <i>take</i> no <i>subcorpus a2</i>	122
5.8.1 Care	124
5.8.2 Off	125
5.8.3 Away	128
5.9 Colocações com o verbo <i>take</i> no <i>subcorpus b1</i>	131
5.9.1 Care	133
5.9.2 Off	134
5.10 Colocações com o verbo <i>take</i> no <i>subcorpus b2</i>	137
5.10.1 Care	138
5.10.2 Off	139
5.10.3 Away	141
5.11 Colocações com o verbo <i>take</i> no <i>subcorpus c1</i>	144
5.11.1 Care	146
5.12 Colocações com o verbo <i>take</i> no <i>subcorpus c2</i>	147
5.12.1 Care	149
5.12.2 Easy	149
5.12.3 Look	151
CONSIDERAÇÕES FINAIS	154
REFERÊNCIAS	166
ANEXO A — Sequência Didática	172

CONSIDERAÇÕES INICIAIS

Primeiramente, é pertinente ressaltar que o estímulo ao emprego da literatura no contexto educacional deve ser um processo contínuo, revelando-se multifacetado em sua aplicação. Nesse sentido, Carvalho (2015, p. 6) advoga em favor da utilização de textos literários, afirmando que "[quanto] mais o aluno ler bons livros, mais ele aprende sobre os mecanismos de funcionamento da língua, tanto escrita quanto falada." Além disso, a literatura não apenas amplia a compreensão do mundo, mas também fomenta a empatia, contrapõe a alienação, proporciona o encontro com diversas identidades, consolida a percepção da realidade, amplifica a capacidade imaginativa, dentre outros benefícios que serão delineados a seguir.

Considerando a literatura como uma forma de se expressar artisticamente, a qual propicia o desenvolvimento do ser humano em diversas áreas do conhecimento, Vieira (2008, p. 442) corrobora a ideia que a literatura, assim como outras formas de expressão artística, "possibilita ao homem a expansão do seu potencial criador e imaginativo, satisfazendo sua necessidade de ficção". Em concordância com essa perspectiva, especialmente no que concerne aos romances, o filósofo Rorty (2001, p. 249) afirma que:

[...] os romances são os principais meios que nos ajudam a imaginar o que é ser um católico de berço perdendo a fé, um fundamentalista caipira levando Jesus ao coração, uma vítima de Pinochet lidando com o desaparecimento de seus filhos, um piloto kamikaze da Segunda Guerra Mundial convivendo com o fato da derrota do Japão, um piloto de bombardeiro que lançou bombas incendiárias em Tóquio lidando com o preço moral da vitória americana, ou um político idealista lidando com as pressões que as corporações multinacionais exercem sobre o processo político.

Efetivamente, a leitura de obras de ficção proporciona um encontro com realidades diversas, conforme expresso por Vieira (2008), permitindo-nos satisfazer a curiosidade, penetrar no universo mental do outro e vivenciar suas tomadas de decisões, angústias, felicidades, paixões, tragédias e reflexões. Todos esses elementos podem se revelar extremamente úteis, uma vez que, enquanto seres humanos, podemos aprimorar nossas habilidades interpessoais e autoconhecimento através da experiência, seja ela vivida na realidade ou vivenciada na ficção. Dessa forma, torna-se desafiador contestar a importância e o impacto positivo que a literatura

exerce sobre os seres humanos, reforçando a necessidade de sua inclusão no currículo escolar, sem qualquer ressalva.

Outro fator relevante sobre o uso da literatura reside no fato de que esta também é uma ferramenta que nos ajuda a refinar o nosso senso crítico ao olharmos para a ficção e nos depararmos com a realidade. Nesse viés, para Todorov (2009), a literatura tem um papel particular, o de libertar o espírito das falsas evidências:

[...] diferentemente dos discursos religiosos, morais ou políticos, ela [a literatura] não formula um sistema de preceitos; por essa razão, escapa às censuras que se exercem sobre as teses formuladas de forma literal. As verdades desagradáveis tanto para o gênero humano ao qual pertencemos quanto para nós mesmos - têm mais chances de ganhar voz e ser ouvidas numa obra literária do que numa obra filosófica ou científica. (Todorov, 2009, p. 80).

Diante desse contexto, torna-se patente a importância da literatura como meio de emancipação intelectual e desenvolvimento de uma perspectiva crítica em relação ao mundo que nos cerca. Sob essa ótica, a promoção da leitura entre os jovens emerge como um potencial catalisador de impactos positivos em nossa sociedade, configurando-se como mais uma razão imperativa para a integração da leitura de obras no contexto educacional. Nesse viés, a discussão acerca do ensino de literatura para adolescentes assume particular relevância, haja vista que, como delineado nessa seção, a prática da leitura fomenta o pensamento crítico, elemento de suma importância na formação de uma consciência política e social. Dessa forma, torna-se incontestável a pertinência da abordagem da literatura no ambiente escolar, dada a sua multiplicidade de contribuições para o desenvolvimento do aluno, em esferas tanto sociais quanto pessoais.

1 OS BENEFÍCIOS DA LITERATURA NO ENSINO DE LÍNGUA

Bozza e Benenito (2019) argumentam que a aula de Língua Inglesa (LI) deve possibilitar ao aluno reconhecimento e a compreensão da diversidade linguística e cultural, desenvolvendo criticamente significados em relação ao mundo em que vive, e afirma ser o texto literário uma ferramenta excelente para que isso possa ocorrer, uma vez que, por meio da literatura, “é possível promover reflexões sobre questões de poder e as ideologias dominantes presentes na sociedade, auxiliando para a formação de um indivíduo crítico capaz de construir significados, respeitar e

reconhecer diferentes culturas” (Bozza; Benenito, 2019, p. 4). Nessa perspectiva, é importante que a literatura seja ensinada com qualidade e com propósito, não só para as crianças que estão ingressando na escola, mas também para os adolescentes que estão se formando no ensino médio.

Reforçando tal ideia, Collie e Slater (1987) sustentam que o uso da literatura nas aulas de LI oferece ao aluno uma visão completa do contexto vivido, no qual personagens de muitas origens sociais podem ser retratados. Com isso, os estudantes podem:

[...] descobrir seus pensamentos, sentimentos, costumes, posses; o que eles compram, acreditam, temem, desfrutam; como eles falam e se comportam a portas fechadas. Este vívido mundo imaginado pode rapidamente dar ao leitor estrangeiro uma ideia dos códigos e preocupações que estruturam uma sociedade real. Ler a literatura de um período histórico é, afinal, uma das formas de que dispomos para nos ajudar a imaginar como era a vida naquele outro território estrangeiro: o passado do nosso próprio país. (Collie; Slater, 1987, p. 6)

Os autores supracitados acreditam que a literatura é vista como um complemento dos materiais pedagógicos, usada para aumentar a visão de mundo do aluno sobre o país cuja língua está sendo aprendida. Em consonância a isso, Corchs (2006) afirma que os textos literários não só dizem sobre a cultura de outros povos, mas também oferecem aos estudantes um cenário e uma linguagem que os incentivam a explorar mais sobre uma cultura ou um grupo específico. Nessa perspectiva, o trabalho com o texto literário possibilita que o aluno desenvolva sua criticidade enquanto sujeito no mundo e vivencie experiências que extrapolam as de sua realidade. Portanto, fica evidente que os benefícios do uso da literatura nas aulas de LI são enriquecedores, sendo assim, o trabalho com os textos literários não deve ser descartado.

São essas contribuições que a literatura nos proporciona que, inicialmente, impulsionaram a realização desta pesquisa. Partindo de uma inquietação pessoal, relacionada à minha percepção como professor, ao constatar a ausência de textos literários nas salas de aula, encontrei na linguística de *corpus* uma abordagem para introduzir a literatura de maneira contextual, inovadora e contemporânea com os alunos.

Portanto, esta pesquisa concilia literatura e ensino de Língua Inglesa, empregando a abordagem *Data-Driven Learning* (DDL) na elaboração de material

didático. Nesse contexto, o estudo busca demonstrar como um *corpus* especializado, constituído por *Graded Readers*, pode ser utilizado para a identificação de colocações relevantes e para o ensino dessas de maneira contextualizada. O objetivo geral deste trabalho consiste na análise das colocações presentes no *corpus*, delineando, assim, os objetivos específicos mencionados a seguir:

a) identificar e selecionar as colocações mais relevantes e exclusivas presentes nos *subcorpora* dos verbos *make* e *take*;

b) identificar as construções gramaticais das colocações mais exclusivas e relevantes de *make* e *take* presentes nos *subcorpora* (verbo + prep, verbo + substantivo etc.), identificando e analisando as semelhanças e diferenças;

c) Elaborar uma atividade didática usando a perspectiva DDL, a partir dos dados encontrados da análise das colocações.

A pesquisa em questão é relevante devido à necessidade de ampliar as práticas literárias nas aulas de LI, visando alcançar os benefícios previamente discutidos nesta introdução. Este estudo também se justifica por proporcionar uma compreensão mais aprofundada das colocações exclusivas de *make* e *take* de cada nível do *corpus* analisado. Adicionalmente, a pesquisa explora o potencial do uso de um *corpus* de textos literários associado à abordagem DDL para o ensino de Língua Inglesa na criação de um plano de aula, utilizando os dados encontrados na análise. Tais iniciativas enriquecem as práticas dos professores de inglês, orientando-os sobre os pontos a serem enfatizados no ensino.

Dessa maneira, após a exposição dos benefícios da integração da literatura e do ensino de Língua Inglesa nesta introdução, dos objetivos e da justificção apresentada, torna-se evidente que a pesquisa detém significativa relevância não apenas para o campo educacional, mas também para a área da linguística de *corpus* aplicada ao ensino de línguas.

1.1 Organização dos capítulos

Este trabalho se apresenta em 5 capítulos.

No segundo capítulo, são apresentados os fundamentos teóricos utilizados nesta pesquisa. Inicialmente, realiza-se uma contextualização sobre linguística de *corpus*, abordando o conceito de *corpus* e delineando os diferentes tipos existentes.

Dentre estes, destaca-se o *corpus* especializado, para, posteriormente, discutir a interação do *corpus* com o ensino de línguas e, em seguida, introduzir a abordagem *Data-driven Learning*. No mesmo capítulo, são discutidos o conceito de colocação, fenômeno linguístico objeto de estudo nesta dissertação, e o conceito de *Graded Readers*, o material coletado para a constituição do *corpus* da pesquisa.

No terceiro capítulo, são abordados os aspectos metodológicos do estudo. Descreve-se o *corpus* utilizado, fornecendo detalhes sobre a sua coleta. Além disso, são apresentados pormenores das ferramentas empregadas para a compilação e análise do *corpus*, o método estatístico adotado e os critérios para a identificação das colocações relevantes.

No quarto capítulo, realiza-se a descrição da análise dos dados coletados, detalhando as colocações relevantes identificadas em todos os *subcorpora*.

No quinto capítulo, a pesquisa é concluída, dissertando acerca das considerações finais. Descreve-se a trajetória realizada na pesquisa, suas limitações e sugestões para trabalhos futuros.

O plano de aula desenvolvido a partir dos dados encontrados na análise está no Anexo desta pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Linguística de *corpus*

A Linguística de *Corpus* (LC), na visão de Sardinha (2004, p. 3), “ocupa-se da coleta e da exploração de *corpora*, ou conjunto de dados linguísticos textuais coletados criteriosamente, com propósito de servirem para a pesquisa de uma língua ou variedade”. Sendo assim, esse campo de pesquisa é empírico e analítico, possibilitando a identificação de diversos fenômenos e padrões linguísticos em uma coleção extensa de textos, isto é, um *corpus*.

Segundo Brezina (2018, p. 5) *corpus* é “[...]uma forma específica de dados linguísticos. É uma coleção de textos escritos ou transcrições de linguagem falada que podem ser pesquisados por um computador por meio de *software* especializado”. Ainda segundo o autor, o *corpus* representa um subconjunto da produção linguística de interesse, chamado de amostra¹. O *corpus* pode ser muito especializado, contendo somente o que for de interesse do pesquisador, ou incluir toda uma população². Existem *softwares* que conseguem armazenar e compilar *corpus* muito extensos, além de proporcionarem ferramentas que auxiliam na exploração dos *corpora* para fins específicos, como, por exemplo, a identificação de colocações.

Os primeiros passos da LC foram dados nas décadas de 1950 e 1960, período em que as tecnologias emergentes possibilitaram o manuseio de grandes quantidades de texto por meio de computadores. No entanto, os avanços na área surgiram a partir de 1964, com o lançamento do *corpus Brown*. A partir disso, se consolidava na área dos estudos linguísticos uma nova forma de investigação da linguagem.

Enquanto área, a LC evoluiu ao longo das décadas com a criação de *corpora* mais abrangentes e representativos como, por exemplo, o *Lancaster-Oslo-Bergan (LOB)* em 1978 com 1 Milhão de palavras, e o *Corpus of Spoken American English* em 1991, com 2 milhões de palavras. Além disso, houve avanços na comprovação dos dados, na utilização e aplicação de métodos estatísticos como o *Log Dice*, *MI-score*, *t-score*, *Delta P1*, *Delta P2*; e foram criados *softwares* mais avançados para

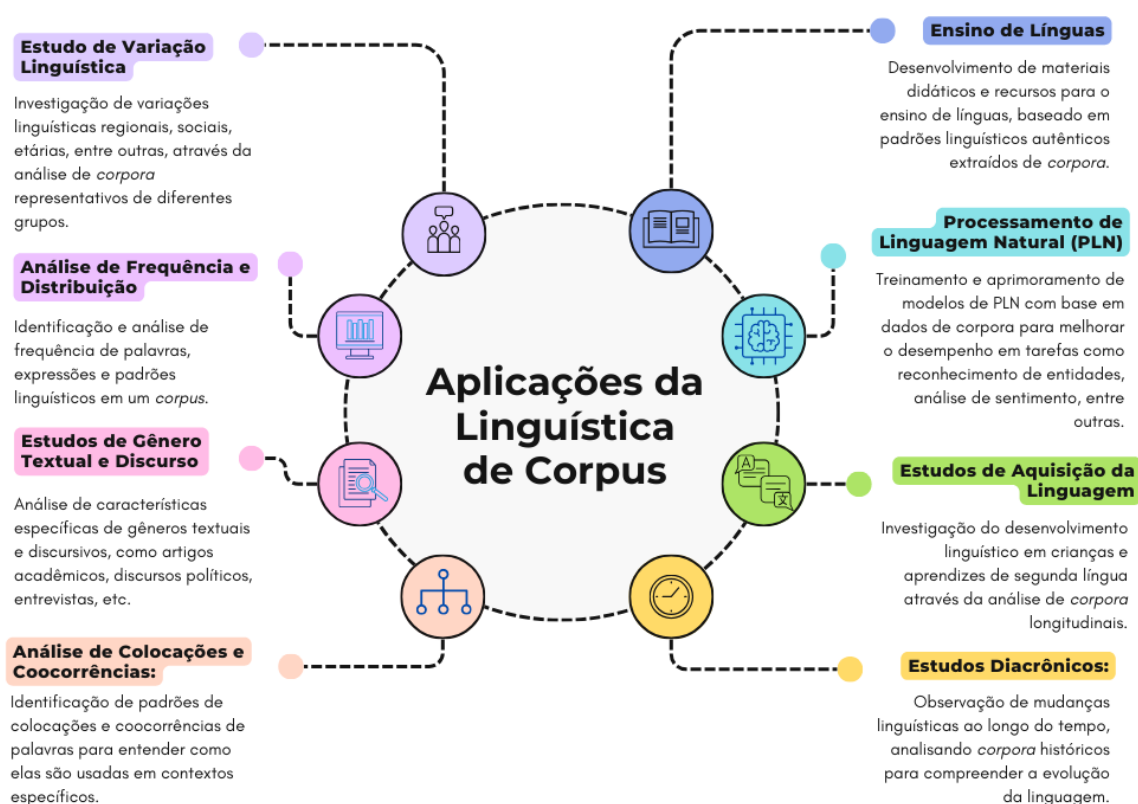
¹Segundo Brezina (2018), amostra é a quantidade de textos selecionados para análise.

²Segundo Brezina (2018), população é toda a linguagem de interesse para o pesquisador.

compilação dos *corpora*, o que tornaram as pesquisas na área mais amplas, consistentes e relevantes para os estudos linguísticos.

No âmbito das aplicações, a LC continuou a evoluir, adaptando-se às mudanças tecnológicas e às demandas crescentes de pesquisas, envolvendo os estudos linguísticos, explorando diversas áreas. Na Figura 1, observa-se alguns dos usos do *corpus* na área da Linguística:

Figura 1 — Aplicações da Linguística de *Corpus*



Fonte: O autor (2024).

Conforme observado na Figura 1, diversas são as possibilidades de exploração de *corpora*, destacando-se duas de relevância significativa para esta pesquisa: a análise de Colocações e Co-ocorrência, bem como a sua aplicação no contexto do Ensino de Línguas. Nessa perspectiva, Almeida (2014, p. 35) afirma que:

[...]evidências provenientes dos *corpora*, como os padrões linguísticos, combinações, colocações e a frequência de ocorrência de determinados itens linguísticos revelam o que pode ser priorizado no ensino e alguns caminhos de como apresentar esses dados aos aprendizes.

Nesse sentido, é relevante haver pesquisas voltadas para a identificação dessas ocorrências que podem ser priorizadas no ensino e levadas para a sala de aula, em forma de dados que possam ser interpretados pelos discentes. Sendo assim, o presente estudo integra os resultados das análises das **colocações** em um **corpus especializado** ao ensino de línguas na abordagem de **ensino movida a dados**. Esses conceitos serão, respectivamente, apresentados com mais detalhes nas seções 2.1.1, 2.2 e 2.2.2.

2.1.1 Tipos de corpus e extensão

Existem vários tipos de *corpora* construídos para distintas finalidades, a partir de diversos materiais. Nesse sentido, de acordo com Martins e Carrilho (2016), o objetivo do *corpus* determinará o tipo de material selecionado, que varia consoante a características como modalidade de produção, gêneros textuais, perfil dos autores/falantes, entre outros.

Em relação à extensão do *corpus*, supõe-se que ele tenha que ser representativo de uma língua ou de uma variedade e, para tal, deve ser vasto, uma vez que há ocorrências mais raras que outras; sendo assim, com o *corpus* extenso, a probabilidade dessas ocorrências aparecerem seria maior. Porém, é difícil estabelecer qual seria o tamanho ideal de um *corpus*, uma vez que, como Sardinha (2004, p. 23) afirma, "o *corpus* é uma amostra de uma população cuja dimensão não se conhece", dessa forma, ao não saber o tamanho da população, não há como definir uma extensão exata para ele.

Em contrapartida, a incerteza relativa ao tamanho ideal do *corpus* pode ser mitigada por meio da consideração da finalidade específica da pesquisa que se pretende realizar. Nessa perspectiva, Brezina (2018, p. 18) sustenta que "o tamanho do *corpus* depende da questão de pesquisa e do tipo de características linguísticas que queremos investigar". Destarte, sua extensão pode variar conforme a abordagem que o pesquisador adota, por isso, cabe ao responsável delimitar o seu fenômeno de estudo, seu objetivo e, a partir de uma abordagem, justificar a extensão do seu *corpus*.

Para ilustrar alguns dos tipos de *corpus* existentes e seus materiais de composição, compusemos a Figura 2:

Figura 2 — Tipos de *Corpus*



Fonte: O autor (2024).

Como visto na Figura 2, os *corpora* são moldados a partir de propósitos. Em consonância a isso, Kennedy (2014) afirma que os fins influenciam no *design*, tamanho e natureza do *corpus*. Desse modo:

[...] *corpora* atuais destinados à pesquisa linguística foram concebidos para fins descritivos gerais - isto é, foram concebidos de modo que possam ser examinados ou arrastados para responder a questões em vários níveis linguísticos sobre a prosódia, o léxico, a gramática, os padrões de discurso ou a pragmática do idioma. Outros *corpora* foram concebidos para fins especializados, tais como descobrir quais palavras e significados das palavras devem ser incluídos num dicionário para alunos; quais palavras ou significados são mais frequentemente utilizados pelos trabalhadores da indústria petrolífera ou da economia; ou que diferenças existem entre os usos de uma língua em diferentes contextos geográficos, sociais, históricos ou relacionados com o trabalho (Kennedy, 2014, p. 3-4).

Dito isso, diante dos objetivos estabelecidos para esta pesquisa e das características dos *corpora* mencionados nessa seção, construímos um *corpus*

especializado, pois o contexto de análise é específico, visto que o *corpus* será utilizado pedagogicamente, ou seja, para fins de ensino.

2.2 *Corpus* e Ensino

Na área da LC, há contribuições importantes que demonstram como o *corpus* pode ser usado por professores. O'keeffe, McCarthy e Carter (2007), por exemplo, tornam acessível as pesquisas com a utilização de *corpus* e explicam como os *corpora* podem ser criados e utilizados, abordando suas contribuições para o ensino de línguas, já Reppen (2010) explica e ilustra como os professores podem usar *corpora* para criar materiais e atividades pedagógicas para atender às necessidades específicas da sala de aula. Sob essa óptica, Oliva (2018, p. 27) alega que:

[o] professor pode criar exercícios baseados em *corpus* para serem apresentados antes ou depois de uma atividade de leitura ou escuta, tendo a nova entrada em vários contextos diferentes com foco em vocabulário, gramática e até mesmo tocar tópicos culturais.

Além dessas contribuições para o ensino citadas anteriormente, há também pesquisas relevantes que mostram o uso do *corpus* para análise de fenômenos linguísticos. Como exemplo disso, Römer (2004) realizou uma análise comparativa de auxiliares modais, sua distribuição, significados e contextos, em dados de *corpus* de inglês britânico falado e em textos selecionados de livros didáticos. Como resultado, o autor argumenta que as evidências do *corpus* podem contribuir para a melhoria dos materiais didáticos e para a observação dos fenômenos frequentes e dos padrões típicos da linguagem utilizada; Frankenberg-garcia (2018) examinou as colocações intuitivamente disponíveis para um grupo de 90 estudantes universitários, tutores de Inglês para Fins Acadêmicos (EAP) e estudantes de doutorado, mestrado e graduação em um teste de preenchimento de lacunas de redação acadêmica onde mais de uma colocação poderia ser usada em cada espaço em branco. Os resultados sugerem que ter o inglês como língua materna não influencia no uso correto das colocações, ao passo de que ter experiência na escrita acadêmica, sim.

Dessa maneira, utilizar ferramentas da LC no ensino ou na sala de aula de línguas pode ser positivo para os estudantes, por oferecer oportunidade aos alunos de se familiarizarem com o uso real da língua, ao invés de contar com modelos

inventados e descontextualizados. Além disso, o *corpus* fornece aos aprendizes acesso a informações e visões sobre o idioma, que podem ser utilizados para auxiliar os estudantes a compreender e a utilizar a língua de maneira mais eficiente. Como prova disso, Crosthwaite e Steeples (2022) conduziram um estudo com foco no desenvolvimento de conhecimento receptivo e produtivo de construções de voz passiva usadas ao escrever relatórios de pesquisa científica para uma aula de física. Os resultados sugerem que a consulta do *corpus* dos alunos foi eficaz para aprimorar o uso da voz passiva na escrita científica dos estudantes. No entanto, as autoras observaram que preferências distintas (com críticas) em relação a certas ferramentas, funções e ao uso do *corpus*, e a adesão contínua após a formação foram relativamente baixas. Em geral, as autoras defendem que as implicações dessas descobertas destacam possibilidades promissoras no que diz respeito ao Aprendizado Baseado em Dados com alunos mais jovens.

De fato, conforme o conjunto de exemplos apresentados, o uso do *corpus* pode ser um grande aliado para os professores e para os aprendizes, podendo ser aplicado de diversas maneiras no ensino por facilitar o aprendizado. Sardinha (2004) resume a exploração desses *corpora* no ensino em quatro áreas de concentração: i) descrição da língua nativa; ii) descrição da linguagem do aprendiz; iii) transposição de metodologias de pesquisa para a sala de aula e iv) desenvolvimento de materiais de ensino, currículos e abordagens.

A primeira área refere-se à descrição da linguagem dos falantes nativos, enquanto a segunda trata-se da descrição da linguagem de falantes não nativos. Nesse campo, há trabalhos relevantes sobre o uso dos *corpora* de aprendizes como de Nesselhauf (2004) e de Gilquin e Granger (2022). No Brasil, temos pesquisas realizadas utilizando os *corpora* de aprendizes de Língua Inglesa como visto em: Almeida, 2014; Silero, 2014; Wright, 2014; De Oliveira, 2015; Da Costa; De Azevedo, 2016; Dutra; Da Silva Queiroz; Alves, 2017; Orfanó; Oliveira; Barbosa, 2019; De Lima Silva, 2021. A descrição de cada trabalho pode ser vista na Figura 3, a seguir:

Figura 3 — Uso dos *corpora* de aprendizes



Fonte: O autor (2024).

A terceira área, transposição de metodologias de pesquisa para a sala de aula, como apontado por Sardinha (2004), diz respeito ao uso do instrumental analítico da Linguística de *Corpus* na sala de aula, como concordância e lista de palavras. Em relação à quarta área, desenvolvimento de materiais de ensino, currículos e abordagens, é referente à criação de metodologias de ensino inspiradas na exploração de *corpora* ou em conceitos da LC. Ele ainda complementa que existem três propostas de utilização no ensino: Currículo Lexical, a Abordagem Lexical e o Ensino Movido a Dados, o qual será descrita com mais detalhes na seção 2.2.2.

2.2.2 Abordagens

Acerca da abordagem, primeiro vale ressaltar que a LC, para esta pesquisa, é vista como sendo mais do que uma metodologia, é uma perspectiva para se chegar à linguagem. Em consonância a isso, Sardinha (2004, p. 37) argumenta que:

[a] Linguística de *Corpus* é uma perspectiva, isto é, uma maneira de se chegar à linguagem, e faz alusão ao conceito de teoria linguística como janela que molda como enxergamos a linguagem (189). Dessa forma, segundo Hoey (109), a Linguística de *Corpus* não seria apenas um instrumental, mas sim uma abordagem. De modo similar, Leech (145, p. 106) a define como "não somente uma nova metodologia emergente para o estudo da linguagem, mas uma nova empreitada de pesquisa e, na verdade, uma nova abordagem filosófica".

Atrelado a esse pensamento, Biber (2015) alega que a LC não é um modelo de linguagem, mas sim, em certo nível, uma abordagem metodológica. Porém, ele complementa dizendo que:

a linguística do *corpus* é muito mais do que uma abordagem metodológica: essas inovações metodológicas permitiram que os pesquisadores fizessem tipos fundamentalmente diferentes de perguntas de pesquisa, às vezes resultando em perspectivas radicalmente diferentes sobre a variação e o uso da linguagem daquelas tomadas em pesquisas anteriores. A pesquisa linguística do *Corpus* oferece um forte apoio para a visão de que a variação da linguagem é sistemática e pode ser descrita usando métodos empíricos e quantitativos (Biber, 2015, p. 2).

Portanto, a LC não se limita em ser um instrumento para se alcançar os objetivos de uma pesquisa, ou mesmo uma teoria para embasar estudos linguísticos, ela é a abordagem para se chegar à compreensão da linguagem. Em relação a isso, nos estudos de *corpus* vêm sendo usadas duas abordagens principais: *Corpus-based* e *Corpus-driven*.

Conforme Tognini-bonelli (2000, p. 81), a abordagem *Corpus-based* "refere-se a um tipo de metodologia em que o compromisso com os dados como um todo não é, em última análise, muito rigoroso ou sistemático". Dessa forma, ainda segundo a autora, os *corpora*, nessa abordagem, são usados para validar categorias existentes ou adicionar uma dimensão probabilística à teoria. Sendo assim, as informações provenientes de padrões sintáticos são mais relevantes, em vez de padrões lexicais.

Biber (2015) complementa dizendo que abordagem *corpus-based* se constitui na validação de formas e estruturas linguísticas derivadas da teoria linguística, tendo como objetivo principal a pesquisa e a análise de padrões sistemáticos de variação e uso para características linguísticas predefinidas. Essa abordagem já identificou que as descrições de variação e uso geralmente são inválidas para a linguagem na totalidade. Isso significa que alguns registros podem apresentar padrões diferentes. Em outras palavras, nessa abordagem, é preciso considerar as diferenças de registro dos *corpora*.

Por outro lado, a abordagem *Corpus-Driven* se difere da abordagem citada anteriormente por permitir que os dados guiem a pesquisa, ou seja, as construções linguísticas não são estabelecidas pela teoria; pelo contrário, elas irão surgir a partir da análise do *corpus*. Nesse panorama, segundo Tognini-bonelli (2000, p. 85):

A linguística baseada em *corpus* rejeita a caracterização da linguística de *corpus* como um método e afirma, em vez disso, que o próprio *corpus* deve ser a única fonte de nossas hipóteses sobre a linguagem. Afirma-se, portanto, que o próprio *corpus* incorpora uma teoria da linguagem.

Sendo assim, o *corpus* para análise na abordagem *Corpus-Driven* é construído consoante o alvo da pesquisa, justamente por essa abordagem ser indutiva aos dados. Assim, como afirma Biber (2015, p. 5) "os padrões linguísticos de uso descritos na análise baseada em *corpus* são generalizáveis, abordando explicitamente questões de validade externa". Portanto, no âmbito dessa abordagem, o *corpus* pode ser especializado, ou seja, construído e utilizado para fins específicos, em destaque na área da educação.

Nessa perspectiva, este estudo se caracteriza por ser *Corpus-Driven*, pois é constituído de um *corpus* especializado, composto por materiais coletados e compilados para análise de um fenômeno linguístico, cujos dados guiaram a análise. Por fim, os dados gerados a partir da análise foram utilizados para a construção de um material pedagógico, a fim de orientar o ensino por meio destes dados. Portanto, a presente pesquisa também se ampara na abordagem *Data-driven learning*, a qual será explicada a seguir.

2.2.3 Data-driven Learning (DDL)

Johns (1991) foi o precursor do termo *Data-driven learning* que, na sua visão, se caracterizava pelo ensino da língua orientada por dados providos de um *corpus*. Em seu artigo, o autor explica que, durante alguns anos, utilizou de um concordaciador (uma técnica de apresentação de dados em um contexto concordanciário, que visa fornecer uma visão rápida e contextualizada de como uma palavra-chave específica é usada em diferentes contextos em um *corpus*) com alguns alunos e chegou à conclusão de que o uso do concordaciador pode ter um impacto significativo no processo de aprendizado de línguas, promovendo a investigação e especulação por parte dos aprendizes.

Ele ainda enfatiza a abordagem indutiva da Linguística de *Corpus*, na qual os dados são primários, diferenciando-se de abordagens tradicionais baseadas em "ocultar regras". O autor, em seu trabalho, já apontava também os efeitos da DDL no papel do professor, que passa a coordenar pesquisas iniciadas pelos alunos, e na reavaliação do papel da gramática no ensino de línguas, permitindo uma conscientização gramatical baseada na descoberta do aprendiz e em evidências autênticas de uso da linguagem. Ele ilustra suas conclusões com exemplos concretos de materiais preparados com base nas consultas dos alunos, destacando a natureza desafiadora e estimulante da DDL.

Indubitavelmente, essa abordagem, introduzida por Johns (*ibid.*), contribuiu para as pesquisas envolvendo o ensino de línguas. Estudada desde a década de 1960, a DDL esteve presente em trabalhos de Antoinette Renouf, Ahmad, Corbett and Rogers (1985); Tribble and Jones (1990) e McEnery (1997), como podemos observar na Figura 4, a qual mostra a linha temporal da história da DDL, com breves descrições dos autores e suas contribuições:

Figura 4 — Linha do Tempo - DDL



Fonte: O autor (2024).

Anteriormente a essa abordagem, como apontado por Chambers (2010), os exercícios que acompanhavam os materiais didáticos dos cursos de inglês baseavam-se em textos inventados e a língua descrita, nesse tipo de material, destoava do uso real da língua. Com os avanços da DDL, criaram-se dicionários e gramáticas que apresentam a língua que realmente é falada e escrita pelos falantes.

Nesse sentido, Chambers (2010, p. 345) sustenta que a DDL "não apenas utiliza dados de *corpus* na preparação de materiais de aprendizagem de línguas, mas também dá aos alunos acesso a quantidades mais substanciais de dados de *corpus* do que os que podem ser encontrados em um dicionário, gramática ou livro didático". Isso não apenas enriquece a experiência de aprendizagem, mas também capacita os alunos a explorarem ativamente o uso da linguagem por meio de suas próprias investigações, promovendo uma abordagem mais dinâmica e participativa no processo de aquisição linguística.

Desde a sua introdução, o significado exato da abordagem continua em debate. Em uma pesquisa recente, Gilquin e Granger (2022) argumentam que a DDL se caracteriza por utilizar as ferramentas e as técnicas da Linguística de *Corpus* para fins pedagógicos. Nesse contexto, o aprendiz é disposto de dados com os quais irá analisar para conseguir chegar a alguma conclusão, por meio investigativo.

A vantagem da utilização dos dados no ensino pode ser vista em Sardinha (2004, p. 292) ao afirmar que "os alunos desenvolvem a habilidade de identificar regularidades (*patterings*) e de fazer generalizações para explicá-las. O professor assume o lugar de orientador ou coordenador de pesquisa". Por esse motivo, essa abordagem se contrasta ao sistema tradicional, pois o ensino gramatical é tirado de foco. Ao invés de um ensino regido por regras, em sua maioria incoerentes, que foram estabelecidas por meio de registros de uma língua canônica, o aluno, por meio de análises de registros reais e atuais da língua, pode identificar padrões e formular uma conclusão plausível para as ocorrências analisadas.

Logo, os alunos podem desenvolver não só autonomia e habilidades cognitivas, mas também conhecimento através da observação de dados, por esse motivo, a DDL é de cunho indutiva. Em consonância a isso, Pérez-Paredes e Mark (2021, p. 9) afirmam que:

O ensino de línguas é mais do que apenas fornecer a língua aos alunos, uma vez que nunca poderemos dar aos alunos toda a língua necessária para cada situação futura. É também importante oferecer-lhes as ferramentas cognitivas e tecnológicas que lhes permitam enfrentar com maior autonomia a imprevisível gama de situações linguísticas que poderão encontrar na sua vida profissional e social.

Nessa abordagem, os *corpora* são processados e analisados por meio de programas de computadores, tais como o *WordSmith*, *Antconc*, *Lancsbox* etc., os quais possuem ferramentas que nos oferecem, entre outros recursos: lista de concordância; frequência das palavras; ocorrência e co-ocorrência de palavras e estruturas. Dessa maneira, a integração entre ferramentas tecnológicas para a produção de conhecimento científico e ensino se faz presente e está em consonância com a ideia dos autores citados anteriormente neste capítulo.

Com a DDL, o aluno se torna o centro da produção científica, enquanto o professor é o mediador dessa construção. Devido a isso, essa abordagem pode oferecer benefícios na exploração de certos fenômenos linguísticos. Posto isso,

segundo Brown (2016), a DDL é especificamente relevante para o conhecimento aprofundado de vocabulário, identificação de colocações, estudos relacionados à Fraseologia, correção de erros comuns ou recorrentes etc.

Por outro lado, há também desafios no uso da abordagem DDL na sala de aula. Nessa perspectiva, McCarthy *et al.* (2021) alegam que os desafios giram em torno da necessidade de os alunos adaptarem suas estratégias de leitura, contextualizarem informações de maneira eficaz, personalizarem seus encontros com a linguagem, perceberem questões de autenticidade e inferência, escolherem materiais de ensino apropriados e considerarem a origem dos erros na aprendizagem de idioma.

Apesar das desvantagens, a abordagem vem sendo, desde a sua criação, em 1991, até os dias atuais, usada em pesquisas. Acerca disso, Boulton (2021) delineou o estado do campo da DDL com base em uma coleção extensa de publicações datadas até 2018. A Figura 5 mostra um gráfico com a produção total de estudos empíricos de DDL.

Figura 5 — Total de estudos empíricos de DDL ao longo dos anos

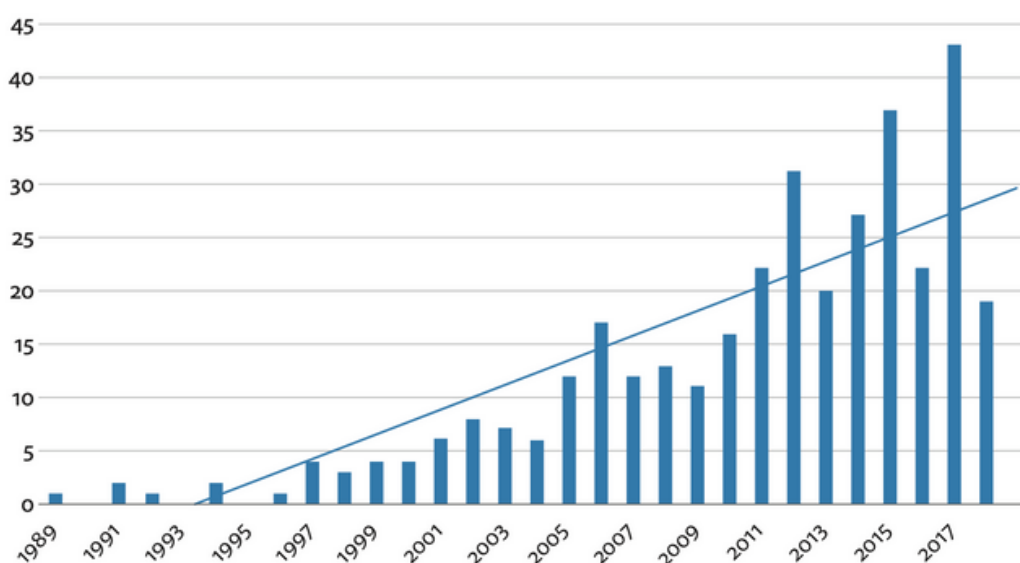


Figure 1.1 Dates of publication of the 351 studies

Fonte: Boulton (2021).

No campo da DDL, conforme delineado por Boulton (2021), os resultados revelam que o inglês foi a única língua-alvo em 88% dos 351 artigos analisados. Notou-se uma escassez de estudos com línguas não europeias, com apenas 2 para

o mandarim e o cantonês. A maioria dos estudos (80%) foi realizada em ambientes onde o inglês não é a língua do país.

A distribuição geográfica mostrou que a Ásia representa 50% dos estudos, com destaque para Taiwan, Japão e China. A África e o Oriente Médio tinham representação limitada. Em termos demográficos, 87% dos estudos foram conduzidos em contextos universitários, principalmente nos Estados Unidos, Canadá e Reino Unido. A maioria dos participantes foi composta por estudantes universitários, predominantemente em programas de graduação.

Quanto à disciplina dos estudantes, 46% estavam envolvidos em estudos de linguagem, enquanto ciência, tecnologia, engenharia, matemática (STEM) e ciências humanas e sociais (HSS) representam 17% cada. Em relação ao propósito da linguagem, 66% dos estudos trataram de linguagem para fins gerais, 17% para propósitos acadêmicos e 15% para propósitos específicos. A proficiência dos aprendizes variou, com a maioria dos estudos focando em níveis intermediários e avançados, apesar de alguns trabalhos empíricos abordarem níveis mais baixos. No entanto, a falta de padronização na descrição da proficiência dos aprendizes representou um desafio para a interpretação dos resultados. Certamente, a pesquisa de Boulton (2020) foi de extrema importância para nos revelar os caminhos que a DDL tomou desde o seu início. Desse modo, a fim de contribuir com a pesquisa dele, ressaltamos alguns trabalhos importantes na área, após 2018, que serão delineados abaixo.

Wicher (2019) aborda como a DDL pode ser integrada na sala de aula moderna de línguas estrangeiras. O estudo avaliou como as atividades, que envolvem a aplicação da DDL, melhor se encaixam nas aulas e como elas podem ser concebidas de forma flexível, conforme as condições de ensino e aprendizagem dos alunos do ensino secundário. O autor concluiu que vincular a DDL às tarefas comunicativas é fundamental para o seu sucesso como metodologia *learner-oriented*.

Lee e Lin (2019) compararam os efeitos das abordagens dedutiva e indutiva no ensino de inglês como língua estrangeira (EFL) baseado em DDL, na aquisição e retenção de vocabulário. Vinte e sete alunos de EFL foram divididos aleatoriamente em grupos dedutivos e indutivos, utilizando o *Corpus of Contemporary American English* (COCA) para aprender oito palavras-alvo. Os resultados indicaram que ambas as abordagens foram igualmente eficazes na facilitação da aquisição e retenção de

vocabulário. Além disso, a abordagem dedutiva mostrou-se tão eficaz quanto a abordagem indutiva, sugerindo que a abordagem dedutiva pode complementar a DDL de maneira mais eficiente, quando a natureza indutiva da DDL impede os alunos de aproveitarem plenamente suas vantagens potenciais.

Muftah (2023) investigou os efeitos a médio prazo das atividades da DDL na escrita de EFL. Utilizando o *BNCweb* como ferramenta de DDL, comparou a eficácia do *BNCweb* com o *Sketch Engine*, uma ferramenta de referência, em um grupo experimental e um grupo de controle de 64 estudantes árabes de graduação em EFL. Os resultados quantitativos revelaram que o grupo experimental, que usou o *BNCweb*, demonstrou uma escrita mais fluida e consistente no pós-teste, em comparação com o grupo de controle. Embora não tenha havido diferença significativa na complexidade da escrita, os resultados qualitativos indicaram atitudes positivas em relação ao uso do *BNCweb*.

Crosthwaite, Luciana e Wijaya (2023) investigaram as tentativas iniciais de professores do ensino secundário indonésio de integrar atividades de consulta de *corpus* em planos de aula de EFL. Os autores concluem que houve limitações em relação à introdução dos estagiários aos *corpora*, além da adaptação ao treinamento *online* devido à COVID-19. Porém, o estudo, embora com uma amostra pequena, respondeu à necessidade de perspectivas qualitativas em contextos de CALL. Os resultados destacaram a importância dos comentários especializados sobre os planos de aula e a orientação de um linguista de *corpus*. Apesar das limitações, o estudo incentivou mais pesquisas sobre a integração do DDL no planejamento de aulas, enfatizando a necessidade de enquadrar o treinamento DDL em teorias estabelecidas de conhecimento do professor para aumentar a aceitação futura.

2.2.4 Teorias de ensino e DDL

Além dos trabalhos citados na seção 2.2.2, há outra obra relevante para a DDL a qual retrata as Teorias de Ensino e DDL. Nesse sentido, em termos de teorias de aprendizagem, O'keeffe (2021) explica duas perspectivas principais, identificadas na pesquisa DDL: a construtivista e a sociocultural. A abordagem construtivista enfatiza a aprendizagem pela descoberta, em que os alunos se envolvem na luta cognitiva com os dados para descobrir padrões regulares. Essa perspectiva destaca a ideia de

que os alunos, ao consultar um *corpus*, estão usando comportamento adaptativo para detectar padrões significativos, em oposição a aprender e aplicar regras dadas explicitamente.

Por outro lado, ainda segundo O'keeffe (2021), a abordagem sociocultural considera a mediação do professor e a aprendizagem entre pares como componentes importantes. Ela explora o grau em que a interação do aluno com os dados é mediada por um professor ou colega, envolvendo considerações sobre autonomia e autorregulação. Elementos da Teoria Sociocultural (TSC), como autorregulação, mediação e *scaffolding*, também desempenham um papel significativo, embora menos explicitamente expressos no discurso da DDL.

Nesse sentido, os estudos relacionados à aprendizagem mediada recaem em dois principais estudiosos, Vygotsky e Feuerstein. Das principais teorias de Vygotsky, uma delas contribui muito para o ensino-aprendizagem de línguas, a Zona de Desenvolvimento Proximal (ZDP), sendo um conceito chave na teoria sociocultural do psicólogo russo. Para o autor, o desenvolvimento cognitivo de uma criança ocorre não apenas mediante processos internos, mas também por meio de interações sociais e culturais. A ZDP refere-se à diferença entre o que uma criança pode fazer de forma independente e o que ela pode realizar com a ajuda de um parceiro mais competente, como visto em Vigotsky (1991, p. 58):

[ZDP] é a distância entre o nível de desenvolvimento real, que se costuma determinar através da solução independente de problemas, e o nível de desenvolvimento potencial, determinado através da solução de problemas sob a orientação de um adulto ou em colaboração com companheiros mais capazes.

Existem dois níveis de desenvolvimento de uma criança, de acordo com Vigotsky (1991): o nível de Desenvolvimento Real, que corresponde à habilidade que uma criança demonstra ao realizar uma tarefa de forma independente, sem a ajuda de outros e a Zona de Desenvolvimento Proximal que é o nível de desenvolvimento no qual a criança é capaz de realizar uma tarefa, com a assistência de um parceiro mais competente. É a diferença entre o que uma criança pode fazer sozinha e o que ela pode fazer com ajuda.

Segundo Vigotsky (1991, p. 59) “A zona de desenvolvimento proximal provê psicólogos e educadores de um instrumento através do qual se pode entender o curso

interno do desenvolvimento”. Esse conceito destaca a importância da interação social e do ambiente cultural no desenvolvimento cognitivo das crianças. O autor enfatiza que o aprendizado não é apenas uma aquisição de conhecimento, mas também um processo social no qual as interações e colaborações desempenham um papel crucial.

Em consonância aos estudos de Vigostky, outra teoria que contribui para o ensino-aprendizagem de línguas é a do psicólogo israelense Reuven Feuerstein (1990), A Experiência de Aprendizagem Mediada (EAM). Ela é um conceito fundamental da abordagem de Feuerstein para a educação e o desenvolvimento cognitivo. Ela se baseia na ideia de que a aprendizagem é mais eficaz quando ocorre em um contexto social, orientado por um “mediador” (um professor, colega ou outra figura de apoio) que fornece suporte e orientação. Conforme o autor, EAM:

é definida como uma qualidade de interação entre o organismo e seu ambiente. Esta qualidade é assegurada pela interposição de um ser humano iniciado e intencional que media os estímulos que incidem sobre o organismo. Este modo de interação é paralelo e qualitativamente diferente da modalidade mais generalizada e mais difundida de interação entre o mundo e o organismo, referida como "exposição direta a estímulos" (Feuerstein, 1990, p. 7).

Dentre os princípios-chave da Experiência de Aprendizagem Mediada, estão a Intencionalidade Pedagógica; o Diálogo Interativo; a Zona de Desenvolvimento Proximal e a Mediação Cognitiva. Basicamente, a aprendizagem é mais eficaz quando há um propósito claro por trás das atividades educacionais, dessa forma, o diálogo entre o mediador e o aluno é visto como uma ferramenta essencial para o desenvolvimento cognitivo, que é trabalhado dentro da ZDP para alcançar maiores avanços. Sendo assim, o mediador é que desempenha um papel ativo na mediação cognitiva, auxiliando o aluno a desenvolver habilidades de pensamento mais avançadas.

Segundo O'keeffe (2021), essas perspectivas teóricas variadas, indo desde construtivista até sociocultural, coexistem na pesquisa DDL. A clareza sobre a posição pedagógica adotada é vital para pesquisas e interpretação de resultados; e a compreensão dessas diversas bases teóricas pode aprimorar o entendimento de como a postura pedagógica impacta os processos de sala de aula e a aprendizagem na DDL. Outrossim, a autora visa estabelecer relações entre as teorias de aquisição de segunda língua (SLA) e a DDL, destacando a importância de uma interação mais

profunda entre essas áreas. Historicamente, as investigações em DDL e SLA seguiram caminhos distantes, apesar do potencial para benefícios mútuos, especialmente no entendimento de como o cérebro opera no processo de aquisição de segunda língua.

Em um estudo anterior, O'Keeffe (2020) argumenta que a DDL deve ser mais consciente de sua posição no debate cognitivo sobre como ocorre a aprendizagem, seja de forma consciente, subconsciente ou uma combinação de ambas. Destaca-se a importância de uma compreensão mais aprofundada da SLA na pesquisa DDL para contribuir com esse debate. O trabalho concentra-se em estabelecer ligações entre a DDL e uma área específica de pesquisa em SLA, a teoria baseada no uso (UB). Essa teoria, originada nos estudos de aquisição da primeira língua (FLA) e, posteriormente, aplicada aos estudos de SLA, enfatiza que o conhecimento linguístico advém da experiência e do uso em um contexto significativo e socialmente comunicativo. A UB sustenta que a exposição frequente a padrões de linguagem é fundamental para a aquisição e que a análise, ao longo da vida das características distributivas do *input* linguístico, leva à emergência de regularidades estruturais.

Nesse sentido, uma forma de efetivar a apresentação recorrente de padrões linguísticos por meio da abordagem DDL é através da investigação do fenômeno das colocações, o qual constitui o foco principal desta pesquisa. Assim, na seção 2.3, a seguir, será oferecida uma descrição mais detalhada desse fenômeno.

2.3 Colocações

De maneira independente, pesquisadores na área da linguística apresentam distintas interpretações e classificações para o conceito de colocação (Halliday, 1966; Sinclair; 1991; Brezina, 2018; Cowie, 1994; Howarth, 1998; Nesselhauf, 2005). Conforme Almeida (2014), dentro das diversas aplicações desse termo, é possível discernir duas perspectivas principais: a abordagem fundamentada na frequência e a abordagem fraseológica. Essas se diferenciam por suas compreensões conceituais do termo "colocação" e pelas metodologias analíticas que preferem, seja uma análise mais voltada para estatísticas ou uma análise de cunho fraseológico.

Apesar de a fraseologia não ser o foco nesta pesquisa, é válido ressaltar que, para Cowie (1994), um dos principais autores que discutem sobre a abordagem,

colocações são como combinações, sendo elas livres, restritas e idiomáticas. Porém, essa abordagem preocupa-se menos com a frequência e a estatística e preocupa-se mais com a transparência e comutatividade nas combinações.

Nesse sentido, Almeida (2014, p. 48) sustenta que:

Transparência está relacionada ao fato dos itens da combinação, e a combinação em si, terem um significado literal ou não literal. Já comutabilidade se refere ao fato de os elementos constituintes da combinação poderem ser substituídos ou não e, em que grau a substituição dos elementos é restrita.

Apesar de Gablasova, Brezina e McEnery (2017) argumentarem que combinar as abordagens, fraseológica e baseada em frequência, reside na complementaridade dessas perspectivas, pois a integração dessas abordagens oferece uma visão mais completa e equilibrada do conhecimento colocacional, permitindo uma avaliação cuidadosa por meio de várias perspectivas analíticas, nesta pesquisa, o termo colocação é usado fundamentado na abordagem da frequência (Halliday, 1966; Sinclair; 1991), ao invés da abordagem fraseologia (Cowie, 1994; Grenger). Portanto, para cumprir os objetivos deste estudo, optamos pela identificação das colocações por meio da frequência, pois as ferramentas que utilizamos para isso cumprem todos os critérios necessários para a investigação das colocações nessa abordagem, os quais serão percorridos na seção 2.3.1, a seguir.

2.3.1 Abordagem baseada na frequência

O conceito de colocação nas discussões linguísticas foi cunhado por Firth (1957) e refere-se à organização e combinação das unidades lexicais em um texto. Em seu capítulo *Mode of Meaning*, ele destaca a importância do significado da palavra por meio da colocação, ou seja, como as relações sistemáticas com outras palavras, que ocorrem antes e depois delas no texto, influenciam o seu significado. Dessa maneira, esse tipo de significado se difere de um significado isolado e conceitual da palavra. Isso significa que, para o autor:

[a] declaração de significado por colocação e várias colocações não envolvem a definição do significado da palavra por meio de outras sentenças em termos deslocados. O significado por colocação é uma abstração no nível sintagmático e não está diretamente relacionado com a abordagem

conceitual ou ideal do significado das palavras. Um dos significados de night é sua colocação com dark e, é claro, a colocação de dark com Night (Firth, 1957, p. 196).

Sendo assim, o autor argumenta que as combinações entre as palavras e suas relações são importantes para o entendimento do significado linguístico. O autor destaca a importância da relação contextual entre palavras para compreender o significado. O exemplo fornecido ilustra como as palavras ganham significado através de suas relações em contextos específicos. Essa perspectiva é fundamental na teoria linguística de Firth (1957). Após ele, outros autores deram continuidade no estudo desse fenômeno, sendo Halliday (1966), o principal destacado neste estudo.

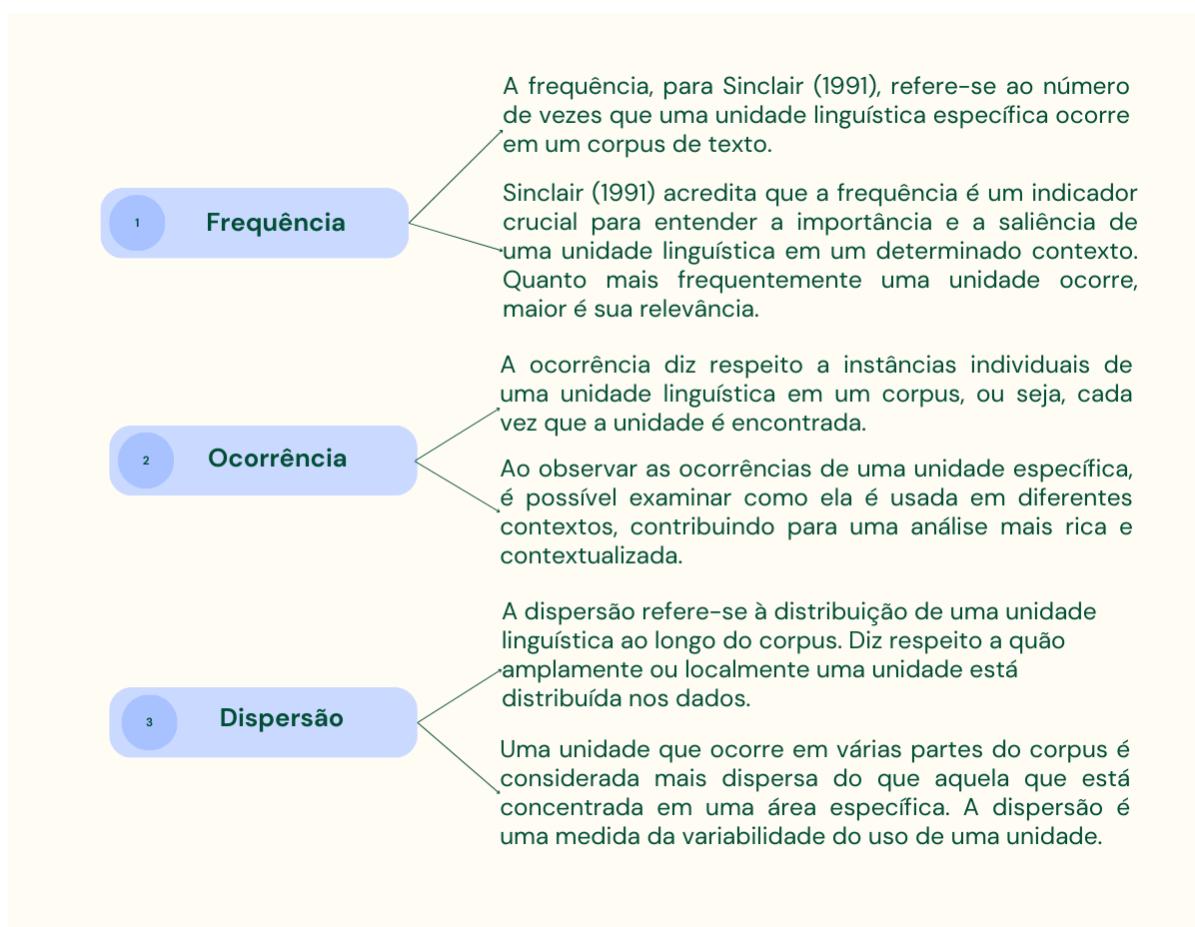
Halliday (1966) tomou uma visão um pouco diferente. Ele cunha as colocações de modo probabilístico. Isto é, conforme o autor, as colocações podem ser quantificadas em termos de sua probabilidade de ocorrência dentro de certa distância um do outro. Isso significa que a probabilidade de uma colocação específica ser usada em uma frase é medida com base na distância entre as palavras na frase. Essa distância é chamada pelo autor de *span*, ou seja, a posição que o termo ocupa na sentença em relação ao **nó**. Sobre o **nó**³, o autor o define como o item lexical estudado. Como item lexical, Halliday (1966) se refere ao lexema e todas as suas formas flexionais. Dessa forma:

[de] acordo com Halliday (1966), a colocação restringe a co-ocorrência de determinados itens lexicais. O termo item lexical é utilizado por Halliday para se referir ao lexema e todas as suas formas flexionais, ou seja, give, gives, given são representações de um mesmo item lexical. Halliday (1966) também introduziu os termos nóculo **[núcleo]** – para se referir ao item lexical estudado, colocado – para se referir ao item que coocorre e, span – para se referir ao ambiente no qual o nóculo e o colocado devem coocorrer. (Almeida, 2014, p. 46, grifo nosso).

Não distante da definição de Halliday, Sinclair (1991) também conceitua as colocações com base na probabilidade da ocorrência. Para o autor, colocação é “a ocorrência de duas ou mais palavras distantes umas das outras em um pequeno espaço de texto” (Sinclair, 1991, p.170). Porém, ele enfatiza o uso de um método estatístico, que considera três dimensões: frequência, ocorrência e dispersão.

³ Para esta pesquisa, traduzimos o termo cunhado de Halliday para "nó"

Figura 6 — Conceitos de Sinclair (1991)



Fonte: O autor (2024).

Gablasova, Brezina e McEnery (2017) ampliam a discussão com três dimensões fundamentais em relação à frequência: dispersão (*dispersion*), exclusividade (*exclusivity*) e direcionalidade (*directionality*). A dispersão refere-se à distribuição uniforme ou variada de uma colocação ao longo de um *corpus*, indicando sua presença em diversos contextos linguísticos. A exclusividade examina o grau em que duas palavras coligadas ocorrem predominantemente juntas em comparação com suas ocorrências separadas, com a aplicação de medidas como o escore de Informação Mútua. Essa dimensão é muito importante para nosso estudo, uma vez que o nosso objetivo é a identificação das colocações exclusivas, portanto, mais relevantes, dos verbos *make* e *take*. A dimensão de direcionalidade postula assimetria na atração entre as palavras de uma colocação, em que uma palavra pode ter maior probabilidade de ocorrer em conjunto do que a outra. Essas dimensões oferecem

pensamentos críticos para compreender a formulaicidade das colocações, influenciando o processamento linguístico e o desenvolvimento do conhecimento colocacional por aprendizes de línguas, ressaltando a importância de considerar tais propriedades ao selecionar métodos ou medidas para a identificação de colocações em estudos de aprendizado de línguas baseados em *corpora*.

Para Brezina (2018, p. 66), "colocações são combinações de palavras que habitualmente coocorrem em textos e *corpora*". O autor ainda argumenta que as colocações podem ser baseadas apenas na frequência ou em uma medida estatística chamada de medida de associação. Sobre as medidas estatísticas para a identificação das colocações, o autor argumenta que os testes podem ser simples ou envolver modelos estatísticos complexos. Acerca disso, ele lista as principais medidas e suas finalidades:

[a]lgumas medidas de colocação destacam a rara exclusividade da relação colocacional, favorecendo colocações que ocorrem quase exclusivamente na companhia do nó... Outras métricas, como Dice e Log Dice, e MI2 favorecem colocações que ocorrem exclusivamente na companhia uma da outra, mas não precisam ser raras. Outros podem considerar a direcionalidade (Delta P) ou a dispersão (d de Cohen) (Brezina, 2018, p. 70, com adaptações).

Por fim, a conceituação do termo "colocação" varia de acordo com cada abordagem e autor e as definições citadas na seção 2.3.1 são mais coerentes para este tipo de pesquisa. Sendo assim, para este estudo, consideramos a abordagem estatística (Sinclair, 1991; Halliday, 1966) para alcançar os objetivos almejados, pois, de acordo com Webb, Newton e Chan (2013) "[a] vantagem de definir a colocação do ponto de vista estatístico é que os pesquisadores podem usar essas medidas para identificar rapidamente as colocações para uma palavra".

2.3.2 Estudos envolvendo as colocações

Vários trabalhos relevantes envolvendo as colocações já foram publicados. Como exemplo, Xiao e McEney (2006) exploraram o comportamento colocacional e a prosódia semântica de sinônimos próximos de uma perspectiva translinguística com base em dados de duas línguas distintas, o inglês e o chinês. Os resultados desse estudo mostraram que o comportamento colocacional e as prosódias semânticas de quase sinônimos são bastante semelhantes nas duas línguas.

Durrant e Schmitt (2010) discutiram a eficácia da exposição de aprendizes adultos às colocações em um ambiente de aprendizagem. Os resultados indicam que os alunos adultos, que aprendem uma segunda língua, retêm informações sobre quais palavras aparecem juntas em sua entrada. Isto sugere que, qualquer deficiência no conhecimento dos não-nativos, sobre associações colocacionais entre palavras, se deve a informações inadequadas, e não a uma abordagem de aprendizagem semelhante à dos não-nativos.

Laufer e Waldman (2011) investigaram o uso de colocações verbo-substantivo em inglês, na escrita de falantes nativos de hebraico, em três níveis de proficiência. Os resultados mostraram que as colocações são problemáticas, mesmo para alunos avançados. Os autores sugerem que o ensino comunicativo e baseado em tarefas deve ser complementado por uma instrução focada na forma.

Brezina, McEnery e Wattam (2015) revisitaram o conceito de redes de colocação, introduzido por Phillippe (1983), e apresentaram o *Graphcoll*, uma ferramenta desenvolvida pelos autores que constrói redes de colocação, a partir de *corpora* definidos pelo usuário. Por meio de estudo de caso, utilizando dados do estudo de McEnery (2006), da *Society for the Reformation of Manners Corpus* (SRMC), os autores demonstraram que as redes de colocação fornecem informações importantes sobre as relações de significado na linguagem. Os autores afirmam que, empregando diferentes medidas de associação, estabeleceram um conjunto estável de colocações que confirmaram as descobertas originais de McEnery (2006), mas também estenderam o âmbito da análise, para além do que era possível no estudo anterior.

Frankenberg-garcia *et al.* (2019) descreveram o desenvolvimento de uma ferramenta de assistência à escrita que ajuda os escritores de Inglês Para Fins Acadêmicos (EAP) a usar colocações adequadas em seus textos. Em conclusão, os autores acreditam que a revisão deles de trabalhos anteriores e a discussão das decisões tomadas na pesquisa levantaram questões importantes sobre a usabilidade de recursos lexicográficos e assistentes de redação em geral, e a integração de informações sobre colocação para usuários de EAP com ambientes de escrita digital.

Conforme mencionado nesta seção, há uma gama de estudos cumprindo diferentes objetivos. Todavia, ainda existe uma lacuna no que se refere à produção

de materiais didáticos aplicáveis no ensino básico por meio da abordagem DDL, que orientem os alunos sobre quais itens lexicais aprenderem primeiro, uma vez que:

[i]ndiscutivelmente, as palavras que ocorrem frequentemente num grande número de contextos são importantes porque é muito provável que sejam encontradas numa variedade de situações comunicativas. Por exemplo, para alunos de uma língua é crucial saber quais itens lexicais devem aprender primeiro (Brezina, 2018, p. 53).

Por isso, esta pesquisa se difere das demais, pois objetiva utilizar a abordagem da DDL para produzir um plano de aula que visa ensinar as colocações encontradas dos verbos *make* e *take* no *corpus* especializado de *Graded Readers*, compilado para essa pesquisa.

2.3.3 Graded readers

Uma excelente maneira de incorporar os textos literários nas aulas de LI é usando os *Graded Readers*, os quais, de acordo com Hill (2008, p.185), “são livros escritos para aprendizes de inglês, usando léxico e sintaxe limitados, o primeiro determinado pela frequência e utilidade e o segundo pela simplicidade”. Esses materiais são livros de leitura de nível de dificuldade variável, categorizados em diferentes níveis conforme a proficiência linguística do aprendiz, apresentando um determinado grau de dificuldade de vocabulário, estruturas gramaticais e temas, criados para estudantes de línguas estrangeiras. Eles são projetados para ajudar os estudantes a melhorar suas habilidades de leitura, vocabulário e compreensão. Por isso, o seu uso é relevante para a sala de aula, não somente pela curadoria do vocabulário, mas também pelas questões mencionadas na introdução desta dissertação. Além disso, outro fator que justifica o uso desses livros no ensino de LI é o histórico-cultural, pois conforme destacado por Hill (2013), ao ensinar línguas estrangeiras, é crucial considerar não apenas a língua em si, mas o olhar para a cultura e a história da comunidade linguística envolvida é, também, de suma importância.

Nessa perspectiva, outro benefício para o uso dos *Graded Readers*, segundo Waring e Takahashi (2000, p 6), é a leitura constante e o processamento rápido de informações devido à facilidade das palavras, pois:

[p]rimeiramente, se a leitura for feita no nível de habilidade de leitura atual do aluno, o aluno processará palavras mais rapidamente e construindo o reconhecimento automático de palavras. Isso permitirá que ele leia mais rápido. À medida que o aluno lê mais rápido, ele começará a ver palavras em grupos ou 'pedaços de linguagem', o que permite ao leitor passar da leitura 'palavra por palavra' para a 'leitura com ideias', ajudando assim na fluência da leitura.

Clandfield e Budden (2020) também elencam razões para o uso dos *Graded Readers*. Entre elas estão: a melhora no vocabulário; na habilidade de fala; de escrita; e na fluência; o desenvolvimento da autonomia, uma vez que a leitura pode ser feita fora da sala de aula; o incentivo da leitura, pois, quanto mais se lê, melhor leitor se é, quanto melhor leitor se é, mais se quer ler, e possibilita a leitura fluida, devido ao fato de os alunos não precisarem recorrer ao dicionário frequentemente, pelo vocabulário cuidadosamente selecionado para cada nível.

Esses livros também podem ser usados para estudar e aprender certos fenômenos linguísticos, como, por exemplo, o das colocações. Nessa perspectiva, um estudo feito por Webb, Newton e Chang (2013) demonstrou como as colocações são aprendidas incidentalmente, durante as atividades de leitura e escuta de *Graded Readers*. Os autores investigaram se a exposição ao contexto e ao uso de colocações é suficiente para promover o aprendizado incidental. O estudo envolveu o uso de um teste de pré e pós-leitura para medir o conhecimento de colocações de estudantes universitários taiwaneses aprendendo inglês como língua estrangeira. Os resultados demonstraram que a leitura de textos em Língua Inglesa, dentre outros benefícios linguísticos, consegue promover o aprendizado de colocações.

Portanto, estudos como esse evidenciam a relevância do uso de livros adaptados nas aulas de LI para o aprendizado das colocações, justificando a investigação desse fenômeno nos *Graded Readers* propostos nesta pesquisa. Ainda, a escassez de estudos baseados em DDL no ensino básico, e o papel das colocações na aquisição e na fluência do aprendiz de línguas estrangeiras justificam a presente pesquisa. Neste contexto, conforme Crosthwaite e Steeples (2022, p.1):

[a]bordagens baseadas em *corpus* para educação em linguagem e alfabetização, comumente conhecidas como aprendizagem movida por dados (DDL), estão aumentando. No entanto, a grande maioria das intervenções de DDL envolve alunos adultos de nível terciário, deixando uma extrema necessidade de estudos abrangentes de DDL para o ensino médio.

Nesse sentido, a DDL pode ser bastante útil, principalmente porque é possível que se crie uma lista de palavras que podem auxiliar os estudantes e orientá-los por onde começar, principalmente os alunos dos níveis iniciais. Em consonância a isso, Nation e Waring (1997) afirmam que a lista de palavras é uma ferramenta de aprendizagem para o uso de qualquer nível de proficiência e vocabulário. Os autores ainda defendem que uma das melhores formas de aprender novas palavras é por meio das habilidades de *Listening* e *Reading*.

Nessa mesma ideia, Hunt e Beglar (2005) entendem que a leitura também é o caminho para a aquisição de vocabulário de forma eficiente; eles examinaram criticamente a pesquisa existente sobre a aquisição de vocabulário de EFL, propondo uma estrutura sistemática para acelerar o desenvolvimento lexical dos alunos. Destacando a importância da aquisição eficaz de vocabulário, para alunos que frequentemente enfrentam léxicos empobrecidos, os autores discutem duas abordagens principais: a promoção de instrução lexical explícita e estratégias de aprendizagem, juntamente com o encorajamento ao uso de instruções lexicais implícitas e estratégias de aprendizagem.

As estratégias cruciais incluem a aquisição de léxico descontextualizado, o uso de dicionários e a inferência a partir do contexto. Enquanto várias formas de instrução lexical implícita são discutidas, a ênfase recai sobre a leitura extensiva como a principal ferramenta para desenvolver o vocabulário de leitura dos alunos de EFL para um nível avançado. A noção central é que o desenvolvimento lexical mais eficaz ocorre em currículos multifacetados que equilibram atividades explícitas e implícitas, abrangendo todos os níveis do desenvolvimento dos alunos de língua estrangeira.

Ainda sobre a aquisição de vocabulário, Brenzina e Gablasova (2015) apresentaram, em um estudo à Nova Lista de Serviço Geral (new-GSL), o resultado de uma comparação robusta de quatro *corpora* linguísticos (LOB, BNC, BE06 e EnTenTen12), totalizando mais de 12 bilhões de palavras em execução. Os *corpora* foram selecionados para representar uma variedade de tamanhos e abordagens em termos de representatividade e amostragem. O estudo investigou a sobreposição lexical entre os *corpora* nas 3.000 palavras principais com base na frequência reduzida média (ARF), uma medida que considera tanto a frequência quanto a dispersão de itens lexicais. Os resultados indicam a existência de um núcleo estável de vocabulário de 2.122 itens (70,7%) entre os quatro *corpora*.

Além disso, esses itens de vocabulário ocorrem com classificações comparáveis nas listas de palavras individuais. Ao criar o *new-GSL*, os itens do núcleo foram combinados com novos itens, que ocorrem com frequência nos *corpora*, que representam o uso atual da língua (BE06 e EnTenTen12). O produto, o *new-GSL*, consiste em 2.494 lemas e abrange entre 80,1 e 81,7% do texto nos *corpora* de origem. Em resumo, o estudo tratou da criação de uma nova lista de vocabulário com base em uma análise abrangente de *corpora* linguísticos, destacando a estabilidade e a abrangência do núcleo vocabular identificado.

Em suma, este capítulo detalhou as teorias em que a pesquisa se embasou. Inicialmente, foi dissertado sobre a LC, o *corpus* e sua extensão e a abordagem DDL. Logo, o capítulo destacou que o estudo das colocações na linguística abrange uma diversidade de abordagens e interpretações, desde enfoques baseados na frequência até análises fraseológicas. Essas perspectivas influenciam não apenas a compreensão conceitual das colocações, mas também as metodologias de análise adotadas. Enquanto alguns estudos exploram o comportamento colocacional em diferentes línguas, outros investigam a eficácia do ensino de colocações para aprendizes de línguas estrangeiras. O uso de *Graded Readers* tem se mostrado promissor no ensino de colocações, oferecendo oportunidades para o aprendizado desses fenômenos linguísticos. Além disso, abordagens baseadas em dados, como a DDL, têm ganhado destaque na educação linguística, preenchendo lacunas na produção de materiais didáticos para o ensino básico. Assim, a integração de práticas de leitura com foco em colocações nos *Graded Readers* pode potencializar significativamente o desenvolvimento vocabular dos alunos de línguas estrangeiras, representando um avanço importante no campo da educação linguística.

Portanto, esta dissertação se posiciona como uma contribuição significativa ao integrar a LC e a DDL, buscando compreender e aplicar efetivamente essas abordagens no contexto do ensino. Nesse âmbito, como mencionado nos objetivos específicos, nesta pesquisa, nos interessa o fenômeno das colocações. Assim sendo, utilizaremos da abordagem DDL para promover recursos para o aprendizado das colocações em aprendizes de LI da educação básica, uma vez que, de acordo com Kreyer (2021), há muitos estudos de desenvolvimento das colocações em alunos avançados e poucos estudos voltados para aprendizes iniciantes. Isto é, a concentração de trabalhos voltados para a educação superior é maior se comparada

com os voltados para a educação básica. Dessa maneira, a fim de diminuir essa discrepância, este estudo se compromete, a partir da identificação das colocações relevantes e exclusivas em um *corpus* de livros literários adaptados, criar uma sequência didática na abordagem DLL para contribuir com o aprendizado das colocações na sala de aula.

3 ASPECTOS METODOLÓGICOS DA PESQUISA

Neste capítulo, a dissertação se concentra na descrição metodológica da pesquisa. Quanto à abordagem, trata-se de uma pesquisa mista, por haver coleta, análise e interpretação de dados. Em relação à pesquisa mista, Creswell (2021) aponta que os procedimentos de métodos mistos se desenvolveram em resposta à necessidade de esclarecer o objetivo de reunir dados quantitativos e qualitativos em um único estudo. Em consonância a isso, Paiva (2019) afirma que o método, comumente chamado de quali-quantitativo, emprega técnicas qualitativas e quantitativas para a busca de informações, a fim de possibilitar um melhor conhecimento do assunto pesquisado. Dessa forma, utilizamos de uma base de dados qualitativa para localizar e descrever o fenômeno linguístico em questão e o analisamos qualitativamente, a fim de apontar suas características na língua.

Inicialmente, é relatado a natureza do *corpus* que se constitui de obras literárias adaptadas, conhecidas como *Graded Readers*, conforme os níveis de proficiência do *Common European Framework of Reference for Languages* (CEFR) em inglês. Esse *corpus* é composto por uma variedade de gêneros literários, desde aventura e biografia até ficção científica e romance, permitindo uma análise abrangente. Apesar de ser classificado como pequeno, em termos de tamanho, sua aplicação é especializada, direcionada para fins pedagógicos em sala de aula e suporte à abordagem DDL. Em seguida, são expostos o *software* de análise linguística empregado e suas respectivas ferramentas. Ademais, é delineado o método estatístico adotado nesta pesquisa e, por fim, são estipulados critérios rigorosos para a identificação das colocações relevantes.

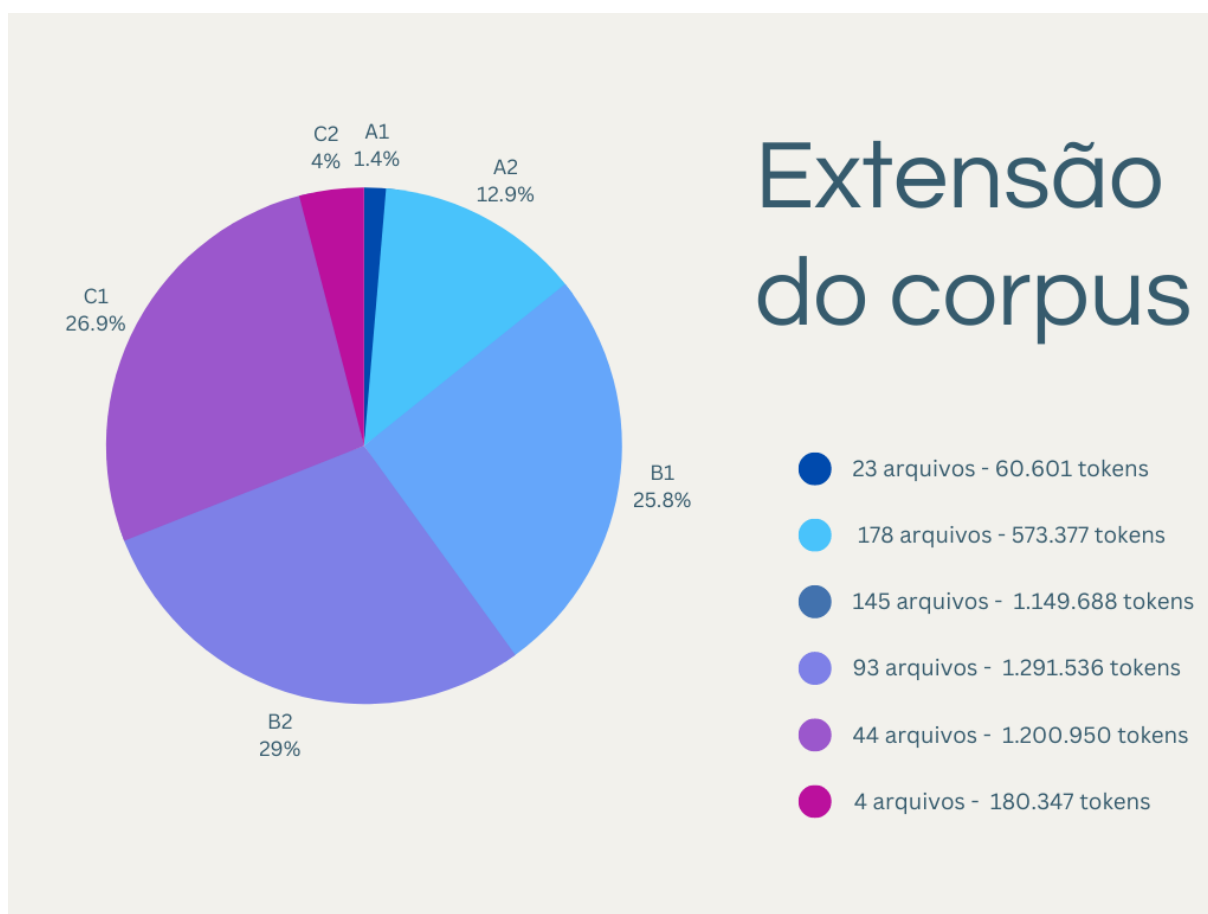
3.1 *Corpus* de estudo

O *corpus*, objeto de estudo, consiste em obras literárias adaptadas de acordo com cada nível de proficiência seguindo CERF, chamados de *Graded Readers*. Trata-se de um *corpus* em Língua Inglesa, de modalidade escrita, composto pelos seguintes gêneros literários, os quais foram rotulados pelo próprio *site*, que disponibiliza o material: aventura; biografia; comédia; crime; cultura; detetive; documentário; drama; conto de fadas; fantasia; ficção; conto popular; Horror gótico; Ficção histórica; Horror;

história de amor; filmes; mistério; romance; novela; sátira; ficção científica; conto; conto de fadas; *thriller*; histórias reais e vida jovem/aventura.

A pesquisa é composta por um *corpus* especializado que, como visto no capítulo 2, é construído e utilizado para fins específicos e se justifica por seus fins serem aplicados à abordagem DDL, a qual o *corpus* para análise é construído conforme o alvo da pesquisa. O presente *corpus* possui 487 arquivos com um total de quatro milhões, quatrocentos e cinquenta e seis mil, quatrocentos e noventa e nove (4.456.499) *tokens*. Os *subcorpora* foram divididos em **A1** (23 arquivos, totalizando 60.601 [sessenta mil e seiscentos e um] *tokens*); **A2** (178 arquivos, totalizando 573.377 [quinhentos e setenta e três mil, trezentos e setenta e sete] *tokens*); **B1** (145 arquivos, totalizando 1.149.688 [um milhão, cento e quarenta e nove mil, seiscentos e oitenta e oito] *tokens*); **B2** (93 arquivos, totalizando 1.291.536 [um milhão, duzentos e noventa e um mil, quinhentos e trinta e seis] *tokens*); **C1** (44 arquivos, totalizando 1.200.950 [um milhão, duzentos mil, novecentos e cinquenta *tokens*) e **C2** (4 arquivos, totalizando 180.347 [cento e oitenta mil, trezentos e quarenta e sete] *tokens*). A extensão do *corpus* pode ser visualizada no gráfico na Figura 7:

Figura 7 — Gráfico da Extensão do *Corpus*



Fonte: O autor (2024).

Devido à sua baixa extensão, o *corpus* se classifica como pequeno. No entanto, o *corpus* se justifica, nesta pesquisa, devido ao fato de ele possuir uma aplicação especializada, como já citado, tendo uso pedagógico em sala de aula e uso na elaboração de material na abordagem DDL. Nesse viés, Flowerdew (2009, p. 328) argumenta que:

[...] corpora muito menores com apenas 100.000 palavras ou menos estão sendo criados o tempo todo para aplicações especializadas. Deve-se ter em mente, no entanto, que, conforme apontado por Gavioli e Aston (2001, p. 238), mesmo os *corpora* muito grandes contêm menos linguagem do que o usuário médio terá experimentado em sua vida diária. [...] Além disso, enquanto cada texto recebe peso igual em um *corpus*, na vida real alguns textos terão mais valor e serão experimentados mais vezes do que outros (poesia e textos religiosos, por exemplo, podem ser altamente valorizados e ouvidos ou lidos muitas vezes).

Para a coleta das obras que compõe o *corpus*, foi feito o *download* dos *Graded Readers* no *site English E-reader* ⁴, o qual disponibiliza gratuitamente os livros adaptados, dessa maneira, facilita futuras pesquisas que objetivem replicar os resultados desta e a replicação das atividades propostas no Capítulo 5. No *site*, foi selecionada na aba “*library*” a opção “*genres*”, desse modo, foi possível filtrar pelos gêneros que consideramos literários, os quais foram citados acima. Os livros foram baixados em formato *TXT* e salvos em pastas nomeadas conforme o nível do livro, o qual é mostrado no canto superior direito do *site* ao selecionar a obra adaptada. Com base nos objetivos da pesquisa, após a fase de armazenamento, o *corpus* não precisou passar por nenhum processo de limpeza. Por fim, com o intuito de alcançar os objetivos de classificação das colocações, foi utilizado o recurso de etiquetagem do *spaCy*, disponível no próprio *software*, utilizado para análise do *corpus*, o qual será descrito a seguir.

3.2 Software

Para a análise do *corpus*, foi utilizado o programa de análise linguística criado pela Universidade de Lancaster chamado *LancsBox* na versão X 3.0.0 (Brezina; Platt, 2023). A versão X do *LancsBox* inclui melhorias e recursos aprimorados em relação às versões anteriores, *#LancsBox v. 4.x* e *#LancsBox v. 5.x*. Esses avanços oferecem maior precisão e eficiência no processamento de texto, além de uma interface mais amigável e intuitiva para facilitar o uso.

Além disso, o *LancsBox X* pode ser personalizado e adaptado, de acordo com as necessidades específicas de cada usuário ou projeto. Ele suporta diferentes idiomas e pode processar eficientemente grandes volumes de texto, os quais podem ter milhões e até bilhões de palavras. O programa foi projetado para ajudar pesquisadores, linguistas e profissionais de análise de texto a realizar uma ampla gama de tarefas. Tais tarefas podem ser executadas, a partir das seguintes ferramentas:

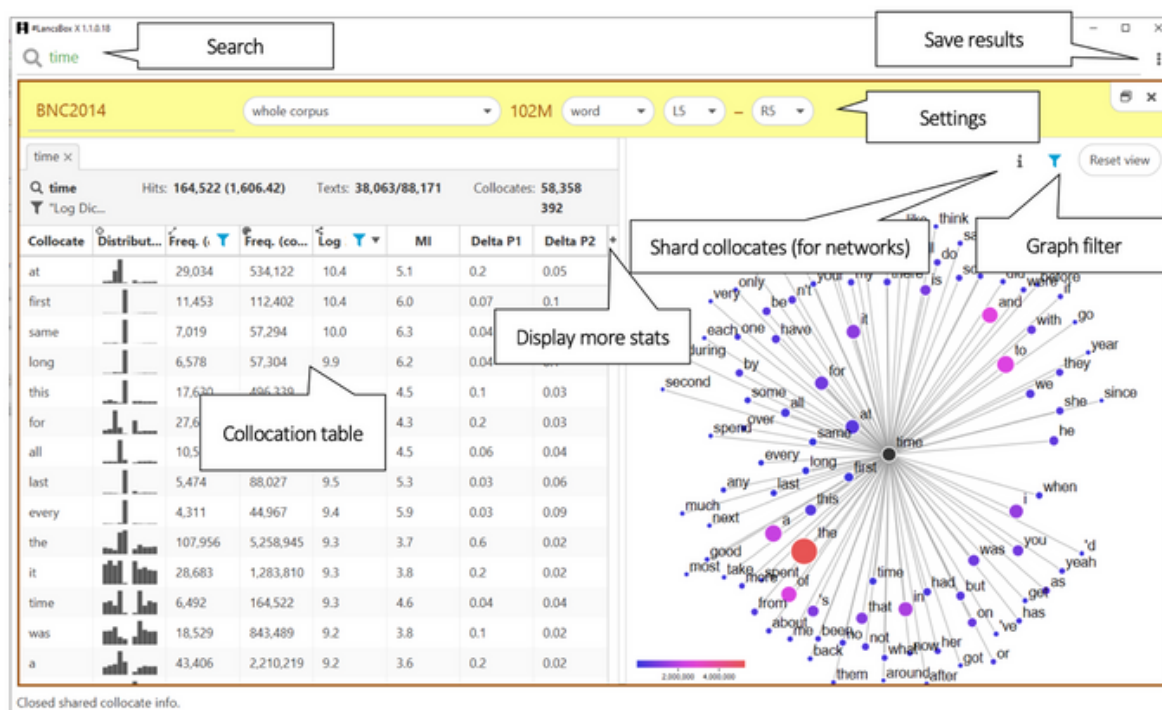
⁴ <https://english-e-reader.net>

- **Key Word in Context (Kwic):** essa ferramenta gera uma lista, em forma de concordância, de todas as instâncias de um termo pesquisado, em um *corpus* selecionado. De acordo com Brezina e Platt (2023), tal ferramenta é indicada para encontrar a frequência de uma palavra ou frase em um *corpus*, a frequência de diferentes classes de palavras, estruturas complexas e comparar múltiplas análises lado a lado.
- **Words:** essa ferramenta permite fazer uma análise aprofundada de frequências de palavras, categorias gramaticais e semânticas, como também fazer comparações de *corpora*, por meio da técnica de palavras-chave. Brezina e Platt (2023) recomendam a ferramenta para calcular medidas de frequência e dispersão e visualizar frequência e dispersão em *corpora*.
- **Text:** essa ferramenta oferece, de acordo com Brezina e Platt (2023), uma visão panorâmica de todos os arquivos do *corpus*, seu tamanho e diversidade lexical e é indicada para analisar o texto de forma quantitativa e visualizar os arquivos do *corpus* antes de analisá-los.
- **Graphcoll:** essa ferramenta identifica as colocações em um *corpus*. Ela será descrita de forma mais detalhada na seção 3.2.1, uma vez que possui grande relevância para a pesquisa, por ser utilizada para análises de colocações.

3.2.1 Graphcoll

Brezina e Platt (2023) explicam que a ferramenta *Graphcoll* encontra colocações de uma palavra ou frase e identifica co-ocorrência de categorias gramaticais em um *corpus* selecionado. Essa funcionalidade é essencial para a análise do *corpus* neste trabalho, uma vez que, ao identificar as co-ocorrência de uma determinada palavra, gera-se uma lista de colocações que são o foco de estudo nesta pesquisa. Tais colocações encontradas são exibidas em uma tabela, contendo informações relevantes que devem ser interpretadas pelo pesquisador, juntamente com seu respectivo gráfico. Na Figura 8, tem-se a visão panorâmica da ferramenta.

Figura 8 — Visão geral do *Graphcoll*



Fonte: Brezina e Platt (2023, p. 13).

Essa ferramenta possui configurações que podem ser alteradas de acordo com os propósitos do pesquisador. Essas configurações ficam localizadas na parte superior da tela e podem ser mudadas a qualquer momento. É possível alterar a unidade (palavra, lema *etc.*) e alterar a extensão da busca, ou seja, quantas palavras à esquerda (L) e quantas palavras à direita (R) do termo de pesquisa, chamado de **nó**, o qual será explicado com mais detalhes na seção 3.3.

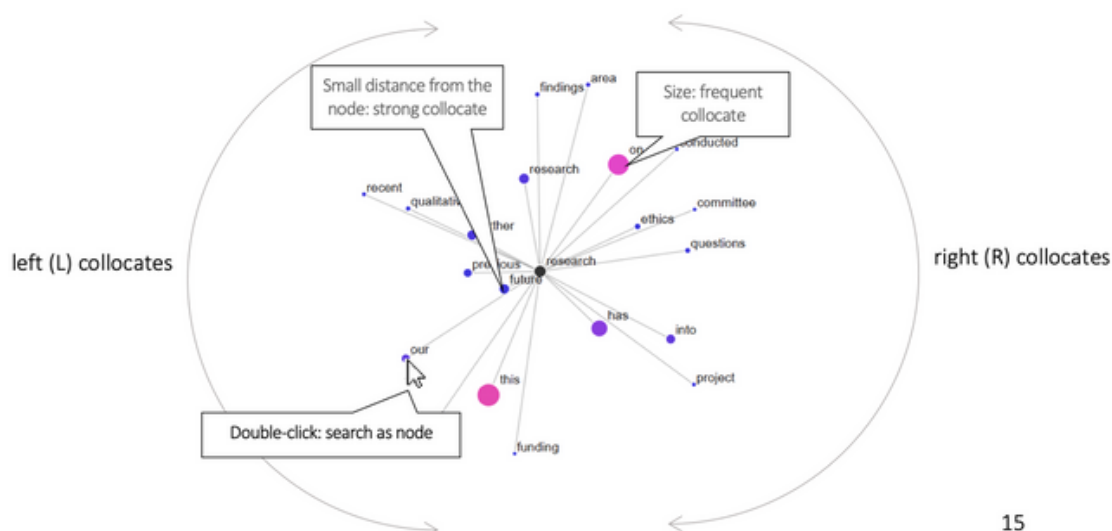
Após selecionar as configurações e digitar a palavra, frase, estrutura *etc.*, na barra de pesquisa, uma tabela será mostrada, com colunas indicando: *collocate* (os colocados encontrados); *distribution* (a localização dos colocados no *corpus* em relação à extensão da busca); *freq(Coll)* (a frequência da colocação em relação ao nó); *freq(corpus)* (frequência do colocado em qualquer lugar no *corpus*) e *Stats* (o valor da medida estatística selecionada, sendo todas as medidas computadas de uma vez só). Em relação aos resultados dos cálculos estatísticos mostrados no *Stats*, o pesquisador pode alterar quais serão mostrados ou não na tabela. A interpretação dos valores estatísticos que aparecem na tabela será explicada na seção 4.3.

A coluna indicando o *freq(Coll)* também pode ser alterada. Ao clicar no filtro, é possível mudar o valor de quantas vezes o colocado precisa aparecer em co-

ocorrência com o **nó** para ser considerado uma colocação. Sobre isso, Brezina, McEnergy e Wattam (2015), ao discutirem alguns critérios de identificação de colocações, sugerem 5 como valor de frequência mínima, ou seja, a palavra precisa, no mínimo, coocorrer 5 vezes com o **nó** para ser considerada uma colocação, por razões de exclusividade. Isso significa que nem sempre uma palavra que coocorre com outra é exclusiva dessa ocorrência, por isso, quanto mais frequente for o colocado em relação ao **nó**, maior é a exclusividade. O pesquisador pode alterar esse valor dependendo da sua pesquisa.

Já os gráficos possuem múltiplas dimensões, sendo elas o comprimento da linha; o tamanho do círculo; a cor do círculo e a posição do colocado, como pode ser visualizado na Figura 9:

Figura 9 — Gráfico do *Graphcoll*



15

Fonte: Brezina e Platt (2023, p. 15).

Como explicado por Brezina e Platt (2023), o comprimento da linha se refere à proximidade do colocado com o **nó**. Quanto mais próximo o colocado estiver do **nó**, mais forte será a associação entre o **nó** e o colocado. O tamanho do círculo se refere à frequência da colocação, de acordo com os valores dados pelo *freq(coll)*. Quanto maior o círculo, maior é a frequência da colocação. A cor do círculo se refere à frequência do colocado em qualquer lugar do *corpus*. Quanto mais próxima da tonalidade escura, mais frequente é o colocado no *corpus*. Por fim, a posição se refere à posição exata dos colocados no texto. Há colocados que aparecem

predominantemente mais à direita e outros à esquerda, e quando o colocado aparece em ambas as posições com frequência, ele aparece mais no centro do gráfico.

3.3 Medidas estatísticas

É importante destacar que a escolha da métrica estatística depende dos objetivos do pesquisador para com o *corpus* e do tipo de colocação que será buscada. Nessa perspectiva, o tipo de colocação que nos interessa nesta pesquisa são as exclusivas. Sob essa ótica, Segundo Brezina (2018), testes como *MI*, por exemplo, destacam a rara exclusividade da relação de colocação, mostrando as colocações que ocorrem quase que exclusivamente com o **nó**, mas outros como o *Dice*, o Log Dice e o *MI2* favorecem a ocorrência exclusiva do colocado com o **nó**, mas não necessariamente de forma rara. Portanto, o uso de métricas com estas são necessárias para a pesquisa.

Como visto anteriormente, na ferramenta do *Lancsbox X*, os resultados dos cálculos estatísticos são mostrados na tabela gerada a partir da busca inicial de uma palavra, frase, estrutura *etc.* Apesar de ser possível editar as medidas que serão apresentadas na tabela, por configuração padrão, o *Graphcoll* mostra o cálculo de quatro medidas estáticas: *Log Dice*, *MI*, *Delta P1* e *Delta P2*. Para esta pesquisa, nos interessam os resultados do *Log Dice*, uma vez que esse método estatístico se mostra mais eficiente para identificar as colocações exclusivas e relevantes para o estudo em questão, como explicado por Brezina (2018). Dessa forma, a seguir, será explicado o que é e como é efetuado o cálculo.

3.3.1 Log Dice

O *Log Dice* é um método estatístico que se baseia no cálculo da frequência do **nó**, das **colocações** (co-ocorrência da colocação com o **nó**) e do **colocado** em todo o *corpus*. Além disso, conforme Gablasova, Brezina e McEnery (2017, p. 164) o *Log Dice*:

é uma medida padronizada que opera em uma escala com um valor máximo fixo de 14, o que torna o Log Dice diretamente comparável entre diferentes *corpora* e um pouco preferível à pontuação MI e MI2, nenhum dos quais tem um valor máximo fixo. Com Log Dice, podemos ver mais claramente do que

com MI ou MI2 o quão longe o valor de uma combinação específica está do máximo teórico, que marca uma combinação totalmente exclusiva.

Em outras palavras, o *Log Dice* se mostra mais eficiente para a busca das colocações nesta pesquisa do que outros métodos como o *MI* e o *MI2*, por ele ter o valor teórico fixo de 14. Isso significa que, quanto mais próximo do número 14 for o resultado do cálculo, mais exclusiva e relevante é a colocação. Além disso, outra vantagem desse método é que sua fórmula não é afetada pelo tamanho do *corpus*, sendo assim, é possível comparar níveis de similaridade entre *corpora* diferentes.

A fórmula do *Log Dice* se apresenta da seguinte maneira:

(1)

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times O_{11}}{R_1 + C_1}$$

O_{11} é a frequência da colocação. R_1 é a frequência do **nó** no *corpus* inteiro, já C_1 é a frequência do colocado no *corpus* inteiro. Para tornar o processo de entendimento do método mais didático, e para entender o que cada nomenclatura significa (**nó**, **colocação** e **colocado**), considere a situação a seguir, a partir do quadro abaixo:

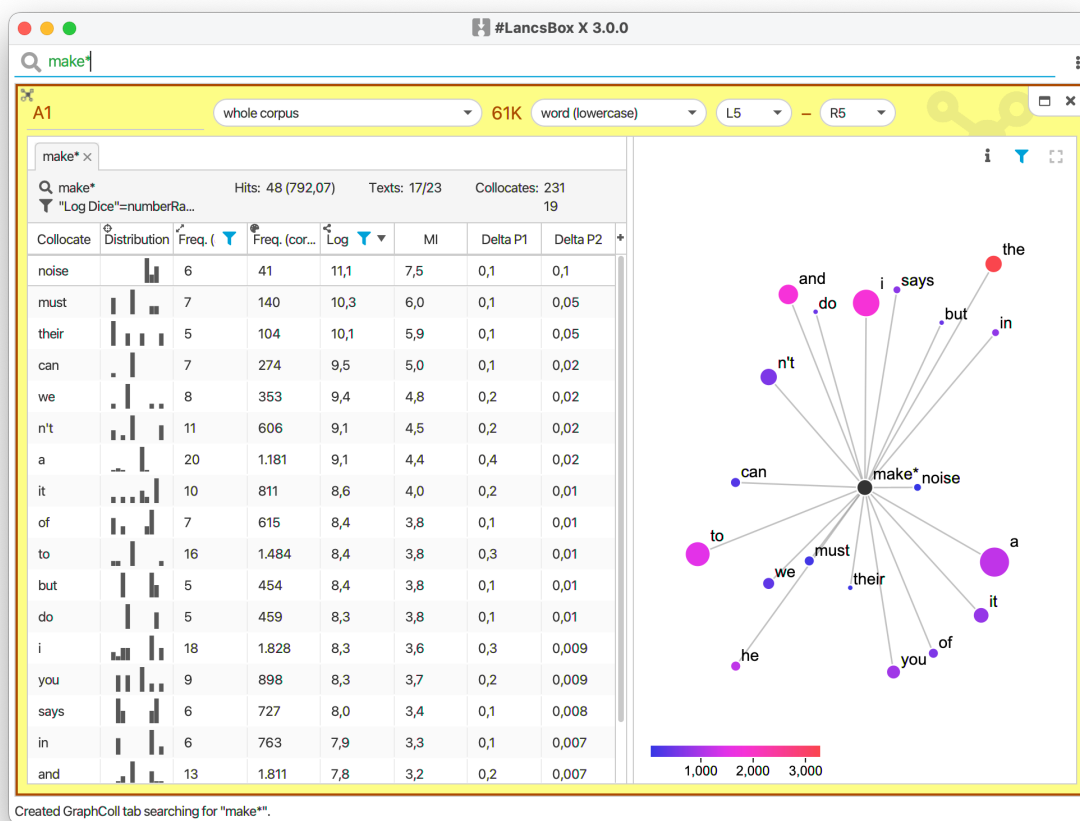
Quadro 1 — Situação exemplo com o verbo *make*

Nomenclatura	Palavra	Fórmula (<i>Log Dice</i>)	Frequência
Nó	<i>make</i>	R_1	36
Colocado	<i>noise</i>	C_1	41
Colocação	<i>make (noise)</i>	O_{11}	7

Fonte: O autor (2023).

Se o verbo, o qual queremos descobrir suas colocações no *corpus*, for *make*, este, então, é denominado **nó**, ou seja, o ponto de partida da busca, a palavra digitada na barra de pesquisa da ferramenta *Graphcoll*, como ilustrado na Figura 10.

Figura 10 — Graphcoll - make



Fonte: O autor (2024).

Após a pesquisa de *make* na ferramenta, será informado que o item ocorre 36 vezes. Esse valor é correspondente ao R_1 .

Também é informada uma lista de palavras que sistematicamente ocorrem com o **nó**. Tais palavras são os **colocados** (*noise, it, a, of, you*), como visto na primeira coluna da Figura 10. Cada colocado possui informações dadas pela ferramenta, como a frequência que eles ocorrem no *corpus* inteiro e a frequência da co-ocorrência com o **nó**, identificada estatisticamente, denominada **colocação**.

É informado pela ferramenta que uma das palavras que coocorrem com *make* é *noise*. É informado também que *noise* ocorre 7 vezes em um campo de distribuição de 5 palavras para direita e cinco para esquerda junto ao **nó**, isso significa que *noise* é uma **colocação** de *make*, ou seja, uma palavra que coocorre com o **nó**. Essa quantidade de vezes que *noise* ocorre em relação ao **nó** é correspondente ao O_{11} .

Além disso, a ferramenta mostra que o colocado *noise* ocorre mais 34 vezes no *corpus*, totalizando 41. Esse valor, 41, é correspondente ao C_1 , ou seja, a quantidade total de vezes que a palavra *noise* ocorre no *corpus* inteiro.

A ferramenta também informa o cálculo desses valores a partir do *Log Dice*. Ao calcular o R_1 , o C_1 e o O_{11} na fórmula, se obtém um valor. A saber:

(2)

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times O_{11}}{R_1 + C_1}$$

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times 7_{11}}{36_1 + 41_1}$$

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times 7}{36 + 41}$$

$$\text{Log Dice} = 14 + \log_2 \frac{14}{77}$$

$$\text{Log Dice} = 14 + \log_2 (0,18)$$

$$\text{Log Dice} = 14 + \log_2 \frac{18}{100}$$

$$\text{Log Dice} = 13 + 2 \log_2 \frac{183}{5}$$

$$\text{Log Dice} = 11,5$$

A partir do cálculo, concluímos que o valor do *Log Dice* é 11,5. Uma vez que o valor máximo é 14, podemos afirmar que 11,5 se aproxima do valor máximo, tornando, assim, *noise* uma colocação relevante de *make* no *corpus* analisado.

É importante ressaltar que todas as palavras que coocorrem com o **nó** são colocações, mas, nem todas elas são colocações relevantes, exclusivas. O que as tornam colocações relevantes é exatamente o cálculo estatístico, justificando o uso do *Log Dice* na pesquisa e sustentando a confiabilidade dos dados obtidos, a partir do uso da ferramenta *Graphcoll*.

Com base no que foi descrito, considere os seguintes dados:

Nó: *make* (frequência: 5);

Colocação: *make noise* (frequência 5);

Colocado: *noise* (frequência: 5).

Se *make* é pesquisada em outro *corpus*, e neste a quantidade de vezes que *ele* ocorre seja 5, e a palavra *noise* ocorra em co-ocorrência com *make* 5 vezes, e a palavra *noise* ocorra 5 vezes no *corpus* inteiro, o valor do cálculo do *Log Dice* será 14, isto é, de todas as vezes que *noise* ocorreu, todas elas foi combinando com *make*, a tornando uma colocação exclusiva de *make* no *corpus* pesquisado.

Aplicando na fórmula, a equação se apresenta da seguinte maneira:

(3)

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times O_{11}}{R_1 + C_1}$$

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times 5_{11}}{5_1 + 5_1}$$

$$\text{Log Dice} = 14 + \log_2 \frac{2 \times 5}{5+5}$$

$$\text{Log Dice} = 14 + \log_2 \frac{10}{10}$$

$$\text{Log Dice} = 14 + \log_2 (1)$$

$$\text{Log Dice} = 14 + 0$$

$$\text{Log Dice} = 14$$

Para reforçar e ilustrar a explicação dada até então, segue a tabela criada por Gablasova, Brezina e McEnery (2017, p. 4) que faz uma comparação de três métodos estatísticos, incluindo o *Log Dice*, com valores hipotéticos:

Figura 11 — Comparação dos resultados dos métodos estatísticos

Comparison of the performance of t-score, MI-score and Log Dice

	O_{11}	R_1	C_1	T- score	MI score	Log Dice
Example 1	50	500	500	7.0	7.6	10.7
Example 2	50	50,000	50	6.7	4.3	5.0
Example 3	50	50	50,000	6.7	4.3	5.0
Example 4	50	50	50	7.0	14.3	14.0
Example 5	5	5	5	2.2	17.6	14.0

Fonte: Gablasova, Brezina e McEnery (2017, p. 4).

Infere-se que quando o O_{11} , o R_1 e o C_1 são valores iguais, como no exemplo dado com *make* acima, o resultado do *Log Dice* é 14, ou seja, a colocação será exclusiva, relevante, forte. Por outro lado, quando o R_1 é extensamente maior que o

O_{11} e o C_1 , ou seja, quando o **nó** ocorre mais vezes no *corpus* inteiro e a palavra que coocorre com ele e o colocado têm ocorrências menores, o valor do *Log Dice* é mais próximo de 0, o que significa que a palavra possui menos exclusividade, é menos relevante no *corpus*. O mesmo acontece quando o C_1 é extensamente maior que o O_{11} e o R_1 . Isso significa que, se o colocado é muito frequente no *corpus*, mas ocorre poucas vezes próximo do **nó**, isso o torna menos exclusivo e relevante. Em contrapartida, quando a frequência do O_{11} (a coocorrência da palavra com o **nó**) é menor, mas a frequência do **nó** é parecida com a frequência do colocado, o valor do *Log Dice* se aproxima mais de 14 do que de 0, configurando a colocação como mais exclusiva e relevante.

4 CRITÉRIOS DE IDENTIFICAÇÃO DE COLOCAÇÕES RELEVANTES

Primeiramente, é importante ressaltar que este estudo visa a identificação das colocações mais exclusivas dos verbos *make* e *take*. Esta escolha é justificada devido a esses verbos serem considerados de alta frequência, isto é, a chance de aparecerem no *corpus* é maior do que quaisquer outros verbos, além de terem relevância maior para aprendizes da língua, por ocorrerem com mais frequência. Nesse sentido, Biber *et al.* (1999) apontam que os mais frequentes verbos lexicais que aparecem no *Longman Spoken and Written English Corpus* (LSWE) são: *say; get; go; know; think; see; make; come; take; want; give e mean*.

Em consonância a isso, Lin e Lin (2019, p. 1) alegam que esses verbos não existem isoladamente, pois “eles ocorrem em vários padrões fraseológicos em inglês, ou seja, colocações”. Isso revela que a aprendizagem dos verbos de alta frequência é importante para o entendimento das colocações, e, por conseguinte, para o aprendizado da Língua Inglesa. Porém, aprendizes de língua estrangeira têm dificuldades no uso desses verbos:

[...] estudos têm mostrado que verbos de alta frequência são, na verdade, usados em excesso por aprendizes de EFL, que em vez disso dependem de "palavras maiores, mais raras e desajeitadas" que tornam sua linguagem "inclinada e desajeitada" (Sinclair, 1991, p. 79), ou usadas em formas que lhes dão "muito pouco significado" (ibidem, p. 147). Examinando casos ou uso excessivo, subuso e uso indevido, fica claro que verbos de alta frequência parecem ser problemáticos para os alunos, particularmente em um contexto de EFL. (Lin; Lin, 2019, p. 1).

Além disso, Juknevičienė (2008) argumenta que os erros de colocação podem ser atribuídos à influência da língua materna. Em seu estudo, a autora analisa as habilidades dos aprendizes italianos de inglês em produzir colocações com os verbos de alta frequência e as compara com um *corpus* de falantes nativos de inglês. Tal estudo concluiu que os alunos, quando recorrem à tradução da língua materna para o inglês, usam colocações diferentes dos nativos. Portanto, a investigação das colocações dos verbos que aparecem com frequência pode ajudar a resolver a dificuldade que os estudantes possam enfrentar em relação ao uso destes.

Dessa forma, como ponto de partida, optamos por mapear as colocações dos verbos frequentes na Língua Inglesa no *corpus* coletado. Uma vez que o *corpus* é composto por textos ficcionais, escolhemos os verbos que aparecem mais vezes nesse gênero, como apontado por Biber *et al.* (1999), os quais são: *make*; *get*; *go*; *give*; *take* e *come*. Nesse âmbito, para mapear a frequência desses verbos no *corpus* desta pesquisa, verificamos a ocorrência deles. Portanto, no Quadro 2 abaixo, é possível observar a relação de ocorrência desses verbos de alta frequência.

Quadro 2 — Relação de ocorrências no *corpus*

Verbo	ocorrência	textos em que ocorre/ número total de textos
<i>make</i> *	3.982	396/487
<i>get</i> *	7.881	447/487
<i>go</i> *	34.721	486/487
<i>give</i> *	3.416	403/487
<i>take</i> *	5.227	440/487
<i>come</i> *	7.640	486/487

Fonte: O autor (2023).

Apesar de os verbos *go*, *get* e *come* serem mais frequentes no *corpus*, optamos por escolher apenas *make* e *take*, pois, ao verificar os colocados no *software*, notamos que as estruturas colocacionais deles são parecidas (*verb + noun*, *verb + adjective etc.*), além de possuírem colocados diversificados e com valores de cálculo estatístico maiores, em comparação ao restante dos verbos. Dessa forma, a comparação entre os colocados dos dois verbos pode ser realizada de forma mais precisa, justamente por compartilharem de estruturas semelhantes, e a ocorrência de colocações mais

exclusivas pode ser mais provável, uma vez que os cálculos do teste estatístico se aproximam mais do 14.

Para esta pesquisa, como forma de identificação das colocações dos verbos *make* e *take*, utilizamos das notações padrões chamados de *Collocation Parameters Notation* (CPN) propostas por Brezina, McEnery e Wattam (2015), uma vez que, segundo Brezina (2018, p. 74):

Por uma questão de replicabilidade dos resultados, todos os principais parâmetros que podem afetar a identificação do colocado devem ser relatados. Para tanto, Brezina et al. (2015) introduzem a notação de parâmetros de colocação (CPN) que captura todos os parâmetros importantes para identificação de colocação.

Dessa forma, utilizaremos do CPN, uma vez que ele contém todos os parâmetros importantes para a identificação de colocações. Esses parâmetros são: a unidade, sendo palavra; a extensão (*span*), sendo de 5 para a direita e 5 para a esquerda, e estipulado como frequência mínima da colocação 5 ocorrências.

De acordo com o que foi discutido na seção anterior, considera-se, nesta pesquisa, uma colocação relevante aquela que coocorre no mínimo 5 vezes com o **nó** e que cujo cálculo estatístico seja aproximado ou igual a 14. Quanto mais próximo do 14, mais exclusiva é a colocação. Em contrapartida, não será considerada uma colocação relevante aquela cujo cálculo estatístico seja mais próximo de 0. Quanto mais próximo do 0, menos exclusiva é a colocação.

Dessa forma, organizamos em 4 níveis o grau de exclusividade das colocações, de acordo com o resultado do cálculo estatístico, partindo do pressuposto de que os dois primeiros níveis representam a primeira metade do valor máximo e os dois outros níveis, a segunda metade desse mesmo valor. O quadro com os níveis pode ser visualizado a seguir:

Quadro 3 – Níveis de exclusividade

Valor do <i>Log Dice</i> (0 - 14)	Grau de exclusividade da Colocação
0 – 4	Não exclusiva
5 – 7	Baixa exclusividade
8 – 10	Média exclusividade

Valor do <i>Log Dice</i> (0 - 14)	Grau de exclusividade da Colocação
11 – 13	Alta exclusividade
14	Exclusiva

Fonte: O autor (2023).

Especificamente, para esta pesquisa, interessam-nos as colocações de exclusividade média e alta, portanto, serão analisados os colocados cujos cálculos estatísticos se enquadrem nos critérios estabelecidos para esses rótulos, de acordo com a tabela.

Sendo assim, os colocados cujos resultados forem menores que 8 não serão listados nessa pesquisa. Uma vez que o nosso objetivo são as colocações mais relevantes, daremos prioridade para analisar as colocações cujos resultados se aproximem mais do 14. Isso significa que, se a ferramenta indicar colocações de média exclusividade e de alta exclusividade no mesmo *subcorpus*, sempre daremos prioridade de escolha para as colocações de alta exclusividade e, mesmo que todas os colocados sejam de média exclusividade ou somente de alta exclusividade, selecionaremos os que possuem maior número, por exemplo, se houver colocados de valores estatísticos 8, 9 e 10, optaremos pela seleção dos colocados de valor estatístico 10. É importante dizer isso, pois, como o cálculo é feito por meio do número de ocorrências em cada texto, é provável que os colocados, em diferentes níveis, apresentem resultados estatísticos diferentes, ou seja, nem sempre teremos colocações de média exclusividade e alta exclusividade no mesmo *subcorpus*.

Outro critério importante a ser considerado é a distribuição (posição do colocado em relação ao **nó**). Mesmo que os cálculos apresentem um número próximo ou igual a 14, é preciso, antes, checar se a presença do colocado tem relação exclusiva com o nó ou se é somente uma palavra que aparece por razões gramaticais com frequência no *corpus*, como é o caso dos artigos, preposições, conectivos *etc.* Por exemplo:

1. *He made a call. The boy talked a lot.* (o autor)

A palavra *the* aparece dentro do campo de extensão de *make* (5 para direita, 5 para esquerda), porém, *the* não agrega sentido exclusivo para *make* como *a call*, por

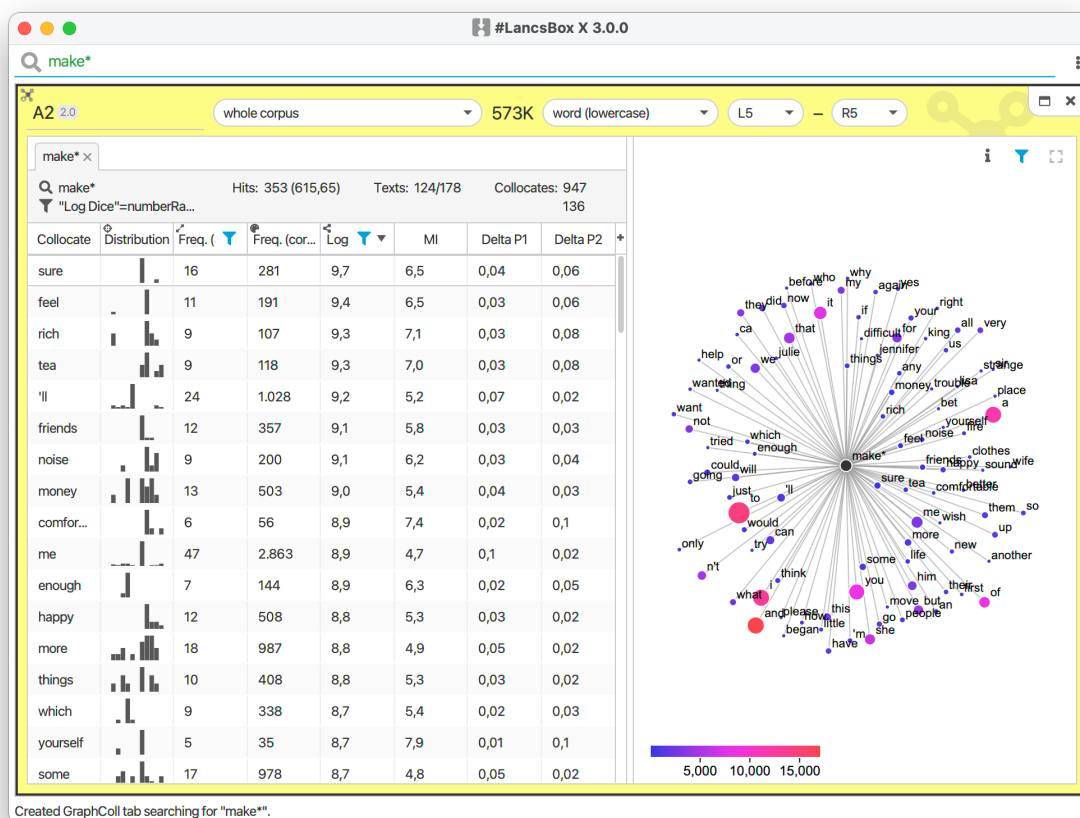
razões gramaticais e textuais. Essa palavra aparecerá como colocação de *make*, não porque é uma palavra exclusiva, mas sim porque sua frequência é grande devido ao fato de ser artigo e essa classe se constituir de palavras muito recorrentes na língua e que ocorrem com um número extenso de palavras.

Isso ocorre porque é uma palavra gramatical e geralmente não podemos juntar frases sem usá-la. Outras palavras específicas de um contexto específico (por exemplo, búzio, hashtag, corpus) serão distribuídas de forma menos uniforme. É importante notar que não existe uma única medida de dispersão, mas sim um conjunto de diferentes medidas que podem ser utilizadas para investigar a variação dentro dos corpora e para destacar diferentes aspectos da dispersão. (Brezina, 2018, p. 47).

Em relação ao significado dos colocados e a sua classificação de nível de proficiência, utilizaremos o Oxford Learner's Dictionaries ⁵. Por fim, após a apresentação dos critérios de seleção das colocações, considere a Figura 12 como exemplo:

⁵ Disponível em: <https://www.oxfordlearnersdictionaries.com>

Figura 12 — Colocados de *make* no *subcorpus* A2



Fonte: O autor (2024).

Todos os colocados mostrados na Figura 12 são de média ou alta exclusividade; ou seja, cumprem com o primeiro critério de seleção. Porém, nem todos cumprem com os outros critérios como: distribuição (posição do colocado na sentença) e palavras que, por razões gramaticais, não agregam sentido exclusivo. Sendo assim, sobram as palavras: *sure*; *feel*; *rich*; *tea*; *friends*; *noise*; *money*; *comfortable*; *happy* e *things*. Como nos interessa os colocados que se aproximam mais do 14, nota-se que os sete primeiros colocados possuem valores estatísticos de 9 ou superior, cumprindo com o critério de aproximação do número estabelecido pelo método estatístico. Matematicamente, 8 é menor que 9, portanto, de acordo com o cálculo do *Log Dice*, os colocados: *sure*; *feel*; *rich*; *tea*; *friends*; *noise* e *money* são mais exclusivos no *subcorpus* A2 do que o restante dos colocados. Dessa forma, daremos prioridade para a análise destes. Essa lógica será aplicada em todas as

pesquisas dos **nós** em todos os *subcorpora*, portanto, não será necessária a descrição dessa etapa no capítulo seguinte.

Este capítulo descreveu o *corpus* de estudo e os procedimentos metodológicos empregados na presente pesquisa. Foram abordados o *software* escolhido para a análise do *corpus*, a ferramenta que será utilizada para identificação das colocações dos verbos *make* e *take*, o método estatístico utilizado e os critérios para selecionar as colocações relevantes. O próximo capítulo abordará a análise e a discussão dos dados.

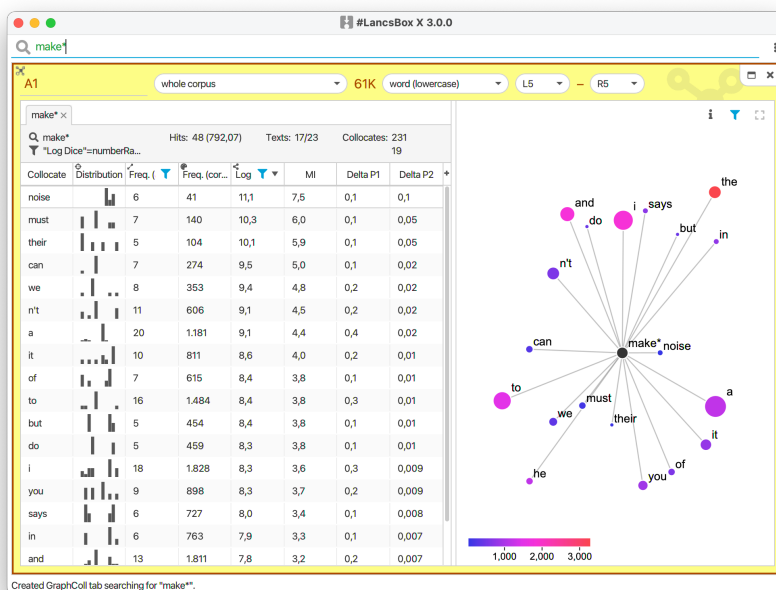
5 DESCRIÇÃO E DISCUSSÃO DA ANÁLISE DOS DADOS

A presente seção tem por desígnio expor o processo de identificação das colocações, seguido pela descrição, classificação e análise dessas estruturas linguísticas. Conforme antecipado na introdução deste estudo, o propósito central consistiu em discernir as colocações mais distintivas associadas aos verbos *make* e *take* em cada um dos *subcorpora* selecionados. O início desta abordagem contemplará as colocações do verbo *make* nos *subcorpora* A1, A2, B1, B2, C1, C2, sendo seguido pela análise das colocações do verbo *take* nos mesmos *corpora*, culminando, assim, na conclusão deste capítulo.

5.1 Colocações com o verbo *make* no *subcorpus* a1

De acordo com as informações disponibilizadas pelo *Oxford's Learners Dictionaries*, constata-se que a palavra *make* é categorizada como parte do vocabulário de nível A1, sendo, do ponto de vista gramatical, classificada como um verbo. A fim de realizar uma análise mais aprofundada, procedemos à seleção de *make* como o **nó**, inserindo-o na barra de pesquisa da ferramenta Graphcoll no *Lancsbox*, com o intuito de obter informações pertinentes relacionadas aos seus **colocados**. Como resultado dessa pesquisa, foram obtidos os seguintes conjuntos de dados mostrados na Figura 13.

Figura 13 — Make - A1



Fonte: O autor (2024).

O referido **nó** ocorre em 48 instâncias no *subcorpus* A1, distribuídas ao longo de 17 de 23 textos, totalizando 231 **colocados**. Após a aplicação do *Collocation Parameters Notation* (CPN), conforme proposto por Brezina, McEnery e Wattam (2015), restaram 19 **colocados**. Em consonância com os critérios anteriormente delineados no Capítulo 3, abrangendo aspectos como distribuição e razões gramaticais, os termos: *their*; *must*; *can*; *we*; *a*; *not*; *but*; *do*; *of*; *to* e *it* foram excluídos como colocações relevantes neste estudo. Dado que nosso intento consiste em identificar as colocações mais exclusivas do verbo em questão, inferimos, com base nos resultados obtidos por meio de cálculos estatísticos e à luz de nossos critérios de seleção, que a colocação *noise* ostenta a máxima exclusividade dentre as 19 ocorrências remanescentes.

No Quadro 4, são mostrados os respectivos valores de co-ocorrência com o **nó**, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para essa pesquisa:

Quadro 4 — Quadro de colocados mais exclusivos

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>noise</i>	6	41	11,1

Fonte: O autor (2024).

5.1.1 Noise

O termo *noise* apresenta uma frequência total de 41 ocorrências no *corpus* e coocorre com o **nó** 6 vezes. O cálculo do índice *Log Dice*, aplicado de acordo com a fórmula estabelecida, resultou em 11,1. Em virtude de o valor de referência preestabelecido no método estatístico ser 14, inferimos que *noise* configura-se como uma colocação relevante, dada sua aproximação ao valor estipulado pela mencionada abordagem estatística. Essa inferência pode, ademais, ser corroborada pela análise do gráfico representado na Figura 3.

Conforme já abordado no Capítulo 3, a proximidade do termo colocado em relação ao **nó** denota uma associação mais robusta entre ambos. Dessa forma, a conclusão acerca da relevância de *noise* como colocação significativa é sustentada tanto pelos resultados quantitativos obtidos no cálculo do índice *Log Dice* quanto pela visualização gráfica apresentada na Figura 3. Além disso, é importante ressaltar que essa argumentação se estende aos próximos colocados analisados neste capítulo. Portanto, a análise quantitativa por meio do cálculo do índice *Log Dice* e a visualização gráfica continuarão a respaldar a relevância dos termos subsequentes da mesma maneira. Conforme as informações fornecidas pelo *Oxford's Learner's Dictionaries*, a palavra *noise* é categorizada como parte do vocabulário de nível A2, sendo gramaticalmente classificada como um substantivo. Com base na etiquetagem do *corpus* em questão, todas as ocorrências do termo colocado foram categorizadas como NN (substantivo, singular).

A fim de explorar mais detalhadamente as estruturas formadas nas seis ocorrências de *make* em conjunto com o **nó** em consideração, realizamos uma análise da etiquetagem, identificando as configurações estruturais presentes. Essas estruturas estão detalhadas no Quadro 5:

Quadro 5 — Estruturas de *Make Noise* - A1

Estruturas
VBZ DT NN
VB DT NN
VBP DT JJ NN
VB DT NN IN NN

Fonte: O autor (2024).

No contexto das obras em que a palavra *noise* emerge como uma colocação vinculada ao verbo *make*, a ferramenta de etiquetagem, *SpaCy*, procedeu à classificação lexical, atribuindo à palavra *make* as categorias gramaticais VBZ (verbo na terceira pessoa do singular), VB (verbo base), e VBP (verbo na primeira pessoa do singular ou plural), enquanto a palavra *noise* foi classificada como NN (substantivo). Consequentemente, simplificando a análise, a combinação *make noise* é constituída por um verbo e um substantivo. É digno de nota que, em todas as sentenças analisadas, foi identificada a presença do determinante (DT) e, em uma delas, observou-se a inclusão de um adjetivo (JJ). Essa constatação sugere que, quando *noise* figura como uma colocação de *make*, a estrutura gramatical permite a co-ocorrência com outros elementos, tais como determinantes (*a, any, some*), assim como a possibilidade de associação com adjetivos, exemplificada pelo termo *lot*, como ilustrado nas linhas de concordância extraídas do *subcorpus* A1, conforme apresentado na Figura 14:

Figura 14 — Linhas de Concordância de *make noise* - A1

N°	Left	Node	Right
1	goes past and makes a	noise. '	Keyaaaa', it cries. Jack smiles
2	They don't make any	noise.	Peter doesn't know where
3	Their swords make a loud	noise.	Suddenly Peter takes Hook's
4	and don't make a	noise.	I want to speak to
5	and make a lot of	noise.	The two men are good
6	will make a lot of	noise	and the men will hear.

Fonte: O autor (2024).

Estes dados podem revelar-se instrumentalmente pertinentes no contexto do ensino da Língua Inglesa, ao oferecerem uma demonstração explícita aos alunos acerca da versatilidade inerente aos determinantes e adjetivos na construção *make noise*.

No que concerne à semântica associada, *Oxford's Learner's Dictionaries* esclarece a existência de expressões idiomáticas relacionadas a *make noise*, as quais podem encapsular conotações de uma personalidade proeminente, manifestações de descontentamento, ou ainda a comunicação indireta de uma ideia. Contudo, após minuciosa análise das sentenças presentes nos *Graded Readers*, constatou-se que todas as ocorrências não se alinham às expressões idiomáticas descritas no referido dicionário, como podemos observar nos excertos a seguir:

Excerto A:

Then they take them to the Jolly Roger. They don't make any noise. Peter doesn't know where they are. He is sad without Wendy, John, Michael, and the Lost Boys. He sits and thinks.

Fonte: Peter Pan, JM Barrie - Nível A.

Excerto B:

Their swords move quickly and make a lot of noise. The two men are good fighters. Zorro is fast and light. Sergeant Gonzales is slow and heavy. Zorro jumps onto a table. Then he jumps onto a chair.

Fonte: Zorro, Sally M - Nível A1.

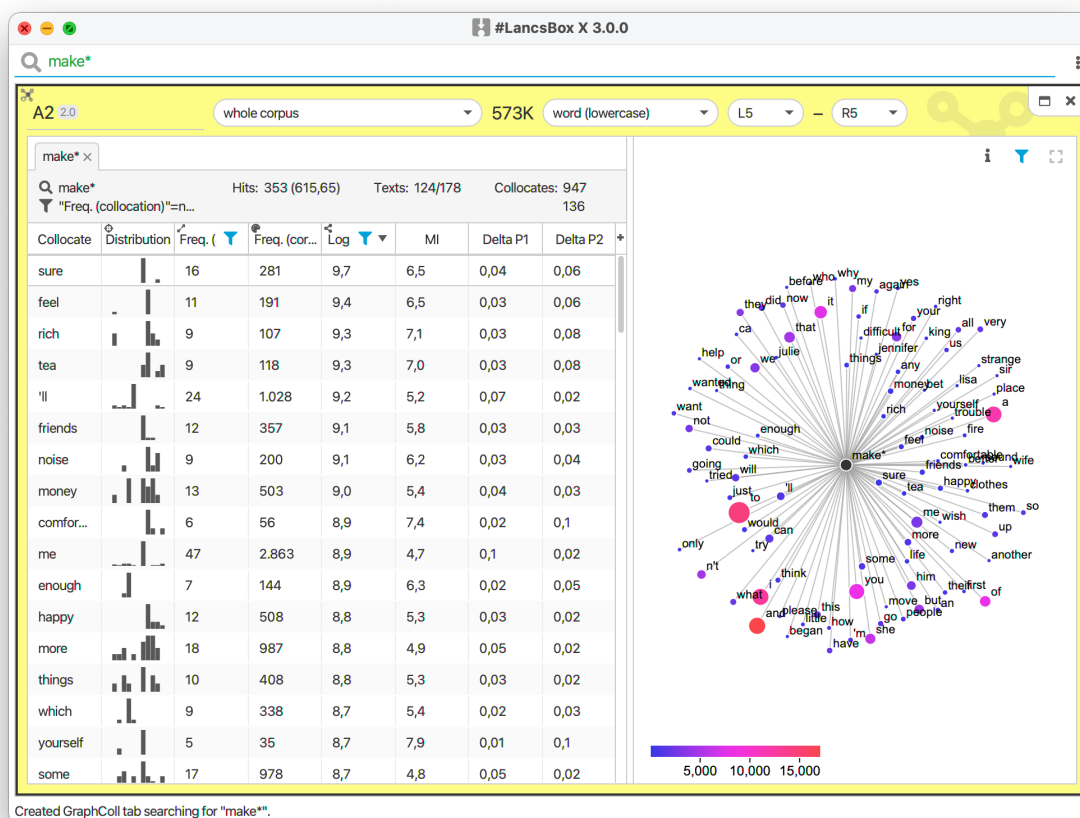
No *subcorpus* A1, *noise* como uma colocação de *make* apresenta um significado estritamente literal, denotando a ação de provocar ou produzir som em todas as instâncias observadas.

Esses dados são importantes para os alunos por uma série de razões. Em primeiro lugar, destacam a importância da compreensão contextual e do significado específico de colocações em diferentes níveis de proficiência linguística. Essa descoberta permite aos alunos reconhecerem a variação na utilização de expressões idiomáticas, mostrando que nem sempre as colocações mantêm significados figurativos ou adicionais. Isso pode sensibilizá-los para a diversidade de usos das colocações em diferentes contextos, estimulando uma abordagem mais crítica na interpretação do significado de expressões idiomáticas. Além disso, essa constatação ressalta a importância de se basear em fontes diversas, como o contexto específico dos *readers*, para compreender plenamente o uso autêntico das colocações. Dessa forma, um *corpus* composto desse material pode fornecer subsídios pedagógicos para serem explorados pelos professores para aprimorar sua compreensão da língua e evitar generalizações inadequadas.

5.2 Colocações com o verbo *make* no *subcorpus* a2

Repetindo o mesmo processo anterior, porém no *subcorpus* A2, digitamos na barra de pesquisa da ferramenta *Graphcoll* o termo *make* para obter os seus **colocados**. Obtivemos os seguintes dados:

Figura 15 — Make - A2



Fonte: O autor (2024).

O **nó** em análise manifesta-se em 353 instâncias no *subcorpus* A2, abrangendo 124 dos 178 textos e apresentando um total de 947 **colocados**. Após a aplicação do CPN, conforme proposto por Brezina, McEnery e Wattam (2015), resultou na retenção de 136 **colocados**. Alinhado aos critérios e objetivos delineados no capítulo 3, identificamos 7 **colocações** que se destacam como mais exclusivas do verbo *make* no *subcorpus* A2, sendo: *sure*; *feel*; *rich*; *tea*; *friends*; *noise* e *money*.

No Quadro 6, são mostrados os respectivos valores de co-ocorrência com o **nó**, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para essa pesquisa:

Quadro 6 — Colocados mais exclusivos de *Make* - A2

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Sure</i>	16	281	9,7
<i>Feel</i>	11	191	9,4
<i>Rich</i>	9	107	9,3
<i>Tea</i>	9	118	9,3
<i>friends</i>	12	357	9,1
<i>noise</i>	9	200	9,1
<i>money</i>	13	503	9,0

Fonte: O autor (2024).

5.2.1 *Sure*

Conforme as informações fornecidas pelo *Oxford's Learner's Dictionaries*, a palavra *noise* é categorizada como parte do vocabulário de nível A1, sendo gramaticalmente classificada como um adjetivo. De acordo com a etiquetagem do *corpus*, todas as ocorrências do termo colocado foram categorizadas como JJ (adjetivo). Procedemos à análise da etiquetagem, examinando todas as estruturas distintas formadas nas 16 ocorrências de *make* com o **nó** em questão. Essas estruturas estão detalhadas no Quadro 7:

Quadro 7 — Estruturas de *make sure* - A2

ESTRUTURAS
VB JJ
VBZ JJ

Fonte: O autor (2024).

Conseqüentemente, nas composições em que *sure* foi identificado como uma colocação associada ao verbo *make*, o *SpaCy* categorizou a palavra *make* como VBZ (verbo na terceira pessoa do singular), VB (verbo base), e *noise* como JJ (adjetivo). Desse modo, de maneira simplificada, a construção *make sure* é composta, em todas

as instâncias, por um verbo e um adjetivo, como pode ser observado nas linhas de concordância na Figura 18, retiradas do *subcorpus A2*:

Figura 16 — Linhas de Concordância de *make sure* - A2

Nº	Left	Node	Right
1	a few minutes, mind! Make	sure	ye stay awake the rest
2	Marishka's blanket, to make	sure	that the ducat was still
3	's why my mother makes	sure	I come to town every
4	and he wanted to make	sure	he would have no more
5	for him. Dunning would make	sure	that his name was not
6	or buses." I'll make	sure	she stays in town - but
7	back several times to make	sure	the picture was really straight.
8	government. We want to make	sure	nothing has happened to her. "
9	still need her to make	sure	her daughter cooperates." Lou tried
10	go back upstairs to make	sure	that he had not left
11	'll fit.' 'We should make	sure,'	he said. Why didn't
12	the heater. 'Just to make	sure	the buttons are right.' She
13	no answer. Just to make	sure,	she rang again, and she
14	I'm going to make	sure	that nothing bad ever happens

Fonte: O autor (2024).

Observamos que, em todas as sentenças analisadas, não se registra a presença de outros elementos gramaticais entre o verbo e o adjetivo, isto é, todas as sentenças são formadas exclusivamente por *make sure*, diferente de outros colocados já vistos neste capítulo que apresentam, como, por exemplo, DT (determinante) entre o verbo e adjetivo, como em *make a noise*. Este tipo de informação pode promover uma análise crítica dos dados linguísticos e suscitar interesse por parte dos alunos, especialmente quando contrastado com o fenômeno observado em relação à colocação *noise*. Sendo assim, os alunos podem ser desafiados a formular hipóteses sobre por que certos elementos gramaticais aparecem em algumas colocações e não em outras, incentivando um pensamento analítico e reflexivo sobre a língua. Ao confrontar esse fenômeno com o observado em relação à colocação *make noise*

(*subcorpus* A1), os alunos desenvolvem habilidades de observação e comparação, fundamentais para a compreensão detalhada da língua e aquisição de novos conhecimentos linguísticos.

Em relação ao significado, o *Oxford's Learner's Dictionaries* informa que existem expressões idiomáticas com *make sure*, classificadas como pertencentes ao nível A2. Esse dado indica a pertinência da ocorrência de *make sure* no *subcorpus* A2, pois em ambos os níveis, o indicado pelo dicionário e o indicado pelo *corpus*, os significados são os mesmos. Ademais, constatamos que *make sure* não se manifesta no A1, reforçando, até então, a coerência entre os *Graded Readers* e seus respectivos níveis de proficiência.

Os significados atribuídos a *make sure*, segundo o dicionário, abrangem a ação de realizar algo para garantir que outra coisa ocorra e a verificação da veracidade ou realização de algo. Ao examinar as sentenças presentes nos *Graded Readers*, constatamos que todas as ocorrências configuram expressões idiomáticas conforme descritas no dicionário, como exemplificado nos excertos A, B e C retirados do *subcorpus* A2:

Excerto A:

Dunning would wait for the train at Croydon West where Harrington would look out for him. Dunning would make sure that his name was not on his luggage and, most importantly, must have the paper with him. On Thursday night, Dunning waited impatiently for the train at Croydon West.
 Fonte: *Casting the Runes*, MR James.

Excerto B:

"No, it isn't." "She must stay in town when she leaves the hospital. No long journeys on trains or buses." "I'll make sure she stays in town - but can I see her now, please?" "Yes, yes, she's waiting for you." Kathy smiled shyly at William.
 Fonte: *Hampton House*, Jenny Dooley.

Excerto C:

One of the pictures on the wall was not hanging straight. He noticed it and went to fix it. He stepped back several times to make sure the picture was straight. Then he gave it a gentle touch with his hand.
 Fonte: *The Californians Tale*, MarkTwain).

Nesse contexto, podemos inferir a consolidação da utilização idiomática de *make sure* nas narrativas analisadas, corroborando a aplicabilidade e a coerência semântica da expressão nos contextos propostos. Ademais, a coerência semântica observada reforça a ideia de que algumas palavras e expressões em inglês têm significados específicos e previsíveis em determinados contextos. Essa consolidação contribui para a construção de um repertório mais robusto e preciso de vocabulário.

5.2.2 Feel

Segundo o *Oxford's Learner's Dictionaries*, a palavra *feel* é categorizada como parte do vocabulário de nível A1, sendo gramaticalmente classificada como um verbo. A etiquetagem do *corpus* revela que todas as ocorrências do termo colocado foram classificadas como VB (verbo, na forma base). Dessa forma, verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 10 ocorrências de *make* com o **nó**. As estruturas podem ser observadas no Quadro 8:

Quadro 8 — Estruturas de *make feel* - A2

ESTRUTURAS
VB PRP VB
VBP PRP VB
VBZ PRP VB

Fonte: O autor (2024).

Nas obras em que *feel* ocorreu como uma colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural) e VBZ (verbo na terceira pessoa do singular), enquanto *feel* foi classificado como VB (verbo base). Dessa forma, simplificando, a construção *make feel* é composta, em todas as ocorrências, por dois verbos.

Observamos que, em todas as sentenças analisadas, existe a presença de outro elemento gramatical entre os dois verbos, o PRP (pronome pessoal), como *me*, *you*, *him* e *themselves*, sendo que, em relação aos outros, o pronome *me* detém 40% das ocorrências. Dessa forma, a existência do pronome pessoal entre os verbos indica que, em todas as ocorrências, a construção analisada tem a conotação de induzir

alguém a sentir alguma coisa ou a se sentir de algum modo, conforme pode ser observado nos excertos A e B dos *readers* extraídos do *subcorpus A2*:

Excerto A:

'Are you feeling all right, Dunn? Have some more whisky.' He opened the other bottle and put some whisky in Dunn's glass. Dunn drank it, and it seemed to make him feel better. *'That was strange,' he said.*

Fonte: The Railway Crossing, John Escott.

Excerto B:

What's happened?" "Too much wine!" "Quiet, Nathan! Luke, you don't look too well. Are you sure you're all right? Here, I have some beef curry for you over there - that should make you feel better." But Luke didn't eat the beef curry, or any of the barbecue food.

Fonte: Lord McDonald Eamonn, Sweeney.

Em conclusão, a presença sistemática do pronome pessoal pode reforçar a conotação de influenciar emocionalmente algum indivíduo, conforme evidenciado nos trechos apresentados. Essa descoberta da presença sistemática do pronome pessoal (PRP) entre os dois verbos nas construções *make feel* pode ser importante para os alunos por várias razões. Primeiramente, ao observar a consistência do pronome pessoal entre os verbos, os alunos podem conseguir identificar padrões específicos na estrutura das frases. Isso possivelmente contribui para uma melhor compreensão da construção verbal específica, nesse caso, *make + PRP + feel*. Segundamente, a presença do pronome pessoal entre os verbos sugere uma relação direta com a influência emocional sobre um indivíduo. Isso pode auxiliar os alunos a interpretarem o significado da expressão *make + PRP + feel* como algo que induz ou causa uma resposta emocional específica em alguém. Além disso, os exemplos fornecidos no *subcorpus A2* ilustram como a construção é aplicada em contextos reais.

5.2.3 Rich

O *Oxford's Learner's Dictionaries* informa que a palavra *rich* é categorizada como parte do vocabulário de nível A1, sendo gramaticalmente classificada como um adjetivo. A etiquetagem do *corpus* revela que todas as ocorrências do termo em questão foram classificadas como JJ (adjetivo). Procedemos, então, à análise das

diversas estruturas formadas nas 7 ocorrências de *make* com o **nó** em questão, com base na etiquetagem. Essas estruturas estão detalhadas no Quadro 9:

Quadro 9 — Estruturas de *make rich* - A2

Estruturas
VB PRP JJ
VB PRP DT JJ
VB PRP RB RB JJ

Fonte: O autor (2024).

Portanto, nas obras em que *rich* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), e *rich* em JJ (adjetivo). Dessa forma, simplificando, a construção, *make rich* é composta, em todas as ocorrências, por um verbo e um adjetivo, além de apresentar outros elementos gramaticais, tais como PRP (pronome pessoal), DT (determinante) e RB (advérbio). Isso revela que, nesse *subcorpus*, a expressão *make rich* está associada a tornar alguém rico, exigindo, portanto, em todos os casos, a presença de um pronome, como podemos ver nos excertos A e B retirados dos *readers* de nível A2:

Excerto A:

I will throw the jar back into the sea. Perhaps somebody will find you in another 300 years.' 'Please open the jar,' called the Giant. 'I will make you a rich man.' 'No. You will kill me,' answered the fisherman.

Fonte: The Fisherman and the Giant, Jane Rollason.

Excerto B:

When I was young, I went looking for gold in California. I never found enough to make me rich, but I did discover a beautiful part of the country. It was called the Stanislaus. The Stanislaus was like heaven on earth.

Fonte: The Californians Tale, Mark Twain.

Na Figura 17, são mostradas as linhas de concordância extraídas do *subcorpus* A2:

Figura 17 — Linhas de Concordância de *make rich* - A2

Nº	Left	Node	Right
1	Now I will make them	rich,	but I will not make
2	I will make you a	rich	man.' 'No. You will kill
3	'll make you very, very	rich.'	The fisherman thought for a
4	it can make me a	rich	man!' 'Oh! You're not
5	found enough to make me	rich,	but I did discover a
6	're going to make us	rich,	Ray,' Charlie said. Raymond looked
7	on someone or make them	rich	to love you. You understand

Fonte: O autor (2024).

Observamos que as ocorrências de *rich* como colocado de *make* estão associadas a diferentes tempos verbais, incluindo o presente simples na linha 5, "*it can make me a rich man*", o futuro com *Will* na linha 2, "*I will make you a rich man.*", e o futuro com *Going To*, na linha 6 "*'re going to make us rich, Ray,' Charlie said.*". Além disso, há exemplos no passado simples, como na linha 5, "*found enough to make me rich.*" Em um *corpus* constituído por *readers* de nível A2, a presença de construções verbais, em diferentes tempos, pode ser um possível indicador das diferenças entre o nível A2 e o A1.

No que diz respeito ao seu significado, o *Oxford's Learner's Dictionaries* informa que a construção de *make something*, no sentido de realizar uma ação para fazer com que algo exista ou aconteça, como exemplificado em *make someone rich*, é classificada como nível A1. Esse dado indica que *make someone rich* está coeso⁶ ao ocorrer no *subcorpus* A2, uma vez que a expressão está um nível abaixo do nível atribuído ao *subcorpus* em questão. No entanto, ao examinarmos o *subcorpus* A1, não se observa a presença dessa construção. Até o momento, essa observação sugere que construções de determinados níveis, como A1, podem ser encontradas em *subcorpora* de níveis superiores, mas não necessariamente se manifestam em seu nível correspondente, como exemplificado pela ocorrência de *make someone rich*, classificada como de nível A1 e encontrada no *subcorpus* A2.

⁶ É importante destacar que coeso se refere à um colocado ocorrer em um *subcorpus* que tenha o seu mesmo nível de proficiência de acordo com o dicionário. Um colocado não coeso seria aquele que ocorresse em um *subcorpus* superior ao seu nível de proficiência.

5.2.4 Tea

De acordo com informações do *Oxford's Learner's Dictionaries*, a palavra *tea* é categorizada como parte do vocabulário de nível A1, sendo gramaticalmente classificada como um substantivo. A etiquetagem do *corpus* revela que todas as ocorrências do termo colocado foram classificadas como NN (substantivo, singular). Procedemos, então, à análise das diversas estruturas formadas nas 9 ocorrências de *make* com o **nó** em questão, com base na etiquetagem. Essas estruturas estão detalhadas no Quadro 10:

Quadro 10 — Estruturas de *make tea* - A2

ESTRUTURAS
VB PRP DT NN IN NN
VBZ DT NN IN NN
VB DT NN
VBP NN
VB NN

Fonte: O autor (2024).

Portanto, nas obras em que *tea* ocorreu como colocação do nó, O *SpaCy* classificou a palavra *make* em VB (verbo base), VBZ (verbo na terceira pessoa do singular) e VBP (verbo na primeira pessoa do singular ou plural), enquanto *tea* foi classificado como NN (substantivo). Dessa forma, simplificando, a construção *make tea* é composta, em todas as ocorrências, por um verbo e um substantivo, além de apresentar outros elementos gramaticais, tais como PRP (pronome pessoal), IN (preposição) e DT (determinante), como evidenciado nas linhas de concordância extraídas do *subcorpus* A2 na Figura 18:

Figura 18 — Linhas de Concordância de *make tea* - A2

N°	Left	Node	Right
1	make her a cup of	tea.	When I got back, she
2	He makes a cup of	tea	and goes into his study.
3	and I'll make some	tea.'	To tell you the truth,
4	the early birds. We make	tea,	then I help her with
5	said. 'I'll make the	tea	first.' 'After,' he said. She
6	the kitchen to make the	tea.	He was very tired the
7	morning at five to make	tea	for you and breakfast. Now

Fonte: O autor (2024).

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que a construção de *make something*, no sentido de criar ou preparar algo, como exemplificado em *make tea*, é classificada como nível A1. Essa informação indica que *make tea* está em conformidade por ter ocorrido no *subcorpus* de A2, uma vez que a expressão está um nível abaixo do nível atribuído ao *subcorpus* em questão. Esta correlação é coincidente com a observação relacionada à colocação *rich*. Dessa maneira, podemos inferir que a distribuição de colocações nas obras analisadas segue padrões de conformidade hierárquica em relação aos níveis de proficiência designados, corroborando a consistência metodológica na análise de construções colacionais em contextos linguísticos específicos.

A descoberta pode ser eficaz para o aprendizado, mostrando aos alunos que a língua é dinâmica e adaptável. Eles podem aprender que certas expressões podem transcender seus níveis iniciais e serem utilizadas em uma variedade de contextos, o que enfatiza a necessidade de compreender não apenas as palavras isoladamente, mas também em seus usos práticos. Ademais, os alunos podem incorporar construções mais simples, como *make tea*, em seus repertórios linguísticos cotidianos, percebendo que expressões, aparentemente básicas, são úteis e aplicáveis, em diversos ambientes comunicativos.

5.2.5 Friends

Conforme informações do *Oxford's Learner's Dictionaries*, a palavra *friends* é categorizada como parte do vocabulário de nível A1, sendo gramaticalmente classificada como um substantivo. A etiquetagem do *corpus* revela que todas as ocorrências do termo colocado foram classificadas como NNS (substantivo, plural). Procedemos, então, à análise das diversas estruturas formadas nas 12 ocorrências de *make* com o **nó** em questão, com base na etiquetagem. Essas estruturas estão detalhadas no Quadro 11:

Quadro 11 — Estruturas *make friends* - A2

ESTRUTURAS
VB NNS
VB DT JJ NNS
VB JJ NNS
VBZ NNS

Fonte: O autor (2024).

Portanto, nas obras em que *friends* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base) e VBZ (verbo na terceira pessoa do singular), enquanto *friends* foi classificado como NNS (substantivo plural). Dessa forma, simplificando, a construção *make friends* é composta, em todas as ocorrências, por um verbo e um substantivo, além de apresentar outros elementos gramaticais, como DT (determinante) e JJ (adjetivo), como evidenciado nas linhas de concordância extraídas do *subcorpus* A2 na Figura 19.

Figura 19 — Linhas de Concordância de *make friends* - A2

Nº	Left	Node	Right
1	clever. She can't make	friends	with a giraffe with twigs
2	think she wants to make	friends	with you. Drop the rope
3	We can make some new	friends	and have a good time
4	and I can make new	friends,"	says Julie. "That sounds fun.
5	I just wanted to make	friends,	I wanted people to like
6	Julie says, trying to make	friends	with Anna. But Anna does
7	start. How can I make	friends	when my roommates all think
8	it. Julie wants to make	friends.	She sees the purse on
9	Why can't I make	friends?	Anna hates me because of
10	'm not surprised, he makes	friends	easily. Everyone likes him, he
11	behind. Julie tries to make	friends	with Anna again. She stops
12	school life and to make	friends.	His life at home, you

Fonte: O autor (2024).

Na análise estrutural, foi constatado que 66,6% das sentenças contendo a expressão *make friends* se enquadram na estrutura de Verbo + Substantivo no plural (VB + NNS), evidenciando a simplicidade gramatical dessas construções. Tal observação se mostra pertinente, considerando que essas sentenças são extraídas dos *readers* classificados no nível A2. Os exemplos fornecidos, extraídos desses *readers*, ilustram que as sentenças nas quais a construção *make friends* é empregada apresentam uma predominância de tempos verbais simples, notadamente no presente e no passado. Essa prevalência da estrutura VB + NNS é atribuída ao fato de essas sentenças funcionarem como complementos de outros verbos, como exemplificado por *began* no excerto A e *wants* no excerto B, abaixo:

Excerto A:

'On the 1st of May young Lord Saltire arrived. He's a nice boy, and he soon began to like school life and to make friends. His life at home, you see, was not very happy - we all know about the Duke and his wife. The Duchess, of course, now lives in the south of France.

Fonte: Sherlock Holmes and the Dukes Son, Arthur Conan Doyle.

Excerto B:

"You can't be clean in the mud! You must be dirty! It's normal." At that moment Spotty's mum stops turning her neck. She starts walking to Bill. 'Bill!' I cry, 'the mother is going to you! I don't think she wants to make friends with you. Drop the rope and get out!' Bill looks at Spotty's mum. He is afraid. He drops the rope and wants to run, but he can't.

Fonte: The Adventures in The Grasslands, John Bookworm.

Excerto C:

Thanks, Anna, for taking the top bunk. You're very kind," Julie says, trying to make friends with Anna. But Anna doesn't reply. "What a stupid thing to tell them, that I'm afraid of heights!" thinks Julie to herself. "This is not a good start. How can I make friends when my roommates all think I'm a scaredy cat?"

Fonte: Mystery of Coniston Water, TerryTomscha.

Essa análise ressalta não apenas a simplicidade gramatical das construções, mas também a sua coerência com o nível de proficiência A2, enfatizando a compreensão dos leitores em relação à utilização de estruturas verbais elementares e a funcionalidade da expressão *make friends* como complemento verbal. Por fim, em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que a construção de *make something*, no sentido de criar ou preparar algo, exemplificada em *make friends* (criar uma amizade, fazer uma amizade), é classificada como nível A1. Esse dado indica que *make friends* está em conformidade por ter ocorrido no *subcorpus* de A2, uma vez que a expressão está um nível abaixo do nível atribuído ao *subcorpus* em questão, seguindo a mesma tendência observada com *rich* e *tea*.

5.2.6 Noise

Dado que a ocorrência em questão foi registrada no nível A1, a classificação lexical correspondente já foi explicitada e mantém sua vigência no nível A2. Sendo *noise*, NN (substantivo, singular), classificado como nível A2. Procedemos, então, à análise das diversas estruturas formadas nas 9 ocorrências de *make* com o **nó** em questão, com base na etiquetagem. Essas estruturas estão detalhadas no Quadro 12:

Quadro 12 — Estruturas *make noise* - A2

ESTRUTURAS
VBZ DT NN

ESTRUTURAS
VBP JJR NN
VB DT NN
VB DT NN IN NN
VB RB JJ NN
VBP DT NN IN NN
NN DT NN VB

Fonte: O autor (2024).

Dessa maneira, nas obras em que *noise* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), VBZ (verbo na terceira pessoa do singular), VBP (verbo na primeira pessoa do singular ou plural), e *noise* em NN (substantivo). Sendo assim, simplificando, *make noise* é constituído, em todas as ocorrências, por um verbo e um substantivo. Além disso, há outros elementos gramaticais que compõem a construção, como DT (determinante), JJ (adjetivo), RB (advérbio) e JJR (adjetivo comparativo), como visto nas linhas de concordância retiradas do *subcorpus A2*, na Figura 20:

Figura 20 — Linhas de Concordância de *make noise* - A2

N°	Left	Node	Right
1	's there.' 'He makes no	noise,	but he's there,' replied
2	engine either. They make more	noise.	He did not know what
3	them, trying to make no	noise,	and keeping close to the
4	and make a lot of	noise.	This time they must send
5	going to make as much	noise	as that? You are not
6	said, "Don't make a	noise!	But I did, and he
7	They make a lot of	noise.	Two miners find David and
8	motorcycles make a lot of	noise.	The race is starting,' David
9	remember something. 'What about the	noise	the cow makes when she

Fonte: O autor (2024).

A análise das estruturas revela que 66,6% das estruturas incorporam um determinante, sendo a configuração mais predominante aquela constituída por um verbo seguido por um determinante e um substantivo. As demais estruturas, desprovidas de determinante, são caracterizadas pela presença de advérbios ou adjetivos. Esse padrão evidencia que a maioria das sentenças, que incluem a expressão *make noise* no nível A2, apresenta elementos gramaticais mais simplificados, notadamente com a utilização de determinantes, como visto nos excertos A, B e C, extraídos dos *readers* de nível A2:

Excerto A:

'He's in the next room, madam. He was in here.' 'But I can't hear him! I don't think he's there.' 'He makes no noise, but he's there,' replied the girl. Suddenly Baptista understood what the girl meant, and a cold hand lay on her heart.

Fonte: A Moment of Madness, Thomas Hardy.

Excerto B:

"What happened?" I asked. "I was in my car," she said. "A man got in and put his hand over my mouth." He said, "Don't make a noise! But I did, and he cut me with his knife." I held Lisa in my arms. "It's going to be all right," I said.

Fonte: Marley and Me, John Grogan.

Excerto C:

Don't move!' The miners run, and the Indians run, too. They make a lot of noise. Two miners find David and catch him. 'They're going to kill me,' David thinks. But then two policemen run to him. 'Stop!' they say.

Fonte: Amazon Rally, Amos Eduardo.

Observamos também que na linha de concordância 9 há uma construção diferente das demais. O substantivo ocorre primeiro na sentença, e o verbo ocorre depois. Ao comparar a sentença da linha 9 com as demais, percebe-se que as sentenças compartilham a intenção de descrever a ação de criar barulho, mas a diferença na posição do verbo e do substantivo destaca variações gramaticais e contextuais na forma como essa ação é expressa. Nas primeiras 6 sentenças, a ênfase está na produção geral de barulho, enquanto na última sentença, há uma ênfase específica no barulho produzido pela vaca. Por fim, em relação ao seu significado, assim como descrito na seção 4.1.1, o termo *noise*, nesse *subcorpus*, mantém o significado literal, referente a fazer barulho. A consistência no significado

pode facilitar a compreensão e a internalização desse termo específico, permitindo aos alunos uma aprendizagem mais eficaz por meio da observação direta de sua aplicação em diferentes contextos.

5.2.7 Money

No *Oxford's Learner's Dictionaries*, o termo *money* está classificado como parte do vocabulário de nível A1 e é gramaticalmente classificado como um substantivo. A etiquetagem do *corpus* revela que todas as ocorrências do termo colocado foram classificadas como NN (substantivo, singular). Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que a construção de *make money* é um *idiom* de nível A2, com o sentido de ganhar muito dinheiro, obter lucro. Esse dado indica que *make money* está coerente ao ter ocorrido no *subcorpus* de nível A2, uma vez que a expressão está no mesmo nível do *subcorpus* em questão. Esse tipo de informação é relevante para contextualizar o entendimento dos alunos sobre a aplicação de expressões idiomáticas, em diferentes níveis de proficiência linguística.

Procedemos, então, à análise das diversas estruturas formadas nas 13 ocorrências de *make* com o **nó** em questão, com base na etiquetagem. Essas estruturas estão detalhadas no Quadro 13:

Quadro 13 — Estruturas de *make money* - A2

ESTRUTURAS
VB JJR NN
VB DT NN IN NN
VB NNS IN NN
VB DT NN
VB JJ JJ NN
VBZ NN
VB NNS IN NN
NN PRP VBP

Fonte: O autor (2024).

Nas obras em que *money* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), VBZ (verbo na terceira pessoa do singular), VBP (verbo na primeira pessoa do singular ou plural), e *money* em NN (substantivo). Sendo assim, simplificando, a construção *make money* é constituída, em todas as ocorrências, por um verbo e um substantivo. Além disso, a construção apresenta outros elementos gramaticais, tais como DT (determinante), JJ (adjetivo), PRP (pronome pessoal), JJR (adjetivo comparativo), IN (preposição) e substantivo, plural (NNS), conforme evidenciado nas linhas de concordância retiradas do *subcorpus* A2 na Figura 21.

Figura 21 — Linhas Concordância de *make money* - A2

N°	Left	Node	Right
1	how he could make more	money.	Sir Daniel was a very
2	'll make a lot of	money	with this marriage,' thought Sir
3	I could make lots of	money	and..." "Please drive until I
4	magazine stories to make some	money.	Suddenly she heard a strange
5	forests to make quick, easy	money.	At last, some Brazilian politicians
6	business or farming. This makes	money	quickly for Brazil and it
7	they only want to make	money.	Some people in the town
8	and I live on the	money	I make typing,' she replied.
9	the movie but Gigli makes	money. "	Are Ben and Jennifer going

Fonte: O autor (2024).

Observamos que no exemplo da linha 8 há uma construção diferente das demais. O substantivo ocorre primeiro na sentença, e o verbo ocorre depois. As primeiras 7 sentenças enfatizam o sentido de ganhar dinheiro, expresso pela construção *make money*. A última sentença, por sua vez, destaca uma forma diferente da utilização da expressão, indicando uma fonte de subsistência relacionada à digitação. Isso mostra aos alunos como uma expressão pode ser adaptada e utilizada de maneiras diversas.

Em suma, ao analisar os colocados do nível A2, constatamos que, nas construções de *make sure*, há a ausência de elementos gramaticais entre o verbo e o

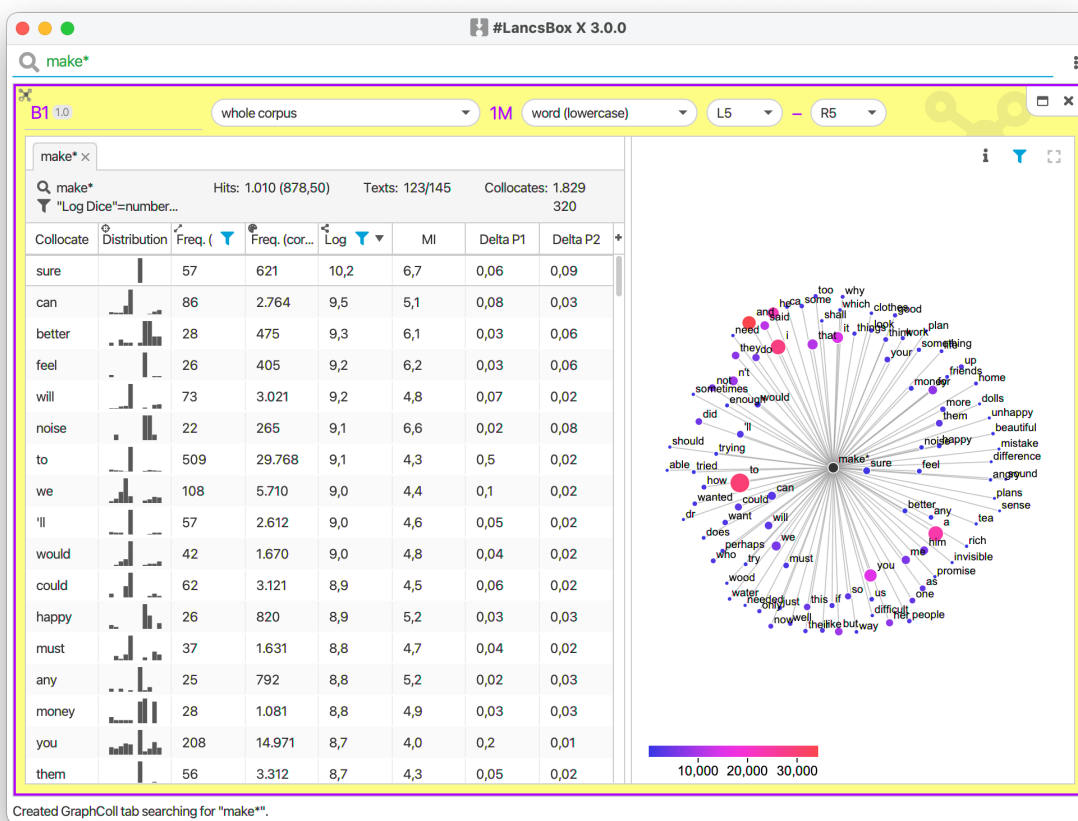
adjetivo em todas as sentenças analisadas, indicando a pertinência dessa expressão no *subcorpus* A2. A consistência de significados entre os níveis A1 e A2, conforme indicado pelo dicionário e pelo *corpus*, reforça a coerência entre os *Graded Readers* e os níveis de proficiência. Além disso, a coerência semântica observada ressalta a previsibilidade de significados específicos em contextos determinados, contribuindo para a construção de um vocabulário mais preciso.

Em suma, em relação a *make feel*, notamos a presença de elementos gramaticais, especialmente pronomes pessoais, entre os dois verbos em todas as sentenças analisadas. Isso indica a intenção dessa construção em induzir alguém a sentir algo ou a se sentir de algum modo. Já a análise das ocorrências de *rich* associadas a *make* revela uma variedade de tempos verbais, sugerindo uma complexidade temporal além das estruturas simples associadas a níveis de proficiência mais baixos. Ao examinar as construções de *make friends*, constatamos que 66,6% das sentenças seguem a estrutura de Verbo + Substantivo no plural (VB + NNS), refletindo a simplicidade gramatical dessas construções. Essa observação é particularmente relevante, considerando que essas sentenças pertencem ao nível A2. Por fim, todas as construções de *make noise* nos níveis A1 e A2 mantêm significados literais, destacando a clareza semântica nas análises. Essas observações ressaltam a simplicidade das estruturas e permitem ao estudante uma leitura fluida e uma assimilação de vocabulário compatível com seu nível linguístico.

5.3 Colocações com o verbo *make* no *subcorpus* b1

Reiterando o procedimento anterior, entretanto, agora no *subcorpus* B1, inserimos na barra de pesquisa da ferramenta *Graphcoll* a palavra *make* para extrair seus colocados. Os resultados obtidos foram os seguintes:

Figura 22 — Make - B1



Fonte: O autor (2024).

O nó em consideração é identificado em 1.010 instâncias no *subcorpus* B1, abrangendo 123 dos 145 textos e apresenta 1.829 **colocados**. Destes **colocados**, após a aplicação do CPN proposto por Brezina, McEnergy e Wattam (2015), restaram 320. Adotando nossos critérios e objetivos, identificamos apenas 1 colocação mais exclusiva do verbo *make* no *subcorpus* B1.

O Quadro 14 exibe os respectivos valores de co-ocorrência com o nó, a frequência de ocorrência no *corpus* e os resultados do cálculo do método estatístico selecionado para esta pesquisa.

Quadro 14 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Sure</i>	57	621	10,2

Fonte: O autor (2024).

5.3.1 *Sure*

Vale ressaltar que, conforme as informações fornecidas na seção anterior, de acordo com o *Oxford's Learner's Dictionaries*, a palavra *sure* é classificada como um vocabulário de nível A1 e, gramaticalmente, é categorizada como um adjetivo. De acordo com a etiquetagem do *corpus*, todas as ocorrências do termo foram classificadas como JJ (adjetivo). A análise das estruturas formadas nas 57 ocorrências de *make* com o **nó** foi realizada a partir da etiquetagem, e as diversas configurações podem ser visualizadas no Quadro 15.

Quadro 15 — Estruturas *make sure* - B1

ESTRUTURAS
TO VB JJ
VBP JJ
VB JJ
VBZ JJ

Fonte: O autor (2024).

Nas composições em que *sure* é utilizada como colocação com o verbo *make*, a classificação realizada pelo *SpaCy* identificou a palavra *make* nas formas VBZ (verbo na terceira pessoa do singular), VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural), e *sure* na forma JJ (adjetivo). Dessa forma, de maneira simplificada, *make sure* é composto, em todas as instâncias, por um verbo e um adjetivo, mantendo consistência com a estrutura observada no nível A2. Até o momento, não foram identificadas diferenças estruturais entre os níveis A2 e B1 em que a expressão *make sure* ocorre.

É relevante observar que, assim como no nível A2, em todas as sentenças analisadas não há a presença de outros elementos gramaticais após o verbo e antes

do adjetivo. Essa consistência pode ser evidenciada nas linhas de concordância apresentadas na Figura 23, retiradas do *subcorpus* B1:

Figura 23 — Linhas de Concordância de *make sure* - B1

Nº	Left	Node	Right
1	down very carefully to make	sure	I don't miss anything.
2	the aisle, checking to make	sure	the compartment doors were closed -
3	plan ready, just to make	sure.' '	The prince will think I
4	her several times to make	sure	that she was dead. 'And
5	it. I wanted to make	sure	you were OK about it.' '
6	in the dark, to make	sure	his mosquito curtains were closed. '
7	with her fingers to make	sure	it had completely gone. 'No,
8	and the engine, to make	sure	the boat could reach its
9	Maybe he wanted to make	sure	of his information. "Oh, look
10	officer, Major Strasser, to make	sure	that Victor and Ilsa boarded
11	for me. I must make	sure	that nobody knows about it!
12	stayed at school to make	sure	that Klyreg went home. When

Fonte: O autor (2024).

Quanto ao seu significado, *make sure* é classificado como um *idiom* de nível A2, conforme já mencionado anteriormente. Essa classificação sugere que a ocorrência de *make sure* no *subcorpus* B1 é coerente, uma vez que a expressão está um nível abaixo do nível atribuído ao *subcorpus* em questão.

Ao examinarmos as sentenças nos *Graded Readers*, verificamos que todas as instâncias de *make sure* são expressões idiomáticas conforme descritas no dicionário, como podem ser evidenciadas nos excertos A, B e C, retirados dos *readers*:

Excerto A:

On Monday I will take it back to the shop. JULY 25. So much happened today. I will have to write it down very carefully to make sure I don't miss anything. When I woke up I put the jacket in my school bag and told my Aunty I'm going to school.

Fonte: The Jacket, Graeme Lay.

Excerto B:

'Olmec planned for Topal to murder me, but he'll have another plan ready, just to make sure.' 'The prince will think I'll return the same way that I left, so I'll enter Tecuiltli from a lower floor,' he thought.

Fonte: Conan the Barbarian Red Nails, Howard Robert.

Excerto C:

Listen, I just want to say that I'm sorry for the other day... the kissing thing. I don't know why I did it. I wanted to make sure you were OK about it.' 'Is that all she wants to say?' William thought.

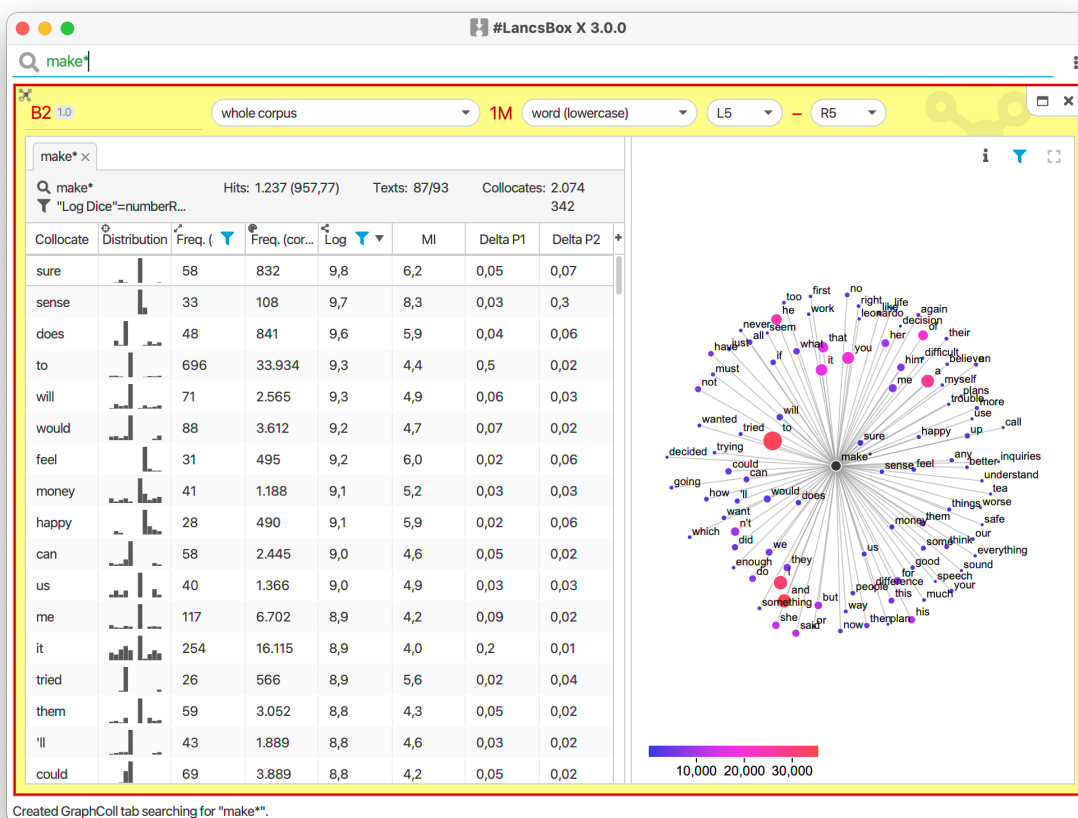
Fonte: Notting Hill, Richard Curtis.

É importante ressaltar que a expressão *make sure* geralmente é empregada para indicar a necessidade de garantir ou assegurar algo, como evidenciado nos excertos a, b e c. No entanto, sua aplicação não se estende a um sentido literal de criar algo fisicamente, ao contrário de expressões como *make a castle*, que teria um significado literal de construir ou criar um castelo. Dessa forma, ao compreendermos a distinção entre o uso idiomático e o sentido literal de expressões, reforçamos a importância da interpretação contextual no aprendizado da língua. Essa distinção sutil nos níveis intermediários pode contribuir para uma compreensão mais refinada e versátil, permitindo que os aprendizes utilizem o idioma de maneira mais precisa e adaptada aos diferentes contextos comunicativos.

5.4 Colocações com o verbo *make* no *subcorpus* b2

Repetindo o mesmo processo anterior, porém agora no *subcorpus* B2, digitamos na barra de pesquisa da ferramenta *Graphcoll* a palavra *make* para obter os seus **colocados**. Obtivemos os seguintes dados:

Figura 24 — Make - B2



Fonte: O autor (2024).

O termo em questão ocorre 1.237 vezes no *subcorpus* B2, presente em 87 dos 93 textos, e exibe 2.074 **colocados**. Após a aplicação do CPN, conforme proposto por Brezina, McEnery e Wattam (2015), restam 342 **colocados**. Nessa análise, considerando nossos critérios e objetivos, identificamos 5 **colocações** mais exclusivas do verbo *make* no *subcorpus* B2, sendo elas: *sure*; *sense*; *feel*; *money* e *happy*. O Quadro 16 apresenta os valores correspondentes de co-ocorrência com o **nó**, a frequência de ocorrência no *corpus* e os resultados dos cálculos estatísticos, fornecendo uma visão abrangente dessas colocações exclusivas e de sua relevância no contexto do *subcorpus* B2.

Quadro 16 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Sure</i>	58	832	9,8
<i>sense</i>	33	108	9,7
<i>Feel</i>	31	495	9,2
<i>money</i>	41	1.188	9,1
<i>happy</i>	28	490	9,1

Fonte: O autor (2024).

5.4.1 Sure

A análise das estruturas formadas nas 50 ocorrências de *make* com o **nó** é apresentada no Quadro 17:

Quadro 17 — Estruturas *make sure* - B2

ESTRUTURAS
VB JJ
VBZ JJ

Fonte: O autor (2024).

Nas obras em que *sure* ocorreu como colocação de *make*, o *SpaCy* classificou a palavra *make* em VBZ (verbo na terceira pessoa do singular), VB (verbo base) e *sure* em JJ (adjetivo). Dessa forma, simplificando, *make sure* é constituído, em todas as ocorrências, por um verbo e um adjetivo, seguindo a mesma estrutura identificada nos níveis A2 e B1. Até o momento, não foi observada nenhuma diferença estrutural entre o nível B2 e os níveis A2 e B1, nos quais a estrutura *make sure* também ocorre.

É relevante notar que, em todas as sentenças, assim como nos níveis A2 e B1, não há a presença de outros elementos gramaticais após o verbo e antes do adjetivo, conforme evidenciado nas linhas de concordância retiradas do *subcorpus* B2, na Figura 25:

Figura 25 — Linhas de Concordância de *make sure* - B2

Nº	Left	Node	Right
1	these. He wanted to make	sure	that everyone in the audience
2	the stairs. 'I'll make	sure	that no one follows you.'
3	away, or I shall make	sure	you go to prison again.
4	staying and I'll make	sure	he gets in touch as
5	took that photo to make	sure.'	Who is it?' 'Anna Scheele.'
6	stayed with him to make	sure	he remained the right way
7	Max. I had to make	sure	that we were not on
8	meet you. I'll make	sure	of that.' 'Good, but take
9	phone his house to make	sure.'	I called the Villa Rotha,
10	through the window to make	sure	that I was still safe,
11	to have a meeting. Make	sure	that they're all there
12	both of them to make	sure	that it didn't result

Fonte: O autor (2024).

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que há expressões idiomáticas com *make sure*, classificadas como nível A2, conforme mencionado na sessão 4.1.1. Essa classificação reforça a coerência da ocorrência de *make sure* no *subcorpus* de B2, uma vez que as palavras dos níveis anteriores são consistentes e se repetem nos níveis subsequentes.

Ao analisar as sentenças nos *Graded Readers*, destacadas nos excertos A, B e C, verificamos que todas as ocorrências são expressões idiomáticas descritas pelo dicionário, seguindo exatamente a mesma tendência observada nos níveis A2 e B1. Essa constatação fortalece a consistência do uso e significado de *make sure*, ao longo dos diferentes níveis de proficiência.

Excerto A:

But Shakespeare always uses simple language in scenes like these. He wanted to make sure that everyone in the audience understood what was happening. And so King Lear, before he dies of a broken heart, says that he will never see his daughter again: 'Never, never, never, never, never.'

Fonte: Shakespeare: His Life and Plays, Will Fowler.

Excerto B:

'Tell the world about it. I don't care. I'm not giving you any help. Now go away, or I shall make sure you go to prison again. And this time you'll never get out.' With those words, Ralph walked off into the rain.

Fonte: Nicholas Nickleby, Charles Dickens.

Excerto C:

'I couldn't believe how much you look like a woman who's been causing us a lot of trouble. I took that photo to make sure.' 'Who is it?' 'Anna Scheele.' Victoria stared at him in surprise. Whatever she had expected, it was not this. 'You mean Anna Scheele looks like me?'

Fonte: They Came to Baghdad, Agatha Christie.

A observação destaca que, em todas as sentenças analisadas, assim como nos níveis A2 e B1, não há a presença de outros elementos gramaticais entre o verbo e o adjetivo em construções com *make sure*, conforme evidenciado nos excertos a, b e c. Isso é um *insight* relevante para o entendimento da consistência sintática dessa expressão. Essa consistência gramatical sugere uma padronização na estrutura da colocação, evidenciando uma característica comum no modo como essa construção é empregada em contextos de proficiência linguística iniciais e intermediários nos *readers* analisados.

5.4.2 Sense

A palavra *sense* está categorizada como vocabulário de nível A2, sendo gramaticalmente classificada como NN (substantivo), conforme indicado pela etiquetagem do *corpus*. Todas as ocorrências do colocado foram identificadas como NN. Ao examinarmos a etiquetagem, observamos as diversas estruturas formadas nas 33 ocorrências de *make* com o **nó**. Essas estruturas são detalhadas no Quadro 18:

Quadro 18 — *make sense*

ESTRUTURAS
VB DT NN
VBZ NN
VB NN

ESTRUTURAS
VBP DT NN
VB JJ NN

Fonte: O autor (2024).

Nas obras em que *sense* ocorreu como colocação de *make*, o *SpaCy* classificou a palavra *make* em VBZ (verbo na terceira pessoa do singular), VB (verbo base), e VBP (verbo na primeira pessoa do singular ou plural) e *sense* em NN (substantivo). Dessa forma, simplificando, *make sense* é constituído, em todas as ocorrências, por um verbo e um substantivo. Ademais, outros elementos gramaticais contribuem para a composição da construção, tais como DT (determinante) e JJ (adjetivo), conforme exemplificado nas linhas de concordância extraídas do *subcorpus* B2, na Figura 26:

Figura 26 — Linhas de Concordância de *make sense* - B2

Nº	Left	Node	Node
1	It doesn't make any	sense	at all.' 'Not at the
2	but Eddie could make no	sense	of them. "Just shut up!
3	those deaths that make no	sense	the kind where the gunman
4	just doesn't make any	sense	at all. But I've
5	him seemed to make no	sense.	He thought she was probably
6	Hopelessly doesn't make much	sense,	does it? Algernon: Cecily. (Merriman

Fonte: O autor (2024).

Esse tipo de dado nos revela que a palavra *make*, em diferentes categorias gramaticais, pode ser crucial para compreender as nuances e variações na utilização da palavra *make* em distintos contextos. Além disso, a presença de determinantes e adjetivos como elementos gramaticais que compõem a construção de *make sense* são fundamentais para compreender a estrutura sintática e semântica das frases em que a expressão é empregada.

Quanto ao seu significado, o *Oxford's Learner's Dictionaries* informa que há expressões idiomáticas com *make sense* classificadas nos níveis A2 e B1. O primeiro nível, classificado como A2, tem o sentido de algo que pode ser compreendido facilmente. Já o segundo nível, B1, tem o sentido de ser uma ação sensata a ser realizada, como evidenciado nos excertos A e B retirados do *subcorpus* B2:

Excerto A:

'I really can't imagine why anyone would want to hurt Arthur Mowbray. It doesn't make any sense at all.' 'Not at the moment, perhaps,' the Inspector agreed. 'But we'll find out everything in the end, I can promise you that. That's why I've been talking to Mr Pryce. He's told me quite a lot about the company.'
 Fonte: Board Games, Butler James.

Excerto B:

Does murder ever make sense? It depended, Mr Entwhistle thought, on the mind of the murderer. They arrived at a very large house which badly needed painting. 'We have no servants,' said Maude slightly angrily as she led the way in. 'Just a couple of women who come in every morning.'
 Fonte: After the Funeral, Agatha Christie.

Diante da análise, conclui-se que o sentido predominante da expressão em questão é o correspondente ao nível A2, com uma frequência de 91,17%, em comparação aos 8,82% do sentido associado ao nível B1. Essa constatação revela que, embora os *Graded Readers* analisados nessa seção sejam B2, existe uma prevalência de expressões A2 (um nível anterior).

5.4.3 Feel

Ao examinarmos a etiquetagem, identificamos todas as estruturas diferentes formadas nas 31 ocorrências de *make* com o **nó**. As estruturas podem ser observadas no Quadro 19.

Quadro 19 — Estruturas com *make feel* - B2

ESTRUTUAS
VBZ PRP VB
VB PRP VB
VB NNS VB
VBZ DT NN VB
VBZ NNP VB
VB PRP\$ NN VB
VB DT JJ NN VB
VB NNP VB

Fonte: O autor (2024).

Nas obras em que *feel* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural), e VBZ (verbo na terceira pessoa do singular), e *feel* em VB (verbo base). Dessa forma, simplificando, *make feel* é constituído, em todas as ocorrências, por um verbo e outro verbo, como pode ser observado nas linhas de concordância a seguir, retiradas do *subcorpus* B2, na Figura 27:

Figura 27 — Linhas de Concordância de *make feel* - B2

Nº	Left	Node	Right
1	girl. This rose makes me	feel	young again." He looked around
2	it - it will make you	feel	stronger again. Promise me you
3	remember. Poison can make people	feel	like that.' Doctor Spencer looked
4	the garden will make you	feel	a lot better.' 'Oh, do
5	eyes. "It makes a person	feel	a strange feeling when he
6	and dull. To make myself	feel	better I took a cold
7	and smiled to make her	feel	better. 'I had to climb
8	different, and everything makes you	feel	more afraid. The door of
9	n't want to make her	feel	any worse than she already
10	mother. Did killing make him	feel	like less of a disappointment?
11	did she always make him	feel	like a clumsy schoolboy? Perhaps

Fonte: O autor (2024).

Ao analisar as linhas de concordância da Figura 27, é notável que, nas construções de *make feel* no *subcorpus* B2, ocorre a presença de diversos elementos gramaticais entre os dois verbos, como PRP (pronome pessoal), NNS (substantivo plural), NN (substantivo singular), DT (determinante), NNP (nome próprio, singular) PPR\$ (pronome possessivo) e JJ (adjetivo). Ao compararmos os padrões gramaticais entre os *subcorpora* A2 e B2, observa-se uma diferença significativa na ocorrência dessas classes gramaticais entre os dois verbos nas sentenças. No nível A2, constatou-se que apenas pronomes pessoais (PRP) ocorreram entre os verbos, como evidenciado nas linhas de concordância a seguir, retiradas do *subcorpus* A2, na Figura 28:

Figura 28 — Linhas de Concordância de *make feel* - A2

N°	Left	Node	Right
1	of you to make me	feel	sorry for you, and then
2	Professor, and music make me	feel	very peaceful." Captain, did you
3	see it would make you	feel	happier,' he said, 'so we
4	my head that makes me	feel	confused. The man must be
5	walk around, it makes you	feel	you are at one with
6	it seemed to make him	feel	better. 'That was strange,' he
7	to try and make themselves	feel	more cheerful. But for the
8	air seemed to make him	feel	better. I turned back before
9	The drink only makes me	feel	okay these days. Still, in
10	there - that should make you	feel	better. But Luke didn't

Fonte: O autor (2024).

Enquanto no nível B2, ocorreram todas as outras classes gramaticais mencionadas, substantivos plural (NNS), substantivos singular (NN), determinantes (DT), nomes próprios singulares (NNP), pronomes possessivos (PPR\$) e adjetivos (JJ), como destacado nas linhas de concordância na Figura 29, retiradas do *subcorpus* B2:

Figura 29 — Linhas de Concordância de *make feel* - B2

N°	Left	Node	Right
1	remember. Poison can make people	feel	like that.' Doctor Spencer looked
2	eyes. "It makes a person	feel	a strange feeling when he
3	time, sir.' 'Makes a fellow	feel	rather silly, doesn't it?
4	of water.' 'It makes Sandy	feel	at home,' said Monkey. 'It
5	history might make his daughter	feel	uncomfortable. Little Dorrit filled her
6	to make this little lad	feel	that he was very small,

Fonte: O autor (2024).

Essa análise destaca que, no *subcorpus* B2, as construções apresentam uma diversidade maior de classes gramaticais, o que sugere uma necessidade de vocabulário mais amplo por parte dos estudantes para compreender plenamente o contexto em que expressão *make feel* ocorre em comparação com os alunos do *subcorpus* A2. Isso sugere que é esperado que os alunos do nível B2 sejam capazes de compreender uma gama mais abrangente de vocabulário e de nuances gramaticais para lidar com as construções que envolvem a expressão *make feel*.

5.4.4 Money

Ao explorar a etiquetagem, identificamos as diversas estruturas que emergiram nas 32 ocorrências de *make* com o **nó**. Essas estruturas estão detalhadas no Quadro 20.

Quadro 20 — Estruturas de *make money* - B2

ESTRUTURAS
VBP DT NN
VB PRP DT NN IN NN
VB DT NN IN NN
VB DT JJ NN IN NN
VB JJ NN
VBP JJR NN
VB NN
VBP NN
VB JJ NN
VB DT NN
VB DT JJ NN
VBZ DT VBN NN
NN PRP VBP

Fonte: O autor (2024).

Nas obras em que *money* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base), VBZ (verbo na terceira pessoa do singular) e VBP (verbo na primeira pessoa do singular ou plural), e *money* em NN (substantivo). Dessa forma, simplificando, *make money* é constituído, em todas as ocorrências, por verbo e um substantivo. Assim como ocorreu no *subcorpus* A2, há outros elementos gramaticais que compõem a construção, como DT (determinante), JJ (adjetivo), PRP (pronome pessoal), e JJR (adjetivo comparativo). Isso pode ser observado nas linhas de concordância do *subcorpus* B2, na Figura 30:

Figura 30 — Linhas de Concordância de *make money* - B2

Nº	Left	Node	Right
1	make me a lot of	money.	My father was not pleased
2	might make a lot of	money	from the other side. You
3	can make a lot of	money	if you put some into
4	can make a lot of	money	by betting on it with
5	make a great deal of	money'.	A few thought that he
6	make a great deal of	money'.	A few thought that he
7	make a great deal of	money	in the end! As for

Fonte: O autor (2024).

Em uma análise comparativa, observa-se que a expressão *make money* comporta-se de maneira semelhante nos *subcorpora* A2 e B2. No entanto, a principal diferença reside nos tipos de sentenças em que essa construção ocorre, o que reflete a complexidade diferenciada entre os níveis A2 e B2.

No nível A2, é comum ocorrer sentenças com a construção PRP RB VBP TO VB NN (Pronome pessoal + advérbio + preposição + verbo substantivo), como pode ser visto no exemplo abaixo retirado do *subcorpus* A2:

Ex. 14: "*they only want to make money.*"

fonte: Jacks Came, Rod Smith.

No nível B2, por outro lado, o colocado apresenta características exclusivas para esse nível. Há estruturas na forma condicional com *if* e estruturas condicionais com *as long as*, como evidenciado nos exemplos abaixo retirados do *subcorpus* B2.

Ex. 15: "*Your father's happy for us to get married, if I can make enough money to look after you.*"

Fonte: A Study in Scarlet, Conan Doyle.

Ex. 16: "*I don't care what kind of shoes we sell, as long as we make money.*"

fonte: Kings Ransom, Ed McBain.

Essa diferenciação entre os níveis A2 e B2 pode sugerir o progresso esperado no aprendizado da língua. Enquanto os alunos de nível A2 lidam com estruturas gramaticais e vocabulário mais básicos, os alunos de nível B2 enfrentam uma variedade mais ampla de construções linguísticas.

5.4.5 Happy

A palavra *happy* está classificada como vocabulário de nível A1 e gramaticalmente se caracteriza como um adjetivo. Segundo a etiquetagem do *corpus*, todas as ocorrências do colocado foram classificadas como JJ (adjetivo). Ao examinarmos a etiquetagem, identificamos as diferentes estruturas formadas nas 25 ocorrências de *make* com o **nó**. As estruturas correspondentes podem ser observadas no Quadro 21.

Quadro 21 — Estruturas de *make happy* - B2

ESTRUTURAS
VB PRP JJ
VB PRP\$ JJ NNP JJ
VB PRP\$ JJ CC JJ
VB PRP\$ NN JJ CC JJ
VBZ PRP JJ
VB PRP RB JJ
VB NNP JJ
VB DT NN JJ
VB PRP RB RB JJ

ESTRUTURAS
VB JJ NNS JJ

Fonte: O autor (2024).

Nas obras em que *happy* ocorreu como colocação do **nó**, o *SpaCy* classificou a palavra *make* em VB (verbo base) e VBZ (verbo na terceira pessoa do singular) e *happy* em JJ (adjetivo). Dessa forma, simplificando, *make happy* é constituído, em todas as ocorrências, por um verbo e um adjetivo. Além disso, há outros elementos gramaticais que compõem a construção, como PRP (pronome pessoal), PRP\$ (pronome pessoal possessivo), NNP (substantivo próprio, singular), CC (conjunção coordenativa), RB (advérbio) e NNS (substantivo comum, plural), como pode ser observado nas linhas de concordância do *subcorpus* B2, na Figura 31:

Figura 31 — Linhas de Concordância de *make happy* - B2

N°	Left	Node	Right
1	I want to make him	happy.	Let me see him!" When
2	to make our dear Clifford	happy.	Why should we talk further
3	make his life warm and	happy.	She opened an old bookcase
4	I want to make him	happy.	Let me see him!" When
5	to make our dear Clifford	happy.	Why should we talk further
6	make his life warm and	happy.	She opened an old bookcase
7	do that truly makes them	happy.	You might hear a man
8	It would make me very	happy.'	I'm not sure which
9	theme. That should make her	happy.	The young investigator who had
10	it would make me very	happy	as I owe you my

Fonte: O autor (2024).

Notamos que, em todas as sentenças, há a presença de outros elementos gramaticais entre os dois verbos, como NN, NNP e PRP. Isso significa que, em todas as ocorrências, a construção *make (someone) happy* tem o sentido de fazer **alguém** feliz. Essa construção se assemelha a de *make someone feel* no *subcorpus* A2, na seção 4.2.1.

Além disso, observamos que a construção Verbo + Pronome Pessoal + Adjetivo é a mais predominante (56%) dentre as outras identificadas, com variações apenas na conjugação verbal, podendo ser VBZ (verbo na terceira pessoa do singular) ou VB

(verbo na forma base), como evidenciado na linha de concordância 1 e 6 acima. Essa consistência na presença do pronome pessoal indica a dinamicidade da língua em usar os pronomes para substituir os substantivos, o que pode permitir uma comunicação mais fluida.

Essa substituição pode ser observada nos excertos A e B retirados do *subcorpus* B2, onde o pronome *him* substitui o substantivo *Clifford* e o pronome *her* substitui o substantivo *daughter*, respectivamente:

Excerto A:

We all love Clifford and I want to help him. He can come to my home in the country and enjoy his life. I can give him all the beautiful things he likes and the best food. I want to make him happy.

Fonte: The House of the Seven Gables, Nathaniel Hawthorne.

Excerto B:

Did you think that I'd be your daughter forever?' I saw that she was angry, and I didn't know how to make her happy.

Fonte: Interview with the Vampire, Anne Rice.

Essa observação pode ser útil para os alunos, pois evidencia a importância dos pronomes na comunicação em inglês. Compreender essa dinâmica pode facilitar a compreensão e a produção de frases de forma mais natural e idiomática. Além disso, os alunos podem perceber como os pronomes são empregados para substituir os substantivos e como podem ampliar sua capacidade de expressão. Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que a expressão *make someone/something/yourself + adj* tem o sentido de fazer com que alguém/algo sinta, mostre ou tenha uma qualidade particular e está classificada como nível A1. Isso nos revela que a colocação está coerente em aparecer no B2, pelos mesmos motivos já citados anteriormente.

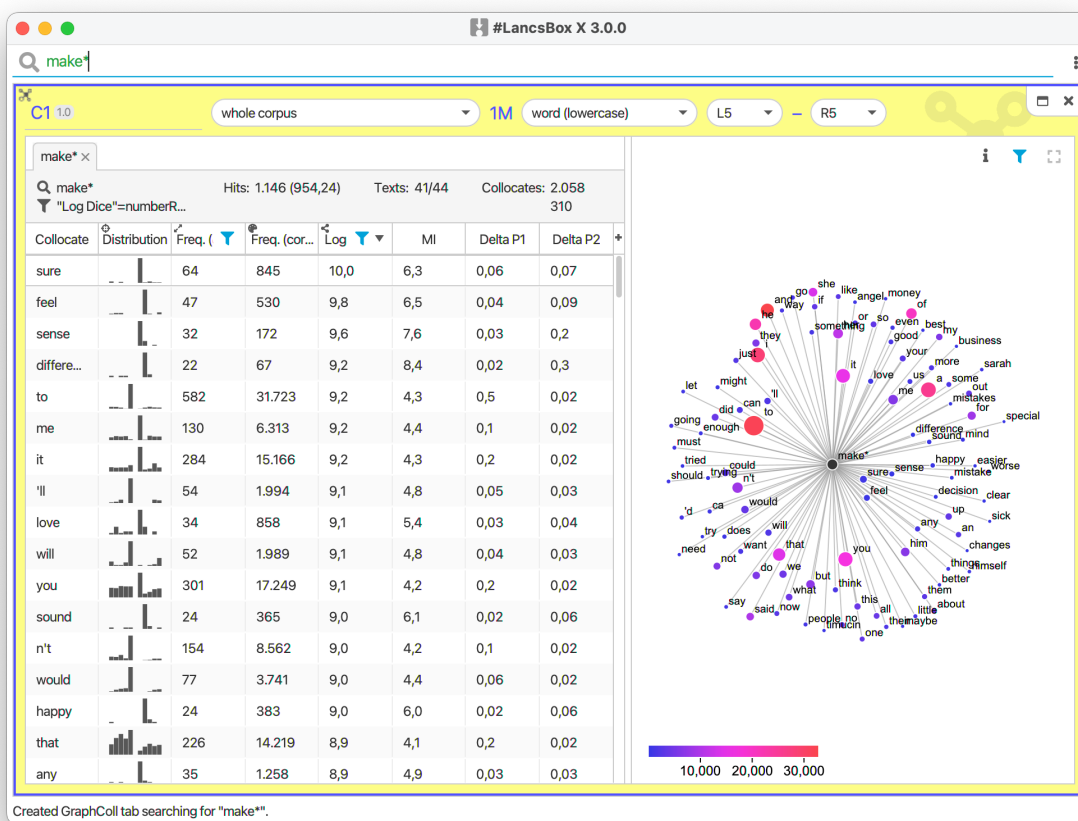
Em resumo, nossas análises revelam que nos *subcorpora* A2 e B1 não há ocorrência de outros elementos gramaticais entre o verbo e o adjetivo em construções de *make sure*. No que diz respeito à expressão *make sense*, constatou-se que seu sentido predominante é de algo que pode ser compreendido facilmente, com uma

frequência de 91,17%, em contraste com os 8,82% associados ao sentido de ser uma ação sensata a ser realizada. Em relação à *make feel* do *subcorpus* B2, notou-se uma maior diversidade de elementos gramaticais entre o núcleo e o complemento, em comparação ao *subcorpus* A2. Quanto a *make money*, embora haja semelhanças entre os *subcorpora* A2 e B2, a diferença reside nos tipos de sentenças em que essa construção ocorre, refletindo as distinções entre os níveis A2 e B2. Por fim, essas observações ressaltam a diversidade da linguagem e como as nuances semânticas podem variar significativamente, dependendo do contexto e do nível de proficiência.

5.5 Colocações com o verbo *make* no *subcorpus* c1

Replicando o procedimento anterior, contudo, desta vez no *subcorpus* C1, inserimos a palavra *make* na barra de pesquisa da ferramenta *Graphcoll* para obter suas colocações. Os resultados obtidos foram os seguintes:

Figura 32 — Make - C1



Fonte: O autor (2024).

O **nó** em questão ocorre 1.146 vezes no *subcorpus* C1, em 41 dos 44 textos e possui 2.058 **colocados**. Desse total de colocados, após a aplicação do CPN, proposta por Brezina, McEnery e Wattam (2015), sobraram 310 **colocados**. Ao levar em consideração os nossos critérios e os nossos objetivos, obtivemos 1 colocação mais exclusiva do verbo *make* no *subcorpus* C1, a qual é *sure*.

No Quadro 22, são mostrados os respectivos valores de co-ocorrência com o **nó**, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para essa pesquisa.

Quadro 22 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>sure</i>	64	845	10,0

Fonte: O autor (2024).

5.5.1 Sure

A análise da etiquetagem revelou diversas estruturas gramaticais formadas nas 58 ocorrências de *make* com o **nó**. Essas estruturas estão detalhadas no Quadro 23.

Quadro 23 — Estruturas de *make sure* - C1

ESTRUTURAS
VB JJ
VBZ PRP RB JJ

Fonte: O autor (2024).

As estruturas gramaticais contendo a expressão *make sure* no *subcorpus* C1 apresentam semelhanças com aquelas encontradas nos demais *subcorpora* (A2, B1 e B2), consistindo predominantemente de um verbo e um adjetivo. Contudo, percebe-se uma diferença na estrutura da expressão, ocorre a inclusão de pronomes pessoais (PRP) e advérbios (RB) entre o **nó** e o colocado. Tal diferenciação é exemplificada nas linhas de concordância, na Figura 33, retiradas do *subcorpus* C1:

Figura 33 — Linhas de Concordância de *make sure* - C1

Nº	Left	Node	Right
1	snapped. "What makes you so	sure,	Marlowe, that this Taylor boy
2	I told him to make	sure	she got there," he said. "

Fonte: O autor (2024).

A sentença da linha 1 representa a estrutura VBZ PRP RB JJ, composta pelo pronome *you* e o advérbio *so* entre o **nó** e o colocado. A sentença da linha 2 representa a estrutura VB JJ, contendo apenas o **nó** e o colocado. Essa observação sugere que, embora a expressão *make sure* retenha seu sentido literal em todos os *subcorpora*, a sua construção sintática no *subcorpus* C1 se destaca pela inserção de elementos adicionais entre o verbo e o adjetivo.

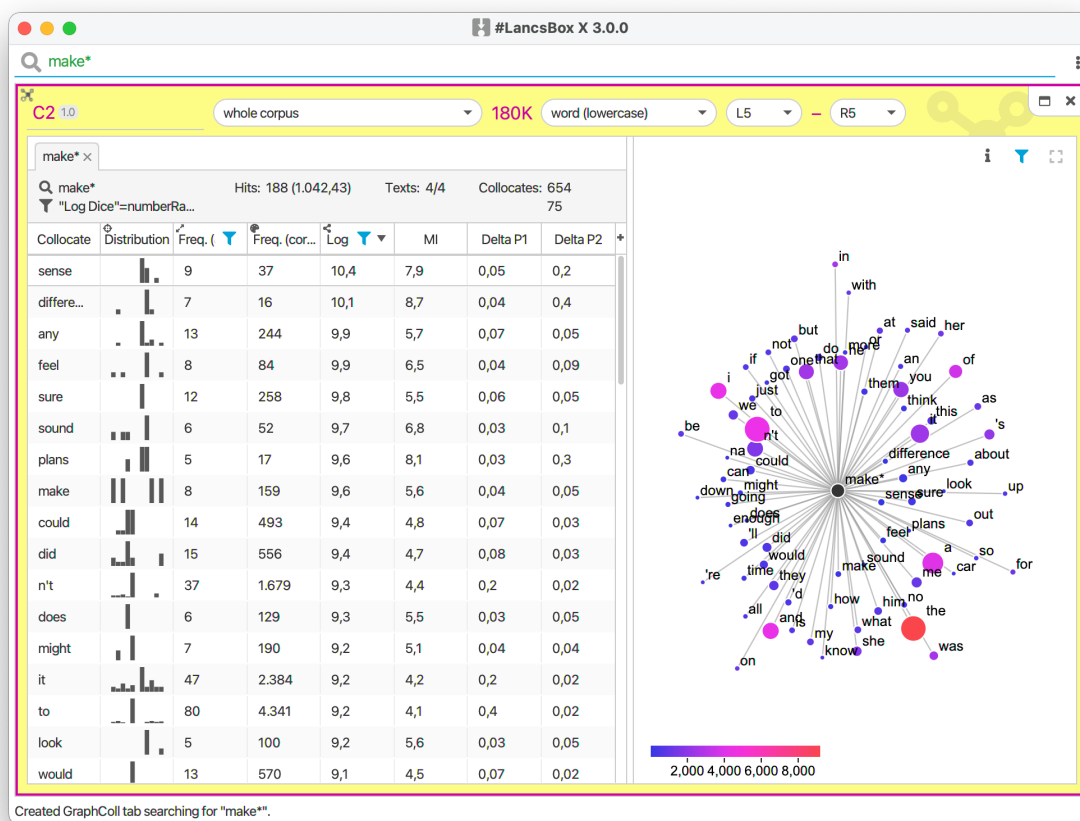
A observação da colocação *sure* até o nível C1 conduziu à conclusão de que, em resumo, a expressão *make sure* mantém seu significado literal em todos os níveis de proficiência linguística analisados. As estruturas sintáticas associadas a ela, caracterizadas pela combinação de um verbo com um adjetivo, revelaram-se similares

em todos os níveis investigados. A distinção significativa foi identificada apenas no nível C1, onde se observou a inclusão de pronomes pessoais e adjetivos entre o verbo e o adjetivo. Essa análise sugere uma variação na estruturação gramatical da expressão *make sure* apenas em níveis mais avançados dos *readers* analisados.

Isso pode indicar algumas implicações para o processo de leitura e interpretação dos textos. Primeiramente, essa consistência no significado literal da expressão ao longo dos diferentes níveis de proficiência sugere que os aprendizes podem confiar na compreensão básica do termo, ao se depararem com ele em textos de diferentes complexidades. Isso pode proporcionar uma sensação de segurança e familiaridade durante a leitura, permitindo que os leitores se concentrem em outros aspectos do texto sem se preocupar com mudanças repentinas de significado. No entanto, a inclusão de pronomes pessoais e adjetivos entre o verbo e o adjetivo no nível C1 sugere uma necessidade maior de conhecimento de vocabulário, o que pode refletir uma expansão do repertório linguístico.

5.6 Colocações com o verbo *make* no *subcorpus c2*

Seguindo o procedimento anterior, contudo, agora no *subcorpus C2*, inserimos na barra de pesquisa da ferramenta *Graphcoll* a palavra *make* para identificar seus **colocados**. Os resultados obtidos foram os seguintes:

Figura 34 — *make* - C2

Fonte: O autor (2024).

O **nó** em análise ocorre 188 vezes no *subcorpus* C2, abrangendo a totalidade de 4 textos, e manifesta-se em 654 contextos específicos. Desses, após a implementação do CPN, proposta por Brezina, McEnery e Wattam (2015), restaram 75 **colocados**. Considerando os parâmetros e os objetivos delineados nesta pesquisa, identificamos 2 colocações particularmente exclusivas do verbo *make* no *subcorpus* C2, nomeadamente *sense* e *difference*.

O Quadro 24 apresenta os respectivos valores de co-ocorrência com o **nó**, a frequência de ocorrência no *corpus* e os resultados calculados, através do método estatístico adotado nesta investigação.

Quadro 24 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Sense</i>	9	37	10,4

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Difference</i>	7	16	10,1

Fonte: O autor (2024).

5.6.1 Sense

A análise das estruturas formadas nas 9 ocorrências de *make* com o **nó** revela diversas configurações, as quais podem ser observadas no Quadro 25.

Quadro 25 — Estruturas *make sense* - C2

ESTRUTURAS
VBZ JJ NN
VBZ DT NN IN NN
VB DT NN
VB NN

Fonte: O autor (2024).

Nas obras em que *sense* ocorreu como **colocação** de *make*, o *SpaCy* classificou a palavra *make* em VBZ (verbo na terceira pessoa do singular) e VB (verbo base) e *sense* em NN (substantivo). Portanto, simplificando, *make sense* é constituído, em todas as ocorrências, por um verbo e um substantivo. Além disso, há outros elementos gramaticais que integram a construção, tais como DT (determinante), JJ (adjetivo), NN (substantivo) e IN (preposição) como pode ser evidenciado nos exemplos retirados do *subcorpus* C2, na Figura 35:

Figura 35 — Linhas de Concordância de *make sense* - C2

Nº	Left	Node	Right
1	you saying, that makes great	sense;	watching a stop-and-go
2	That makes a lot of	sense,	Tina." Fling! She thought angrily.
3	It doesn't make any	sense,"	Zachariah said. "No sense at
4	road, which would make no	sense,	as it is a secret

Fonte: O autor (2024).

A presença de outros elementos gramaticais, como DT, JJ, IN e NN enriquece a complexidade sintática dessa construção, evidenciando a variedade de formas gramaticais que podem acompanhar *make sense*. Essas estruturas compostas por esses termos gramaticais são bem recorrentes no *subcorpus C2*, apresentando 44,55% das ocorrências. Em contrapartida, a estrutura com apenas o **nó** e o colocado, Verbo + Substantivo, apresenta 55,55% das ocorrências. Esse equilíbrio na distribuição das estruturas ressalta a diversidade estrutural na expressão *make sense* em níveis avançados.

Além disso, foi observado que todas as ocorrências da expressão *make sense* no *subcorpus* analisado possuem um significado intrinsecamente ligado a algo facilmente compreensível, conforme exemplificado nos excertos A e B, extraídos dos *readers* do *subcorpus C2* abaixo. Essa característica difere de *make sense* no *subcorpus B2*, onde além de denotar algo compreensível, também pode transmitir o sentido de uma ação sensata a ser realizada.

Excerto A:

That was scary because it didn't make sense. A stranger? It must be. Michael was the only person who had ever blamed her for Danny's death.

Fonte: The Breathing Method, Stephen King.

Excerto B:

"Maybe not." Elliot put his hand on her shoulder and squeezed gently. "Tina, you know this theory doesn't make sense. Danny is dead." "See? You do think I'm crazy."

Fonte: The Eyes of Darkness, Koontz Dean.

5.6.2 Difference

A palavra *difference* é classificada como parte do vocabulário de nível A1 e, gramaticalmente, se enquadra como um substantivo. Conforme a etiquetagem do *corpus*, todas as ocorrências do termo colocado foram classificadas como NN (substantivo). A análise da etiquetagem permitiu a identificação de diversas estruturas diferentes formadas nas 7 ocorrências de *make* com o **nó**, as quais podem ser observadas no Quadro 26.

Quadro 26 — Estruturas *make difference* - C2

ESTRUTURAS
VBZ DT NN
VB DT NN
VB DT JJ NN
JJ CD NN MD VB

Fonte: O autor (2024).

Na literatura em que *difference* ocorreu em conjunto com *make*, a ferramenta *SpaCy* classificou a palavra *make* em VBZ (verbo na terceira pessoa do singular) e VB (verbo base) e a palavra *difference* em NN (substantivo). Dessa forma, simplificando, a estrutura *make difference* é constituída, em todas as ocorrências, por um verbo e um substantivo. Destaca-se que, semelhantemente a outros colocados analisados nesta pesquisa, algumas sentenças apresentam variação na posição do verbo em relação ao substantivo, conforme exemplificado na Figura 36:

Figura 36 — Linha de Concordância de *make difference* - C2

N°	Left	Node	Right	POS Left
1	was astonished at how much	difference	one year could make in	VBD VBN IN WRB JJ

Fonte: O autor (2024).

Adicionalmente, em relação à estrutura, foi constatada a presença de determinantes com sentido de negação, como *no* e *any* em 71,42% das sentenças analisadas, destacando a frequência significativa desses elementos nos *readers* examinados.

No que concerne ao significado da expressão, o *Oxford's Learner's Dictionaries* informa que existem idiomas com *make difference*, classificados no nível B1, com o sentido de ter efeito/nenhum efeito sobre alguém/algo. Ao examinarmos as sentenças no *subcorpus*, confirmamos que todas elas estão em conformidade com o sentido descrito pelo dicionário, como podem ser observadas nos excertos A e B, retirados do *subcorpus* abaixo:

Excerto A:

Then, too, the dealer just happened to have a spare set of keys for it." "What if the guy's changed locks?". "It won't make any difference. With the tools I carry, I can take any car, locked or not, in about 30 seconds.

Fonte: The Thrill of the Hunt, Darryl Matter.

Excerto B:

"No, the car won't be camouflaged for our purposes. That's for sure." Wes laughed, then got serious. "It won't make any difference about the paint job, though, 'cause we're not gonna drive that car again anyway, anywhere.

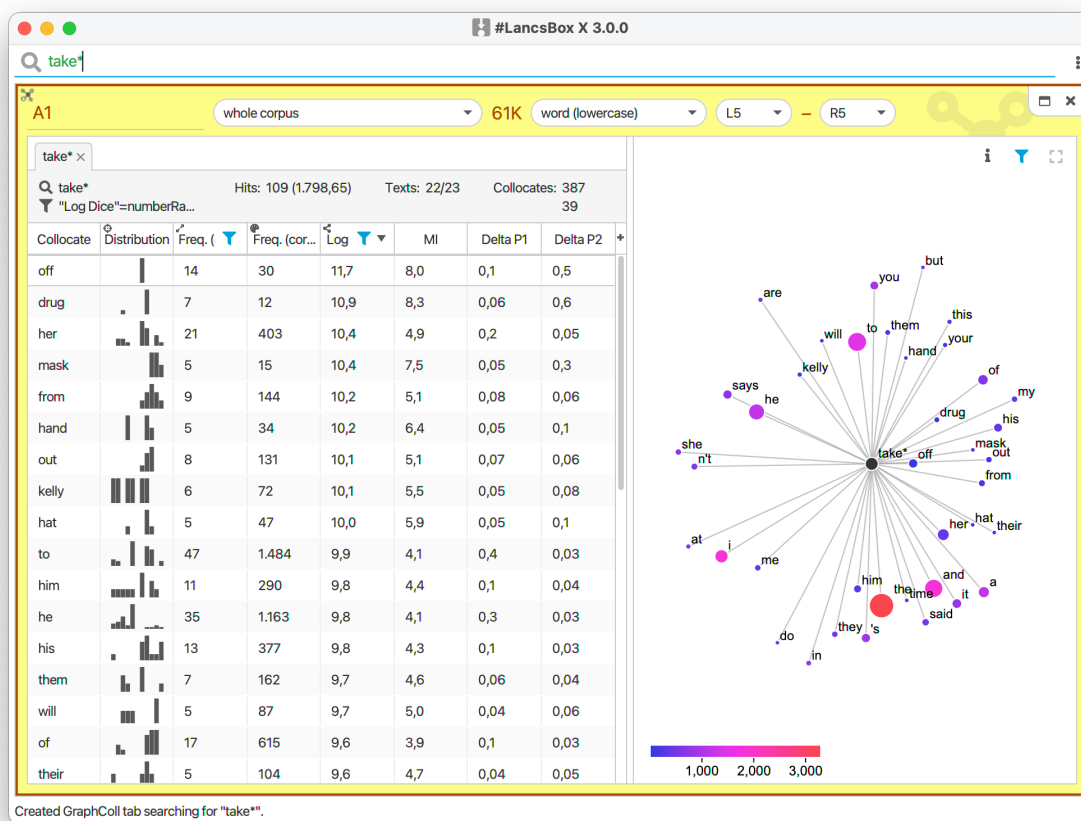
Fonte: The Thrill of the Hunt, Darryl Matter.

Portanto, o fato de *make difference* ser um item relevante no *subcorpus C2* é coerente, uma vez que as palavras dos níveis anteriores se repetem nos níveis seguintes. A próxima seção abordará os resultados da análise do verbo *take*, nesta pesquisa. Logo após, serão apresentadas as considerações finais com a apresentação geral de todos os resultados obtidos neste estudo.

5.7 Colocações com o verbo *take* no *subcorpus a1*

Conforme os registros do *Oxford's Learners Dictionaries*, a palavra *take* é categorizada como parte do vocabulário nos níveis A1, A2 e B1, variando de acordo com seu significado. Gramaticalmente, *take* é classificado como um verbo. Ao considerar *take* como o **nó** de interesse, inserimos o termo na barra de pesquisa da ferramenta *Graphcoll* no *Lancsbox* para identificar seus **colocados**. Os resultados obtidos foram os seguintes:

Figura 37 — Take - A1



Fonte: O autor (2024).

O nó em questão ocorre 48 vezes no *subcorpus* A1, em 22 dos 23 textos, possuindo 387 colocados. Desse total de colocados, após a aplicação do CPN, propostas por Brezina, McEnery e Wattam (2015), sobraram 39 colocados. Uma vez que nosso objetivo é identificar as colocações mais exclusivas do verbo, concluímos que, de acordo com o resultado do cálculo estatísticos e os nossos critérios de seleção, o colocado *off* é o mais exclusivo dentre os 39.

No Quadro 27, são mostrados os respectivos valores de co-ocorrência com o nó, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para esta pesquisa.

Quadro 27 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Off</i>	14	30	11,7

Fonte: O autor (2024).

5.7.1 Off

Quando analisamos a co-ocorrência de *off*, em conjunto com o **nó** em questão, observamos que ocorre em 14 instâncias, todas elas relacionadas ao verbo. Sua frequência no *corpus* é de 30 vezes e o cálculo na fórmula do *LogDice* resulta em 11,7. Ao considerar o valor de referência de 14, concluímos que *off* é uma colocação relevante por se aproximar do limiar estatístico estabelecido. Essa conclusão também pode ser sustentada pela análise do gráfico presente na Figura 19.

A palavra *off* é classificada como parte do vocabulário de nível A1 e, gramaticalmente, é categorizada como um advérbio. A classificação gramatical das ocorrências de *off* associadas ao **nó** foi identificada como RP (advérbio, partícula), conforme a etiquetagem do *corpus*. Ao examinar as diversas estruturas formadas nas 14 ocorrências de *take* com o **nó** em questão, podemos observar as respectivas configurações no Quadro 28.

Quadro 28 — Estruturas *take off* - A1

ESTRUTURAS
VBZ RP
VB RP
VBP RP

Fonte: O autor (2024).

No contexto das obras em que a colocação de *off* ocorreu com o verbo *take*, a classificação realizada pelo *SpaCy* apontou que a palavra *take* é categorizada como VBZ (verbo na terceira pessoa do singular), VB (verbo base), e *off* como RP (advérbio, partícula). Simplificadamente, *take off* é composto, em todas as instâncias, por um verbo e um advérbio.

Quanto ao seu significado, *take off* é identificado como um *phrasal verb* de nível A2, apresentando múltiplos significados, incluindo a remoção de algo, especialmente uma peça de roupa do corpo próprio ou de outrem. Esse contexto é corroborado pelas ocorrências observadas no *corpus* A1, conforme exemplificado nas linhas de concordância retiradas do *subcorpus* A1, na Figura 38:

Figura 38 — Linhas de Concordância de *take off* - A1

Nº	Left	Node	Right
1	who are you?' She takes	off	The Cat's mask. 'You!'
2	Wendy asks. The boy takes	off	his cap and asks, 'What
3	her cold hand. She takes	off	Kelly's mask and wig. "
4	's sit down and take	off	our masks. We want to
5	the alien?" The alien takes	off	her mask. "It's Megan!"
6	Bill. The old lady takes	off	her mask. Everyone screams. It
7	her watch. Then she takes	off	her Walkman and starts to
8	window and started to take	off	her dress. 'Let's go
9	in the canteen and take	off	their heads. Sally is thirsty.
10	the people say. Zorro takes	off	his black mask. It's
11	right, Sally,' he says. 'Take	off	your heads!' Mr Parry says

Fonte: O autor (2024).

A discrepância na classificação do dicionário, indicando nível A2, enquanto a expressão ocorre no *corpus* de nível A1, pode ser atribuída à forma introdutória de ambos os níveis. Dada a similaridade entre os níveis A1 e A2, em termos de proficiência e familiaridade com a língua, é razoável que a expressão *take off* seja apropriada para ambos.

Ao analisar as estruturas em que *take off* ocorre no *subcorpus* A1, observamos que 85,71% das ocorrências estão em sentenças no presente simples, conforme exemplificado nos excertos A, B e C, retirados dos *readers*:

Excerto A:

They are free. Everyone is happy and cheers. The Governor says, "Now that you are free, show us your face!". "Yes! Yes!" the people say. Zorro takes off his black mask. "It's Don Diego Vega!" exclaims Sergeant Gonzales. Everyone is very surprised.

Fonte: Zorro, Sally M Stockton.

Excerto B:

But then the mice hear a noise. 'Bing bong!' The American woman looks at her watch. Then she takes off her Walkman and starts to put it in her bag. 'Hey!' she says. 'What...? Two mice! What's this?'

Fonte: Marcel and the Shakespeare Letters, Stephen Rable.

Excerto C:

We want to see the person behind the mask!" Nick asks, "Who is the alien?" The alien takes off her mask. "It's Megan!" Everyone laughs. "And who is the ghost? It's Susan!"

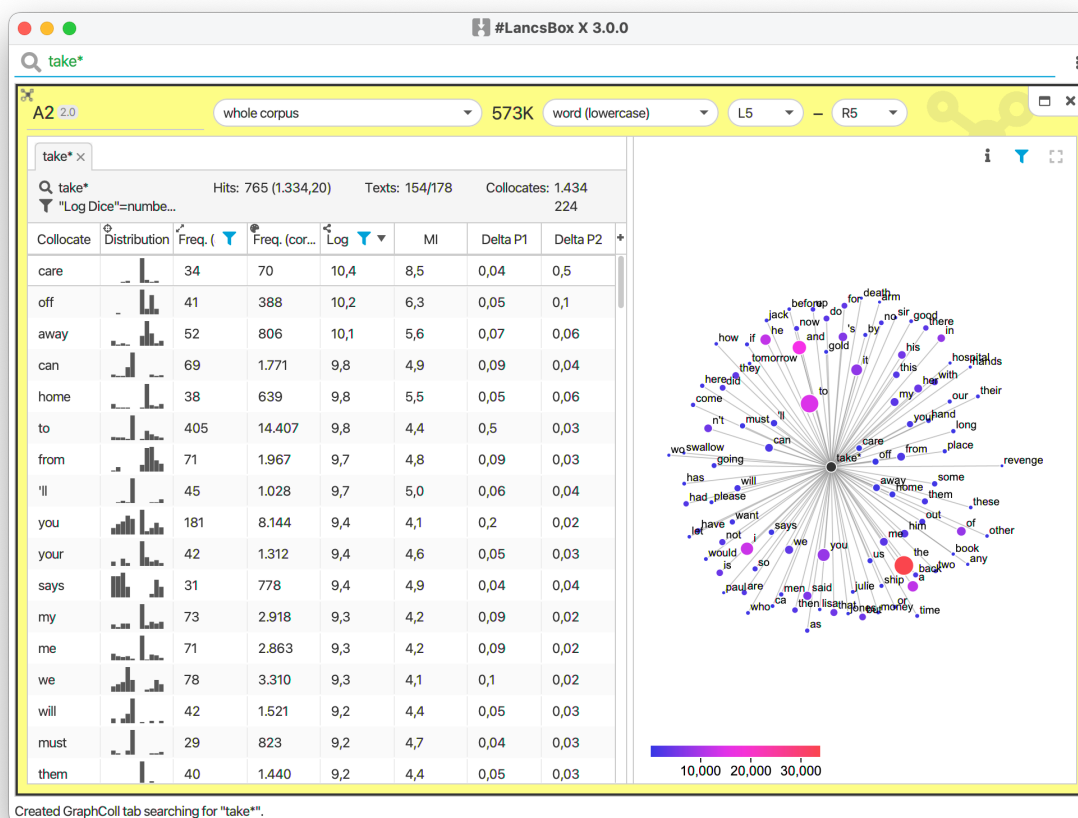
Fonte: Halloween Horror, Clemen D B Gina.

Esse dado sugere que a expressão *take off* é predominantemente utilizada para descrever ações ou eventos habituais no contexto do nível de proficiência A1 nos *readers* analisados.

5.8 Colocações com o verbo *take* no *subcorpus a2*

Repetindo o procedimento anterior, porém agora no *subcorpus A2*, inserimos na barra de pesquisa da ferramenta *Graphcoll* a palavra *take* para obter seus colocados. Os resultados obtidos foram os seguintes:

Figura 39 — Take - A2



Fonte: O autor (2024).

O **nó** em questão ocorre 765 vezes no *subcorpus* A2, em 154 dos 178 textos, possuindo 1.434 **colocados**. Desse total de colocados, após a aplicação do CPN, propostas por Brezina, McEnery e Wattam (2015), sobraram 224 **colocados**. Uma vez que nosso objetivo é identificar as colocações mais exclusivas do verbo, concluímos que, de acordo com o resultado do cálculo estatístico e os nossos critérios de seleção, os colocados *care*, *off* e *away* são os mais exclusivos dentre os 224.

No Quadro 26, são apresentados os valores correspondentes de co-ocorrência com o **nó**, a frequência de ocorrência no *corpus* e os resultados do cálculo do método estatístico adotado nesta pesquisa.

Quadro 29 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Care</i>	34	70	10,4
<i>Off</i>	41	388	10,2
<i>Away</i>	52	806	10,1

Fonte: O autor (2024).

5.8.1 Care

A palavra *care* é categorizada como parte do vocabulário de nível A2 e, gramaticalmente, classificada como um substantivo. A etiquetagem do *corpus* indica que todas as ocorrências do termo colocado foram classificadas como NN (substantivo). Ao examinar as estruturas resultantes das 32 ocorrências de *make* com o **nó**, conforme a etiquetagem, observamos diversas formas, como evidenciado no Quadro 30.

Quadro 30 — Estruturas *take care* - A2

ESTRUTURAS
VB NN
VBP JJ NN
VBN NN
VBZ RB JJ NN

Fonte: O autor (2024).

Nas obras em que *care* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em (Verbo na 3ª pessoa do singular), VB (Verbo na forma base), e VBP (Verbo na 1ª e 2ª pessoa do singular e no plural) e *care* em NN (substantivo). Dessa forma, simplificando, *take care* é constituída, em todas as ocorrências, por um verbo e um substantivo. Além disso, foram identificados outros elementos gramaticais entre o **nó** e o colocado, como JJ (adjetivo) e RB (advérbio). Apesar de haver ocorrências com JJ e RB, as estruturais mais comuns nesse *subcorpus* são aquelas sem esses elementos gramaticais. Dessa maneira, 78,1% das estruturas são compostas por Verbo + Substantivo.

No que diz respeito ao significado, *take care* é um *idiom* de nível A2, abrangendo diversos significados, tais como tomar cuidado, cuidar de alguém, algo ou de si mesmo, e ter cautela em relação a algo. A análise das sentenças no *subcorpus* em questão confirmou a presença de *take care* com esses significados, conforme ilustrado nas linhas de concordância retiradas do *subcorpus* A2, na Figura 40:

Figura 40 — Linhas de Concordância de *take care* - A2

N°	Left	Node	Right
1	sitting by the fire! Take	care,	sir, take care.' 'What do
2	said. 'I normally take good	care	of my girlfriends.' 'I see,'
3	left alone again to take	care	of the sheep. He tried
4	was no one to take	care	of me. Then my mother

Fonte: O autor (2024).

Essa conformidade reforça a pertinência de *take care* no contexto do *subcorpus* A2, alinhando-se com as expectativas do nível de proficiência linguística correspondente.

5.8.2 Off

A palavra *off* é classificada como um vocabulário de nível A1 e, gramaticalmente, é categorizada como um advérbio. Segundo a etiquetagem do *corpus*, todas as ocorrências do colocado foram classificadas como RP (advérbio, partícula). A análise detalhada das estruturas formadas nas 39 ocorrências de *make* com o **nó** revela diversas configurações, conforme evidenciado no Quadro 31.

Quadro 31 — Estruturas *take off* - A2

ESTRUTURAS
VB PRP\$ NN IN
VB RP
VB PRP\$ NNS IN
VB DT NN RP
VB PRP RP

ESTRUTURAS
VBZ RP
VB PRP\$ NN RP
VBP RP
VBP DT NN IN
VB DT NN NN IN
VBP DT JJR NN RP
VB DT JJ NNS RP
VB PRP\$ NNS RP
VB PRP IN

Fonte: O autor (2024).

Nas obras em que *off* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base) e VBP (verbo na primeira pessoa do singular ou plural), e *off* em RP (advérbio, partícula) e IN (preposição). Dessa forma, simplificando, *take off* é constituído, em algumas ocorrências, por um verbo e um advérbio e, em outras, por um verbo e uma preposição. Além disso, observamos que outros elementos gramaticais ocorrem entre o **nó** e o colocado, tais como PRP\$ (pronome pessoal possessivo), PRP (pronome pessoal), NN (substantivo comum, singular), NNS (substantivo comum, plural), DT (determinante), JJ (adjetivo), JJR (adjetivo comparativo).

A expressão *take off* é um *phrasal verb* de nível A2, possuindo diversos significados, sendo um deles remover algo, especialmente uma peça de roupa do corpo seu/de alguém, e remover algo/alguma coisa de algum lugar. Esses dois significados podem ser encontrados nas ocorrências do *subcorpus* A2, como visto nas linhas de concordância, na Figura 41, retiradas do *subcorpus* A2:

Figura 41 — Linhas de Concordância de *take off* - A2

N°	Left	Node	Right
1	you to take your eyes	off	him, then made his move. '
2	stupid!' He started to take	off	the veil from his face. '
3	The bride cannot take	off	her veil until after the
4	and he started to take	off	the veil. 'No!' shouted Loki,
5	cannot take their eyes	off	her. Connie says, 'He says

Fonte: O autor (2024).

A conclusão decorrente da presente análise é que *take off* exibe variações em sua estrutura gramatical e significado, dependendo do contexto em que é empregada. Em algumas obras sujeitas à análise do *SpaCy*, a palavra *take* é categorizada como verbo (VBZ, VB, VBP), enquanto *off* é classificada, em alguns casos, como advérbio (RP):

Figura 42 — Linhas de Concordância com POS de *take off* - A2

N°	Left	Node	Right	POS Left	POSN	POS Right
1	stupid!' He started to take	off	the veil from his face. '	JJI' PRP VBD TO VB	RP	DT NN IN PRP\$ NN. '
2	The bride cannot take	off	her veil until after the	DT NN MDRB VB	RP	PRP\$ NN IN IN DT
3	and he started to take	off	the veil. 'No!' shouted Loki,	CC PRP VBD TO VB	RP	DT NN. 'UH!' VBD NNP,
4	n't sleep. Take the spell	off.	I don't want to	RB VB. VB DT NN	RP.	PRP VBPRB VB TO
5	I'll take the spell	off	immediately, my love.' She went	PRPMD VB DT NN	RP	RB, PRP\$ NN.' PRP VBD
6	n't sleep. Take the spell	off.	I don't want to	RB VB. VB DT NN	RP.	PRP VBPRB VB TO

Fonte: O autor (2024).

Em outros casos, *off* é classificado como preposição (IN):

Figura 43 — Linhas de Concordância com POS de *take off* - A2

N°	Left	Node	Right	POS Left	POSN	POS Right
1	cannot take their eyes	off	her. Connie says, 'He says	MDRB VB PRP\$ NNS	IN	PRP. NNP VBZ, 'PRP VBZ
2	my commander, take my gun	off	my shoulder, and hold it	PRP\$ NN, VB PRP\$ NN	IN	PRP\$ NN, CC VB PRP
3	you to take your eyes	off	him, then made his move. '	PRP TO VB PRP\$ NNS	IN	PRP, RB VBD PRP\$ NN. '
4	're doing? Take your hands	off	me!" "It's about your	VBP VBG? VB PRP\$ NNS	IN	PRP!" "PRPVBZ IN PRP\$
5	I take the paper cover	off	the wool, find one end	PRP VBP DT NN NN	IN	DT NN, VB CD NN
6	Can you take a week	off	work?' he asked. 'But Dad!	MD PRP VB DT NN	IN	NN?' PRP VBD. 'CC NNP!
7	the wizard. Take your hands	off	her, Rocford. Rocford let go	DT NN. VB PRP\$ NNS	IN	PRP, NNP. NNP VBD VB

Fonte: O autor (2024).

Isso sugere que, em determinadas instâncias, *take off* é composto por um verbo e um advérbio (20,5% das ocorrências), ao passo que, em outras, é composto por um verbo e uma preposição (79,48% das ocorrências). Essa distinção ocorre

principalmente quando a próxima palavra depois do colocado *off* é um pronome, como vista na linha de concordância I "*take your eyes off him*" (tire os olhos dele).

Adicionalmente, ao examinarmos os *Readers*, associados aos níveis de proficiência linguística (A1 e A2), evidenciamos variações nos significados atribuídos à expressão. No nível A1, o significado se restringe predominantemente (79,48%) à ideia de remover uma peça de roupa do corpo. Em contrapartida, no nível A2, além desse significado, emerge a interpretação de remover algo de algum lugar/de alguém. Essa diferença de significados está intrinsecamente relacionada à estrutura gramatical da expressão em cada contexto específico: quando associada à remoção de uma peça de roupa, a estrutura consiste em um verbo e um advérbio; enquanto na remoção de algo de um local específico, a estrutura é configurada por um verbo e uma preposição. Portanto, a inferência resultante é que a expressão *take off* nos *readers* analisados, revela-se flexível em sua estrutura gramatical, sendo capaz de assumir diferentes significados, conforme o contexto e o nível de proficiência linguística.

5.8.3 Away

A palavra *away* está classificada como um vocabulário de nível A1 e gramaticalmente se classifica como um advérbio. Segundo a etiquetagem do *corpus*, todas as ocorrências do colocado foram classificadas como RP (advérbio, participio). Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 43 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 32:

Quadro 32 — Estruturas *take away* - A2

ESTRUTURAS
VB DT NN RB
VB RB
VB PRP\$ NN RB
VB DT NN-NN RB
VB DT NNPOS NNS RB

ESTRUTURAS
VBN PRP RB
VB PRP RB
VBP PRP RB
VBN RB
VBZ PRP\$ NN RB
VB PDT DT JJ NNS RB

Fonte: O autor (2024).

Nas obras em que *away* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural) e *away* em RB (advérbio). Dessa forma, simplificando, *take away* é constituído, em algumas ocorrências, por um verbo e um advérbio e em outras por um verbo e uma preposição.

A análise linguística realizada revela nuances na utilização da palavra *take* em diferentes contextos, conforme classificado pelo *SpaCy*. O estudo identifica uma variedade de estruturas linguísticas, envolvendo adjetivos, preposições, pronomes possessivos, substantivos e determinantes, como observado nos excertos A, B e C, retirados dos *readers*:

Excerto A:

I know he has the same dream every night. In his dream the rebels attack his home and take him away from his mother. That's what happened to all his brothers and sisters.

Fonte: Remember Atita, Jackee Budesta Batanda.

Excerto B:

Now, mister, you can take all the damn things away,' she said breathlessly, 'because I have no money to buy anything.' Again, the packman did not hear her, but said tiredly, 'This you like, good lady.'

Fonte: Mr Sing My Hearts Delight, BrianFriel.

Excerto C:

I see it is not a man, but a boy. He is about seventeen years old. I take my hand away from his arm. 'Who are you? And what are you doing here?' I ask.
 Fonte: Jacks Came, Rod Smith.

Essa versatilidade de elementos gramaticais destaca a gama de vocabulários que podem ser aprendidos com os *readers* do nível A2.

Quanto ao seu significado, conforme *Oxford's Learner's Dictionaries*, salienta-se que *take away* é classificado como um *phrasal verb*. Este termo denota a ação de dissipar um sentimento ou removê-lo. Tal conotação é perceptível nas instâncias presentes no *subcorpus* A2, como ilustrado nas linhas de concordância na Figura 44:

Figura 44 — Linhas de Concordância de *take away* - A2

N°	Left	Node	Right
1	came to take the horse	away,	and the Duke and Duchess
2	afternoon and I'll take	away	the spell. Tell all the
3	you, who comes to take	away	my daughter, even in my
4	the sea will take us	away."	We were very close to
5	told to take his wife	away	for a change of air. '
6	you to take the necklace	away	from Freyja!' 'Very well,' said
7	him. 'Take the wood-cutter	away,'	he said to the dog
8	Take the king's men	away,'	Hans said to the dog
9	his home and take him	away	from his mother. That's

Fonte: O autor (2024).

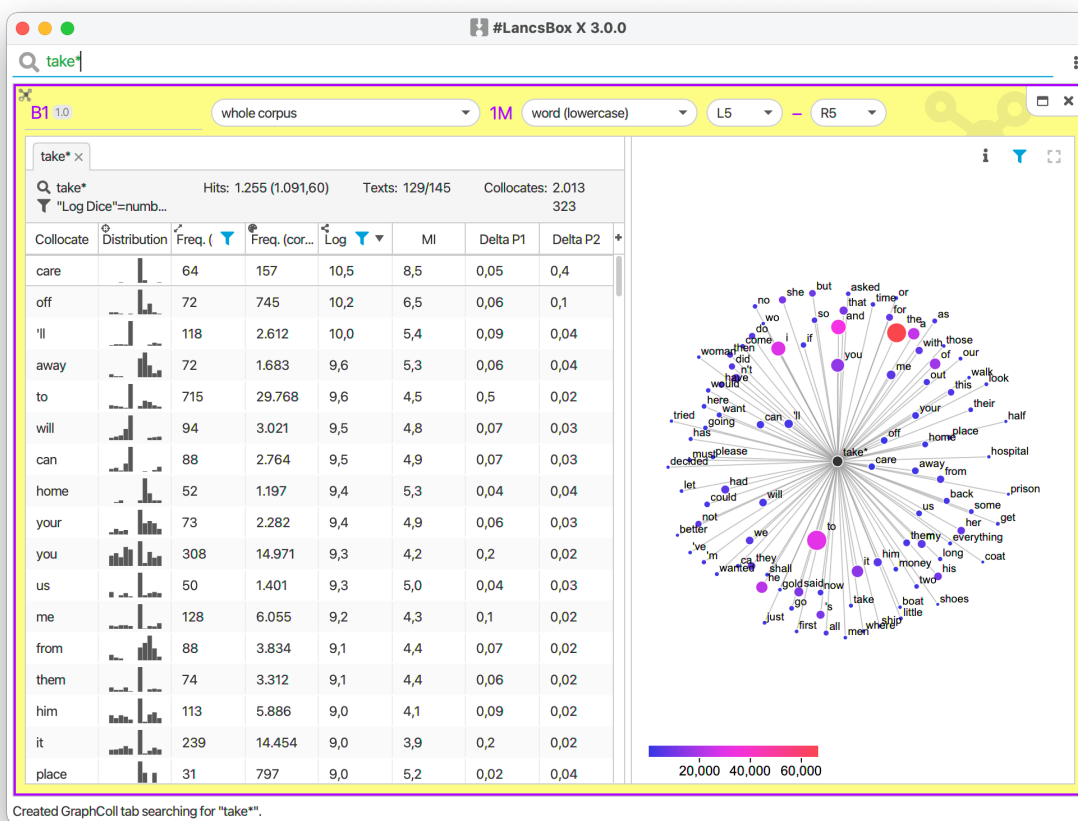
Portanto, o significado atribuído ao *phrasal verb take away* sugere uma ação de remover ou eliminar algo, seja físico ou emocional. Esse sentido parece ser consistente com os exemplos fornecidos, como retirar um cavalo, desfazer um feitiço, ou afastar alguém para uma mudança de ar. Dessa forma, a variedade de contextos nos quais essa expressão é empregada destaca sua versatilidade e adaptabilidade em diferentes situações. Além disso, a observação de que essa definição é encontrada nas ocorrências do *subcorpus* A2 sugere que esse uso específico do *phrasal verb* pode ser mais comum em contextos linguísticos de nível de proficiência mais básico. Isso pode indicar que a expressão *take away* é acessível e utilizada em contextos mais simples de comunicação, o que tem implicações interessantes para o ensino e aprendizado de inglês como língua estrangeira.

Em suma, no nível A1, *take off* é usado principalmente para descrever ações habituais e significa remover uma peça de roupa (79,48%). Já no nível A2, além desse significado, *take off* pode significar remover algo de um lugar/de alguém. Essa mudança de significado está relacionada à estrutura gramatical: verbo + advérbio para remover roupa e verbo + preposição para remover algo de um local. *Take care*, por outro lado, apresenta a mesma estrutura e significados em ambos os níveis: cuidar de alguém/algo/si mesmo e ter cautela. A principal diferença é que, no nível A2, além da estrutura verbo + substantivo, há ocorrências com adjetivos e advérbios entre o verbo e o substantivo, porém com menor frequência. Por fim, as análises dos *subcorpora* A1 e A2 demonstram que os usos de *take off* e *take care* nos *readers* analisados variam de acordo com o nível de proficiência linguística e o contexto específico.

5.9 Colocações com o verbo *take* no *subcorpus* b1

Repetindo o mesmo processo, porém agora no *subcorpus* B1, digitamos na barra de pesquisa da ferramenta *Graphcoll* a palavra *take* para obter os seus **colocados**. Obtivemos os seguintes dados:

Figura 45 — Take - B1



Fonte: O autor (2024).

O **nó** em análise ocorre 1.255 vezes no *subcorpus* B1, distribuído em 129 dos 145 textos, com um total de 2.013 **colocados**. Após a implementação do CPN, proposto por Brezina, McEnery e Wattam (2015), restaram 323 **colocados**. Considerando que nosso intento consiste em identificar as colocações mais exclusivas do verbo, deduzimos que, com base nos resultados dos cálculos estatísticos e seguindo nossos critérios de seleção, os **colocados** *care* e *off* são os mais distintivos.

O Quadro 33 exhibe os respectivos valores de co-ocorrência com o **nó**, sua frequência de ocorrência no *corpus* e os resultados do cálculo do método estatístico adotado para esta pesquisa.

Quadro 33 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>care</i>	64	157	10,5
<i>Off</i>	72	754	10,2

Fonte: O autor (2024).

5.9.1 Care

Através da etiquetagem, examinamos todas as estruturas distintas formadas nas 62 ocorrências de *take* com o **nó** em questão. Tais estruturas podem ser observadas no Quadro 34.

Quadro 34 — Estruturas *take care* - B1

ESTRUTURAS
VB NN
VBN NN
VB JJR NN
VB JJ NN
VBP JJ NN
VBN JJ NN
VBN NN

Fonte: O autor (2024).

Nas obras em que *care* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural) e *care* em NN (substantivo). Dessa forma, simplificando, *take care* é constituída, em todas as ocorrências, por um verbo e um substantivo.

A análise das estruturas formadas com o **nó** *care* em combinação com o verbo *make* revela diversas construções gramaticais, incluindo diferentes formas de verbos (VB, VBN, VBP) combinadas com adjetivos (JJR, JJ) e o substantivo *care* (NN). Essa variedade de estruturas destaca a flexibilidade sintática da palavra *care* nos *readers*

analisados, permitindo sua combinação com diferentes elementos gramaticais, como visto nas linhas de concordância, na Figura 46, retiradas do *subcorpus* B1:

Figura 46 — Linhas de Concordância com POS de *take care* - B1

N°	Left	Node	Right			
1	Our nurses can take better	care	of them there.' They did	PRP\$ NNS MD VB JJR	NN	IN PRP RB.' PRP VBD
2	and artless eye. "Take good	care	of my house. Alas! Only	CC JJ NN. "VB JJ	NN	IN PRP\$ NN. UH! RB

Fonte: O autor (2024).

Apesar de ocorrer essas estruturas com adjetivo entre o verbo e o colocado, elas integram apenas 9,67% das ocorrências, sendo a estrutura Verbo + Substantivo a mais comum, com 90,32% das ocorrências. Esse dado sugere que a expressão mais comum no *subcorpus* B1 é *take care*, podendo ser enfatizada, embora não frequente, por adjetivos, como *better* e *good*, como mostrado nos exemplos I e II.

5.9.2 Off

Por meio da etiquetagem, examinamos todas as estruturas distintas formadas nas 63 ocorrências de *take* com o **nó** em questão. Tais estruturas podem ser observadas no Quadro 35.

Quadro 35 — Estruturas *take off* - B1

ESTRUTURAS
VB RP
VB PRP\$ NNS RP
VBP RP
VB PRP\$ NNS RB
VB DT NN RP
VBN DT NN IN
VB PRP\$ NNS IN
VB DT NN IN
VBN PRP\$ NNS RP
VB PRP RP
VB PRP\$ JJ NNS RP

ESTRUTURAS
VB PRP\$ NN RP
VBN RP
VBN DT NNS RP
VB DT NN NNS IN
VBZ RP

Fonte: O autor (2024).

Nas obras em que *off* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural) e *off* em RP (advérbio) e IN (preposição). Dessa forma, simplificando, *take off* é constituído, em algumas ocorrências, por um verbo e um advérbio e, em outras, por um verbo e uma preposição.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take off* é um *phrasal verb* de nível A2, possuindo diversos significados, sendo um deles remover algo, especialmente uma peça de roupa do corpo seu/de alguém, e remover algo/algumas coisas de algum lugar. Esses dois significados podem ser encontrados nas ocorrências do *subcorpus* B1, como visto nas linhas de concordância na Figura 47:

Figura 47 — Linhas de Concordância de *take off* - B1

Nº	Left	Node	Right
1	Don't take your glasses	off	until I tell you,' he
2	last week. If I take	off	the skin, it'll be
3	don't take my clothes	off -	well, not all of them.
4	Could you take the shaft	off	underwater?' she asked. 'And then
5	Pa had taken the canvas	off	the top of the house.
6	could not take her eyes	off	Dr Tan. He talked with
7	to his hair. 'Take it	off!	Take it off!' cried Fritz.
8	places." "Take your dirty hands	off	me, you ugly old lesbian,"
9	the room. 'Take his head	off	outside!' shouted the Queen. Two
10	me jump! Take my ring	off!	It's not yours any
11	recognized him. He had taken	off	his tie, and round his
12	aunt had taken the covers	off	the furniture and put flowers
13	to take the supper things	off	the table, 'Leave the coffee,
14	couldn't take his eyes	off	the picture. Even when the
15	that it was. The take-	off	was good, and soon I

Fonte: O autor (2024).

A mudança de significado está intrinsecamente relacionada com a estrutura do colocado. No *subcorpus* B1, a expressão *take off* ocorre 85,24% das vezes com a estrutura Verbo + Advérbio, possuindo o sentido de retirar uma peça de roupa, e 14,75% das vezes com a estrutura Verbo + Preposição, com o sentido de retirar algo de algum lugar. As duas estruturas podem ser observadas, respectivamente, nos excertos A e B, retirados do *corpus* a seguir.

Excerto A:

She was a little afraid of him out there in the darkness. In the hall, he sat down heavily. As he bent down to take off (VERBO + ADVÉRBIO) his boots, he didn't seem blind.

Fonte: The Blind Man, D H Lawrence.

Excerto B:

And they've been very hungry. How did they find enough food to stay alive?" "Hunting? Stealing maybe?" Nell suggested. "I saw the old dog take some food off (VERBO + PREPOSIÇÃO) a plate on the table after breakfast!"

Fonte: The Incredible Journey, Burnford Sheil.

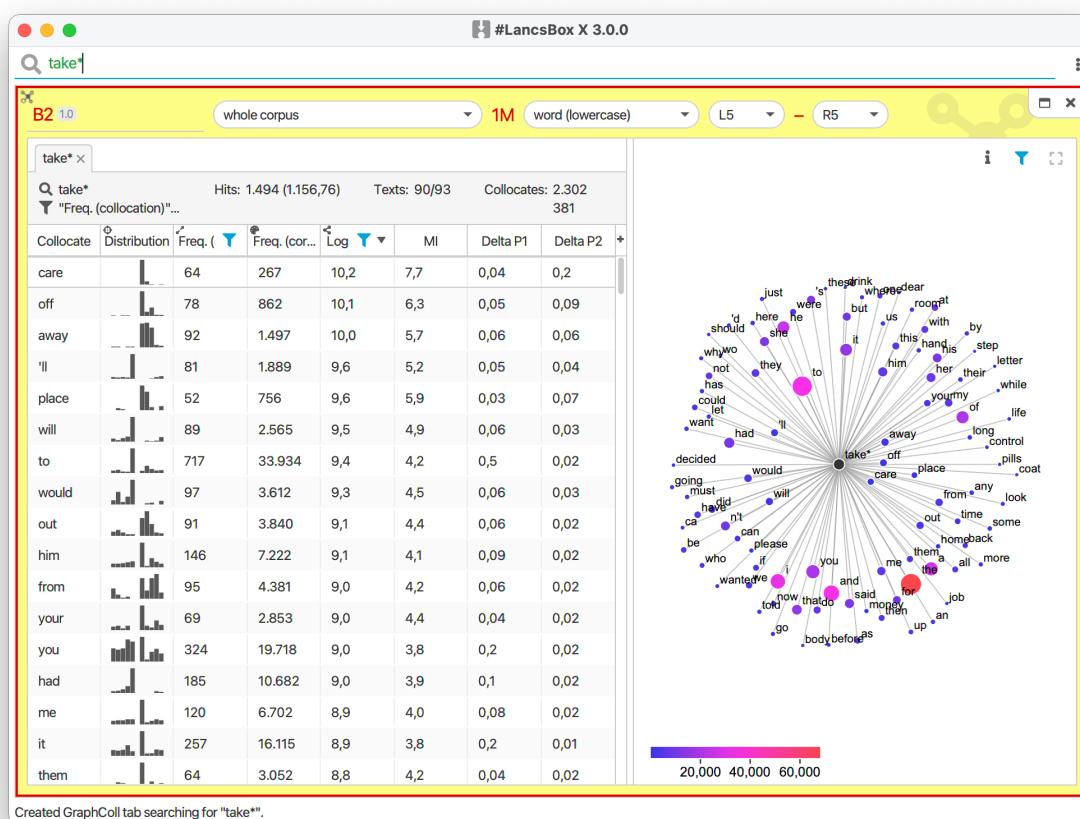
Isso indica que a escolha da preposição influencia significativamente o significado da expressão. No primeiro caso, a ausência da preposição indica que o

foco está na ação de tirar a roupa do corpo, enquanto no segundo caso, a preposição *off* indica a origem de onde algo está sendo retirado.

5.10 Colocações com o verbo *take* no subcorpus *b2*

Repetindo o processo, porém agora no *subcorpus* B2, digitamos na barra de pesquisa da ferramenta *Graphcoll* a palavra *take* para obter os seus **colocados**. Obtivemos os seguintes dados:

Figura 48 — *Take* - B2



Fonte: O autor (2024).

O **nó** em questão ocorre 1.494 vezes no *subcorpus* B2, em 90 dos 93 textos, possuindo 2.302 **colocados**. Desse total de colocados, após a aplicação do CPN, propostas por Brezina, McEnery e Wattam (2015), sobraram 381 **colocados**. Uma vez que nosso objetivo é identificar as colocações mais exclusivas do verbo,

concluimos que, de acordo com o resultado do cálculo estatísticos e os nossos critérios de seleção, os colocados, *care*, *off* e *away* são os mais exclusivos.

No Quadro 36, são mostrados os respectivos valores de co-ocorrência com o **nó**, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para essa pesquisa.

Quadro 36 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Care</i>	64	267	10,2
<i>Off</i>	78	862	10,1
<i>Away</i>	92	1.497	10,0

Fonte: O autor (2024).

5.10.1 Care

A palavra *care* está classificada como um vocabulário de nível A2 e gramaticalmente se categoriza como um substantivo. Segundo a etiquetagem do *corpus*, todas as ocorrências do colocado foram classificadas como NN (substantivo). Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 63 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 37:

Quadro 37 — Estruturas de *take care* - B2

ESTRUTURA
VB NN
VB JR NN
VBP NN
VBZ NN
VBN NN
VB JJ NN
VB DT NN
VBP JJ JJ NN

Fonte: O autor (2024).

Nas obras em que *care* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base), VBP (verbo na primeira pessoa do singular ou plural) e *care* em NN (substantivo). Dessa forma, simplificando, *take care* é constituída, em todas as ocorrências, por um verbo e um substantivo.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take care* é um *idiom* de nível A2, possuindo diversos significados como tomar cuidado, e cuidar de alguém/algo/de si mesmo e ter cuidado com algo, como já mencionado. Ao verificar as sentenças no *subcorpus* em questão, constatou-se a ocorrência de *take care* com esses significados, como pode ser conferido nas linhas de concordância, na Figura 49, retiradas do *subcorpus* B2:

Figura 49 — Linhas de Concordância de *take care* - B2

N°	Left	Node	Right
1	even more reason to take	care	of his wife and daughters.
2	dramatist did not take more	care	of his work. To understand
3	right that their children take	care	of them, and make their
4	people believe that he takes	care	of the island and those
5	her. 'Now everything's taken	care	of,' she told him. She
6	I can't even take	care	of my own son?' 'Well,
7	dog with you and take	care	of her." With that they

Fonte: O autor (2024).

A análise das estruturas formadas com o **nó** *care* em combinação com o verbo *take* revela uma construção gramatical comum, sendo Verbo + Substantivo, com 87,30% das ocorrências, como visto na linha de concordância 1, e construções menos frequentes incluindo adjetivos entre o **nó** e o colocado, com 12,69% das ocorrências, como visto na linha de concordância 2. Além disso, observamos que a expressão *take care* ocorre em construções no *subcorpus* B2 que não são comuns nos *subcorpora* A1 e A2, como evidenciado na linha de concordância 5. Nesse exemplo, o **nó** está classificado como verbo no participípio em uma sentença no *present perfect*, caracterizada pelo uso do auxiliar *has* com um verbo no *past participle*.

5.10.2 Off

Através da etiquetagem, examinamos todas as estruturas distintas formadas nas 74 ocorrências de *take* com o **nó** em questão. Tais estruturas podem ser observadas no Quadro 38.

Quadro 38 — Estruturas de *take off* - B2

ESTRUTURAS
VB RP
VCN IN DT NN NN IN
VCN PRP DT IN
VCN RP
VB DT NNS IN
VCN IN
VBP PRP\$ NNS RP
VBZ PRP\$ NN RP
VB PRP VB RP
VB DT NN RP
VBP PRP RP
VB PRP\$ JJ NN IN
VB PRP IN
VB PRP\$ NN RP
VB PRP\$ NNS IN
VB NN IN
VCN PRP\$ NN RP
VB PRP RP
VCN PTD DT NNS IN
VB NNS RP

Fonte: O autor (2024).

Nas obras em que *off* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VBZ (verbo na terceira pessoa do singular), VB (verbo base) e VBP (verbo na primeira pessoa do singular ou plural) e *off* em RP (advérbio) e IN

(preposição), como pode ser observado nas linhas de concordância, na Figura 50, retiradas do *subcorpus* B2:

Figura 50 — Linhas de Concordância de *take off* - B2

N°	Left	Nod	Right
1	rest.' Newman made Nicholas take	off	his coat and washed the
2	taken to a waiting room	off	the corridor, which ran straight
3	you've taken us all	off	your list?' She suddenly saw
4	as your plane had taken	off,	we arrested the French woman
5	began to take the clothes	off	her still body. The younger
6	soon as he is taken	off	the road. Behind the wheel
7	now. We take our uniforms	off	and put them in the
8	As she takes her coat	off,	Branwell watches her. She has
9	so she takes her jacket	off,	and in the mirror she
10	rest. Now help me take	off	the lid."
11	They helped him to take	off	his armor, and to carry

Fonte: O autor (2024).

Assim como no *subcorpus* anterior, a expressão *take off* é formada tanto por um verbo e um advérbio, quanto por um verbo e uma preposição, possuindo também as mesmas características em relação ao seu significado. O que difere a expressão do *subcorpus* B1 do B2 é somente a porcentagem de ocorrência das estruturas com advérbio (72,58%) e com preposição (27,41%). Ainda assim, esse colocado se comporta de maneira bem semelhante, pois até mesmo a porcentagem das ocorrências é equiparada ao nível B1 (85,24% e 14,75%).

5.10.3 Away

Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 43 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 39:

Quadro 39 — Estruturas de *take away* - B2

ESTRUTURAS
VB PRP RB
VB DT RB
VB DT NN RB
VB PRP\$ NNS RB
VCN RB
VBP DT RB
VB PRP\$ NN RB
VB-RP
VCN NNPOS NN RB
VB NNP RB
VBZ DT NN RB

Fonte: O autor (2024).

No contexto específico das ocorrências da expressão *take away* no *corpus* analisado, observa-se que o *SpaCy* classificou *take* como VBZ (verbo na terceira pessoa do singular), VB (verbo base) e VBP (verbo na primeira pessoa do singular ou plural) e *away* como RB (advérbio). Essa classificação indica que *take away* é constituída por um verbo e um advérbio.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take away* é um *phrasal verb*, possuindo o significado de fazer algo ou algum sentimento desaparecer, ou ser retirado. Esse significado pode ser encontrado nas ocorrências do *subcorpus* B2, como destacado nas linhas de concordância, na Figura 51:

Figura 51 — Linhas de Concordância de *take away* - B2

Nº	Left	Node	Right
1	I can't take you	away,	I can't leave here
2	Now, Captain Marvin, take them	away.'	Scanlan had delivered a note
3	terrible,' she complained. Take it	away.	I want a glass of
4	with the breakfast.	away,'	That girl want to
5	calmly. 'We'll take that	away,	Constable. We'll take those
6	come to take the body	away.	Now then... is there a
7	him to take the boy	away	from London as soon as
8	you to take the body	away	quietly.'
9	could not take my eyes	away	from these creatures.
10	the man outside had taken	away	all his powers of thought.
11	brightly. The Martians had taken	away	the digging-machine and apart

Fonte: O autor (2024).

A análise das estruturas nos possibilitou perceber que a expressão *take away*, do *subcorpus* B2, se assemelha a do *subcorpus* A2. A maior diferença está no tempo verbal que o **nó** ocorre nas sentenças. Em específico, o **nó** no *subcorpus* B2 ocorre no *present perfect*, caracterizado pelo uso do auxiliar *Have/has* e o verbo no participípio. Tais estruturas podem ser observadas nos excertos A, B e C retirados do *subcorpus* B2:

Excerto A:

a) *Perhaps this attack from Mars will be helpful to us in the end. It has taken away our confidence in the future, which was making us soft; it has given great help to science, and it has made us think of human beings as one family.*

Fonte: The War Of The Worlds, H G Wells.

Excerto B:

b) *And it was a beautiful and interesting world. 'But probably the machine had only been taken away. I must be calm, find its hiding-place and get it back by force and cleverness. I stood up and looked around me, wondering where I could wash.*

Fonte: Time Machine, H G Wells.

Excerto C:

c) *She waited for him to say something, to become angry; then she could explain why the chicken was taken away 'for his good'. But Conradin said nothing.*

Fonte: Wicked and Humorous, Tales Sak.

A estrutura no *present perfect* possui 29,54% das ocorrências, o que sugere que, ocasionalmente, pode ser encontrada pelos alunos nesse *subcorpus* em questão.

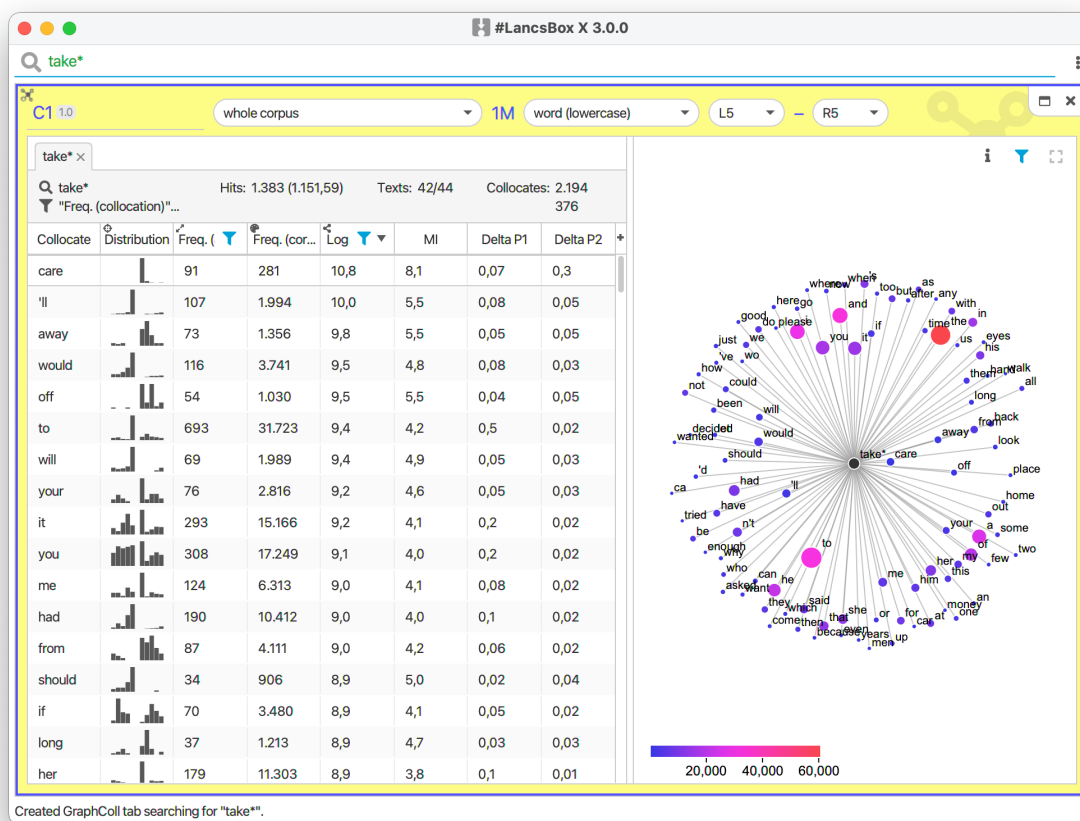
Em suma, a análise das estruturas envolvendo as expressões *take care*, *take off* e *take away* nos *subcorpora* B1 e B2 revela padrões distintos de uso e significado nos *readers* analisados. No *subcorpus* B1, a expressão *take care* é predominantemente empregada na estrutura Verbo + Substantivo, com uma incidência significativa de 90,32%, sugerindo uma ênfase na ação de cuidar, ocasionalmente intensificada por adjetivos como *better* e *good*. Por sua vez, *take off* é utilizada em duas construções distintas: Verbo + Advérbio, denotando a remoção de vestimentas e Verbo + Preposição, indicando a retirada de objetos de um lugar específico.

No entanto, no *subcorpus* B2, a expressão *take care* é também predominante na estrutura Verbo + Substantivo, porém com a adição de construções menos frequentes que incluem adjetivos, bem como ocorrências em contextos não comuns nos *subcorpora* A1 e A2. Quanto a *take away*, sua estrutura e significado se assemelham aos do *subcorpus* A2, mas com uma maior incidência no *present perfect*. Portanto, as análises apontam para uma variação tanto na estrutura quanto no contexto de uso dessas expressões entre os *subcorpora* B1 e B2, o que pode indicar uma progressão na competência linguística dos aprendizes, ao longo dos níveis de proficiência.

5.11 Colocações com o verbo *take* no *subcorpus* c1

Seguindo o procedimento previamente estabelecido, porém agora no *subcorpus* C1, inserimos na barra de pesquisa da ferramenta Graphcoll a palavra *take* com o intuito de obter seus termos associados. Os resultados obtidos foram os seguintes:

Figura 52 — Take - C1



Fonte: O autor (2024).

O nó em questão ocorre 1.383 vezes no *subcorpus* C1, em 42 dos 44 textos, possuindo 2.194 **colocados**. Desse total de colocados, após a aplicação do CPN, propostas por Brezina, McEnery e Wattam (2015), sobraram 376 **colocados**. Uma vez que nosso objetivo é identificar as colocações mais exclusivas do verbo, e como elas se comportam nos diferentes níveis de proficiência estabelecidos pelos *readers*, considerando seu significado e sua estrutura léxico-gramatical, concluímos que, de acordo com o resultado do cálculo estatístico e dos nossos critérios de seleção, o colocado *care* é o mais exclusivo dentre os 376.

No Quadro 40, são mostrados os respectivos valores de co-ocorrência com o nó, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para esse estudo.

Quadro 40 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Care</i>	91	281	10,8

Fonte: O autor (2024).

5.11.1 Care

Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 90 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 41:

Quadro 41 — Estruturas de *take care* - C1

ESTRUTURAS
VCN NN
VB NN
VBP NN
VB JJ NN
VB DT JJS NN
VBZ NN
VB IN NN
VCN JJR NN
VB JJ NN
VB JJR NN

Fonte: O autor (2024).

Nas obras em que *care* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em (Verbo na 3ª pessoa do singular), VB (Verbo na forma base), e VBP (Verbo na 1ª e 2ª pessoa do singular e no plural) e *care* em NN (substantivo). Dessa forma, simplificando, *take care* é constituída, em todas as ocorrências, por um verbo e um substantivo. Assim como no *subcorpus* A2, foi identificado também outros elementos gramaticais entre o **nó** e o colocado, como JJ (adjetivo) e RB (advérbio), que também aparecem em pequenas proporções, com o total de 6,66% das

ocorrências. Dessa maneira, 93,33% das estruturas são compostas por Verbo + Substantivo.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take care* é um *idiom* de nível A2, possuindo diversos significados. Ao verificar as ocorrências no *subcorpus* C1, encontramos dois sentidos: tomar cuidado e cuidar de alguém/algo/de si mesmo. Os exemplos correspondentes estão, respectivamente, apresentados nas linhas de concordância, na Figura 53, retiradas do *subcorpus* C1:

Figura 53 — Linhas de Concordância de *Take care* - C1

N°	Left	Node	Right
1	her, panting. "Good-bye. Take	care,"	he said, stopping by the
2	me?" "Inspector Rock will take	care	of that," I said. "It

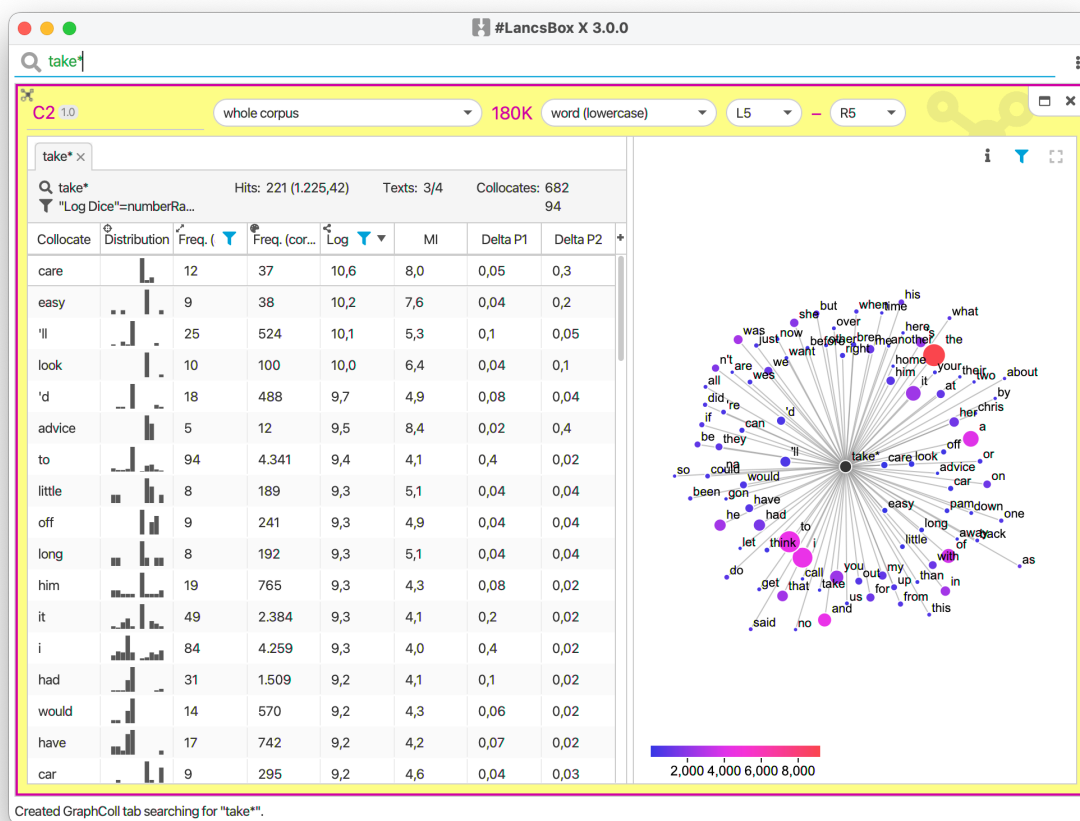
Fonte: O autor (2024).

Em resumo, a expressão *take care* do *subcorpus* C1 se assemelha com a do *subcorpus* A2.

5.12 Colocações com o verbo *take* no *subcorpus* c2

Repetindo o procedimento anterior, contudo, agora no *subcorpus* C2, inserimos a palavra *take* na barra de pesquisa da ferramenta *Graphcoll* para obter seus **colocados**. Os dados obtidos são os seguintes:

Figura 54 — Take - C2



Fonte: O autor (2024).

O nó em questão ocorre 221 vezes no *subcorpus* C2, em 3 dos 4 textos, possuindo 682 **colocados**. Deste total de colocados, após a aplicação do CPN, propostas por Brezina, McEnery e Wattam (2015), sobraram 94 **colocados**. Uma vez que nosso objetivo é identificar as colocações mais exclusivas do verbo, concluímos que, de acordo com o resultado do cálculo estatísticos e os nossos critérios de seleção, os **colocados** *care*, *easy* e *look* são os mais exclusivos.

No Quadro 35, são mostrados os respectivos valores de co-ocorrência com o nó, sua frequência de ocorrência no *corpus* e o resultado do cálculo do método estatístico escolhido para essa pesquisa.

Quadro 42 – Frequência da colocação

Colocado	frequência da colocação	frequência no <i>corpus</i>	<i>Log Dice</i>
<i>Care</i>	12	37	10,6
<i>Easy</i>	9	38	10,2
<i>Look</i>	10	100	10,0

Fonte: O autor (2024).

5.12.1 Care

Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 12 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 43:

Quadro 43 — Estruturas de *take care* - C2

ESTRUTURAS
VB NN
VBN NN
VB DT JJS NN
VB JJ NN
VBP NN

Fonte: O autor (2024).

Em comparação com as ocorrências de *take care* nos *subcorpora* A2 e C1, as estruturas de *take care* no C2 são semelhantes, tanto no significado quanto em suas construções com Verbo + Substantivo e Verbo + Adjetivo + Substantivo, possuindo também uma maior porcentagem de sentenças com a construção Verbo + Substantivo (75% das ocorrências). Esses dados nos revelam a constância do significado e de classes gramaticais presentes na expressão *take care* no *corpus* construído para esta pesquisa.

5.12.2 Easy

Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 9 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 44.

Quadro 44 — Estruturas de *take easy* - C2

ESTRUTURAS
VB PRP RB

Fonte: O autor (2024).

Nas obras em que *easy* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VB (verbo base) e *easy* em RB (advérbio). Dessa forma, simplificando, *take easy* é constituída por um verbo e um advérbio.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take it easy* é um *idiom*, possuindo o significado de dizer a alguém para não ficar preocupado ou zangado. Ao verificar as sentenças no *subcorpus* em questão, constatou-se a ocorrência de *take it easy* com esses significados, como pode ser visto nas linhas de concordância na Figura 55:

Figura 55 — Linha de Concordância de *take easy* - C2

N°	Left	Node	Right
1	take it real nice and	easy."	Elliot backed up against the
2	Hall -" "Awright. Awright. Take it	easy,"	McCormick interrupted. "I'll put
3	Morgan wants to take it	easy,"	Alexander said, "then he should
4	Hammit. "Hi, Wes! Take it	easy	on the shoulder, huh?" "Sorry,
5	words tumbling out. "Take it	easy,	Joan. You want him and
6	'll have to take it	easy	for awhile, and Wes got
7	it wasn't. Take it	easy."	It wasn't much to

Fonte: O autor (2024).

Além disso, em todas as ocorrências houve a presença da preposição *it*, formando assim, *take it easy*. A observação da presença sistemática dessa preposição nas construções reforça a ideia de que algumas expressões idiomáticas têm elementos fixos. Isso pode incentivar os alunos a identificar e compreender padrões linguísticos.

5.12.3 Look

Segundo a etiquetagem do *corpus*, todas as ocorrências do colocado foram classificadas como NN (substantivo). Isso ocorre devido à posição do *look* depois de um artigo, sendo assim, ele está exercendo a função de um substantivo na estrutura *take a Look*. Verificamos, a partir da etiquetagem, todas as estruturas diferentes que foram formadas nas 9 ocorrências de *take* com o **nó**. As estruturas podem ser observadas no Quadro 45:

Quadro 45 — Estruturas de *take look* - C2

ESTRUTURAS
VBP DT NN
VB DT NN

Fonte: O autor (2024).

Nas obras em que *look* ocorreu como colocação de *take*, o *SpaCy* classificou a palavra *take* em VB (verbo base) e VBP (verbo na primeira pessoa do singular ou plural) e *look* em NN (substantivo). Dessa forma, simplificando, *take look* é constituída por um verbo e um substantivo.

Em relação ao seu significado, o *Oxford's Learner's Dictionaries* informa que *take a look* está classificado como nível B1, usado com substantivos para dizer que alguém está fazendo algo, realizando uma ação *etc.* Ao verificar as sentenças no *subcorpus* em questão, constatou-se a ocorrência de *take look* com esses significados, como pode ser visualizado nas linhas de concordância na Figura 56:

Figura 56 — Linhas de Concordância de *take look* - C2

N°	Left	Node	Right
1	seventy. You better take a	look	at it." "You must have
2	asked me to take a	look-	see. Here's what I
3	Now, let's take a	look	at something else." Bren led
4	mind. "Let's take a	look	around the area before we
5	When can I take a	look	at the car, anyway?" Wes
6	grinned. "Out to take a	look	at a car." "A Porsche?" "
7	and we'll take a	look-	see at what's going
8	and let him take a	look	at that package before you
9	want us to take a	look	for him, and I'm

Fonte: O autor (2024).

A classificação da palavra *look* como pertencente ao vocabulário de nível A1 e B1 e sua categorização gramatical como um verbo indicam que essa palavra é amplamente acessível e usada em diferentes níveis de proficiência em inglês. No entanto, no contexto específico da expressão *take a look*, *look* está sendo usado como um substantivo singular (NN), exercendo a função de um objeto direto.

A análise das estruturas formadas com o **nó** *look* em combinação com o verbo *take* revela que as ocorrências no *corpus* se limitam principalmente a duas construções gramaticais: "VBP DT NN" e "VB DT NN". Ao identificar as estruturas gramaticais comuns associadas à colocação *take a look*, os alunos ganham *insights* sobre como certas palavras, nesse caso, *look*, são frequentemente utilizadas após determinados verbos. Isso contribui para uma compreensão mais abrangente dos padrões gramaticais na Língua Inglesa. Dessa forma, podemos concluir que *take a look* é uma expressão comum que indica a ação de observar ou inspecionar algo, e sua estrutura gramatical envolve um verbo seguido por um substantivo, destacando a natureza visual ou observacional da ação.

Em resumo, a análise das estruturas verbais envolvendo as expressões *take care*, *take it easy* nos *subcorpora* C1 e C2 revela uma consistência significativa tanto no significado quanto na construção gramatical dessas expressões em níveis de proficiência avançados. Tanto em C1 quanto em C2, a expressão *take care* é predominantemente utilizada na estrutura Verbo + Substantivo, com uma ocorrência mínima de elementos gramaticais adicionais como adjetivos e advérbios, refletindo

uma semelhança notável com seu uso no *subcorpus* A2. Além disso, a presença sistemática da preposição *it* na expressão *take it easy* no *subcorpus* C2 destaca a natureza fixa de certas expressões idiomáticas.

No desfecho desse capítulo de análise, exploramos as colocações mais significativas dos verbos *take* e *make* nos *subcorpora* A1, A2, B1, B2, C1 e C2. A investigação desses colocados proporcionou uma melhor compreensão das nuances e variações de uso das expressões envolvendo os colocados nos *subcorpora*, revelando padrões gramaticais e semânticos iguais ou distintos em cada nível de proficiência linguística. Os dados analisados podem constituir um bom acervo para a elaboração de materiais didáticos, especialmente voltados para o ensino de colocações em sala de aula. Como prova disso, no Anexo 1 desta pesquisa, se encontra uma sequência didática que pode ser aplicada na sala de aula, a qual abrange a abordagem DDL. O enfoque na abordagem DDL emerge como uma estratégia pedagógica eficaz, permitindo que os educadores incorporem esses dados linguísticos de maneira pragmática e direcionada, enriquecendo assim a experiência de aprendizado dos alunos. A aplicação prática desses resultados na criação de materiais didáticos pode potencialmente aprimorar a competência linguística dos aprendizes, proporcionando uma abordagem mais contextualizada e envolvente no ensino de colocações. No próximo capítulo, iremos abordar como os resultados atendem aos objetivos da pesquisa e às perguntas que estabelecemos neste trabalho.

CONSIDERAÇÕES FINAIS

Como delineado nos capítulos iniciais, autores como Vieira (2008), Todorov (2009) e Rorty (2001) possibilitaram o entendimento de que a inclusão da literatura no contexto educacional é essencial, proporcionando não apenas uma compreensão aprofundada da linguagem, mas também cultivando empatia, combatendo a alienação e ampliando horizontes imaginativos. Dessa forma, a pesquisa se dispôs a integrar a literatura no ensino de Língua Inglesa, para que tais benefícios pudessem ser alcançados, pois, conforme argumentam Bozza e Benenito (2019) e Collie e Slater (1987), a literatura permite aos alunos compreenderem a diversidade cultural e a própria língua alvo, desenvolvendo criticamente significados em relação ao mundo. Sendo assim, nosso intuito foi produzir um material pedagógico que pudesse aproximar a literatura das aulas de LI. Nessa perspectiva, mostramos que a LC e seu instrumental têm sido essenciais para pesquisas e estudos na área de ensino de línguas, oferecendo aos professores ferramentas valiosas para desenvolver materiais e atividades pedagógicas adaptadas às necessidades específicas da sala de aula. Ressaltamos que, conforme Oliva (2018), a utilidade dos exercícios baseados em *corpus* pode ampliar a compreensão dos alunos sobre o uso real da língua e abordar aspectos de vocabulário, gramática e cultura.

Pontuamos também que pesquisas como as de Römer (2004) e Frankenberg-garcia (2018) demonstram como o *corpus* pode ser utilizado para análise de fenômenos linguísticos, contribuindo para o desenvolvimento de materiais didáticos mais eficazes e fornecendo *insights* sobre o uso da linguagem em contextos específicos. Em resumo, discutiu-se como o uso do *corpus* no ensino de línguas oferece uma gama de benefícios tanto para professores quanto para alunos, facilitando o aprendizado e a compreensão da Língua em diferentes contextos.

Como abordagem de ensino, buscamos na LC o apoio teórico para a investigação de um *corpus* pedagógico. Essa fundamentação justificou a escolha deste estudo, que integrou a análises de colocações em um *corpus* especializado, composto de *Graded Readers*, ao ensino de línguas, visando priorizar as colocações relevantes de dois verbos de alta frequência em textos literários: *make* e *take*. Além disso, foi destacado que a pesquisa está amparada na abordagem DDL, que se caracteriza pelo ensino da língua orientada por dados provenientes de um *corpus*. Foi

discutido também que a DDL utiliza ferramentas da LC para fins pedagógicos, permitindo que os alunos analisem dados linguísticos reais e cheguem a conclusões por meio de investigações. Dessa maneira, foi investigado o fenômeno das colocações nos *Graded Readers*. Foi exposto que vários pesquisadores da linguística têm diferentes interpretações sobre colocação e optamos por uma abordagem estatística (Sinclair, 1991; Halliday, 1966) para uma identificação mais precisa das colocações. Portanto, concluímos que, ao combinar práticas de leitura com a abordagem DDL, os alunos podem aprender e ampliar mais vocabulário em Língua Inglesa. Dessa maneira, a nossa pesquisa sobre *Graded Readers* e colocações pode preencher a lacuna existente em relação às pesquisas aplicadas na educação básica, proporcionando benefícios tangíveis aos alunos no processo de aprendizado da Língua Inglesa.

A metodologia empregada na pesquisa constituiu da análise e identificação das colocações relevantes nos *subcorpora* A1, A2, B1, B2, C1, C2. Os resultados respondem aos objetivos que nortearam esta pesquisa, os quais foram:

- a) identificar e selecionar as colocações mais relevantes e exclusivas presentes nos *subcorpora* dos verbos *make* e *take*;
- b) identificar as construções gramaticais das colocações mais exclusivas e relevantes de *make* e *take* presentes nos *subcorpora* (verbo + 155rep., verbo + substantivo, etc.), identificando e analisando as semelhanças e diferenças;
- c) Elaborar uma atividade didática usando a perspectiva DDL a partir dos dados encontrados da análise das colocações.

Em relação ao primeiro objetivo, o *software LancsBox X* possibilitou identificar e selecionar as colocações mais relevantes e exclusivas, levando em consideração os nossos critérios de seleção e o cálculo estatístico estabelecido. Isso proporcionou uma melhor compreensão das colocações associadas aos verbos *make* e *take* em um *corpus* especializado. No que se refere ao verbo *make*, o termo ocorre em todos os *subcorpora*, cada um apresentando uma distribuição variada de colocados. Após a aplicação do *Collocation Parameters Notation* (CPN), proposto por Brezina, McEnery e Wattam (2015), foram retidos apenas os colocados considerados mais relevantes para o estudo. Os resultados podem ser observados na Figura 58:

Figura 57 — Resultados de *Make* obtidos com o *LancsBox*

RESULTADOS DE MAKE			
<i>Subcorpus</i>	Nº de ocorrência e total de colocados	Distribuição no <i>corpus</i>	Colocações Relevantes
A1	48 ocorrências e 231 colocados	17 dos 23 textos	noise
A2	353 ocorrências e 947 colocados	124 dos 178 textos	sure, feel, rich, tea, friends, noise e money
B1	1.010 ocorrências e 1.829 colocados	123 dos 145 textos	sure
B2	1.237 ocorrências e 2.074 colocados	87 dos 93 textos	sure, sense, feel, money e happy
C1	1.146 ocorrências e 2.058 colocados	41 dos 44 textos	sure
C2	188 ocorrências e 654 colocados	4 dos 4 textos	sense e difference

Fonte: O autor (2024).

O verbo *take* também ocorre em todos os *subcorpora*, apresentando os seguintes resultados observados na Figura 59:

Figura 58 — Resultados de *take* obtidos com o *LanCSBox*

RESULTADOS DE TAKE			
<i>Subcorpus</i>	Nº de ocorrência e total de colocados	Distribuição no <i>corpus</i>	Colocações Relevantes
A1	48 ocorrências e 387 colocados	22 dos 23 textos	off
A2	765 ocorrências e 1.434 colocados	154 dos 178 textos	care, off e away
B1	1.255 ocorrências e 2.013 colocados	129 dos 145 textos	care e off
B2	1.494 ocorrências e 2.302 colocados	90 dos 93 textos	care, off e away
C1	1.383 ocorrências e 2.194 colocados	42 dos 44 textos	care
C2	221 ocorrências e 682 colocados	3 dos 4 textos	care, easy e look

Fonte: O autor (2024).

Em relação ao segundo objetivo, os resultados da pesquisa também permitiram identificar as construções gramaticais das colocações mais exclusivas e relevantes dos verbos *make* e *take*, presentes nos *subcorpora*. De forma geral, ao analisarmos os colocados do verbo *make* nos *subcorpora*, observamos algumas tendências distintas, mencionadas abaixo:

1. **Make sure**: a expressão mantém seu significado literal do nível A2 ao B2, com estruturas sintáticas semelhantes, como visto nos exemplos abaixo retirado do *subcorpus* A1 e B2:

Ex.1: "I'll **make sure** she stays in town".

Fonte: Hampton House, Jenny Dooley (A1).

Ex.2. I'm going to **make sure** that nothing bad ever happens.

Fonte: Life Exchange, Jenny Dooley (A1)

Ex.3. I'll **make sure** that no one follows you.

Fonte: Nicholas Nickleby, Charles Dickens (B2).

Já no nível C1, há inclusão de pronome pessoal e advérbio entre o verbo e o adjetivo, como visto no exemplo retirado do *subcorpus* C1, sendo **you** o pronome pessoal e **so** o advérbio:

Ex. 4: "What makes **you so** sure?";

Fonte: The Big Sleep, Raymond Chandle.

2. **Make feel**: os resultados mostram que no nível A2 apenas pronomes pessoais ocorrem entre os verbos, como evidenciado no exemplo retirado do *subcorpus* A2:

Ex. 5: "The drink only makes **me** feel okay these days."

Fonte: Lord Mcdonald, Eamonn Sweeney.

Ex.6: walk around, it makes **you** feel you are at one with.

Fonte: Lisa in China, Mitchell HQ.

No nível B2, ocorrem substantivos, determinantes, nomes próprios, pronomes possessivos e adjetivos entre os verbos, como evidenciado em um alguns dos exemplos retirados do *subcorpus* B2:

Ex. 7: "Makes a fellow feel rather silly, doesn't it?".

Fonte: Jeeves and Friends, P G Wodehouse.

Ex. 8: "to make this little lad feel that he was very small."

Fonte: Little Lord Fauntleroy, Frances Eliza Burnett.

3. **Make friends**: no nível A2, a estrutura "Verbo + Substantivo no plural" é predominante (66,6% das sentenças), refletindo a simplicidade gramatical dessas

construções. A estrutura pode ser encontrada no exemplo abaixo, retirado do *subcorpus* A2:

Ex. 9: *she wants to **make friends** with you.*

Fonte: The Adventures in The Grasslands, John Bookworm.

4. **Make noise**: todas as construções nos níveis A1 e A2 mantêm significados literais, destacando a clareza semântica, como evidenciado nos exemplos abaixo, retirado do *subcorpus* A1 e A2:

Ex. 10: "*Their swords **make a loud noise**.*"

Fonte: Peter Pan, JM Barriet.

Ex. 11: They **make** a lot of **noise**. Two miners find David and.

Fonte: Amazon Rally, Amos Eduardo.

5. **Make sense**: predomina o sentido de algo facilmente compreensível, evidenciado no exemplo retirado do *subcorpus* B2 abaixo:

Ex. 12: "*It doesn't make any sense at all.*"

Fonte: Board Games, Butler James.

Porém, a expressão corresponde ao significado de “fazer sentido”, como visto no exemplo abaixo retirado no *subcorpus* B2:

Ex. 13: "*Does murder ever make sense.*"

Fonte: After the Funeral, Agatha Christie.

6. **Make money**: No nível A2 é comum ocorrer sentenças com a construção: PRP RB VBP TO VB NN (Pronome pessoal + advérbio + preposição + verbo substantivo), como visto no exemplo abaixo retirado do *subcorpus* A2:

Ex. 14: "*they only want to make money.*".

Fonte: Jacks Came, Rod Smith.

No nível B2, por outro lado, o colocado apresenta características exclusivas para o nível B2. Há estruturas na forma condicional com *if* e estruturas condicionais com *as long as*, como evidenciado nos exemplos abaixo retirados do *subcorpus B2*

Ex. 15: "*Your father's happy for us to get married, if I can make enough money to look after you.*".

Fonte: A Study in Scarlet, Conan Doyle.

Ex. 16: "*I don't care what kind of shoes we sell, as long as we make money.*".

Fonte: Kings Ransom, Ed McBain.

7. Make happy: o estudo mostra que *make happy* consiste em um verbo e um adjetivo, porém, com a presença de outros elementos gramaticais entre o verbo e o adjetivo, como PRP (pronome pessoal); PRP\$ (pronome pessoal possessivo); NNP (substantivo próprio, singular); CC (conjunção coordenativa); RB (advérbio) e NNS (substantivo comum, plural). Devido a ocorrências de pronomes, essa construção tem sempre um significado de "fazer **alguém** feliz", não obstante, o estudo também revelou que a construção: Verbo + Pronome Pessoal + Adjetivo é a mais predominante (56%) dentre as outras identificadas, como evidenciado no exemplo a seguir, retirado do *subcorpus B2*:

Ex. 17: *I didn't know how to make her happy.*

Fonte: Interview with the Vampire, Anne Rice.

Ex. 18: *She wanted to make you happy.*

Fonte: Interview with the Vampire, Anne Rice.

Identificamos que há a ocorrência de pronomes ou substantivos próprios em todas as sentenças, tornando a expressão com o sentido de fazer alguém feliz em

todas as ocorrências, facilitando a compreensão para o leitor, como evidenciado no exemplo retirado do *subcorpus* B2:

Ex. 19: *I want to make him happy.*

Fonte: *The House of the Seven Gables*, Nathaniel Hawthorne.

Ex.20: *I'll do anything to make Gwendolen happy.*

Fonte: *Importance of Being Earnest*, Oscar Wilde.

Do ponto de vista pedagógico, é interessante observar como as diferentes construções com o verbo *make* podem ser utilizadas para ensinar e consolidar habilidades linguísticas em diferentes níveis de proficiência. Por exemplo, no nível A1, as construções mais simples como *make noise* podem ser ensinadas para introduzir o vocabulário básico e ensinar estruturas gramaticais simples. Enquanto isso, no nível C2, a análise das construções com mais elementos gramaticais como determinantes, advérbios, preposições, como ocorre em estruturas envolvendo *make sense*, como *mostrado no Ex.21*, têm o potencial de auxiliar os alunos a expandir seu vocabulário ao expressar quantidade, ao usar as palavras *total*, *nenhum*, *muito etc.* em inglês.

Ex.21: *"That makes **a lot of** sense, Tina."*

Fonte: *The Eyes of Darkness*, Koontz Dean.

Do ponto de vista linguístico, é interessante estudar como as construções com *make* variam em frequência, forma e tipologia em diferentes níveis de proficiência. Por exemplo, a predominância de estruturas como: "Verbo + Substantivo" no nível A2 pode indicar uma preferência por construções gramaticais diretas e facilmente compreensíveis, como evidenciado no exemplo abaixo retirado do *subcorpus* A2:

Ex. 22: *"they only want to make money."*

Fonte: *Jacks Came*, Rod Smith.

Por outro lado, a presença de construções nas formas condicionais como em "IF + Verbo + Pronome + Modal + Verbo + Substantivo", presentes no nível B2 pode refletir

uma habilidade maior em interagir e compreender a linguagem de forma mais elaborada, como evidenciado no exemplo retirado do nível B2 abaixo:

Ex. 23: "*Your father's happy for us to get married, **if I can make enough money to look after you.***".

Fonte: A Study in Scarlet, Conan Doyle.

Da mesma maneira, ao analisarmos os colocados do verbo *take* nos *subcorpora*, foram observadas algumas tendências distintas, listadas abaixo:

1. No nível A1, **take off** é usado principalmente para descrever ações habituais, significando remover uma peça de roupa, relacionado à estrutura gramatical: "verbo + advérbio", como mostrado nos exemplos abaixo retirado do *subcorpus* A1:

Ex. 24: "*Zorro takes **off** his black mask.*".

Fonte: Zorro, Sally MStockton.

Ex. 25: "*The old lady takes **off** her mask.*"

Fonte: Halloween Horror, Clemen D B Gina.

Por outro lado, no nível A2 e B2, além do significado anterior, *take off* pode significar remover algo de um lugar/de alguém, relacionado à estrutura gramatical "verbo + preposição", como visto nos exemplos abaixo, retirado do *subcorpus* B2:

Ex. 26: "*You've taken us all **off** your list?*".

Fonte: Death in the Clouds, Agatha Christie.

Ex. 27: "*I want to be taken **off** the case.*"

Fonte: Seven Bruno, Anthony.

2. **Take care**: apresenta a mesma estrutura e significados em ambos os níveis "cuidar de alguém/algo/si mesmo e ter cautela", com algumas ocorrências adicionais de

adjetivos e advérbios no nível A2, B1 e B2, como visto no exemplo abaixo, retirados do *subcorpus* B2:

Ex. 28: "*Reynolds said, 'and **take good care** of that nice sweater'.*"

Fonte: Kings Ransom, Ed McBain.

Porém, *take care* é predominantemente empregado na estrutura: "Verbo + Substantivo", como evidenciado no exemplo abaixo, retirado do *subcorpus* B2:

Ex. 29: "*people believe that he **takes care of** the island.*"

Fonte: This Rough Magic, Mary Stewart.

3. **Take away** é usado de forma semelhante nos níveis A2 e B2, porém, com maior incidência de sentenças no *present perfect* no B2, como visto no exemplo abaixo retirado do *subcorpus* B2:

Ex. 30: "*then she could explain why the chicken was **taken away** 'for his good'.*"

Fonte: Wicked and Humorous, Tales Saki.

Em suma, os resultados da pesquisa revelaram que os *subcorpora* A1 e A2 apresentam colocados com sentenças diretas e simples, com poucas classes gramaticais entre o nó e o colocado, geralmente possuindo apenas um significado; os *subcorpora* B1 e B2 apresentam colocados com sentenças com diversidade em conjugação verbal, com mais classes gramaticais entre o nó e o colocado, geralmente apresentando mais de um significado; os *subcorpora* C1 e C2 apresentam colocados com sentenças mais complexas, como condicionais com *IF* e sentenças no *present perfect*. Tal análise dos colocados das estruturas verbais que envolvem as expressões *take care*, *take off* e *take away* feita nesta pesquisa, em diferentes níveis de proficiência, pode ser útil tanto do ponto de vista linguístico quanto pedagógico. Do ponto de vista linguístico, como exemplificado na listagem acima, essa análise revelou padrões distintos de uso e significado, desde níveis mais básicos até níveis mais avançados, fornecendo informações sobre essas expressões que aparecem nos textos literários adaptados, como, por exemplo, ao examinar *take off*, observou-se

uma progressão no significado e na estrutura gramatical, conforme avançamos nos níveis de proficiência. No nível A1, o foco está na remoção de uma peça de roupa, enquanto no nível A2, além desse significado, há uma expansão para incluir a ideia de retirar algo de um lugar ou de alguém. Por exemplo, "*Zorro takes off his black mask*" (Ex.24) versus "*You've taken us all off your list?*" (Ex.25). Essa distinção destaca a evolução na compreensão e aplicação das estruturas gramaticais: "verbo + advérbio" e "verbo + preposição" pelos aprendizes de inglês.

No contexto de ensino de línguas, com base nas tendências observadas nos diferentes níveis de proficiência linguística, em relação ao uso do verbo *take* e seus colocados, podemos sugerir atividades que podem ser apropriadas para cada nível, como, por exemplo: no nível A1 a expressão *take off* pode ser ensinada com o significado principal de "remover uma peça de roupa", como mencionado no Ex.24. Já a expressão *take care*, pode ser introduzida com o significado básico de "cuidar de alguém/algo/si mesmo", como mostrado no Ex.28 e Ex.29. No nível A2, a expressão *take off*, além do significado de "remover uma peça de roupa", pode ser introduzida com o uso de remover algo de um lugar ou de alguém, como visto no Ex.25. No nível B2, a expressão *take away* pode ser explorada com o uso predominante do "*present perfect*" com *take away* e ensinada com essa estrutura gramatical em contextos retiradas dos *readers*, como visto no Ex.30.

Em relação ao terceiro objetivo, o plano de aula proposto (ver anexo A) foi elaborado, tendo como referência os alunos do ensino médio, com foco na introdução e utilização do *software LancsBox* e suas ferramentas, *KWIC* e *Graphcoll*. Por meio de uma atividade prévia de familiarização com essas ferramentas, os alunos podem explorar e compreender melhor os recursos disponíveis para análise linguística. O ponto central do plano é direcionar os alunos para trechos literários selecionados do *subcorpus* B2, com especial atenção para as colocações dos verbos *make* e *take*. Esse enfoque visa não só ampliar a compreensão textual dos alunos, mas também aprimorar sua capacidade de interpretação ao identificar e analisar padrões de uso desses verbos de alta frequência. Os objetivos de aprendizagem delineados incluem a familiarização efetiva com as ferramentas de análise linguística, a compreensão aprofundada das colocações relevantes dos verbos estudados, a melhoria da compreensão dos *readers* e a reflexão sobre o próprio processo de aprendizagem. Com a utilização das ferramentas como *KWIC* e *Graphcoll*, os alunos podem não

apenas identificar as colocações, mas também compreender as nuances de significado e uso, contribuindo, assim, para uma melhoria na capacidade de interpretação e leitura. Ademais, o plano buscou mostrar como a abordagem *Data-driven Learning* (DDL) pode ser altamente eficaz, proporcionando aos estudantes uma compreensão mais precisa das colocações dos verbos *make* e *take*. Ao explorar um *corpus* pedagógico específico, os alunos são capazes de contextualizar as colocações identificadas e aplicá-las de forma significativa em sua compreensão textual. Consequentemente, essa abordagem não apenas amplia a competência linguística dos alunos, mas também enriquece sua experiência com a literatura ao capacitá-los a interpretar textos em inglês. Em suma, o plano de aula proporciona o aprimoramento das habilidades de leitura e interpretação dos alunos.

Por fim, é inevitável afirmar que a pesquisa pode trazer benefícios tanto para os pesquisadores interessados em conhecer mais sobre a ferramenta de pesquisa *Lancsbox*, uma vez que o Capítulo 3 oferece um compêndio de informações relevantes sobre como utilizar o *software*, quanto para pesquisadores que desejem explorar as colocações relevantes de outros tipos de *corpora*. Além disso, a presente pesquisa pode servir de base para análise de outros verbos de alta frequência, que não foram contemplados devido ao escopo deste estudo. Não obstante, a análise dos outros colocados menos exclusivos e até mesmo outros verbos de alta frequência como o *get*, *go*, *give* e *take*, pode ser abordada em pesquisas futuras. Outro desdobramento importante é verificar a eficácia de atividades desenvolvidas sob a perspectiva DDL, em contexto real de ensino na escola básica. Tal empreitada também pode ser explorada em estudos futuros.

REFERÊNCIAS

- ALMEIDA, Valdênia C. **Investigando colocações em um corpus de aprendiz**. Tese - Universidade Federal de Minas Gerais, 2014.
- ALMEIDA, Valdênia C. **Investigando colocações em um corpus de aprendiz**. Minas Gerais. Universidade Federal de Minas Gerais, 2014.
- ALMEIDA, Valdênia C. **Investigando colocações em um corpus de aprendiz**, 2014.
- BIBER, Douglas *et al.* **Grammar of Spoken and Written English**, 1999.
- BIBER, Douglas. Corpus-based and corpus-driven analyses of language variation and use. **The Oxford Handbook of Linguistic Analysis**, 2015.
- BOULTON, Alex. Research in data-driven learning. *In*: PÉREZ-PAREDES, Pascual; MARK, Geraldine. **Beyond Concordance Lines: Corpora in language education**. John Benjamins Publishing Company, f. 134, 267 p. cap. 1, p. p. 9-34, 2021.
- BOZZA, Morgana Cristina; BENENITO, Calixto. A importância do texto literário nas aulas de Língua Inglesa no ensino médio. **Portal Dia a Dia Educação** (Seed-PR), 2019. Disponível em: <http://www.diaadiaeducacao.pr.gov.br/portals/pde/arquivos/809-4.pdf>. Acesso em: 21 set. 2023.
- BREZINA, Vaclav. **Statistics in Corpus Linguistics: A Practical Guide**. Cambridge University Press, 2018.
- BREZINA, Vaclav; GABLASOVA, Dana. Is there a core general vocabulary? Introducing the new general service list. **Applied Linguistics**, v. 36, n. 1, p.1-22, 2015.
- BREZINA, Vaclav; MCENERY, Tony; WATTAM, Stephen. Collocations in context: A new perspective on collocation networks. **International Journal of Corpus Linguistics**, v. 20, n. 2, p. 139-173, 2015.
- BREZINA, Vaclav; PLATT, Willian. **#LancsBox X 2.0 [software package]**, 2023. Disponível em: <https://lancsbox.lancs.ac.uk>. Acesso em: 1 abr. 2023.
- BROWN, M. H. **Driven Learning: Designing Activities for Novices (slides)**, 2016. Disponível em: https://www.academia.edu/20810999/Data_Driven_Learning_Designing_Activities_for_Novices_slides_. Acesso em: 26 mar. 2024.
- CARVALHO, Damiana Maria. A importância da leitura literária para o ensino. **EntreLetras**, v. 6, n. 1, p. 6-21, 2015.
- CHAMBERS, Angela. What is data-driven learning?. *In*: O'KEEFFE, Anne; MCCARTHY, Michael. **The Routledge Handbook of Corpus Linguistics**. Routledge, v. 1, f. 356, 711 p, p. 345-358, 2010.

CLANDFIELD, Lindsay; BUDDEN, Jo. **Teaching materials: using readers in the EFL/ESL classroom 1**. Macmillan Education. UK, 2020. Disponível em: Macmillan Education. Acesso em: 2 jan. 2023.

COLLIE, Joanne; SLATER, Stephen. **Literature in the Language Classroom: A Resource Book of Ideas and Activities**. Cambridge University Press, v. 1, f. 138, 276 p., 1987.

CORCHS, Margaret. **O uso de textos literários no ensino de Língua Inglesa**. Portal da UECE. Ceará, 2006. Disponível em: <https://www.uece.br/posla/wp-content/uploads/sites/53/2009/12/MargaretCorchs.pdf>. Acesso em: 21 set. 2023.

COWIE, Anthony Paul (Ed.). **Phraseology: Theory, analysis, and applications**. OUP Oxford, 1994.

CRESWELL, John W.; CRESWELL, J. David. **Projeto de pesquisa - 2.ed.:** Métodos qualitativo, quantitativo e misto. Penso Editora, v. 3, f. 132, 264 p., 2021

CROSTHWAITE, Peter. **Data-Driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners**. Routledge, v. 3, f. 110, 219 p., 2019.

CROSTHWAITE, Peter; LUCIANA; WIJAYA, David. Exploring language teachers' lesson planning for corpus-based language teaching: A focus on developing TPACK for corpora and DDL. **Computer Assisted Language Learning**, v. 36, n. 7, p. p. 1392-1420, 2023.

CROSTHWAITE, Peter; STEEPLES, Brett. Data-driven learning with younger learners: exploring corpus-assisted development of the passive voice for science writing with female secondary school students. **Computer Assisted Language Learning**, p. 1-32, 2022.

CROSTHWAITE, Peter; STEEPLES, Brett. Data-driven learning with younger learners: exploring corpus-assisted development of the passive voice for science writing with female secondary school students. **Computer Assisted Language Learning**, p. 1-32, 2022.

DA COSTA, Priscilla Tulipa; DE AZEVEDO, Adriana Maria Tenuta. Multi-word verbs na escrita acadêmica de aprendizes do inglês: um estudo baseado em corpus de estudantes brasileiros. **Scripta**, v. 20, n. 38, p. 215-244, 2016.

DE LIMA SILVA, Stefânia Aparecida *et al.* **O particípio presente como modificador do sintagma nominal da Língua Inglesa: uma análise de dados de um corpus de aprendizes brasileiros**, 2021.

DE OLIVEIRA, Janaina Henriques. **O uso de advérbios intensificadores na escrita de aprendizes brasileiros de inglês como L2: um estudo baseado em corpus**, 2015.

DURRANT, Phillip; SCHMITT, Norbert. Adult learners' retention of collocations from exposure. **Second language research**, v. 26, n. 2, p. 163-188, 2010.

DUTRA, Deise Prina; DA SILVA QUEIROZ, Jessica Maria; ALVES, Jessica Ceritello. Adding information in argumentative texts: a learner corpus-based study of additive linking adverbials. **Revista Estudos Anglo-Americanos**, 2017.

FEUERSTEIN, R.; FEUERSTEIN, S. Mediated Learning Experience: A Theoretical Review. *In*: FEUERSTEIN, Reuven; KLEIN, Pnina S.; TANNENBAUM, Abraham J. **Mediated Learning Experience (MLE): Theoretical, Psychosocial and Learning Implications**. Freund Publishing House Ltd., v. 1, f. 209, 418 p., 1990.

FIRTH, John Rupert. **Papers in Linguistics, 1934-1951**. London: Oxford University Press, 1957.

FLOWERDEW, Jonh. Corpora in Language Teaching. *In*: LONG, Michael H.; DOUGHTY, Catherine J. **The Handbook of Language Teaching**. John Wiley & Sons, v. 1, f. 418, cap. 19, p. 328-350, 2009.

FRANKENBERG-GARCIA, Ana. Investigating the collocations available to EAP writers. **Journal of English for Academic Purposes**, v. 35, p. 93-104, 2018.

FRANKENBERG-GARCIA, Ana *et al.* Developing a writing assistant to help EAP writers with collocations in real time. **ReCALL**, v. 31, n. 1, p. 23-39, 2019.

GABLASOVA, Dana; BREZINA, Vaclav; MCENERY, Tony. Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence. **Language Learning**, v. 67, n. S1, p. 155-179, 2017.

GABLASOVA, Dana; BREZINA, Vaclav; MCENERY, Tony. Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. **Language learning**, v. 67, n. S1, p. 155-179, 2017.

GILQUIN, Gaëtanelle; GRANGER, Sylviane. 30 Using data-driven learning in language teaching. *In*: O'KEEFFE, Anne; MCCARTHY, Michael. **The Routledge Handbook of Corpus Linguistics**. Routledge, p. 430-442, 2022.

GRIES, Stefan, Th. Phraseology and linguistic theory. *In*: GRANGER, Sylviane; MEUNIER, Fanny. **Phraseology: An Interdisciplinary Perspective**. John Benjamins Publishing, v. 1, f. 228, p. 3-25, 2007.

HALLIDAY, Michael AK. Lexis as a linguistic level. *In*: **MEMORY of JR Firth**, v. 148, 162 p., 1996.

HILL, David R. Graded readers in English. **ELT journal**, v. 62, n. 2, p. 184-204, 2008.

HILL, David R. Graded readers. **ELT journal**, v. 67, n. 1, p. 85-125, 2013.
HOWARTH, Peter. Phraseology and second language proficiency. **Applied linguistics**, v. 19, n. 1, p. 24-44, 1998.

HUNT, Alan; BEGLAR, David. A framework for developing EFL reading vocabulary. **Reading in a foreign language**, v. 17, n. 1, p. 23-59, 2005.

JOHNS, Tim. **Should you be persuaded: Two samples of data-driven learning materials**, 1991.

JUKNEVIČIENĖ, Rita. Collocations with high-frequency verbs in learner English: Lithuanian learners vs native speakers. **Kalbotyra**, v. 59, p. 119-127, 2008.

KENNEDY, Graeme. **An Introduction to Corpus Linguistics**. Routledge, v. 3, f. 167, 334 p., 2014.

KREYER, Rolf. Collocations in learner English. *In*: PÉREZ-PAREDES, Pascual; MARK, Geraldine. **Beyond Concordance Lines: Corpora in language education**. John Benjamins Publishing Company, f. 134, 267 p. 97 p., 2021.

LAUFER, Batia; WALDMAN, Tina. Verb-noun collocations in second language writing: A corpus analysis of learners' English. **Language learning**, v. 61, n. 2, p. 647-672, 2011.

LEE, Pinshuan; LIN, Huifen. The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. **System**, v. 81, p. 14-25, 2019.

LIN, Ching-Hsi; LIN, Yen-Liang. Grammatical and Lexical Patterning of Make in Asian Learner Writing: A Corpus-Based Study of ICNALE. **3L: Southeast Asian Journal of English Language Studies**, v. 25, n. 3, 2019.

MARTINS, Ana Maria; CARRILHO, Ernestina. **Manual de linguística portuguesa**. Walter de Gruyter GmbH & Co KG, f. 357, 714 p., 2016.

MCENERY, Tony; WILSON, Andrew. Teaching and language corpora (TALC). **ReCALL**, v. 9, n. 1, p. 5-14, 1997.

MCCARTHY, Michael *et al.* Looking back on 25 years of TaLC: In conversation with Profs Mike McCarthy and Tony McEnery. *In*: PÉREZ-PAREDES, Pascual; MARK, Geraldine. **Beyond Concordance Lines: Corpora in language education**. John Benjamins Publishing Company, f. 134, 267 p., 2021.

MUFTAH, Muneera. Data-driven learning (DDL) activities: do they truly promote EFL students' writing skills development?. **Education and Information Technologies**, p. 1-27, 2023.

NATION, Paul; WARING, Robert. Vocabulary Size, Text Coverage and Word Lists. *In*: SCHMITT, Norbert; MCCARTHY, Michael. **Vocabulary: Description, Acquisition and Pedagogy**. Cambridge University Press, v. 14, p. 6-19, 1997.

NESSSELHAUF, Nadja. **Collocations in a Learner Corpus**. John Benjamins Publishing, v. 1, f. 176, 352 p., 2004.

O'KEEFFE, Anne; MCCARTHY, Michael. **The Routledge Handbook of Corpus Linguistics**. Routledge, v. 1, f. 356, 711 p., 2010.

O'KEEFFE, Anne; MCCARTHY, Michael; CARTER, Ronald. **From Corpus to Classroom: Language Use and Language Teaching**. Cambridge University Press, v. 1, f. 159, 317 p., 2007.

OLIVA, Katherine Nunes Pereira. **Integrating corpora with language classroom: online corpus as na editing tool**, 2018.

ORFANÒ, Bárbara Malveira; OLIVEIRA, Ana Larissa Adorno Marciotto; BARBOSA, Spencer. The role of pragmatic markers in academic spoken interlanguage: a corpus-based study of a group of B Brazilian efl university students. **Diacrítica**, 2019.

O'KEEFFE, Anne; MCCARTHY, Michael; CARTER, Ronald. **From corpus to classroom: language use and language teaching**. Cambridge University press, v. 1, f. 159, 317 p., 2007.

O'KEEFFE, Anne. Data-driven learning, theories of learning and second language acquisition. *In*: PÉREZ-PAREDES, Pascual; MARK, Geraldine. **Beyond Concordance Lines: Corpora in language education**. John Benjamins Publishing Company, f. 134, 267 p. cap. 2, p. 35-55, 2021.

PAIVA, Vera Lúcia Menezes de Oliveira. **Manual de pesquisa em estudos linguísticos**. 1 ed. São Paulo: Parábola, 2019.

PÉREZ-PAREDES, Pascual; MARK, Geraldine. **Beyond Concordance Lines: Corpora in language education**. John Benjamins Publishing Company, f. 134, 267 p., 2021.

PHILLIPS, M. K. **Lexical macrostructure in science text** (Unpublished doctoral dissertation). University of Birmingham, Birmingham, UK, 1983.

REPPEN, Randi. **Using Corpora in the Language Classroom**. Cambridge University Press, 2010.

RORTY, Richard. Redemption from egotism: James and Proust as spiritual exercises. **Telos: Revista de Estudos Interdisciplinarios En Ciencias Sociales**, v. 3, n. 3, p. 243-263, 2001.

RÖMER, Ute. A corpus-driven approach to modal auxiliaries and their didactics. **How to use corpora in language teaching**, p. 185-199, 2004.

SARDINHA, Tony Berber. **Linguística de corpus**. Editora Manole Ltda, 2004.

SHI, Jing. The application of data-driven learning in Business English teaching for MBA students in China: A critical assessment. **International Journal of English Linguistics**, v. 3, n. 1, p. p. 26-30, 2013.

SILERO, Rejane Wagner Protzner. **Os quantificadores a few e few: questões de interlíngua e prosódia semântica em corpus de aprendizes**, 2014.

SINCLAIR, John. **Corpus, Concordance, Collocation**. Oxford University Press, 1991.

SÁNCHEZ, Aquilino et al. Definición e historia de los corpus. **CUMBRE–Corpus lingüístico de español contemporáneo**, Madrid, p. 7-24, 1995.

TRIBBLE, Chris; JONES, Glyn. *Concordances in the classroom: a resource guide for teachers*. Athelstan, 1997.

TODOROV, Tzvetan. **A literatura em perigo**. Tradução Caio Meira, 2009. *La Littérature en péril* [Trad.].

TOGNINI-BONELLI, Elena. **Corpus Linguistics at Work**. John Benjamins Publishing, v. 1, f. 119, 237 p., 2000.

VIEIRA, Alice. Formação de leitores de literatura na escola brasileira: caminhadas e labirintos. **Cadernos de pesquisa**, v. 38, n. 134, p. 441-458, 2008.

VIGOTSKY, Lev Semenovich. **A formação social da mente**: o desenvolvimento dos processos psicológicos superiores. 4 ed. São Paulo: Martins Fontes, 1991.

WARING, Rob; TAKAHASHI, Sachiko. **The Oxford University Press guide to the 'why and 'how' of using graded readers**. Tokyo: Oxford University, 2000.

WEBB, Stuart; NEWTON, Jonathan; CHANG, Anna. Incidental learning of collocation. **Language learning**, v. 63, n. 1, p. 91-120, 2013.

WICHER, Oliver. A critical evaluation from the perspective of foreign language didactics. *In*: CROSTHWAITE, Peter. **Data-Driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners**. Routledge, v. 3, f. 110, 219 p. cap. 3., 2019.

WRIGHT, Vanessa Cristina Oliveira. **Os sintagmas nominais do inglês**: um estudo comparativo de corpus entre o uso de determinantes nas produções escritas de falantes nativos do inglês e aprendizes brasileiros, 2014.

XIAO, Richard; MCENERY, Tony. Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective. **Applied linguistics**, v. 27, n. 1, p. 103-129, 2006.

ANEXO A — Sequência Didática

1 INTRODUÇÃO

O plano de aula proposto é destinado aos alunos do Ensino Médio. Por meio da introdução ao *software LancsBox* e suas ferramentas *KWIC* e *Graphcoll*, os alunos serão guiados em uma atividade prévia para familiarização do uso das ferramentas. O ponto principal do plano é o contato com trechos literários selecionados do *subcorpus B2*, com foco nas colocações dos verbos *make* e *take*, visando ampliar a compreensão e a interpretação textual. Os objetivos de aprendizagem englobam a familiarização com ferramentas de análise linguística e de colocações relevantes de verbos de alta frequência, bem como a melhoria da compreensão textual e a reflexão sobre o próprio processo de aprendizagem. A estrutura geral do plano de aula está representada na Figura 1.

Figura 1 - Visão Geral do Plano de Aula.

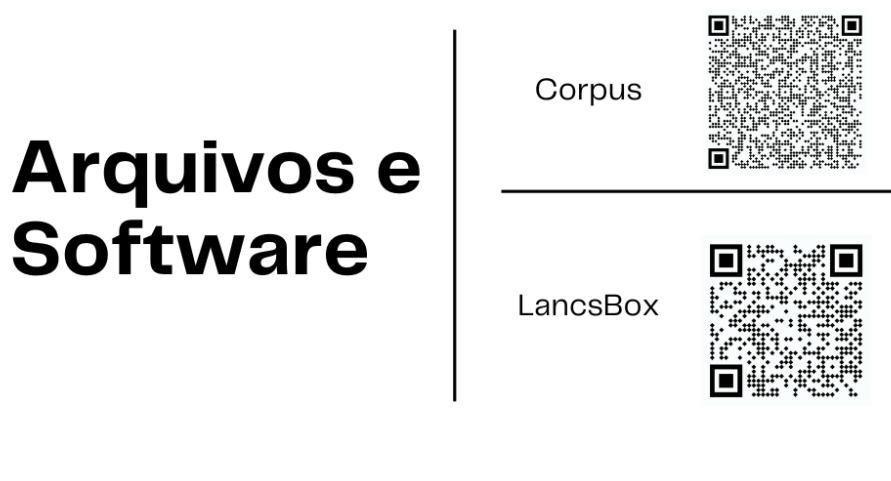
Visão geral do Plano De aula		
<p>O presente plano de aula:</p> <ul style="list-style-type: none"> • É adequado para alunos do Ensino Médio. • Utiliza os seguintes recursos: laboratório de informática, impressões A4 e projetor/Televisão. • Possui o tempo de execução de 2 aulas de 50 minutos. 		
<ul style="list-style-type: none"> • Introduz o software LancsBox e as funções básicas das ferramentas <i>KWIC</i> e <i>Graphcoll</i>. inclui uma pré-tarefa para praticar o uso das ferramentas <i>KWIC</i> e <i>GraphColl</i> na sala de aula. 	<ul style="list-style-type: none"> • Apresenta uma tarefa principal orientada por trechos literários retirados do subcorpus B2, para que os alunos explorem as colocações mais relevantes dos verbos de alta frequência “make” e “take” para facilitar a compreensão delas nos diferentes contextos dos readers. 	<ul style="list-style-type: none"> • Finaliza com três questões de revisão que incentivam os alunos a refletir sobre o uso do corpus especializado composto por readers para aprenderem as colocações.
<p>À luz do exposto, emergem os seguintes objetivos de aprendizagem:</p> <ul style="list-style-type: none"> • Pesquisar por linhas de concordância na ferramenta <i>KWIC</i>. • Verificar as colocações dos verbos de alta frequência como <i>Make</i> e <i>Take</i> na ferramenta <i>Graphcoll</i>. • Comunicar como colocados específicos são usados com os verbos <i>Make</i> e <i>Take</i>, consultando um corpus pedagógico. • Melhorar a habilidade de leitura por meio da compreensão dos colocados dos verbos de alta frequência. 		

Fonte: O Autor (2024).

2 CORPUS E FERRAMENTAS

O recurso usado para essa lição foi um *corpus* pedagógico compilado para essa pesquisa composto por *Graded Readers* de nível B2, possuindo 93 arquivos, totalizando 1.291.536 [um milhão, duzentos e noventa e um mil, quinhentos e trinta e seis] *tokens*. O *software* utilizado para compilação foi o programa de análise linguística, criado pela Universidade de Lancaster, chamado *LancsBox* na versão X 3.0.0 (Brezina; Platt, 2023). As ferramentas que serão utilizadas durante a aula pelos alunos serão o *Graphcoll* e o *KWIC*. Na Figura 1, são mostrados os *QR codes* para o *download* do *software* e do *corpus* para a aplicação desta SD.

Figura 2 - *QR codes* para *download* do material



Fonte: O autor (2024)

3 APLICAÇÃO DA SQ

Na Figura 3, é possível visualizar a estrutura das aulas.

Figura 3 - Estrutura das aulas

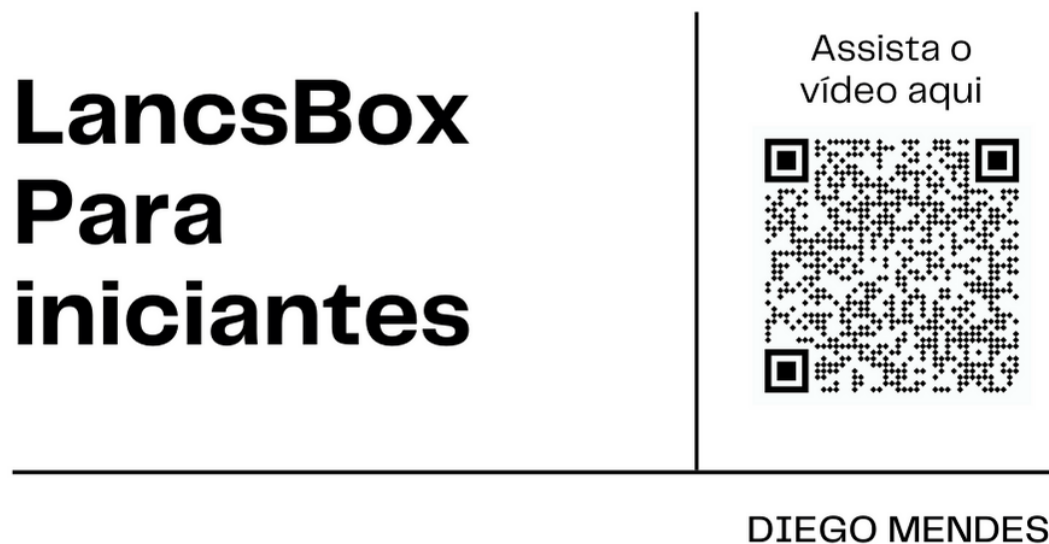
Visão geral do Plano De aula		
<p>O presente plano de aula:</p> <ul style="list-style-type: none"> • É adequado para alunos do Ensino Médio. • Utiliza os seguintes recursos: laboratório de informática, impressões A4 e projetor/Televisão. • Possui o tempo de execução de 2 aulas de 50 minutos. 		
<ul style="list-style-type: none"> • Introduz o software LancsBox e as funções básicas das ferramentas KWIC e Graphcoll. inclui uma pré-tarefa para praticar o uso das ferramentas KWIC e GraphColl na sala de aula. 	<ul style="list-style-type: none"> • Apresenta uma tarefa principal orientada por trechos literários retirados do subcorpus B2, para que os alunos explorem as colocações mais relevantes dos verbos de alta frequência “make” e “take” para facilitar a compreensão delas nos diferentes contextos dos readers. 	<ul style="list-style-type: none"> • Finaliza com três questões de revisão que incentivam os alunos a refletir sobre o uso do corpus especializado composto por readers para aprenderem as colocações.
<p>À luz do exposto, emergem os seguintes objetivos de aprendizagem:</p> <ul style="list-style-type: none"> • Pesquisar por linhas de concordância na ferramenta KWIC. • Verificar as colocações dos verbos de alta frequência como Make e Take na ferramenta Graphacoll. • Comunicar como colocados específicos são usados com os verbos Make e Take, consultando um corpus pedagógico. • Melhorar a habilidade de leitura por meio da compreensão dos colocados dos verbos de alta frequência. 		

Fonte: O autor (2024)

3.1 Instruções para a Atividade Prévia

1° Passo: Introduza o programa de análise linguística *Lancsbox*, exibindo o vídeo disponível no *QR code*, na Figura 4.

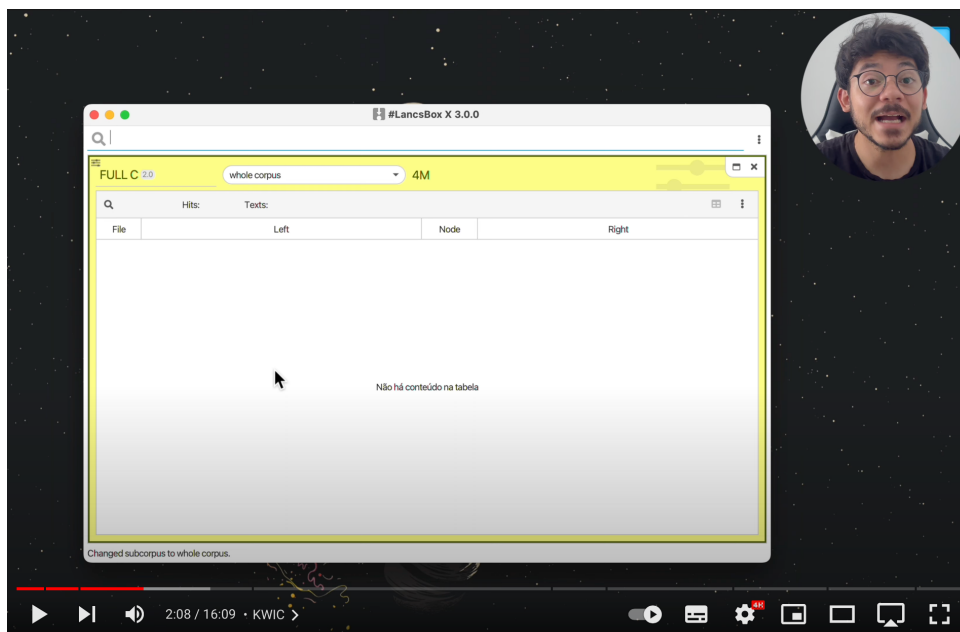
Figura 4 - QR Code do Vídeo de Introdução do *LancsBox*



Fonte: O autor (2024)

As instruções do vídeo abordam as funções gerais do *software LancsBox*, com enfoque no uso das ferramentas *KWIC* e *Graphcoll*. No vídeo, é mostrado como a interface das ferramentas funcionam com um exemplo que enfoca os diferentes usos do colocado *off* em combinação com o verbo *Take*.

Figura 5 - Tela de exibição do vídeo



Fonte: O autor (2024)

Ao final do vídeo, os alunos devem ser capazes de:

- a) Identificar as ferramentas *KWIC* e *Graphcoll* no *Software*;
- b) Usar a barra de pesquisa do *Software*;
- c) Reconhecer as informações disponíveis das palavras pesquisadas no *KWIC* e no *Graphcoll*;
- d) Ler e interpretar os gráficos da ferramenta *Graphcoll*;
- e) Usar as colocações identificadas no *Graphcoll* para fazer pesquisas no *KWIC*.

2º Passo: Separe os estudantes em duplas⁷ e distribua a atividade da Figura 6 (*Task 1*) para os alunos.

⁷A vantagem de realizar atividades em dupla, conforme a perspectiva de Lev Vygotsky, está intrinsecamente relacionada ao conceito de Zona de Desenvolvimento Proximal (ZDP). Segundo Vygotsky, a ZDP representa a distância entre o nível de desenvolvimento real de uma pessoa e o nível de desenvolvimento potencial, que pode ser alcançado com o auxílio de um parceiro mais experiente ou com a interação social. Portanto, ao realizar atividades em dupla, os estudantes têm a oportunidade de colaborar e interagir entre si, proporcionando um ambiente propício para o desenvolvimento cognitivo. VYGOTSKY, L. S. **Mind in Society: The Development of Higher Psychological Processes**. Harvard University Press, 1978.

Figura 6 - Task 1

Task 1

Searching for patterns of language use

This task demonstrates how corpus data can be used to search for patterns of language use and test our intuitions about language. First, answer two questions drawing on your experience with the English language. Then, use corpus evidence by carrying out searches in the **B2 SUBCORPUS** to check whether your intuitions were correct.

1.a What particles (e.g. up, out) are typically used with the verb come? Make a list:

.....

1.b What other words are used with the verb come? List them below.

.....

Now let's check your answers using corpus data:

1.c First, in the KWIC, search for [come* PARTICLE] in the Search Query. What particles are used with the verb come? List them below and see if they match your list in Task 1a:

.....

1.d Second, in the GraphColl, search for [come*] in the Search Query. According to the Graphafic, what other relevant words are used with the verb come? List them below and see if they match your list in Task 1b

.....

A *Task 1* objetiva propiciar aos alunos a prática dos conhecimentos adquiridos no vídeo disponibilizado. O foco da atividade está no manuseio da ferramenta, uma vez que, para a realização da tarefa principal, os alunos precisam ter compreendido bem a interface. Portanto, a tarefa demonstra, por meio das ferramentas, como os dados do *corpus* podem ser usados para procurar padrões de uso da linguagem e testar nossas intuições sobre a linguagem.

Primeiro, os alunos precisam responder a duas perguntas com base nas próprias experiências com a Língua Inglesa. Em seguida, utilizando as evidências do *corpus*, eles precisam realizar as pesquisas no *subcorpus* B2 para verificar se as intuições dos estudantes estavam corretas. Dessa forma, ao realizarem essa verificação, eles praticarão o uso das ferramentas, possibilitando, assim, a fixação do conteúdo aprendido no vídeo.

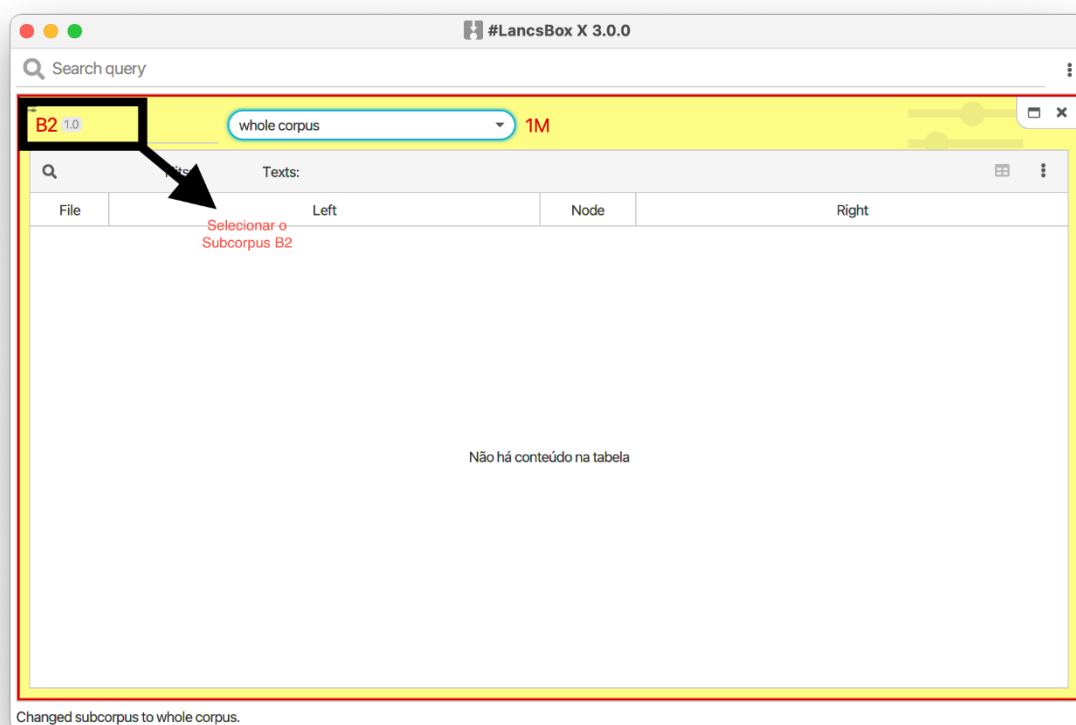
Em relação à primeira pergunta (1a), os alunos são questionados sobre quais partículas são comumente relacionadas com o verbo *come*. Solicita-se que eles façam uma lista no espaço disponibilizado na folha. O intuito de questioná-los sobre as partículas, é, além de fazê-los refletir sobre o uso da expressão: COME + PARTICLES, é também proporcioná-los a oportunidade de usarem e explorarem especificamente as funções e a interface da ferramenta *KWIC*, ao fazerem a atividade 1c.

As possíveis respostas da atividade 1a são: ***up, out, in, down etc.***

Para obter a resposta da atividade 1c, os alunos precisam seguir os seguintes passos:

1° passo: Abrir o *LancsBox*, escolher a ferramenta *KWIC* e selecionar o *subcorpus* B2, como mostrado na Figura 7:

Figura 6 - KWIC 1



Fonte: O autor (2024)

2º passo: Pesquisar por $[come^* PARTICLE]$ na barra de pesquisa, como pode ser visualizado na Figura 7.

Figura 7 - KWIC 2

The screenshot shows the LancsBox X 3.0.0 interface. At the top, the search input field contains "come* PARTICLE" and the search button is labeled "Pesquisar por come* PARTICLE". Below the search bar, the interface displays "B2 1.0" and "whole corpus" with a "1M" indicator. The search results are shown in a table with the following columns: File, Left, Node, and Right. The table contains 15 rows of search results. At the bottom of the window, it says "Searched KWIC for 'come* PARTICLE'".

File	Left	Node	Right
Deadlock-1	and then went outside again. '	Come on,	Martin. Explain things to me.'
Deadlock-1	don't, she'll only	come in	uninvited and search the house
Shakespea	people that we love. Lear	comes on	to the stage at the
Deadlock-1	ship. I saw Captain Bemis	come down	from the bridge. 'Martin. Our
Deadlock-1	I persuaded the watchman to	come up	and let me in. I
Deadlock-1	voice, 'Who asked you to	come down	to the Port?' 'No one.
The_House	a week or two." "Please	come in,	little Phoebe," said Hepzibah. "You
The_House	Hepzibah, but he did not	come in.	It was her cousin, Judge
The_House	so that the sunshine can	come in.	He likes the light! I
Shakespea	At this point another messenger	comes in	to say that the young
Deadlock-1	deck. Then footsteps. Someone had	come on	to the boat. I put
A_Kiss_Be	hour, but she didn't	come out	again. "A business meeting!" he
The_Valley	recommended to me.' 'Well then	come in,	sir. I'm Ettie Shafter,
A_Kiss_Be	get your letter. But please	come in	and sit down for a

Fonte: O autor (2024)

3º passo: Observar na coluna do "node" as ocorrências das partículas com o verbo *come*.

Figura 8 - KWIC 2

Search results for "come* PARTICLE" in the whole corpus (1M hits).

File	Left	Node	Right
Deadlock-5	and then went outside again. '	Come on, ...	in. Explain things to me.'
Deadlock-5	don't, she'll only	come in	uninvited and search the house
Shakespea	people that we love. Lear	comes on	to the stage at the
Deadlock-5	ship. I saw Captain Bemis	come down	from the bridge. 'Martin. Our
Deadlock-5	I persuaded the watchman to	come up	and let me in. I
Deadlock-5	voice, "Who asked you to	come down	to the Port?' 'No one.
The_House	a week or two." "Please	come in,	little Phoebe," said Hepzibah. "You
The_House	Hepzibah, but he did not	come in.	It was her cousin, Judge
The_House	so that the sunshine can	come in.	He likes the light! I
Shakespea	At this point another messenger	comes in	to say that the young
Deadlock-5	deck. Then footsteps. Someone had	come on	to the boat. I put
A_Kiss_Be	hour, but she didn't	come out	again. "A business meeting!" he
The_Valley	recommended to me.' 'Well then	come in,	sir. I'm Ettie Shafter,
A_Kiss_Be	get your letter. But please	come in	and sit down for a

Searched KWIC for "come* PARTICLE".

Fonte: O autor (2024)

4º passo: Escrever na folha de resposta todas as diferentes partículas que ocorrem com *come*.

Figura 9 - Resposta da atividade 1c

1.c

First, in the KWIC, search for [come* PARTICLE] in the Search Query. What particles are used with the verb come? List them below and see if they match your list in Task 1a:

in, on, up, down, up, over, along, across

Fonte: O autor (2024)

5° passo: Comparar as respostas da atividade 1a com a 1c para verificar se a intuição se assemelha com as ocorrências, levando assim, a reflexão de como o *corpus* pode ajudar na comprovação de intuições/teorias sobre a língua.

Em relação à segunda pergunta (1b), os alunos são questionados sobre quais outras palavras (não especificamente as partículas como na atividade 1a) são comumente relacionadas ao verbo *come*. É pedido que eles façam uma lista no espaço disponibilizado na folha. O intuito desse questionamento é, além de fazê-los refletir sobre o uso de *come* com outros tipos de palavras que são diferentes das partículas, também proporcionar a eles a oportunidade de usarem e explorarem especificamente as funções e a interface da ferramenta **Graphcoll** ao fazerem a atividade 1d.

As possíveis respostas da atividade 1b são: ***here, back, with me etc.***

Para obter a resposta da atividade 1d, os alunos precisam seguir os seguintes passos:

1° passo: Abrir o *LancsBox*, escolher a ferramenta **Graphcoll** e selecionar o *subcorpus* B2, como mostrado na Figura 10:

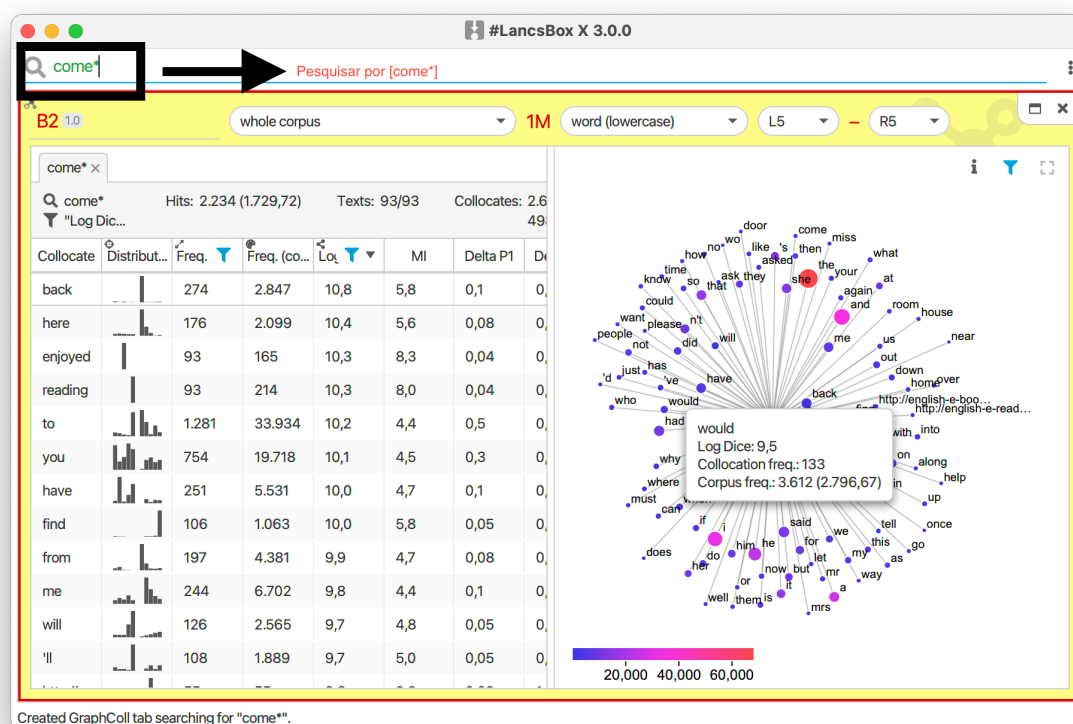
Figura 10 - Graphcoll 1



Fonte: (2024)

2° passo: Pesquisar por [come*] na barra de pesquisa, como pode ser visualizado na Figura 11:

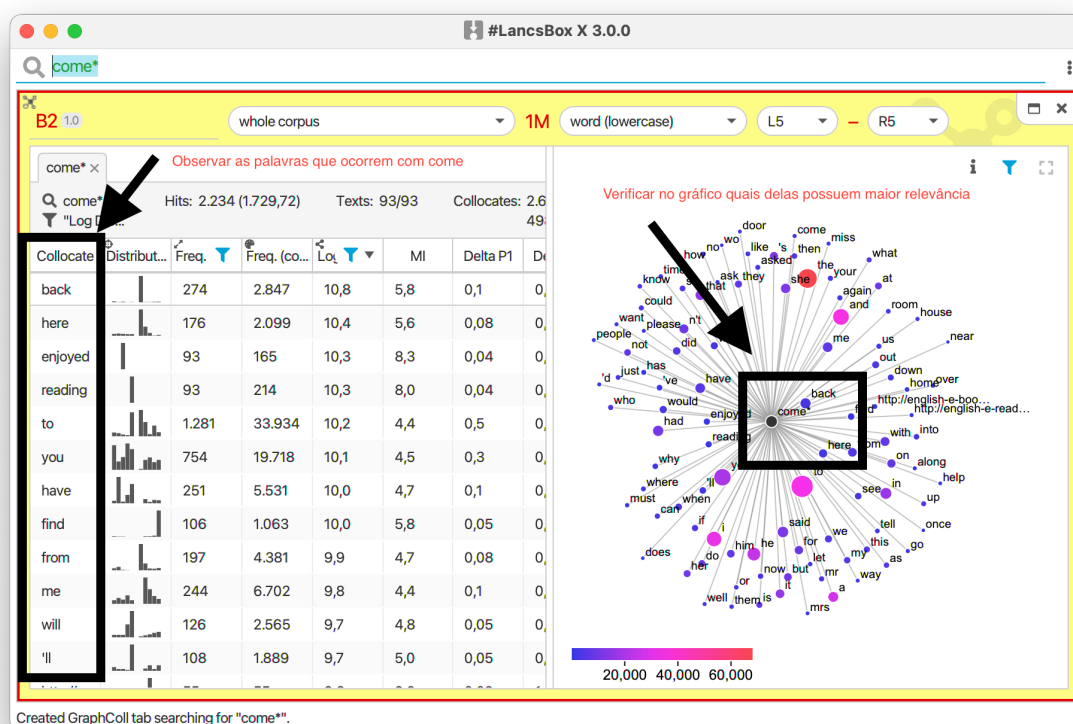
Figura 11 - Graphcoll 2



Fonte: O autor (2024)

3º passo: Observar na coluna de colocados as palavras que se relacionam com *come*, em seguida, verificar no gráfico de colocações as palavras que estão mais próximas do **nó**, uma vez que a distância do colocado com o **nó** indica a relevância deste colocado, portanto, as palavras mais próximas de *come*, são colocações relevantes.

Figura 12 - Graphcoll 3



Fonte: O autor (2024)

4º passo: Escrever na folha de resposta todas as diferentes palavras que ocorrem com *come*. As respostas podem ser conferidas na Figura 13:

Figura 13 - Respostas da atividade 1d

1.d Second, in the GraphColl, search for [come*] in the Search Query. According to the Graphafic, what other relevant words are used with the verb come? List them below and see if they match your list in Task 1b

back, find, here, from, home

Fonte: O autor (2024)

5° passo: Comparar as respostas da atividade 1b com a 1d para verificar se a intuição se assemelha com as ocorrências, levando assim, a reflexão de como o *corpus* pode ajudar na comprovação de intuições/teorias sobre a língua.

Ao término da atividade preparatória, espera-se que os alunos estejam mais capacitados para utilizar os recursos das ferramentas KWIC e Graphcoll na realização da atividade principal. Embora o objetivo primário da atividade preparatória seja o desenvolvimento da habilidade de manuseio do software, os alunos também terão a oportunidade de examinar as partículas que ocorrem com o verbo "come", bem como de identificar outros tipos de palavras que se relacionam com esse verbo.

3.2 Instruções para a primeira parte da Atividade Principal

1° Passo: Separe os alunos em duplas (as mesmas duplas da *Task 1*) e distribua a atividade da Figura 14 (*Task 2*) para os alunos.

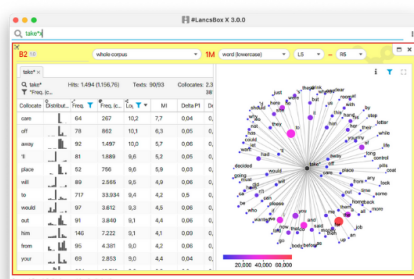
Figura 14 - Task 2

Task 2

Searching for patterns of language use

This task demonstrates how corpus data can be used to search for patterns of language use and test our intuitions about language. First, search for the collocates in KWIC. Then, use corpus evidence by carrying out searches in the B2 SUBCORPUS to answer the questions.

1. "When searching for [take*] in the GraphColl tool, it is possible to notice that the verb take has the words care, off, and away as relevant collocations. As seen in the image below:"



According to this information, open KWIC in LancsBox and search for [take*] and filter by care, off and away. Observe the concordance lines generated and answer the questions:

- a) The expression take off does not have the same meaning in all concordance lines. What different meanings did you identify? Write two examples for each meaning.

.....

.....

.....

.....

- b) The expression take care also has different meanings, identify them, explain them and write two examples.

.....

.....

.....

.....

- c) According to Oxford's Learner's Dictionaries, take away is classified as a phrasal verb. This term denotes the action of dispelling a feeling or removing it. Does it have the this meaning in all concordance lines?

.....

.....

.....

.....

2. Open the GraphColl tool and search for [make*]. Look at the collocated column and check the graph to see which are the most relevant collocations of the verb make. Then, open the KWIC tool and check whether or not these collocations have different meanings within the context of the readers.

.....

.....

.....

.....

A *Task 2* objetiva propiciar aos estudantes a prática pela busca de padrões linguísticos para adquirirem conhecimento prévio para realizarem a *Task 3*. O foco da atividade está em identificar os diferentes significados das expressões *take off*, *take care* e *take away* e a compreensão das expressões *make sure*, *make sense*, *make feel* e *make happy*, a fim de preparar os alunos para a leitura dos *readers*. Portanto, a tarefa demonstra, por meio das ferramentas, como os dados do *corpus* podem ser usados para procurar padrões de uso da linguagem para compreenderem que há diferentes significados das expressões com o verbo *take* e perceber a consistência de significado com algumas expressões com o verbo *make*.

Primeiramente, os alunos devem abrir a ferramenta *KWIC* no *LancsBox* e pesquisar por [*take**]. Em seguida, eles precisam filtrar pelos colocados *off*, *care* e *away*, como demonstrado nas Figuras 15, 16 e 17.

Figura 15 - Take Off 1

The screenshot shows the LancsBox X 3.0.0 interface. The search query is 'take*' and the corpus is 'whole corpus'. The search results are displayed in a table with the following columns: File, Left, Node, and Right. The results are as follows:

File	Left	Node	Right
Shakespea	a romantic comedy, a comedy	taken	from Plautus, and one that
Shakespea	him but the young man	takes	a dark-haired woman from
Shakespea	had even more reason to	take	care of his wife and
Shakespea	's greatest dramatist did not	take	more care of his work.
Shakespea	's death that Octavius will	take	her to Rome and show
Shakespea	never have. Shakespeare does not	take	sides in the argument. He
Shakespea	of the scenes, which clearly	take	place in the England of
Shakespea	who could sing well, had	taken	his place. Shakespeare wrote the
Shakespea	afterwards the man refused to	take	it. He preferred to keep
Shakespea	and queen. James decided to	take	personal control of the theatre
Shakespea	not mind when another girl	takes	her place. We cannot
Shakespea	please him. The story is	taken	from a history of Scotland,
The_Crocc	is weak and sick, they	take	her out into the countryside
Shakespea	and comedy. These plays were	taken	from popular stories, like Shakespeare

The filter dialog box is open on the right, showing a list of matching results (R1-R9) with checkboxes. The 'off' filter is selected, and R1-R5 are checked. The 'Apply' and 'Delete' buttons are visible at the bottom of the dialog box.

Fonte: O autor (2024)

Figura 16 - Take Care 1

#LancsBox X 3.0.0

take*

B2 1.0 whole corpus 1M

take* Hits: 1.494 (1.156,76) Texts: 90/93

File	Left	Node	Right
Shakespea	a romantic comedy, a comedy	taken	from Plautus, and one that
Shakespea	him but the young man	takes	a dark-haired woman from
Shakespea	had even more reason to	take	care of his wife and
Shakespea	's greatest dramatist did not	take	more care of his work.
Shakespea	's death that Octavius will	take	her to Rome and show
Shakespea	never have. Shakespeare does not	take	sides in the argument. He
Shakespea	of the scenes, which clearly	take	place in the England of
Shakespea	who could sing well, had	taken	his place. Shakespeare wrote the
Shakespea	afterwards the man refused to	take	it. He preferred to keep
Shakespea	and queen. James decided to	take	personal control of the theatre
Shakespea	not mind when another girl	takes	her place. We cannot
Shakespea	please him. The story is	taken	from a history of Scotland,
The_Croco	is weak and sick, they	take	her out into the countryside
Shakespea	and comedy. These plays were	taken	from popular stories, like Shakespeare

care

Matching within:

- R1
- R2
- R3
- R4
- R5
- R6
- R7
- R8
- R9

Apply Delete

Searched KWIC for "take*".

Fonte: O autor (2024)

Figura 17 - Take Away 1

The screenshot shows the LancsBox X 3.0.0 interface. The search query is "take*" and the corpus is "whole corpus". The results table is as follows:

File	Left	Node	Right
Shakespea	a romantic comedy, a comedy	taken	from Plautus, and one that
Shakespea	him but the young man	takes	a dark-haired woman from
Shakespea	had even more reason to	take	care of his wife and
Shakespea	's greatest dramatist did not	take	more care of his work.
Shakespea	's death that Octavius will	take	her to Rome and show
Shakespea	never have. Shakespeare does not	take	sides in the argument. He
Shakespea	of the scenes, which clearly	take	place in the England of
Shakespea	who could sing well, had	taken	his place. Shakespeare wrote the
Shakespea	afterwards the man refused to	take	it. He preferred to keep
Shakespea	and queen. James decided to	take	personal control of the theatre
Shakespea	not mind when another girl	takes	her place. We cannot
Shakespea	please him. The story is	taken	from a history of Scotland,
The_Croco	is weak and sick, they	take	her out into the countryside
Shakespea	and comedy. These plays were	taken	from popular stories, like Shakespeare

The matching window on the right shows the word "away" and a list of rules (R1-R9) with checkboxes. The rules are:

- R1
- R2
- R3
- R4
- R5
- R6
- R7
- R8
- R9

Buttons for "Apply" and "Delete" are visible at the bottom of the matching window.

Fonte: O autor (2024)

Em relação à pergunta 1a, os alunos são questionados sobre quais os sentidos que a expressão *take off* possui. Para obter a resposta da pergunta 1a, os alunos devem analisar as linhas de concordância geradas, como visto na Figura 18.

Figura 18 - Take off 2

Search results for "take*" in the whole corpus (1M documents). Hits: 1,494 (1,156,76); Texts: 90/93. Filter: "Right"=context("off", ...). 75 (58,07) hits; 30/93 texts.

File	Left	Node	Right
Black_Bea	above me and luggage was	taken	off the cab, but it
N_or_M-A	voice, 'I am about to	take	the gag off. You will
Farewell_r	there? How did these guys	take	it off you?' 'They must
N_or_M-A	Sprot appeared and Betty was	taken	off to be dressed for
Farewell_r	if he was dreaming. 'And	take	your dirty hand off my
Jeeves_an	Is there a cover? Well,	take	it off, Beach.' 'Very good,
Jeeves_an	it would be sensible to	take	off his top hat before
Seven-Bru	argue. 'I want to be	taken	off the case," he said. "
A_Study_ir	to court.' Jefferson Hope was	taken	off to a prison cell.
The_Invisi	But she saw he had	taken	off his glasses; they were
The_Invisi	to bed. When he had	taken	off his coat and shirt,
Importance	on a silver tray. Algernon	takes	it immediately. Lane goes off)
The_Monk	cap that I cannot	take	off my head? It gives

Searched KWIC for "take*".

Fonte: O autor (2024)

Ao analisar as linhas de concordância, os alunos podem visualizar um contexto maior das sentenças, ao posicionarem o cursor na linha desejada, como visto na Figura 19. Isso deve ajudá-los a terem uma melhor compreensão dos significados das expressões.

Figura 19 - Contexto Expandido, exemplo

File	Left	Node	Right
Black_Beat	above me and luggage was	taken	off the cab, but it
N_or_M-Ai	voice, 'I am about to	take	the gag off. You will
Farewell_rr	there? How did these guys	take	it off you? 'They must
N_or_M-Ai	Sprot appeared and Betty was	taken	off to be dressed for
Farewell_rr	if he was dreaming. 'And	take	your dirty hand off my
Jeeves_an	Is there a cover? Well,	take	it off, Beach.' 'Very good,
Jeeves_an	it would be sensible to	take	off his top hat before
Seven-Brui	argue. 'I want to be	taken	off the case," he said. "
A_Study_ir	to court.' Jefferson Hope was	taken	off to a prison cell.
The_Invisit	But she saw he had	taken	off his glasses; they w
The_Invisit	to bed. When he had	taken	off his coat and shirt,
Importance	on a silver tray. Algemon	takes	it immediately. Lane g
The_Monk	cap that I cannot	take	off my head? It gives

had gone upstairs. When Mrs Hall took his dinner in to him, he did not hear her until she had cleared away most of the paper and had put the food on the table. Then he half turned his head, and turned it away again. But she saw he had taken off his glasses; they were beside him on the table, and he seemed to her to have no eyes. He put on his glasses again, and then turned and faced her. She was about to complain about the paper on the floor, but he spoke first. 'I wish you would

Searched KWIC for "take*".

Fonte: O autor (2024)

A resposta esperada pelos alunos da atividade 1a da *Task 2*, pode ser visualizada na Figura 20:

Figura 20 - Resposta da Atividade 1a

a) The expression ***take off*** does not have the same meaning in all concordance lines. What different meanings did you identify? Write two examples for each meaning.

The expression take off has the meaning of removing a piece of clothing and also the meaning of removing something/someone from somewhere.

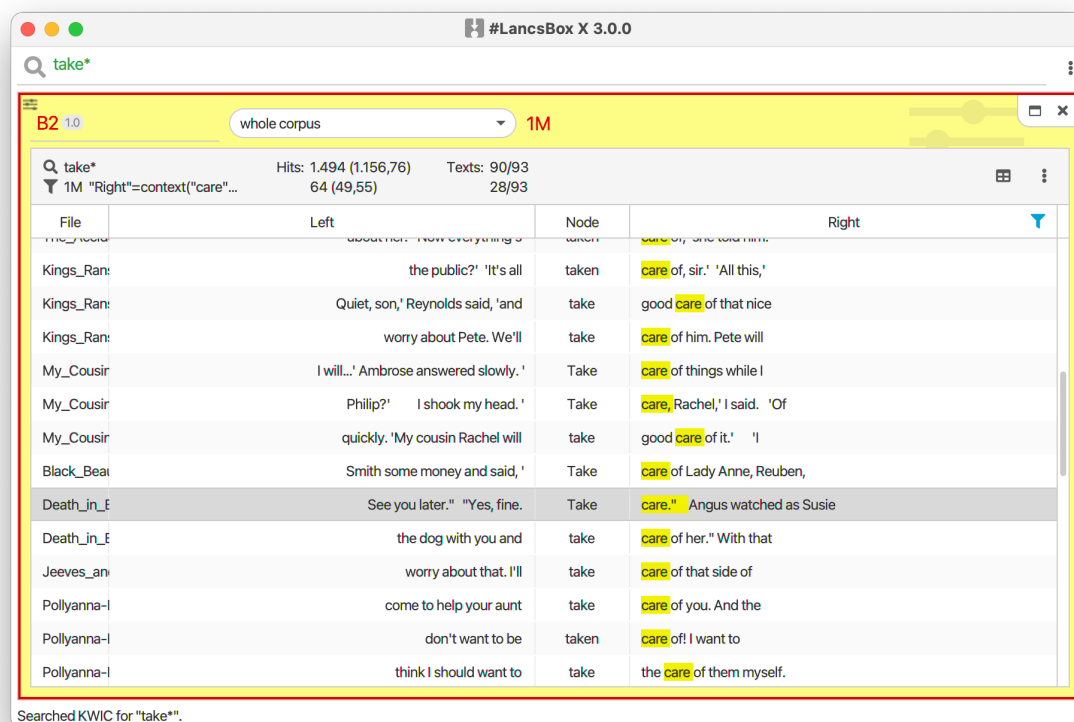
Ex1: "it would be sensible to take off his top hat before" and "But she saw he had taken off his glasses; they were".

Ex2: "Sprot appeared and Betty was taken off to be dressed" and " "Don't take your eyes off your opponent".

Fonte: O autor (2024)

Em relação à pergunta 1b, os alunos são questionados sobre quais os sentidos que a expressão *take care* possui. Para obter a resposta da pergunta 1b, os alunos devem analisar as linhas de concordância geradas, como visto na Figura 21:

Figura 21 - Take care 2



File	Left	Node	Right
Kings_Ran:	the public?' 'It's all	taken	care of, sir.' 'All this,'
Kings_Ran:	Quiet, son,' Reynolds said, 'and	take	good care of that nice
Kings_Ran:	worry about Pete. We'll	take	care of him. Pete will
My_Cousir	I will...' Ambrose answered slowly. '	Take	care of things while I
My_Cousir	Philip?' I shook my head. '	Take	care, Rachel, I said. 'Of
My_Cousir	quickly. 'My cousin Rachel will	take	good care of it.' 'I
Black_Bea	Smith some money and said, '	Take	care of Lady Anne, Reuben,
Death_in_E	See you later." "Yes, fine.	Take	care." Angus watched as Susie
Death_in_E	the dog with you and	take	care of her." With that
Jeeves_an	worry about that. I'll	take	care of that side of
Pollyanna-I	come to help your aunt	take	care of you. And the
Pollyanna-I	don't want to be	taken	care of! I want to
Pollyanna-I	think I should want to	take	the care of them myself.

Searched KWIC for "take*".

Fonte: O autor (2024)

A resposta que se espera que os alunos deem da pergunta 1b da *Task 2* pode ser visualizada na Figura 22.

Figura 22 - Resposta da atividade 1b

b) The expression take care also has different meanings, identify them, explain them and write two examples.

The expression take care has the meaning of supervising/protecting someone/something/oneself and to be careful.

Ex1. "had even more reason to take care of his wife" and "have to leave college to take care of the baby."

Ex2. "one of the local doctors. 'Take care, my dear,' she said," and "See you later." "Yes, fine. Take care." Angus watched as Susie"

Fonte: O autor (2024)

Em relação à pergunta 1b, os alunos são questionados se a expressão *take away* possui sentidos diferentes. Para obter a resposta da pergunta 2c, os alunos devem analisar as linhas de concordância geradas, como visto na Figura 23.

Figura 23 - Take Away 2

Search results for "take*" in the whole corpus (1M hits).

File	Left	Node	Right
Sink_or_Sv	could undermine his plans or	take	away the credit. He set
Wicked_an	to tell the gardener to	take	away his hen. Then there
Wicked_an	explain why the chicken was	taken	away for his good'. But
Wicked_an	will tell the gardener to	take	them away'. The woman then
Wicked_an	gardener and tell him to	take	away my wonderful god, who
Murder_in_	about her sofa, her Chinese	take-	away and her bed. After
Count_Vla	my men and have him	taken	away." During this confrontation, Ernest
The_Proble	prison," said The Thinking Machine. "	Take	this one away and kill
My_Cousin	Kendall would like you to	take	them away.' I'll speak
Black_Bear	when a man came to	take	me away to Mr Gordon
N_or_M-A	Oh, my dear.' 'They've	taken	him away. I'll never
N_or_M-A	mother says so. They'll	take	him away and lock him
Death_in_E	sitting on a bench with	take-	away rolls from the deli,

Searched KWIC for "take*".

Fonte: O autor (2024)

A resposta que se espera que os alunos deem da pergunta 1c da *Task 2* pode ser visualizada na Figura 24.

Figura 24 - Resposta da atividade 1c

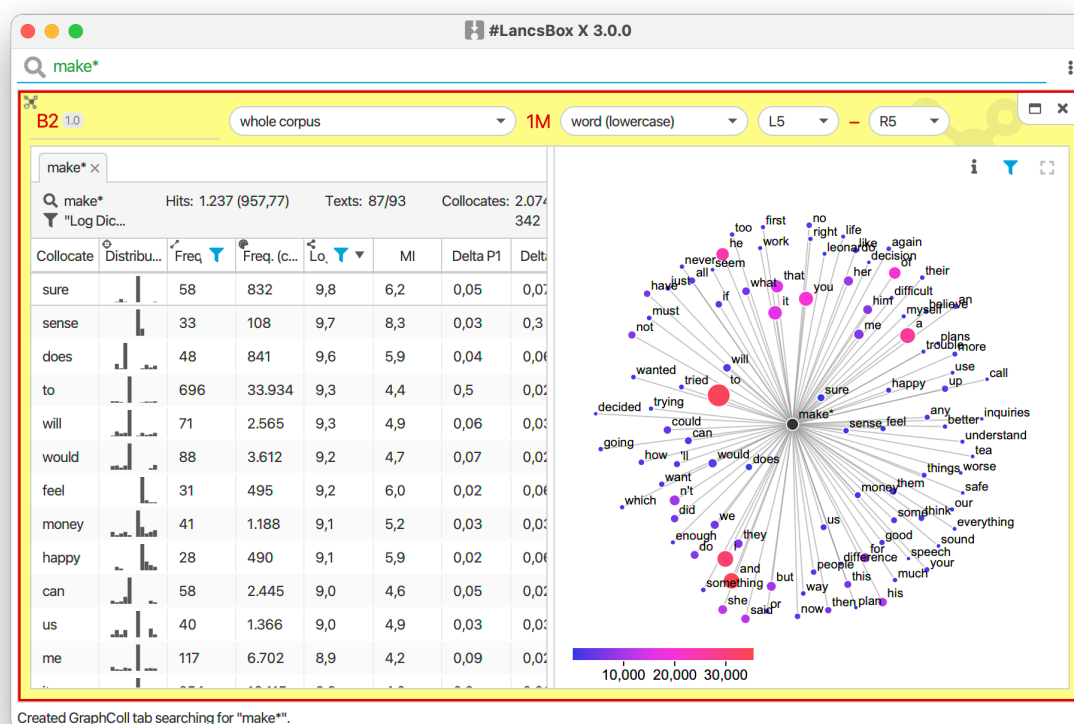
c) According to Oxford's Learner's Dictionaries, take away is classified as a phrasal verb. This term denotes the action of dispelling a feeling or removing it. Does it have the same meaning in all concordance lines?

No, in the concordances line the expression take away has the meaning of removing or eliminating something, whether physical or emotional. This shows that the variety of contexts in which this expression is used highlights its versatility and adaptability in different situations.

Fonte: O autor (2024)

Em relação à segunda pergunta (2), os alunos são questionados sobre quais são os colocados relevantes do verbo *make* e se possuem significados diferentes no *subcorpus* B2. Para obter a resposta da pergunta 2, os alunos devem seguir os seguintes passos:

1° passo: Abrir a ferramenta *Graphcoll* no *LancsBox* e pesquisar por [*make**], como mostrado na Figura 25.

Figura 25 - Colocados de *Make*

Fonte: O autor (2024)

2º Passo: Os alunos devem observar a coluna dos colocados e verificar no gráfico quais são os colocados mais relevantes de *make*. Espera-se que os alunos respondam que os colocados mais relevantes são: *sure*, *sense*, *feel* e *happy*.

3º Passo: A seguir, devem abrir a ferramenta *KWIC* e pesquisar por [*make**]. Logo, os alunos devem filtrar os colocados *sure*, *sense*, *feel* e *happy* para obter as suas respectivas linhas de concordância, como mostrado nas Figuras 26, 27, 28 e 29.

Figura 26 - Make sure

Searched KWIC for "make*".

File	Left	Node	Right
Shakespea	like these. He wanted to	make	sure that everyone in the
Nicholas_N	of the stairs. 'I'll	make	sure that no one follows
Nicholas_N	go away, or I shall	make	sure you go to prison
They_Cam	are staying and I'll	make	sure he gets in touch
They_Cam	I took that photo to	make	sure.' Who is it?' 'Anna
This_Rougl	I stayed with him to	make	sure he remained the right
This_Rougl	'I'll phone his house to	make	sure.' I called the Villa
This_Rougl	looked through the window to	make	sure that I was still
This_Rougl	will meet you. 'I'll	make	sure of that.' 'Good, but
This_Rougl	as Max. I had to	make	sure that we were not
Board_Gar	going to have a meeting.	Make	sure that they're all
Sink_or_Sv	with both of them to	make	sure that it didn't
Time_Macl	and shoulders. They wanted to	make	sure that I was real. '

Fonte: O autor (2024)

Figura 27 - Make sense

#LancsBox X 3.0.0

Q make*

B2 1.0 whole corpus 1M

Q make* Hits: 1.237 (957,77) Texts: 87/93
 1M "Right"=context("sens... 33 (25,55) 19/93

File	Left	Node	Right
Board_Gar	Arthur Mowbray. It doesn't	make	any sense at all.' 'Not
Sink_or_Sv	trophy of some sort. 'It	makes	sense , Eddie. Really," said Neal. "
Sink_or_Sv	some words but Eddie could	make	no sense of them. "Just
Time_Macl	in time?' 'It doesn't	make	sense ,' said Filby. 'Possibly not,'
Murder_in_	is something that does not	make	sense . The hammering noise. They
The_Accid	one of those deaths that	make	no sense the kind where
Kings_Ran	yours, too. It doesn't	make	sense .' 'Doesn't it?' Diane
N_or_M-A	nice when things begin to	make	sense ? Now we can move
N_or_M-A	Tommy. It just doesn't	make	any sense at all. But
N_or_M-A	he was saying did not	make	sense to her. She had
Seven-Bru	did it. It doesn't	make	sense . He doesn't feel
A_Study_ir	from a cigar. That all	makes	sense ,' I admitted, 'but I
Importance	devotedly, hopelessly. Hopelessly doesn't	make	much sense , does it? Algernon:

Searched KWIC for "make*".

Fonte: O autor (2024)

Figura 28 - Make feel

The screenshot shows the LancsBox X 3.0.0 interface. At the top, the search query is "make*" and the corpus is "whole corpus" with 1M documents. The search results are displayed in a table with the following columns: File, Left, Node, and Right. The word "feel" is highlighted in yellow in the Right column for each row.

File	Left	Node	Right
The_House	kind young girl. This rose	makes	me feel young again." He
A_Taste_o	walk in the garden will	make	you feel a lot better.' '
A_Devoted	must take it - it will	make	you feel stronger again. Promise
A_Taste_o	now I remember. Poison can	make	people feel like that.' Doctor
The_House	kind young girl. This rose	makes	me feel young again." He
The_War_c	extremely hot and dull. To	make	myself feel better I took
Time_Mad	the top and smiled to	make	her feel better. 'I had
Murder_in_	strange and different, and everything	makes	you feel more afraid. The
The_Tham	he didn't want to	make	her feel any worse than
The_Tham	stupid. Why did she always	make	him feel like a clumsy
The_Tham	and his mother. Did killing	make	him feel like less of
Count_Vlar	like to avoid it myself.	Makes	me feel rather insecure." "The
The_Accid	city seemed, and how to	make	themselves feel they were back

Searched KWIC for "make*".

Fonte: O autor (2024)

Figura 29 - Make happy

Search results for "make*" in the whole corpus (1M hits).

File	Left	Node	Right
The_House	best food. I want to	make	him happy . Let me see
The_House	together and work together, to	make	our dear Clifford happy . Why
The_House	and now she wanted to	make	his life warm and happy .
Death_in_t	job to do that truly	makes	them happy . You might hear
The_House	best food. I want to	make	him happy . Let me see
The_House	together and work together, to	make	our dear Clifford happy . Why
The_House	and now she wanted to	make	his life warm and happy .
This_Rough	said Sir Julian. 'It would	make	me very happy .' 'I'm
The_Tham	that desert theme. That should	make	her happy . The young investigator
Count_Vlar	to come, but it would	make	me very happy as I
The_Intelliç	for you, Mother? What will	make	you happy ? Seeing my children
Importance	of course, do anything to	make	Gwendolen happy . Lady Bracknell: My
Pollyanna-l	last cent, will go to	make	you happy ." Pollyanna looked shocked. "

Searched KWIC for "make*".

Fonte: O autor (2024)

Passo 4: Os alunos devem observar as linhas de concordância das expressões e verificar se elas possuem o mesmo sentido em todas as sentenças. Espera-se que cheguem à conclusão de que todas as expressões mantêm o sentido em todas as linhas de concordância.

A resposta esperada da atividade 2 da *Task 2* pode ser observada na Figura 30.

Figura 30 - Resposta esperada da atividade 2

2.

Open the GraphColl tool and search for [make*]. Look at the collocated column and check the graph to see which are the most relevant collocations of the verb make. Then, open the KWIC tool and check whether or not these collocations have different meanings within the context of the readers.

All expressions maintain the same meaning in all concordance lines. This shows the constancy of meanings in certain expressions.

Fonte: O autor (2024)

Ao final da *Task 2*, os alunos podem ser capazes de refletir sobre os diferentes significados das expressões *take off*, *take care* e *take away* e sobre a constância dos significados das expressões *make sure*, *make sense*, *make feel*, *make happy*. Esse conhecimento pode ajudá-los a realizarem a próxima tarefa (*Task 3*), a qual prepara os estudantes para uma melhor compreensão dos *readers*.

3.2 Instruções para a segunda parte da Atividade Principal

1° Passo: Distribua a atividade da Figura 31 (*Task 3 - Side A*) e da Figura 32 (*Task 3 - Side B*) para as duplas formadas na realização da *task* anterior.

Figura 31 - Task 3 - Side A

Task 3 - Side A

This activity demonstrates how prior knowledge of words that co-occur with the verbs take and make helps students deduce from the context which words to use. As students have previously observed in the Key Word In Context (KWIC) tool, the potential occurrences of various words alongside the verbs have been elucidated. With this enhanced understanding and interpretive capacity, students might choose the appropriate word." Solving this type of exercise can facilitate the comprehension of new words, especially collocations, which can improve the student's ability to understand when reading.

1. Rewrite the excerpts below taken from B2 Subcorpus, replacing the highlighted expressions with expressions with take and make, preserving the same meaning.

- a)** And now he's dead.' 'Of course she did,' said Marcus. 'Nobody likes the police. I do not like the police. But I have to keep good relations with them because of my hotel. You want me to give them money?' 'We want you to **remove** the body quietly.' 'That is very nice, I agree. I, too, I do not want a body in my hotel. But it is not so easy to do.' 'I think it could be,' said Dakin. 'You've got a doctor in your family, haven't you?' (They Came to Baghdad, Agatha Christie)
- b)** Marion Kingship didn't really love anybody. But she had her apartment, and she loved that. Every Saturday, she spent the day cleaning it. And she often dreamed that one day, a good, kind man would visit her there - someone who would love her and **treat her nice**. "Will he ever come?" (A Kiss Before Dying, Ira Levin)
- c)** The snakes had gone now that the monster was dead, and the water no longer burned. Beowulf swam strongly for the land, pulling his heavy load behind him. The Geats were full of joy when they saw their hero reach the land. They helped him to **remove** his armor, and to carry the monster's head. They put the head on a spear, and they set off for Hrothgar's Hall. (Beowulf, Robert_Hill)
- d)** "What about me? How can I be 'glad' about anything without you? But if I had you for my own little girl, I'd be glad for anything; and I'd try to make you glad, too, my dear. All my money, to the last cent, will go to **bring happiness to** you." Pollyanna looked shocked. "Aunt Polly has been so good to me," she began; but the man interrupted her sharply. The old irritability had come back to his face. (Pollyanna, Eleanor H Porter)
- e)** 'Oh. Well.' 'You think I'm that stupid I can't remember even the simplest thing? So useless I can't even **protect** of my own son?' 'Well, I wasn't sure, you see, that-' 'A fine one you are! You desert that child without a goodbye and then you call me up to tell me I'm not a good mother!' 'No, wait, Muriel -' 'Dominick is dead,' she said. 'What? Dominick?' (The Accidental Tourist, Anne Tyler)
- f)** 'Where's Velma?' That's all he said. The big black man nearly laughed. 'Velma? No Velma here, white boy. She's not in the business anymore, maybe.' 'Velma used to work here,' the big man said. He spoke as if he was dreaming. 'And **don't touch** my shirt with your dirty hand.' That annoyed the black. People didn't speak like that to him, not in his job, throwing drunks out of the club. (Farewell my lovely, Raymond Chandler)
- g)** With the magnifying glass, I saw that the wall was scratched, meaning that he had a long fingernail. Do you remember the grey dust I put in the envelope?' I nodded. 'The ash from a cigar. That all **seem logical**,' I admitted, 'but I still feel confused. How did one man make the other take the poison? Where did the blood come from? What was the reason for the murder? Why was the woman's ring there?' (A Study in Scarlet, Conan Doyle)

Figura 32 - Task 3 - Side B

Task 3 - Side B

h) Newman asked when he saw Nicholas's blood-covered face. 'Don't worry about me,' Nicholas replied. 'I'm not badly hurt. But I want you to tell me about my sister. I've learnt part of the story tonight. You must tell me the rest.' Newman made Nicholas **remove** his coat and washed the blood from his face. While he was doing this, he told him everything about his uncle, Sir Mulberry Hawk and his sister. (Nicholas Nickleby, Charles Dickens)

i) 'Good morning, Sergeant,' he said briskly. 'It's going to be a busy day. I want you to ask all the directors to come into the dining room. We're going to have a meeting. **Ensure** that they're all there in five minutes.' 'Yes, sir,' the Sergeant said. 'I'll tell them now.' 'One other thing, Sergeant,' the Inspector told him. 'When everybody's here, I want you to do something for me.' He handed the Sergeant a piece of paper. (Board Games, Butler James)

j) For years Hepzibah had looked forward to spending her life with the brother she loved so dearly. She had waited patiently for him and now she wanted to **bring** warm and **joy** to his life. She opened an old bookcase and took out several books that she had enjoyed reading many years before. She tried reading to Clifford, but he wasn't interested in the books and he did not like the sound of her hoarse voice. (The House of the Seven Gables, Nathaniel Hawthorne)

k) 'But we can't move at all in time. We are always in the present moment.' 'That is at the centre of my great discovery. Why can a modern man not hope that one day he might travel in time?' 'It doesn't **looks logical**,' said Filby. 'Possibly not,' said the Time Traveller. 'But now you begin to see the reason for my work on the geometry of four dimensions. Long ago I had an idea for a machine that can travel in any direction of space and time, as the driver wants.' (Time Machine, H G Wells)

l) "Cousin, here's a rose from the garden," she said, giving him a red flower. "Ah, let me hold it!" he said eagerly. "Thank you, my kind young girl. This rose **cause me the feeling of** young again." He looked around and saw the painting of old Colonel Pyncheon on the wall. "Hepzibah!" he cried. "Why do you keep that horrible portrait on the wall? (The House of the Seven Gables, Nathaniel Hawthorne)

m) Lear's best-loved child is dead, and he can do nothing to change that. At such moments, most dramatists of Shakespeare's time wrote long speeches for the actor. But Shakespeare always uses simple language in scenes like these. He wanted to **ensure** that everyone in the audience understood what was happening. And so King Lear, before he dies of a broken heart, says that he will never see his daughter again: 'Never, never, never, never, never.' (Shakespeare- His Life and Plays, Will Fowler)

n) Little Johnny Bold had been troubled for the last few days with his first tooth, and Miss Thorne was shocked to find that Eleanor was giving him some dreadfully modern medicine, recommended by one of the local doctors. '**Be careful**, my dear,' she said, looking very serious, 'that that man doesn't harm your little boy. But then,' speaking more in pity than in anger, 'I don't know which doctor you can trust now. Poor dear old Dr Bumpwell, of course-' (Barchester Towers, Anthony Trollope)

o) I shook her off and in another moment I was inside the well. I saw her frightened face over the top and smiled to **comfort her**. 'I had to climb down a well perhaps two hundred metres deep. There were metal bars all the way down, but these were made for a person much smaller and lighter than I was and I was quickly tired by the climb. And not simply tired! (Time Machine, H G Wells)

p) I sat down. The silence was worse than at lunchtime. Mrs Blakemore only ate half of her scrambled eggs. 'If you don't eat your food, you won't get better,' I said. 'I hate to waste things, but this tastes terrible,' she complained. **Put it in other place**. I want a glass of hot milk.' I went back to the kitchen and I told Bernard. 'Oh, she always says my food tastes terrible,' he said. 'She wants a glass of hot milk,' I told him. 'Does she?' he said. 'Well, I'm not making it for her. (A Taste of Murder, Sue Arengo)

2º Passo: Peça aos alunos para que realizem a *Task 3*.

Para responder às questões da *Task 3*, os alunos precisam ler os trechos retirados dos *readers* e identificar a parte destacada no texto. Eles devem substituir essas partes em negrito por expressões com *make* e *take*, aprendidas na *Task 2*. Dessa forma, com o conhecimento prévio adquirido na atividade anterior, os alunos podem apresentar uma facilidade em resolver os exercícios, verificando, assim, seu progresso de aprendizagem.

As respostas da *Task 3* podem ser observadas na Figuras 34 e 35.

Figura 34 - Respostas da Task 3 - Side A

Task 3 - Side A - Answers

This activity demonstrates how prior knowledge of words that co-occur with the verbs take and make helps students deduce from the context which words to use. As students have previously observed in the Key Word In Context (KWIC) tool, the potential occurrences of various words alongside the verbs have been elucidated. With this enhanced understanding and interpretive capacity, students might choose the appropriate word." Solving this type of exercise can facilitate the comprehension of new words, especially collocations, which can improve the student's ability to understand when reading.

1. Rewrite the excerpts below taken from Subcorpus B2, replacing the highlighted expressions with expressions with take and make, preserving the same meaning.

a) And now he's dead.' 'Of course she did,' said Marcus. 'Nobody likes the police. I do not like the police. But I have to keep good relations with them because of my hotel. You want me to give them money?' 'We want you to **take the body away** quietly.' 'That is very nice, I agree. I, too, I do not want a body in my hotel. But it is not so easy to do.' 'I think it could be,' said Dakin. 'You've got a doctor in your family, haven't you?' (They Came to Baghdad, Agatha Christie)

b) Marion Kingship didn't really love anybody. But she had her apartment, and she loved that. Every Saturday, she spent the day cleaning it. And she often dreamed that one day, a good, kind man would visit her there - someone who would love her and take **care of her**. "Will he ever come?" (A Kiss Before Dying, Ira Levin)

c) The snakes had gone now that the monster was dead, and the water no longer burned. Beowulf swam strongly for the land, pulling his heavy load behind him. The Geats were full of joy when they saw their hero reach the land. They helped him to **take off his armor**, and to carry the monster's head. They put the head on a spear, and they set off for Hrothgar's Hall. (Beowulf, Robert_Hill)

d) "What about me? How can I be 'glad' about anything without you? But if I had you for my own little girl, I'd be glad for anything; and I'd try to make you glad, too, my dear. All my money, to the last cent, will go to **make you happy**." Pollyanna looked shocked. "Aunt Polly has been so good to me," she began; but the man interrupted her sharply. The old irritability had come back to his face. (Pollyanna, Eleanor H Porter)

e) 'Oh. Well.' 'You think I'm that stupid I can't remember even the simplest thing? So useless I can't even **take care** of my own son?' 'Well, I wasn't sure, you see, that-' 'A fine one you are! You desert that child without a goodbye and then you call me up to tell me I'm not a good mother!' 'No, wait, Muriel -' 'Dominick is dead,' she said. 'What? Dominick?' (The Accidental Tourist, Anne Tyler)

f) 'Where's Velma?' That's all he said. The big black man nearly laughed. 'Velma? No Velma here, white boy. She's not in the business anymore, maybe.' 'Velma used to work here,' the big man said. He spoke as if he was dreaming. 'And **take your dirty hand off** my shirt.' That annoyed the black. People didn't speak like that to him, not in his job, throwing drunks out of the club. (Farewell my lovely, Raymond Chandler)

g) With the magnifying glass, I saw that the wall was scratched, meaning that he had a long fingernail. Do you remember the grey dust I put in the envelope?' I nodded. 'The ash from a cigar. That all **makes sense**,' I admitted, 'but I still feel confused. How did one man make the other take the poison? Where did the blood come from? What was the reason for the murder? Why was the woman's ring there?' (A Study in Scarlet, Conan Doyle)

Figura 35 - Respostas da Task 3 - Side B

Task 3 - Side B - Answers

h) Newman asked when he saw Nicholas's blood-covered face. 'Don't worry about me,' Nicholas replied. 'I'm not badly hurt. But I want you to tell me about my sister. I've learnt part of the story tonight. You must tell me the rest.' Newman made Nicholas **take off** his coat and washed the blood from his face. While he was doing this, he told him everything about his uncle, Sir Mulberry Hawk and his sister. (Nicholas Nickleby, Charles Dickens)

i) 'Good morning, Sergeant,' he said briskly. 'It's going to be a busy day. I want you to ask all the directors to come into the dining room. We're going to have a meeting. **Make sure** that they're all there in five minutes.' 'Yes, sir,' the Sergeant said. 'I'll tell them now.' 'One other thing, Sergeant,' the Inspector told him. 'When everybody's here, I want you to do something for me.' He handed the Sergeant a piece of paper. (Board Games, Butler James)

j) For years Hepzibah had looked forward to spending her life with the brother she loved so dearly. She had waited patiently for him and now she wanted to **make his life warm and happy**. She opened an old bookcase and took out several books that she had enjoyed reading many years before. She tried reading to Clifford, but he wasn't interested in the books and he did not like the sound of her hoarse voice. (The House of the Seven Gables, Nathaniel Hawthorne)

k) 'But we can't move at all in time. We are always in the present moment.' 'That is at the centre of my great discovery. Why can a modern man not hope that one day he might travel in time?' 'It doesn't **make sense**,' said Filby. 'Possibly not,' said the Time Traveller. 'But now you begin to see the reason for my work on the geometry of four dimensions. Long ago I had an idea for a machine that can travel in any direction of space and time, as the driver wants.' (Time Machine, H G Wells)

l) "Cousin, here's a rose from the garden," she said, giving him a red flower. "Ah, let me hold it!" he said eagerly. "Thank you, my kind young girl. This rose **makes me feel** young again." He looked around and saw the painting of old Colonel Pyncheon on the wall. "Hepzibah!" he cried. "Why do you keep that horrible portrait on the wall?" (The House of the Seven Gables, Nathaniel Hawthorne)

m) Lear's best-loved child is dead, and he can do nothing to change that. At such moments, most dramatists of Shakespeare's time wrote long speeches for the actor. But Shakespeare always uses simple language in scenes like these. He wanted to **make sure** that everyone in the audience understood what was happening. And so King Lear, before he dies of a broken heart, says that he will never see his daughter again: 'Never, never, never, never, never.' (Shakespeare- His Life and Plays, Will Fowler)

n) Little Johnny Bold had been troubled for the last few days with his first tooth, and Miss Thorne was shocked to find that Eleanor was giving him some dreadfully modern medicine, recommended by one of the local doctors. '**Take care**, my dear,' she said, looking very serious, 'that that man doesn't harm your little boy. But then,' speaking more in pity than in anger, 'I don't know which doctor you can trust now. Poor dear old Dr Bumpwell, of course-' (Barchester Towers, Anthony Trollope)

o) I shook her off and in another moment I was inside the well. I saw her frightened face over the top and smiled to **make her feel better**. 'I had to climb down a well perhaps two hundred metres deep. There were metal bars all the way down, but these were made for a person much smaller and lighter than I was and I was quickly tired by the climb. And not simply tired!' (Time Machine, H G Wells)

p) I sat down. The silence was worse than at lunchtime. Mrs Blakemore only ate half of her scrambled eggs. 'If you don't eat your food, you won't get better,' I said. 'I hate to waste things, but this tastes terrible,' she complained. **Take it away**. I want a glass of hot milk.' I went back to the kitchen and I told Bernard. 'Oh, she always says my food tastes terrible,' he said. 'She wants a glass of hot milk,' I told him. 'Does she?' he said. 'Well, I'm not making it for her.' (A Taste of Murder, Sue Arengo)

Ao final da atividade, os alunos terão verificado como o conhecimento prévio de palavras que coocorrem com os verbos *take* e *make* e de seus possíveis sentidos podem ajudá-los a deduzir pelo contexto quais palavras usar para realizar a *Task 3*. Com essa compreensão, a capacidade interpretativa pode ser aprimorada. Dessa forma, a resolução desse tipo de exercício com os trechos dos *readers* pode facilitar a compreensão de novas palavras, principalmente colocações, o que pode melhorar a capacidade de compreensão do aluno, durante a leitura completa dos *readers*.

3.3 Instruções para a atividade de reflexão

1º Passo: Peça aos alunos que, individualmente, criem um infográfico com o intuito de representar o progresso de cada um, a partir das respostas pessoais das seguintes perguntas:

a) Na primeira aula, foi ensinado o uso e o manuseio das ferramentas de análise linguística. Na sua opinião, qual a importância de aprender novos recursos e como eles podem ser úteis para o aprendizado de um novo idioma?;

b) Na sua opinião, qual a melhor maneira de testar um conhecimento novo adquirido? De que forma as atividades da *Task 1* te auxiliaram na prática dos seus novos saberes?;

c) De que forma esse conhecimento prévio te auxiliou no aprendizado de novas expressões com o verbo *take* e *make*? Com base na sua experiência, como a aplicação desse conhecimento te ajudou a melhorar a compreensão dos *readers*?

Para essa atividade, os alunos devem lembrar do passo a passo que realizaram, desde a primeira aula, até o final da última. Eles precisam pensar nas tarefas solicitadas e no que foi aprendido. Espera-se que os estudantes cheguem à conclusão de que o uso de dados facilita a compreensão do uso das colocações dos verbos *make* e *take*, e, por conseguinte, que isso pode facilitar a compreensão de futuras leituras.

O infográfico pode ser feito tanto em folha A4, como em programas *online* no computador no laboratório de informática. Na Figura 36, é mostrado um modelo de como os alunos podem fazer esse infográfico no computador. As informações contidas no infográfico podem variar de aluno para aluno, uma vez que as perguntas serão respondidas de acordo com a experiência de aprendizagem de cada um.

Figura 36 - Modelo do Infográfico



Fonte: O autor (2024)

Ao utilizar ferramentas como *KWIC* para identificar linhas de concordância e *Graphcoll* para verificar as colocações de verbos frequentes como *Make* e *Take*, juntamente com a análise de um *corpus* pedagógico, para compreender como colocados específicos são empregados com esses verbos, os alunos podem aprimorar suas habilidades de leitura. Através da abordagem *Data-driven Learning*, os estudantes podem não apenas entender melhor as colocações, mas também captar nuances de significado e uso que contribuem para uma compreensão mais profunda do idioma. Essa abordagem pode ampliar a competência linguística, capacitando os aprendizes a interpretar textos de forma mais precisa e eficaz, o que, por conseguinte, enriquece sua experiência com a literatura.