

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS**

ANNALENA DE SOUZA GUEDES

**VERBOS DO INGLÊS ACADÊMICO ESCRITO E SUAS COLOCAÇÕES: UM
ESTUDO BASEADO EM UM *CORPUS* DE APRENDIZES BRASILEIROS DE
INGLÊS**

**BELO HORIZONTE
2017**

ANNALENA DE SOUZA GUEDES

VERBOS DO INGLÊS ACADÊMICO ESCRITO E SUAS COLOCAÇÕES: UM ESTUDO BASEADO EM UM *CORPUS* DE APRENDIZES BRASILEIROS DE INGLÊS

Tese 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 Doutora em Linguística Aplicada.

Área de concentração: Linguística Aplicada.
Linha de Pesquisa: Ensino-Aprendizagem de Línguas Estrangeiras.
Orientadora: Profa. Dra. Deise Prina Dutra.

**Belo Horizonte
Faculdade de Letras da UFMG
2017**

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca F.ALE/UFMG

G924v Guedes, Annallena de Souza.
 Verbos do inglês acadêmico escrito e suas colocações [manuscrito].: um estudo baseado em um Corpus de aprendizes brasileiros de inglês / Annallena de Souza Guedes. – 2017.
 199 f., enc.: il., grafs, (color)

Orientadora: Deise Prina Dutra.

Area de concentração: Linguística Aplicada.

Linha de Pesquisa: Ensino/Aprendizagem de Línguas Estrangeiras.

Tese (doutorado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f.162-170.

Apêndices: f.171-197.

Anexos: f.198.

1. Língua inglesa – Estudo e ensino – Falantes estrangeiros – Teses. 2. Língua inglesa – Verbos – Teses. 3. Aquisição de segunda linguagem – Teses. 4. Linguística de corpus – Teses. 5. Língua inglesa – Retórica – Teses. 6. Redação acadêmica – Teses. I. Dutra, Deise Prina. I. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 420.7



FOLHA DE APROVAÇÃO

VERBOS DO INGLÊS ACADÊMICO ESCRITO E SUAS COLOCAÇÕES: UM ESTUDO BASEADO EM UM CORPUS DE APRENDIZES BRASILEIROS DE INGLÊS

ANNALENA DE SOUZA GUEDES

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

Aprovada em 18 de dezembro de 2017, pela banca constituída pelos membros:

Prof(a). Deise Prina Dutra - Orientador
UFMG

Prof(a). Valdênia de Carvalho e Almeida
Universidade Federal de Viçosa

Prof(a). Patrícia Pereira Bértoli
UERJ

Prof(a). Heliana Ribeiro de Mello
UFMG

Prof(a). Bárbara Malveira Orfanó
UFMG

Belo Horizonte, 18 de dezembro de 2017.

À minha amada filha Isa Bella e aos meus pais Aleide e Ionildo.
À minha Bahia, minha terra, meu encanto.

AGRADECIMENTOS

Ao meu amado e soberano Deus por me permitir, com Sua infinita bondade, trilhar um caminho de tantas lutas e conquistas. Esse sonho, antes de ser meu, foi Dele! “Quem primeiro lhe deu alguma coisa, para que Ele lhe recompense? Portanto, Dele, por Ele e para Ele são todas as coisas” (Romanos 11, 35-36). Obrigada, Senhor, por dizer tão fortemente ao meu coração: “Não temas, porque eu sou contigo; não te assombres, porque eu sou teu Deus; eu te fortaleço, e te ajudo, e te sustento com a destra da minha justiça” (Isaías 41, 11).

À Isa Bella, minha obra-prima, joia rara e preciosa, por me ensinar, pelo olhar doce e sorriso cativante, que a vida é simples e, ao mesmo tempo, tão linda...Renasci quando você nasceu e, assim, cresceu minha coragem, astúcia e força de vontade para batalhar ainda mais pelos meus sonhos. Você me encanta, me inspira, me completa, me renova e me fortalece a cada milésimo de segundo. Tudo isso é pra você e por você! *From the bottom of my heart, I love you!*

Aos meus pais Aleide e Ionildo, minha rocha firme, meu porto seguro, toda a gratidão desse universo! Obrigada pelo amor, carinho, apoio, incentivo, por não medirem esforços para me educar e fazer de mim quem sou e, principalmente, por cuidarem tão bem de minha Bella para que eu pudesse estudar. Nada disso teria sentido sem vocês dois... Amo vocês!

À minha orientadora, Profa. Dra. Deise Prina Dutra, pelo inigualável exemplo e brilhantismo como pessoa e profissional. Minha profunda gratidão pela majestosa orientação, pela oportunidade de aprender tanto contigo e, sobretudo, por toda a confiança que me foi depositada. *Wholeheartedly, you are quite special for me, dear Deise!*

Às professoras que compuseram a Banca de Qualificação, Profa. Dra. Bárbara Malveiro, Profa. Dra. Heliana Mello e Profa. Dra. Valdênia Almeida pelas valiosas contribuições sugeridas no Exame de Qualificação.

À Profa. Dra. Ursula Wingate, pelo aceite e pela gentileza em coorientar esta pesquisa de forma tão competente e cuidadosa durante o período de estágio doutoral na King’s College – University of London, em Londres, no Reino Unido. Sinto-me honrada por tamanho

privilégio!

Ao Prof. Dr. Christopher Tribble, por tão gentilmente ter me permitido participar de suas aulas durante o período de estágio doutoral na King's College London. Obrigada pelas sugestões preciosas e discussões tão profícuas que ajudaram a alinhar os rumos desta tese. De certo, tento, neste trabalho, deixar pequenas marcas da excelência de suas contribuições.

Ao meu grande e amado amigo Urbano Cavalcante, pela amizade sincera, parceria, cumplicidade e imensa ajuda ao longo da minha caminhada desde a graduação. Obrigada pelas inúmeras leituras dos meus textos e por todas as vezes que me dissestes “Vai, Anna! Você consegue!”. Certamente, se não fossem suas palavras de motivação, não teria aceitado tantos desafios, tampouco teria vivido tantas experiências acadêmicas e profissionais de forma tão intensa. És exemplo e força pra mim, meu irmão!

Ao meu amado Homero Andrade, pelo carinho e incentivo, pela leitura acurada de alguns dos meus textos, pela despreziosa sugestão da epígrafe de abertura desta tese e, especialmente, por sempre vibrar e brindar minhas vitórias. Sua ajuda, paciência e tranquilidade para comigo, durante a etapa final da escrita deste texto, foram extremamente importantes. Te amo muito...

Às amigas Karina Pitta, Mariluce Silva e Neidinha Almeida, minhas irmãs de alma e coração, pela amável e marcante presença na minha vida, nos momentos mais difíceis e também, nos mais felizes. Vocês são anjos sem asas...

Às minhas queridas Andreisa Queiroz, Carolina Calmon, Dona Marlene, Geisa Fróes, Girlene Écio, Ianna Cerqueira, Juliana Menezes, Katyane Alves, Laura Helena, Leonor Guimarães, Mércia Pimentel, Núbia Enedina, Roseane Leite, Selma Bezerra, Tatiana Polliana, pela amizade e toda a ajuda que me dedicam.

À Catiane Rocha e Marijane Correia, pelo incentivo ao ingresso no doutorado, pelas discussões acerca do pré-projeto de pesquisa e pelo apoio que me foi dado, principalmente, durante o período de participação como ouvinte no Programa de Pós-Graduação em Língua e Cultura, na Universidade Federal da Bahia. De coração, muito obrigada, meninas!

À minha doce irmã Patrícia Bonfim, por me ensinar a ver o amor de Deus nas pequenas coisas

e por ter sempre uma palavra de conforto quando me falta o chão. Suas orações e amizade me sustentam e me fazem seguir sempre em frente.

Aos amigos do Instituto Federal da Bahia, Christian Ricardo, Danilo Souza, João Pedro Valadares e Thiago Barbosa, por toda a atenção e apoio ao longo desses anos.

À minha orientadora durante o período de mestrado, Profa Dra. Inez Mattoso Silveira, pelo incentivo constante e pelo legado da sua grandiosa contribuição na minha vida acadêmica.

À Profa. Dra. Élide Ferreira e à Profa Dra. Patrícia Rosa que, carinhosamente, há alguns anos, me despertaram para a possibilidade de ser pesquisadora e muito me auxiliaram para que essa tarefa se concretizasse. Quanto brilho e humanidade emanam de vocês!

Às minhas tias Dora, Leda, Lilian, Maria, Mene, Mira, Nete, Sônia, Vera, Yolanda e Yone. Às minhas primas, companheiras desde a infância e parceiras de todas as horas, Giovanna e Renata. À minha prima sempre presente Cristiane, pelo afeto, atenção, torcida e orações destinadas a mim durante todo esse percurso. Gratidão também aos meus primos Helinho, Júnior e Lismar, e aos meus tios queridos Carlinhos, Lan, Mazinho e Yonaldo (*in memoriam*) pelo cuidado e preocupação que sempre tiveram para comigo.

Às meninas da República Vaidanadas, Amanda Soares, Caroline Duarte, Jéssica Prates e Luíza Miligrana, pela companhia e amizade durante todo o tempo de convívio em Belo Horizonte.

Aos colegas do Poslin que se tornaram amigos: Caroline Martins, Claudiene Diniz, Jackson Vitória, Joana Costa, Maria José, Nara Nogueira, Natália Giarola, Patrícia Leite, Tâmara Milhomem e Waldemar Duarte. Que grande presente a Universidade Federal de Minas Gerais (UFMG) me deu!

Aos colegas do Grupo de Pesquisa Learner Corpora Research, da UFMG, pelas discussões significativas, intercâmbio de ideias e por toda a ajuda durante esse percurso.

Aos colegas do Grupo de Pesquisa Linguagem e Representação do IFBA –*campus* Salvador, pelo apoio e pela compreensão à minha ausência.

À Profa. Dra. Bárbara Malveira, Jéssica Ceritello e a Rodrigo Castro, pela disposição e auxílio na compilação dos *corpora* utilizados nesta pesquisa. Sou imensamente grata a vocês!

Aos meus professores da UFMG, toda a minha admiração e gratidão: Profa. Dra. Adriana Pagano, Prof. Dr. Brian Street (*in memoriam*), Prof. Dr. César Nardelli, Profa. Dra. Delaine Cafiero, Profa. Dra. Heliana Mello, Profa. Dra. Maria Lúcia Castanheira e Prof. Dr. Sérgio Cirino. Em especial, agradeço à Profa. Dra. Laura Miccoli que, através das lentes fenomenológicas, me provocou uma incessante reflexão sobre o papel fundamental do professor de inglês na escola pública, que constitui meu lugar de fala e atuação. De certo, Laura, suas aulas mexeram profundamente com meu modo de pensar o ensino de línguas e acabei tornando-me uma pessoa e profissional mais consciente da tarefa de refletir e educar, pensando no que mais se ensina quando se ensina inglês.

Ao C.C.A.A de Ibicaraí, em especial, a Petronila Nascimento que, por ter sido a minha primeira professora de inglês, foi a principal inspiradora do meu amor pelo ensino de línguas. Obrigada também, Nila, por ter me convidado, ainda tão menina, a começar a engatinhar pelos caminhos da docência. *I'm greatly thankful!*

Ao Instituto Federal da Bahia, pela concessão da licença para estudos, o que me possibilitou a dedicação integral ao doutoramento.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pela concessão da bolsa de estágio doutoral na King's College London- University of London, em Londres, no Reino Unido.

Aos amigos da Casa Londres, Carla Semedo, Clóvis Lima e Júlio Frantz, pelas conversas, parceria e toda a ajuda durante o período do estágio nas terras da rainha.

A todos os professores que passaram pela minha história e a todos os estudantes que a vida me presenteou ao longo dessa caminhada como professora e, mais particularmente, aos meus queridos do Instituto Federal da Bahia – *campus* Ilhéus e *campus* Porto Seguro. De fato, vocês foram as molas propulsoras do meu desejo em tornar-me pesquisadora e, dessa forma, poder contribuir de modo mais significativo, para o ensino e a aprendizagem de inglês no

nosso país.

À universidade pública, gratuita e de qualidade, da qual sou fruto, que me oportunizou, desde a graduação, reconhecer que, através do conhecimento e da luta, é possível mudar o rumo das nossas vidas, da nossa história. Em especial, à Universidade Estadual de Santa Cruz, à Universidade Federal de Alagoas e à Universidade Federal de Minas Gerais, a quem serei eternamente grata por tudo o que lá vivi. Em mim, reside a esperança de mais pessoas com direito de acesso e permanência em universidades públicas de qualidade neste país, especialmente aquelas que nunca sequer sonharam ou se imaginaram universitárias, mas que trazem consigo a vontade, o vigor e a paixão por aprender.

Ao Ministério Universidades Renovadas (MUR-BH), da Renovação Carismática Católica, mais particularmente, ao Grupo de Oração FAFICH, por fazer de uma sala de aula da UFMG um palco de encontro pessoal com Deus. O sorriso de cada um dos servos na acolhida, a palavra aberta e sincera, cantada e proferida, e o abraço da paz me serviam de antídoto e não me deixavam desanimar. De verdade, o GOU “é um sonho de amor para o mundo”.

How many roads must a man walk down
Before you can call him a man?
Yes and how many seas must a white dove sail
Before she sleeps in the sand?
Yes and how many times must cannonballs fly
Before they're forever banned?

The answer, my friend, is blowin' in the wind
The answer is blowin' in the wind

Yes, how many years must a mountain exist
Before it is washed to the sea
Yes and how many years can some people exist
Before they're allowed to be free?
Yes and how many times must a man turn his
head
And pretend that he just doesn't see?

The answer, my friend, is blowin' in the wind
The answer is blowin' in the wind

Yes, how many times must a man look up
Before he can really see the sky?
Yes and how many ears must one man have
Before he can hear people cry?
Yes and how many deaths will it take till he
knows
That too many people have died?

(Bob Dylan, "Blowin' in the Wind")

RESUMO

Estudos com *corpora* de aprendizes (HUNSTON, 2002) têm contribuído para a pesquisa na área de ensino e aprendizagem de língua adicional, principalmente ao utilizar as abordagens de análise linguística, tal como a Análise Contrastiva de Interlíngua (ACI) (GRANGER, 1996) que, dentre outros objetivos, busca comparar produções escritas de aprendizes a *scorpora* de falantes cuja primeira língua (L1) não é o inglês, inseridos em contextos acadêmicos. Considerando os princípios da ACI, esta pesquisa tem como propósito analisar a escrita de estudantes universitários brasileiros no gênero redação argumentativa, em disciplinas de Inglês para Fins Acadêmicos, verificando o uso dos Verbos do Inglês Acadêmico Escrito (VIAEs), aqui entendidos como verbos mais tipicamente usados no contexto acadêmico (GRANGER; PAQUOT, 2009) escrito. Decorrente desse objetivo central, este estudo busca ainda analisar as mais frequentes colocações formadas por “verbo + advérbio-ly”, dada a importância dessas unidades fraseológicas para a compreensão dos padrões linguísticos. Parte-se do pressuposto de que há uma tendência de os falantes cuja L1 não é o inglês fazerem determinadas escolhas lexicais (GILQUIN; PAQUOT, 2007) que se diferem do uso dos falantes de inglês como L1. Diante desse contexto, este estudo originou-se do desejo de analisar o uso dos Verbos do Inglês Acadêmico Escrito, utilizados pelos aprendizes. Nesse sentido, analisamos as redações argumentativas do Corpus do Inglês para Fins Acadêmicos (CorIFA), compilado na Universidade Federal de Minas Gerais, e comparamos com as redações argumentativas do British Academic Written English (BAWE), compilado no Reino Unido. Assim, utilizam-se as ferramentas da Linguística de Corpus para possibilitarem a evidência dos padrões de uso da língua em relação às categorias gramaticais. Os resultados desta pesquisa mostraram que existe uma grande variação lexical de verbos e de advérbios -ly em ambos os subcorpora, no entanto, no subcorpus do CorIFA, há uma variação lexical menor no que se refere aos VIAEs, o que demonstra uma preferência dos aprendizes de uso dos mesmos verbos, talvez, por uma questão de proficiência. Os resultados apontaram para a preferência pelo uso de verbos, classificados semanticamente como verbos de ação (e.g. *apply* e *provide*). Quanto às colocações, não foi possível detectar entre os subcorpora a força de atração entre os verbos e advérbios -ly por conta da sua baixa frequência. Conclui-se que a análise do subcorpus do CorIFA revela que os aprendizes universitários de inglês se apoiam no uso do léxico que lhes é mais familiar, no caso dos VIAEs estudados, bem como nas combinações verbo-advérbio (-ly). Este estudo pode ser base para futuras intervenções pedagógicas centradas nas necessidades de aprendizes do mesmo nível que os participantes do

subcorpus do CorIFA.

Palavras-chaves: *Corpus* de aprendizes. Verbos do Inglês Acadêmico Escrito. Padrões de uso.

ABSTRACT

Studies with learner corpora (HUNSTON, 2002) have contributed to the research in the foreign language teaching and learning area, mainly when linguistic analysis approaches, such as the Contrastive Interlanguage Analysis (CIA) (GRANGER, 1996), are used. CIA aims at comparing learners production to native speakers corpora, embedded in academic contexts. Considering the principles of the Contrastive Interlanguage Analysis (CIA), this research has the goal to analyze the writing of Brazilian university students in argumentative essays, verifying the use and the patterns of use of „Verbs from the Written Academic ENGLISH“ (VWAE), understood here as more typical verbs of the academic context (GRANGER & PAQUOT, 2009). Due to this central objective, this study also aims at analyzing the most frequent collocations (verb + *-ly* adverbs) due to the importance of these phraseological units for the linguistic pattern comprehension. We have assumed that there is a tendency to speakers whose first language (L1) is not English to make certain lexical choices (GILQUIN; PAQUOT, 2007) that differ from the native speakers“ patterns of use. In the light of the context the corpora analyzed were compiled, this study is originated from the wish of analyzing the frequency, use and patterns of use of „Verbs from the Written Academic English“, used by university students. For such, we analysed the argumentative essays in CorIFA (Corpus of English for Academic Purposes), compiled at the Federal University of Minas Gerais and we compared it to the argumentative essays of BAWE (British Academic Written English), compiled in the United Kingdom. Hence, we use Corpus Linguistics tools, since they allow us to see the use of grammatical categories. The results of this research showed that there is a great lexical variation of verbs and *-ly* adverbs in both subcorpora; however, in CorIFA subcorpus, there is less lexical variation in relation to VWAEs, which demonstrates learner“s preference to use the same verbs, maybe, due to a proficiency issue. The results pointed out to the preference of using the verbs, semantically classified as action verbs (e.g apply and provide). In relation to collocations, it was not possible to identify the associate strength between verbs and *-ly* adverbs in both subcorpora because of their low frequency. We conclude that the analysis of CorIFA subcorpus shows that the Brazilian university learners trust the use of lexis that is more familiar to them, in the case of the studied VWAEs, as well as in the combinations „verb + *-ly* adverb“. This study may be the basis for future pedagogical interventions focused on learners' needs in the same level as the CorIFA subcorpus participants.

Key-words: Learner corpus. Verbs of the Written Academic Discourse. Patterns of use.

LISTA DE FIGURAS

Figura 1 – Tela inicial do <i>website</i> do CorIFA	887
Figura 2 – Tela do AntConc 3.4.3 com as linhas de concordância do verbo <i>represent</i> no CorIFA.....	890
Figura 3 – Interface do BAWE.....	932
Figura 4 – Nuvem de palavras gerada a partir da ferramenta Thesaurus	976
Figura 5 – Exemplo de distribuição de dados para obtenção do LL	1043
Figura 6 – Distribuição de dados para obtenção do LL.....	1054
Figura 7 – Distribuição de dados para obtenção do LL – <i>acquire</i>	1054

LISTA DE GRÁFICOS

Gráfico 1 – Número de ocorrências dos verbos lexicais mais frequentes no subcorpus do CorIFA.....	910
Gráfico 2 – Distribuição geral dos tipos de verbos do subcorpus do CorIFA.....	1087
Gráfico 3 – Frequência normalizada dos VIAEs presentes no subcorpus do CorIFA	11009
Gráfico 4 – Frequência normalizada dos VIAES no subcorpus do BAWE	1110
Gráfico 5 – Distribuição dos VIAEs ($f < 20$) no CorIFA por domínio semântico (BIBER et al., 1999).....	1210
Gráfico 6 – Distribuição dos valores de <i>log-likelihood</i> (LL) dos VIAEs mais frequentes no subcorpus do CorIFA.....	1243
Gráfico 7 – VIAEs de maior valor de LL.....	1254
Gráfico 8 – VIAEs de menor valor de LL	1254

LISTA DE QUADROS

Quadro 1 –Estrutura do subcorpus do CorIFA.....	78
Quadro 2 – Estrutura do CorIFA: número de textos, gênero e nível dos alunos.....	80
Quadro 3 – Número mínimo de palavras das tarefas dos textos do subcorpus do CorIFA.....	81
Quadro 4 –Número de textos em cada família de gênero no BAWE.....	92
Quadro 5 – Distribuição de redações argumentativas por área	8394
Quadro 6 – Quadro descritivo do teste piloto para a contagem de <i>tokens</i>	97
Quadro 7 – Quadro descritivo para a contagem de <i>tokens</i>	98
Quadro 8 – Frequência normalizada de verbos lexicais comuns de cada domínio semântico (sem as formas lematizadas) baseado em Biber et al. (1999).....	107
Quadro 9 – Frequência absoluta e frequência normalizada dos verbos do inglês acadêmico escrito mais frequentes no subcorpus do CorIFA e do BAWE	110
Quadro 10 –Distribuição da frequência dos VIAEs no subcorpus do CorIFA ($f < 20$)	117
Quadro 11 – Valor de <i>log-likelihood</i> e indicação de sobreuso ou subuso.....	1232
Quadro 12 – Síntese dos resultados dos VIAEs – subcorpus do CorIFA e subcorpus do BAWE	1265
Quadro 13 – Número de ocorrências dos advérbios de posicionamento epistêmico	13029
Quadro 14 – Distribuição dos advérbios (número de ocorrências) do tipo maximizadores e seus colocados verbais.....	13029
Quadro 15 – Frequência para verbo <i>adopt</i>	1354
Quadro 16 – Frequência para verbo <i>adopt</i>	1354
Quadro 17 – Frequência para verbo <i>develop</i>	1365
Quadro 18 – Frequência para verbo <i>develop</i>	13837
Quadro 19 – Frequência para verbo <i>improve</i>	13938
Quadro 20 – Frequência para verbo <i>increase</i>	14544
Quadro 21 – Frequência para verbo <i>increase</i>	1465
Quadro 22 – Frequência para verbo <i>provide</i>	1476
Quadro 23 – Frequência para verbo <i>provide</i>	14847
Quadro 24 – Distribuição dos advérbios do tipo intensificadores.....	15049
Quadro 25 – Verbos do CorIFA com frequência mínima de cinco ocorrências (organizado por ordem decrescente de frequência)	1710
Quadro 26 – Advérbios <i>-ly</i> utilizados com os VIAEs em ambos os subcorpora	1932
Quadro 27 – Análise de <i>Mutual Information</i> dos cinco VIAEs mais frequentes do subcorpus	

do CorIFA..... 196

LISTA DE TABELAS

Tabela 1 –Principais características das abordagens de ensino de inglês relacionadas à escrita acadêmica	47
Tabela 2 – Classificação e caracterização dos verbos lexicais (BIBER et al, 1999)	60
Tabela 3 – As sete principais categorias semânticas de advérbios.....	68
Tabela 4 – Gêneros textuais solicitados para os estudantes do IFA.....	
Tabela 5 – Instruções das tarefas dos textos do subcorpus do CorIFA.	82
Tabela 6 – Instruções para textos do IFA– 2013.2.	83
Tabela 7 – Exemplo de código de texto do CorIFA e suas especificações.	85
Tabela 8 – Advérbios <i>-ly</i> comuns ao subcorpus do CorIFA e ao subcorpus do BAWE.	127

LISTA DE ABREVIATURAS

ACI	Análise Contrastiva de Interlíngua
ALESS	Active Learning of English for Science Students
ASL	Aquisição de Segunda Língua
AVL	Academic Vocabulary List
CAE	Cambridge English: Advanced
CAPES	Coordenação de Aperfeiçoamento de Pessoal
CLEC	Chinese Learner English Corpus
COCA	Contemporary Corpus of American English
CORIFA	Corpus do Inglês para Fins Acadêmicos
CORISF	Corpus do Inglês sem Fronteiras
EAP	English for Academic Purposes
ESP	English for Specific Purposes
FCE	Cambridge English: First
ICLE	International Corpus of Learner English
IELTS	International English Language Testing System
IFA	Inglês para Fins Acadêmicos
ISF	Inglês sem Fronteiras
L1	Língua 1
L2	Língua 2
LC	Linguística de Corpus
LGSWE	Longman Grammar of Spoken and Written English
LGSWEC	Longman Grammar of Spoken and Written English Corpus
LL	Log-likelihood
LOCNESS	Louvain Corpus of Native English Essays
MCTI	Ministério da Ciência, Tecnologia e Inovação
MEC	Ministério da Educação
NICLE	Norwegian subcomponent of International Corpus of Learner English
TOEFL ITP	Test of English as a Foreign Language – Institutional Testing Program
TOEFL iBT	Test of English as a Foreign Language – Internet-based Testing
UFMG	Universidade Federal de Minas Gerais
VIAE	Verbos do Inglês Acadêmico Escrito

SUMÁRIO

INTRODUÇÃO.....	22
CAPÍTULO 1.....	34
FUNDAMENTAÇÃO TEÓRICA.....	34
A Linguística de Corpus.....	35
<i>Corpus</i> de aprendizes	36
A interlíngua de aprendizes de inglês e a Análise Contrastiva de Interlíngua (ACI)	
.....	38
1.4 O discurso acadêmico escrito	45
Escrita acadêmica e abordagens para o ensino de inglês	47
O Inglês para Fins Acadêmicos.....	49
<i>O gênero acadêmico “redação argumentativa”</i>	52
<i>As listas de vocabulário acadêmico: a Academic Word List (AWL), a New Academic Word List (NAWL) e a New Academic Vocabulary List (AVL)</i>	55
Verbos lexicais: caracterização e classificação	59
1.7.1 <i>Os verbos lexicais no discurso acadêmico</i>	63
Os advérbios no discurso acadêmico	66
As colocações e a fraseologia no inglês acadêmico escrito	72
CAPÍTULO 2	77
PROCEDIMENTOS METODOLÓGICOS	77
Perfil dos aprendizes	77
Caracterização e compilação do <i>corpus</i> de estudo.....	78
<i>O subcorpus do CorIFA</i>	79
<i>Dados sobre as instruções das tarefas do subcorpus do CorIFA</i>	80
<i>Compilação do CorIFA</i>	85
2.3 Organização dos dados do subcorpus de estudo.....	88
2.3.1 <i>O programa AntConc 3.4.3: um dispositivo de extração de dados</i>	89
Distribuição dos verbos lexicais mais frequentes no subcorpus do CorIFA.....	90
O British Academic Written English Corpus (BAWE).....	92
A AVL e o BAWE	96
2.6.1 <i>O subcorpus de referência: compilação e descrição</i>	97
A lista de referência.....	100
<i>Instrumentos</i>	100

<i>Procedimentos de análise</i>	101
Descrição do teste estatístico para extração de dados: o <i>log-likelihood</i> (LL)	103
CAPÍTULO 3	106
RESULTADOS E ANÁLISE DE DADOS.....	106
Introdução.....	106
Resultados gerais.....	106
<i>Frequência geral</i>	107
<i>Distribuição de domínios semânticos no CorIFA e no BAWE</i>	108
<i>Frequência normalizada dos VIAEs mais frequentes no subcorpus do CorIFA e do BAWE</i>	109
<i>Características compartilhadas entre os subcorpora do CorIFA e do BAWE</i>	113
<i>Características não compartilhadas entre os subcorpora do CorIFA e do BAWE</i>	115
Resultados do teste <i>log-likelihood</i> com os verbos do inglês acadêmico escrito do CorIFA	122
Síntese dos resultados.....	126
As colocações no subcorpus do CorIFA e do BAWE: descrição das colocações formadas por advérbios <i>-ly</i> do tipo maximizadores	127
Resultados –colocações com <i>adopt, develop, improve, increase e provide</i>	134
Os advérbios e sua classificação semântica	149
Conclusão do capítulo	156
CONCLUSÃO.....	156
Limitações do estudo	159
APÊNDICES	171
ANEXOS	198

INTRODUÇÃO

Atualmente, a importância e as demandas de comunicação entre povos distintos ou não linguisticamente é evidente. Nesse patamar, o inglês se configura como língua de extrema relevância na disseminação do conhecimento acadêmico-científico, considerando o número de publicações nessa língua que circulam por todo o mundo. Desse modo, pensar nesse contexto implica pensar nas práticas acadêmicas e, mais particularmente, no discurso acadêmico escrito, nas suas especificidades e, do ponto de vista pedagógico, exercendo o papel de professor, nas formas de compreender as escolhas linguísticas do aprendiz e de contribuir para a sua aprendizagem.

Trazendo essas reflexões para o ensino de inglês, vale lembrar que, nos bastidores de quem ensina, cabem muitas histórias. A que será brevemente contada aqui foi protagonizada pela autora e, desse modo, com a licença da palavra, ela autoriza-se a brevemente narrá-la, nesta introdução, em primeira pessoa. Como professora de inglês, minha experiência se iniciou desde a mais tenra idade. Em 1990, comecei a estudar inglês em uma turma para crianças que se chamava Children's Course. Com a ansiedade própria da idade, cada vez mais me encantava com a língua. Lembro-me dos personagens das lições dos livros (Jean, Daniel, Sally, Leroy, Doug, Mr. and Mrs. Strickland) e tantos outros cujas vozes, até hoje, ressoam aos meus ouvidos. Parece que eles nunca se calaram... Ao concluir o nível básico, em 1995, nessa época com 12 anos, fui convidada pela proprietária do curso de línguas que, na ocasião, era também minha professora de inglês, a trabalhar como monitora de uma turma de alunos de 7 anos de idade. Minha função era ajudá-los a realizar as atividades orais e escritas. Participei do encontro pedagógico e logo precisei, também, aprender a lidar com retroprojetor e gravador de fita cassete, afinal, em algumas aulas, eu ocuparia o lugar do professor quando ele precisasse se ausentar.

De fato, aquele contexto me instigava, fazendo-me pensar quão interessante era poder ajudar aquelas crianças a aprender o que eu sabia da língua pela qual havia me apaixonado. E isso me fazia querer aprender mais para continuar ensinando. No ano seguinte, por conta das demandas escolares, não pude continuar como monitora na escola de línguas. Com o passar do tempo, comecei a ensinar em algumas turmas da educação infantil no colégio da minha família. Ingressei na graduação em Letras no ano 2000 e, em 2003, um ano antes de finalizar o curso, passei a trabalhar com turmas do Ensino Fundamental I e II e Ensino Médio, bem como passei pela experiência de

ensinar inglês em uma escola pública estadual, em turmas do Ensino Fundamental II.

Com relação à escrita, desde que comecei minha experiência docente, fui sempre provocada a compreender o porquê da resistência a esta por parte de muitos dos meus estudantes, mesmo se tratando de textos curtos. Alguns estudantes escreviam o texto solicitado, inicialmente, na sua L1, o português, e depois faziam a tradução, alegando que essa era uma prática recorrente durante seu processo de aprendizagem de inglês e, apenas assim, conseguiriam escrever. Ainda, por vezes, notava que os estudantes que já escreviam logo em inglês, tinham dificuldades em combinar palavras da língua, principalmente ao trazerem para o texto combinações que muito se assemelhavam à estrutura da sua L1. Talvez por falta de experiência naquela época, eu acreditava que esse problema acontecia, dentre outras razões, pelo fato de serem ainda estudantes do Ensino Básico (única modalidade de ensino na qual havia atuado até então) com pouco conhecimento lexical e, por isso, desconheciam as combinações fraseológicas do inglês. Ledo engano. Anos depois, em um contexto acadêmico, ao ingressar como professora substituta na Universidade Estadual de Santa Cruz (UESC) no sul da Bahia, mesma universidade em que cursara a graduação em Letras, deparei-me com a produção escrita de estudantes do curso de graduação em Línguas Estrangeiras Aplicadas às Negociações Internacionais, que traziam inadequações linguísticas, se comparadas às normas do inglês padrão. Parei para refletir, então, sobre o que os estudantes ainda precisavam aprender. Cheguei à conclusão de que as combinações de palavras no ensino de uma língua adicional¹, como o inglês, devem ser motivo de preocupação que anteceda o ensino superior. Nessa instância, além de necessitar conhecer as combinações básicas da língua, ao entrar no nível superior, o estudante precisa compreender e produzir textos mais longos e de diversos gêneros acadêmicos, bem como saber lidar com combinações de palavras frequentes em sua área técnico-científica.

Desde então, comecei a inquietar-me com essa realidade, uma vez que meus estudantes do Ensino Médio da Educação Profissional e Tecnológica, modalidade de ensino em que atuo há 10 anos, buscam ingressar na universidade e, sendo assim,

¹ De acordo com Schlatter e Garcez (2009), o motivo pelo qual preferem o uso do termo “língua adicional” em detrimento de “língua estrangeira” se deve, principalmente, ao fato de a disciplina trazer acréscimo a quem dela se ocupa, nesse caso, os estudantes que já são falantes da Língua Portuguesa e, muitas vezes, podem já conhecer outras línguas. Judd et al (2003) também preferem o termo “língua adicional”, uma vez que pode ser aplicado a quaisquer outras línguas. Com base nisso, decidimos, nessa pesquisa, pelo uso do termo “língua adicional”, em substituição ao termo “língua estrangeira” que, conforme Oss (2013, p. 689) implica em uma conotação de “estranheza” ou “distanciamento”.

precisarão participar das práticas acadêmicas de leitura e escrita em inglês sempre que requisitados. Entendemos, então, que, para isso, o modo como as combinações de palavras precisam ser abordadas nas salas de aula desde o Ensino Básico podem auxiliar os estudantes às práticas de letramento acadêmico quando chegam na universidade.

Mais recentemente, a partir da experiência docente, no segundo semestre de 2015, em uma turma de Inglês para Fins Acadêmicos II (nível B1, de acordo com o Quadro Comum Europeu de Referência para Línguas²), ofertada pela Faculdade de Letras da Universidade Federal de Minas Gerais (UFMG), pude perceber que alguns estudantes apresentavam certa dificuldade na combinação de palavras, principalmente no que diz respeito aos verbos lexicais. Tive a oportunidade de trabalhar com estudantes de diversos cursos da graduação, tais como Música, Administração, Matemática, Física, Engenharia de Minas, Gestão Pública, e também da pós-graduação em Geografia, História, Ciências da Computação, Engenharia Química e Engenharia de Alimentos. No decorrer das aulas, percebia o quanto a heterogeneidade da turma pode auxiliar na construção do conhecimento linguístico dos estudantes, na medida em que cada um traz para as atividades em sala de aula, experiências acadêmicas específicas de suas áreas de estudo. Além disso, era notório o interesse dos estudantes pela aprendizagem da língua, principalmente os estudantes da pós-graduação, uma vez que participavam frequentemente de eventos internacionais com apresentações de comunicações individuais em inglês e muitos deles são, de igual modo, exigidos pelos seus respectivos programas de pós-graduação à publicação de artigos em inglês em revistas indexadas. Parece ser exatamente esse o objetivo dos estudantes em cursar disciplinas do Inglês para Fins Acadêmicos (IFA): buscar o aperfeiçoamento da língua e a aprendizagem de práticas acadêmicas, nas quais seja demandada o uso do inglês – a apresentação de trabalhos acadêmicos em eventos científicos, a publicação de textos e a participação em programas de mobilidade internacional e intercâmbio.

Ao me debruçar sobre os textos que constituem o *corpus*³ de estudo desta pesquisa, pude perceber que há um uso frequente de verbos lexicais⁴, contudo há pouca

² A distribuição das disciplinas do Inglês para Fins Acadêmicos se aporta nos níveis estabelecidos no Quadro Comum Europeu de Referência para Línguas (European Framework of Reference for Languages). Na seção 2.2.1, apresentamos tal distribuição.

³ O *corpus* de estudo ao qual nos referimos aqui é o CorIFA, sobre o qual detalhamos no Capítulo 2 desta tese.

⁴ Os verbos lexicais são tidos como verbos que operam como verbos principais, podendo apresentar características de natureza morfológica, sintática e semântica. Mais adiante, apresentamos a caracterização e classificação dos verbos lexicais, na seção 1.7.

variação lexical⁵. Além disso, quando comparamos ao subcorpus do BAWE⁶, notamos que há pouco uso dos verbos mais frequentes do discurso acadêmico escrito de falantes proficientes do inglês em contextos também universitários. É notório que existe uma tentativa dos aprendizes na utilização de verbos mais comuns e, por isso, não muito recorrentes em textos acadêmicos. Ainda, os textos parecem discorrer sobre uma determinada temática, sem, no entanto, preocupar-se com o uso de vocabulário acadêmico⁷, mais apropriado para o contexto em que foram produzidos.

Ao ingressar como estudante de doutoramento na Universidade Federal de Minas Gerais, encontrei na Linguística de Corpus, doravante LC, o respaldo teórico-metodológico que pudesse me ajudar a responder às inquietações de outrora, uma vez que a LC, como disciplina empírica, tenciona não somente investigar fenômenos linguísticos por meio de dados reais de língua, mas possibilitar que a linguagem seja entendida como a integração do léxico e da gramática, dada a natureza formulaica da língua (BIBER et al., 1999).

O que se percebe nos estudos sobre escrita acadêmica, mais particularmente, de aprendizes de inglês como língua adicional, é justamente uma forte tendência a escolhas léxico-gramaticais que desconsideram a perspectiva fraseológica (NESSELHAUF, 2005; GILQUIN et al., 2007; ALLEN, 2009), como se as palavras pudessem ser usadas randomicamente. Tal constatação, a nosso ver, pode estar relacionada ao estágio de conhecimento linguístico desses aprendizes ou de sua interlíngua⁸, entendendo-a como um sistema linguístico usado pelos aprendizes de língua adicional que se encontram no processo de aprendizagem.

É nosso desejo ressaltar que, neste estudo, concordando com Tribble (2017), decidimos pelo não uso da dicotomia “falante nativo” e “falante não nativo” (apesar de, em alguns momentos deste texto, fazermos referência ao posicionamento de determinados pesquisadores que defendem tal binômio), pois entendemos que nem todo falante de uma L1 pode servir como referência de proficiência linguística demandada nos contextos acadêmicos. Para os fins desta pesquisa, o “falante nativo” seria apenas

⁵ A noção de variação lexical sobre a qual nos referimos, nesta pesquisa, corresponde à variedade de vocabulário utilizado nos textos e, portanto, não relaciona-se à ideia de mudança linguística.

⁶ *British Academic Written English Corpus*, compilado no Reino Unido, e que nos serviu de base para organizarmos um subcorpus de referência para esta pesquisa. Explicitamos sobre o BAWE na seção 2.5.

⁷ Tendo em vista o pressuposto teórico de Hyland e Tse (2007), compreendemos por “vocabulário acadêmico” um conjunto de palavras com alta frequência no registro acadêmico em inglês.

⁸ Com base em Selinker (1972, p. 214, tradução nossa), entendemos por interlíngua “[...] um sistema linguístico separado, baseado em *output* observável, que resulta da tentativa dos aprendizes de produção da língua-alvo”. Esse conceito será melhor discutido no Capítulo 2 do presente texto.

aquele inserido no contexto acadêmico e, portanto, imerso nas práticas acadêmicas, orais ou escritas. A título de ilustração, um britânico que nunca adentrou num curso universitário, apesar de ser “falante nativo”, pois sua L1 é o inglês, pode, muitas vezes, ser considerado pouco proficiente na sua própria língua por não ter familiaridade com as práticas no contexto acadêmico.

Recentemente, diversos estudos com *corpora* de aprendizes têm se voltado para analisar o uso dos padrões gramaticais e têm se tornado importante recurso para a pesquisa na área de ensino e aprendizagem de inglês (GRANGER, 2003). Há pesquisadores, tais como Gilquin e Paquot (2007), que defendem que os “falantes não-nativos” tendem a evitar o uso de estruturas do inglês que não tenham correspondente direto na sua L1. Isso pode se justificar pelo fato de, muitas vezes, apresentarem um repertório lexical restrito quando comparado a “falantes nativos” (GILQUIN, GRANGER; PAQUOT, 2007). A partir da comparação entre *corpora* de aprendizes e um *corpus* de falantes de inglês como L1, é possível identificar padrões de uso (sobreuso e subuso⁹) de categorias gramaticais que podem ajudar a descrever a interlíngua de aprendizes.

Estudos baseados em *corpus* também investigam associações de estruturas gramaticais com outras (CONRAD, 2010) como, por exemplo, os colocados, a fim de conhecer em que medida uma característica gramatical tende a ocorrer com outras características específicas. Tais estudos são importantes para a área de aquisição de segunda língua¹⁰, principalmente por conta das contribuições das ferramentas e métodos advindos da pesquisa com *corpus* de aprendizes para que possam representar “[...] mais línguas, mais modos comunicativos, mais tipos de textos e gêneros, mais dados de língua longitudinais e mais variáveis” (MEUNIER, 2011, p. 467).

Biber et al. (1998) travam uma discussão sobre como a gramática pode ser analisada, através da abordagem baseada em *corpus*¹¹, e, de como essas investigações podem contribuir para o ensino e a aprendizagem de inglês. Nesse sentido, a preocupação primordial constitui os principais interesses dos estudos gramaticais cuja natureza é essencialmente descritiva. De acordo com os autores, tradicionalmente, os

⁹ “Os termos sobreuso e subuso são entendidos como medidas quantitativas de diferenças linguísticas e não como julgamentos qualitativos da interlíngua do aprendiz” (GRANGER, 1998, p. 18).

¹⁰ O termo inglês como “segunda língua” refere-se, tradicionalmente, à língua falada por uma comunidade cuja primeira língua não é o inglês (GUNDERSON, 2009). Grande parte dos estudos da área de Linguística Aplicada utiliza o termo “segunda língua” e não “língua adicional”. Contudo, pelas razões que já elencamos em nota anterior, na página 12, priorizamos o uso do termo “língua adicional”.

¹¹ Em inglês, *corpus-based approach*. No Capítulo 2 desta tese, tratamos dessa abordagem.

estudos gramaticais objetivam fornecer formas na língua e uma descrição das regras de como combinar tais formas. Esse modo de analisar a língua não leva em consideração o uso padronizado das características gramaticais nos textos, tampouco a variação no uso da língua. Novas ferramentas e técnicas permitem o estudo das formas padronizadas e possibilitam a investigação da distribuição da frequência de uma variedade de construções, os padrões de associação entre as estruturas gramaticais e outras características linguísticas e não linguísticas.

Nessa direção e, considerando o papel relevante da categoria “verbos” para a arquitetura de uma língua, esta pesquisa se preocupa em analisar o uso do que intitulamos “verbos do inglês acadêmico escrito”, doravante, VIAEs em redações argumentativas produzidas por alunos da UFMG, matriculados em disciplinas de Inglês para Fins Acadêmicos. Partindo da categorização defendida por Granger e Paquot (2009) e observando que são pouquíssimos os estudos que se dedicam a investigação sobre os verbos típicos do discurso acadêmico, decidimos adotar a terminologia VIAEs aos verbos mais específicos e mais frequentes nos registros¹² acadêmicos que desempenham uma função retórica específica na escrita. Tais “registros” são entendidos por Biber e Conrad (2009) como “[...] um termo geral que abrange qualquer variedade de linguagem definida em termos situacionais, incluindo o propósito do interlocutor na comunicação, o tópico, a relação entre o interlocutor e o ouvinte, o modo escrito ou falado, e as circunstâncias de produção” (BIBER; CONRAD, 2009, p.10, tradução nossa)¹³. A fim de contribuir para preencher essa lacuna e reconhecendo a necessidade de os estudantes universitários produzirem textos em inglês durante o período da graduação, bem como em estudos de pós-graduação, reiteramos a necessidade do conhecimento e uso desses verbos e das características próprias de gêneros acadêmicos. Além disso, pensamos na necessidade de colaborar com a pesquisa sobre escrita acadêmica e acreditamos que os resultados dessa investigação podem informar as características gramaticais, retóricas e fraseológicas dos textos dos aprendizes (BIBER;

¹² Biber (2006) esclarece que o termo “registro”, bem como o termo “gênero” são essenciais para a pesquisa sobre linguagem acadêmica. Embora muitos pesquisadores adotem os termos indistintamente, há outros que preferem um em detrimento do outro. A posição de Biber (2006), apesar de relatar a dificuldade em delimitar as diferenças, pode ser explicada ao considerar “registro” como um termo mais geral, associado a um domínio de uso, ao passo que “gênero” refere-se a uma mensagem do tipo convencionalizada culturalmente. O autor dá preferência ao uso do termo “registro”, alegando fazer referência às variedades de características léxico-gramaticais descritas em seus estudos.

¹³ No original: “The term register here is a cover term for any language variety defined in situational terms, including the speaker’s purpose in communication, the topic, the relationship between speaker and hearer, spoken or written mode, and the production circumstances” (BIBER; CONRAD, 2009, p. 10).

GRAY, 2016).

Considerando esse escopo, acreditamos ser de grande relevância estudar tal elemento gramatical, bem como reconhecermos que os verbos não servem apenas para nomear ações, mas também para predizer sobre um conjunto de participantes de uma determinada situação (FILLMORE, 1977; 1978), o que reforça a sua importância na construção de textos em gêneros específicos.

Considerando que muitos trabalhos têm focado na escrita de aprendizes e, mais especificamente, na escrita do gênero “redações argumentativas” (NESSELHAUF, 2005; WAIBEL, 2008), este estudo busca analisar o uso dos verbos do inglês acadêmico escrito¹⁴ mais frequentes em redações argumentativas no CorIFA, bem como analisar seu uso e suas colocações (advérbio-*ly* + verbo) mais frequentes. A escolha por esse gênero tem duas principais razões: a) constitui um dos mais requisitados na universidade, o que reforça sua importância no âmbito da disseminação de conhecimento acadêmico; b) requer do estudante uma postura crítica ao construir e sustentar argumentos de forma coerente (NESI; GARDNER, 2012), e isso, a nosso ver, consiste em uma prática acadêmica essencial para a formação do estudante, pois concordamos com a premissa de que a “[...] argumentação envolve a habilidade de relatar e criticamente avaliar o trabalho de outros, de modo que sustente um macroargumento ou tese” (NESI; GARDNER, 2012, p.94, tradução nossa)¹⁵.

Os cursos de Inglês para Fins Acadêmicos, nessa perspectiva, precisam voltar especial atenção ao desenvolvimento de vocabulário dos estudantes e, particularmente, ao uso dos verbos mais comumente usados em registros acadêmicos escritos. Portanto, voltar o olhar investigativo para o uso desses verbos e seu uso constitui um trabalho pertinente porque possibilitará a compreensão dos aspectos da interlíngua dos aprendizes, autores das produções escritas aqui em questão. Ainda, os resultados desse estudo poderão informar futuros estudos sobre o ensino de Inglês para Fins Acadêmicos.

Diante disso, acreditamos que este estudo gera contribuições importantes para a área de ensino de inglês como língua adicional, uma vez que permitirá que os

¹⁴ Neste estudo foram selecionados os verbos mais frequentes disponíveis na Academic Vocabulary List¹⁴ (AVL) (GARDNER; DAVIES, 2013), uma lista com verbos que ocorrem nos domínios acadêmicos do Contemporary Corpus of American English(COCA), mais bem descritos mais adiante. Discutimos sobre o que chamamos de “verbos do inglês acadêmico escrito” com mais detalhes no Capítulo 3 desta tese.

¹⁵ No original: “[...] argumentation involves the ability to report and critically evaluate the work of others in ways that support a macroargument or thesis” (NESI; GARDNER, 2012, p. 94).

professores possam refletir sobre o ensino dessa língua nas escolas brasileiras, principalmente no ensino de aspectos gramaticais e lexicais, na tentativa de minimizar as dificuldades desses estudantes na produção de textos e auxiliá-los na tarefa de produzir textos que levem em consideração a necessidade de compreender as palavras como itens constituintes de uma fraseologia¹⁶, podendo ocorrer com mais frequência com determinadas palavras.

Através desta pesquisa, esperamos provocar algumas reflexões sobre o ensino da gramática do inglês, especialmente em relação à necessidade de tornar os estudantes conscientes do uso dos elementos gramaticais não como unidades individuais, mas como parte de uma fraseologia, bem como o uso dos VIAEs em seus textos. Importante ressaltar que este estudo toma como escopo de análise os princípios da abordagem da Análise Contrastiva de Interlíngua¹⁷ (ACI), que trouxe uma mudança de paradigma¹⁸ em relação às construções linguísticas que se distanciam do uso mais frequente feito por falantes da língua-alvo.

Fazer essa reflexão nos leva a crer que os erros produzidos pelos estudantes em seus textos são tratados a partir de uma perspectiva que privilegia o processamento de linguagem gradual desses aprendizes, bem como o sistema linguístico que criam, mas que, ao mesmo tempo, permitem que hipóteses sobre a língua adicional sejam testadas, considerando as dificuldades na escrita do inglês como um fenômeno próprio do processo de aquisição de línguas. De todo modo, há a necessidade de os estudantes produzirem textos acadêmicos, mas muitas vezes isso se torna um desafio devido à sua dificuldade com o conhecimento sistêmico e lexical da língua.

Considerando o exposto, concordamos com o pressuposto de que “[...] os estudos em *corpus* têm também começado a dar contribuições consideráveis para a nossa compreensão de variedades específicas da gramática do inglês no mundo”¹⁹ (CONRAD, 2010, p.238, tradução nossa). Também é possível afirmar que ensinar gramática para a escrita acadêmica não pode se desvincular das características lexicais e discursivas do texto, uma vez que, no discurso acadêmico, os tempos verbais, por exemplo, são determinados pelo tipo de contexto em que são usados (CONRAD, 2010).

¹⁶Tomamos fraseologia como combinações de palavras. Sinclair (1991) acredita que o significado de uma palavra pode ser determinado a partir dessas combinações.

¹⁷Essa abordagem metodológica será melhor discutida no Capítulo 2.

¹⁸A mudança de paradigma da ACI à qual nos referimos será discutida no Capítulo 1.

¹⁹No original: “Corpus studies are also beginning to make considerable contributions to our understanding of the Grammar of specific varieties of English in the world and to English used as a lingua franca” (CONRAD, 2010, p. 238).

Portanto, reiteramos a necessidade de realização de estudos que se interessem pela análise de padrões de uso do inglês, através do qual possam emergir caminhos que deem conta de analisar determinadas características nas construções gramaticais e, assim, colaborar com os rumos do ensino e aprendizagem da gramática do inglês, nos mais diversos contextos, entendendo ser a Linguística de Corpus uma importante disciplina empírica para os avanços significativos ao estudo da interlíngua. Nessa direção, podemos afirmar que o uso de *corpus* de aprendizes e a pesquisa em interlíngua têm se tornado um dos mais promissores desenvolvimentos nos estudos de língua baseado em *corpus* (McENERY; XIAO, 2010).

Em estudos como os mencionados aqui, podemos encontrar referencial teórico basilar para a presente pesquisa. Assim, vimos defendendo a ideia de que a visão de língua trazida à baila a partir da Linguística de Corpus, através da qual é possível analisar dados reais de língua em grandes quantidades, a fim de se verificar o uso que os aprendizes fazem de determinados padrões léxico-gramaticais podem colaborar para o ensino e a aprendizagem de inglês nos mais diversos contextos. Ainda, apesar de compreendermos que *corpora* com grande quantidade de palavras podem ser mais representativos de um determinado grupo de falantes, nos informando, por exemplo, sobre as escolhas lexicais e os usos fraseológicos dos aprendizes, não podemos desmerecer a relevância de estudos com *corpus* com menor número de palavras (PAQUOT; GRANGER, 2012), como é o caso desta nossa pesquisa. Isso se justifica pelo fato de os fenômenos linguísticos encontrados em nosso *corpus* de estudo nos fornecerem importantes subsídios que nos possibilitem analisar o uso que os aprendizes universitários de inglês fazem dos verbos e de suas colocações.

Dessa forma, esta pesquisa assume o compromisso de contribuir para a área de ensino e aprendizagem de línguas, considerando o potencial do campo da Linguística de Corpus, tanto para a compreensão dos fenômenos linguísticos quanto para o aperfeiçoamento das práticas pedagógicas que circundam as salas de aulas de línguas. No papel de professora de inglês, reconheço os efeitos desta pesquisa sobre a qual nos debruçamos para a minha abordagem de ensino. A base metodológica da LC descortinou-me a necessidade de identificar e analisar as características específicas da produção dos meus estudantes, bem como refletir acerca das motivações retóricas e semânticas que determinam suas escolhas linguísticas. Hoje, percebo que a LC me permitiu repensar sobre a minha prática pedagógica, principalmente ao enfatizar o papel preponderante da língua em uso. Em outras palavras, defendemos que a LC não coloca a

frequência de determinado item lexical como preocupação primordial, na tentativa de estabelecer uma relação dicotômica entre a produção de “falantes nativos” e “falantes não nativos”, mas, sobretudo, volta o olhar para o “como” a língua é usada em dadas comunidades linguísticas, em determinados contextos de produção, a fim de atender a diversos propósitos comunicativos.

Tomamos como base, dentre outras, as seguintes perspectivas teóricas: a) os estudos sobre interlíngua (SELINKER, 1972; TARONE, 1983, 2014; ELLIS, 1992); b) a Linguística de Corpus (BERBER SARDINHA, 2000; CHARLES, HUNSTON; PECORARI, 2009; CONRAD, 2010; PAQUOT, 2014); c) pesquisas com *corpora* de aprendizes (PAQUOT, 2008); d) a Análise Contrastiva de Interlíngua (GRANGER, 1996, 2015; JOHANSSON, 2008); e) os estudos linguísticos sobre verbos lexicais em inglês (BIBER et al., 1999; GRANGER; PAQUOT, 2009; BIBER; GRAY, 2016); f) estudos das colocações (SINCLAIR, 1991; NESSELHAUF, 2005). Esses referenciais teóricos serão apresentados no decorrer do Capítulo 1.

As questões que norteiam a presente pesquisa traduzem-se em:

- a) Quais verbos são mais frequentes no subcorpus do CorIFA?;
- b) Quais desses verbos estão na Academic Vocabulary List (AVL)²⁰;
- c) Como se dá o uso dos verbos do inglês acadêmico escrito no CorIFA?;
- d) Com quais advérbios *-ly* esses verbos co-ocorrem?

Diante disso, a presente pesquisa teve como objetivo geral:

- Analisar o uso e padrões de uso dos verbos do inglês acadêmico escrito em redações argumentativas de alunos universitários brasileiros, aprendizes de inglês.

Os objetivos específicos que motivaram este estudo consistem em:

1. Analisar a frequência e o uso dos verbos do inglês acadêmico escrito (VIAEs²¹) no subcorpus do CorIFA²²;

²⁰ A Academic Vocabulary List (AVL) é uma lista de vocabulário acadêmico, originada de um subcorpus de 120 milhões de palavras do Contemporary Corpus of American English (COCA). Mais adiante, no Capítulo 2, apresentamos mais detalhes sobre a AVL.

²¹ Intitulamos aqui de VIAEs os verbos do inglês acadêmico escrito, ou seja, verbos que são mais

2. Identificar se os VIAEs encontrados no subcorpus do CorIFA, como mais frequentes, aparecem na nova AVL;
3. Contrastar os resultados do subcorpusdo CorIFA com o subcorpuscompilado a partir das redações argumentativas do BAWE;
4. Verificar e comparar o uso das colocações do tipo “advérbio *-ly* + verbo” com os VIAEs, no subcorpus do CorIFA em comparação às colocações do subcorpus do BAWE.

Desse modo, esta tese, além deste capítulo introdutório e da conclusão, é constituída de mais três capítulos, que descrevemos a seguir.

No **Capítulo 1**, mostramos as reflexões teóricas que embasam este estudo. Num primeiro momento, apresentamos historicamente a Linguística de Corpus, discutindo seu relevante papel para a descrição de línguas. Em seguida, tratamos sobre *corpus* de aprendizes e relatamos as contribuições de alguns estudos, além de destacarmos a importância da Análise Contrastiva de Interlíngua e seus impactos em pesquisas que se voltam para o processo de aquisição de língua adicional. Além disso, discutimos sobre escrita acadêmica e as abordagens de ensino de inglês e, em seguida, destacamos a abordagem de Inglês para Fins Acadêmicos, o gênero redação argumentativa e discutimos acerca das listas de vocabulário acadêmico. Considerando a classificação semântica de Biber et al. (1999), trazemos as contribuições teóricas sobre verbos lexicais e advérbios e seu uso no registro acadêmico. Por fim, nos debruçamos sobre as colocações e a fraseologia na escrita acadêmica, principalmente no que se refere ao uso de verbos.

No **Capítulo 2**, descrevemos a metodologia utilizada na presente pesquisa, discorrendo sobre o subcorpus de estudo, o subcorpus de referência e a lista de referência, além de apresentar os procedimentos de coleta de dados.

No **Capítulo 3**, apresentamos a análise e discussão dos dados, assim como discutimos os resultados do estudo. Finalizamos com uma síntese do capítulo.

Por fim, na Conclusão, refletimos sobre a validade dos métodos e dos resultados obtidos. Concluimos discutindo as limitações desta pesquisa, além de apontarmos alguns desdobramentos possíveis e as implicações de estudos como esse para o

tipicamente usados em textos escritos, produzidos por falantes proficientes do inglês, ou por falantes cuja L1 seja o inglês.

²²*Corpus* de textos produzidos por estudantes de Inglês para Fins Acadêmicos, que tem sido compilado desde 2013, na Universidade Federal de Minas Gerais.

ensino/aprendizagem de inglês como língua adicional.

CAPÍTULO 1

FUNDAMENTAÇÃO TEÓRICA

More and more people in every branch of information science are coming to realize that a corpus as a sample of the living language, accessed by sophisticated computers, opens new horizons.

J. SINCLAIR (1987, p.79)

Este capítulo tem como objetivo abordar e discutir os dispositivos teóricos que norteiam as análises desta pesquisa, mais particularmente, como os aprendizes brasileiros de inglês usam os verbos do inglês acadêmico escrito e as colocações com advérbios-*ly*. Para dar conta desses objetivos, este capítulo está dividido em nove seções.

No primeiro momento, apresentamos um breve histórico sobre a Linguística de Corpus como uma abordagem metodológica para a investigação e descrição de línguas. Na segunda seção, trazemos à baila as contribuições dos estudos com *corpora* de aprendizes para a área de ensino-aprendizagem de inglês como língua adicional e, em seguida, na terceira e quarta seções, abordamos acerca da interlíngua de aprendizes, bem como da Análise Contrastiva de Interlíngua, respectivamente, discutindo os resultados e impactos de alguns estudos que nela se aportaram e entendendo-a como uma abordagem teórico-metodológica que vem contribuindo amplamente para o desvelamento de características linguísticas distintivas, principalmente entre *corpora* de aprendizes e de falantes de inglês como L1. Ainda, na quinta seção, tratamos sobre a escrita acadêmica e as abordagens de ensino de inglês. Na sexta seção, dedicamos atenção especial ao Inglês para Fins Acadêmicos, abordagem que fez emergir o contexto de produção dos textos que compõem o subcorpus do CorIFA. Esta seção foi subdividida em duas outras, nas quais apresentamos as especificidades do gênero acadêmico “redação argumentativa” (gênero utilizado como base de análise nesta pesquisa), além de discorrermos sobre as listas de vocabulário acadêmico (Academic Word List, New Academic Word List e New Academic Vocabulary List), sendo esta última, nossa lista de referência nesta pesquisa. Em seguida, na sétima seção, considerando a caracterização e classificação de Biber et al. (1999), apresentamos e discutimos acerca dos verbos lexicais e seu uso no registro acadêmico. Na oitava seção,

trazemos as características e classificações dos advérbios, de igual modo, à luz de Biber et al. (1999) e, por fim, na nona seção, discutimos sobre as colocações e a fraseologia no inglês acadêmico escrito.

A Linguística de Corpus

A Linguística de Corpus se ocupa da “coleta e exploração de *corpora*, ou conjuntos de dados linguísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística” (BERBER SARDINHA, 2000, p. 325). Como disciplina empírica, a LC começa a ganhar força na década de 1990, trazendo à tona investigações empíricas que suscitam questões de pesquisa nunca antes tratadas (BIBER, 1998) e vem impactando, consideravelmente, diversas áreas da pesquisa em Linguística, principalmente por se basear em dados reais de língua e considerar o que está no *corpus* uma amostra da língua em uso.

A disponibilidade de computadores e a grande quantidade de dados têm ajudado os pesquisadores a realizar estudos sobre as formas padronizadas nas quais os falantes usam a língua (BIBER; CONRAD; REPPEN, 1998). Nesse sentido, é possível investigar como as línguas são padronizadas em termos de aspectos e categorias gramaticais, usando, para tanto, as ferramentas da Linguística de Corpus.

O crescimento da LC se deve não apenas ao seu caráter empírico, mas por gerar evidências importantes e singulares sobre a língua, tais como a frequência de palavras e a ocorrência de determinadas palavras com outras. Além disso, “[...] as técnicas da linguística de *corpus* nos permitem determinar escolhas comuns e incomuns e observar padrões que revelam o que é típico ou atípico em contextos particulares”²³ (CONRAD, 2010, p.228, tradução nossa), ou seja, a identificação do que é típico ou não na língua (SINCLAIR, 1987) pode ser evidenciada através das ferramentas da LC.

As investigações no âmbito da LC, disciplina empírica na qual este estudo se insere, têm contribuído para a pesquisa na área de ensino e aprendizagem de inglês como língua adicional, pautando-se em questões como o uso de *corpora* de aprendizes e de *corpora* de “falantes nativos” para o ensino de línguas (GRANGER, 1998, 2009; MEUNIER, 2011, REPPEN, 2010), além da interface da escrita acadêmica com as

²³No original: “Corpus linguistics techniques allow us to determine common and uncommon choices and to see patterns that reveal what is typical or untypical in particular contexts” (CONRAD, 2010, p.228).

ferramentas de análise de *corpus* (CHARLES; HUNSTON; PECORARI, 2009; PAQUOT, 2014).

Há duas principais abordagens para a pesquisa em *corpus*: a abordagem movida por *corpus* e a abordagem baseada em *corpus* que correspondem, em inglês, a *corpus-driven* e *corpus-based*. A primeira abordagem movida por *corpus* refuta a hipótese da Linguística de Corpus como um método e argumenta ser o *corpus* a fonte das hipóteses que levantamos sobre a língua (TOGNINI-BONELLI, 2001) para ter acesso aos padrões de uso. Por sua vez, a abordagem baseada em *corpus* usa dados de um determinado *corpus* para explorar uma hipótese, no intuito de validá-la ou rejeitá-la, como é o caso da investigação tratada por esta tese, ou seja, nos baseamos em dados de um *corpus* de aprendizes para comparar com dados de um *corpus* de referência. Essa abordagem tornou possível identificar e analisar padrões de associação, como é o caso, por exemplo, das colocações, que serão apresentadas e discutidas ao final deste capítulo.

Como nossa pesquisa se insere na segunda abordagem, ou seja, constitui um estudo baseado em *corpus*, trataremos apenas dessa abordagem. Uma importante contribuição desse tipo de estudo reside nos estudos da fraseologia, que entende que as palavras não ocorrem aleatoriamente na língua (HUNSTON, 2009), o que implica numa visão que contraria a ideia de palavras como unidades individuais e isoladas e textos que funcionam como conjunto de frases superpostas. Nessa direção, ver a língua a partir de uma perspectiva fraseológica significa rejeitar a visão tradicional de léxico e gramática como entidades separadas, uma vez que a linguagem fraseológica pode contribuir para uma visão mais abrangente do uso linguístico de aprendizes de língua adicional (MARTINEZ; SCHIMTT, 2012).

Na seção seguinte, apresentamos alguns estudos com *corpus* de aprendizes, desenvolvidos no intuito de investigar como os aprendizes usam a língua, nos mais diferentes registros²⁴ e, mais precisamente, no discurso acadêmico.

***Corpus* de aprendizes**

Corpus de aprendizes são coletâneas de textos produzidos por aprendizes de

²⁴Há na literatura uma distinção entre os termos “registros” e “gêneros”. Na perspectiva da Linguística Sistemico-Funcional, por exemplo, os registros correspondem ao contexto de situação, ao passo que o gênero associa-se ao contexto de cultura. O conceito que tomamos está em Gardner (2012), que define registros como uma percepção léxico-gramatical que permite a realização dos gêneros (MARTIN, 1992).

uma determinada língua. Para Cotos (2014, p.203, tradução nossa)²⁵, os *corpora* de aprendizes são “coleções eletrônicas de textos autênticos produzidos por aprendizes de L2, podem ajudar a revelar aquelas dificuldades e entender as diferenças entre a produção do aprendiz e as características que perfazem o uso de língua dos nativos”.

No fim dos anos 1980, a pesquisa com *corpus* de aprendizes começou a entrar em vigor, uma vez que os pesquisadores passaram a pensar em evidências e características da língua que poderiam ser geradas para melhor compreender e descrever a língua de aprendizes de inglês (GRANGER, 2015). O primeiro *corpus* de aprendizes a ser compilado e publicado foi o ICLE (*International Corpus of Learner English*) (GRANGER et al., 1998). A partir dele, inúmeros trabalhos de natureza descritiva e pedagógica foram gerados. Sua primeira versão consistiu de 2.5 milhões de palavras extraídas de redações argumentativas escritas por estudantes universitários de vários países. Outra versão mais expandida do ICLE foi publicada em 2009, contendo 3.7 milhões de palavras de redações argumentativas produzidas por estudantes de 16 L1 distintas.

Os estudos com *corpora* têm se preocupado em mostrar como a língua é usada por seus usuários (falantes da língua como L1 e não falantes da língua como L1), tornando possível a associação entre contexto e uso. Há, em certa medida, a ideia de “falante nativo” como modelo de referência de proficiência (GILQUIN, 2016) em relação ao aprendiz desde o auge da abordagem da Análise de Erro, mesmo não sendo objetivo da Análise Contrastiva de Interlíngua (GRANGER, 1996) – que será mais bem discutida na subseção 1.3 – desmerecer o uso dos “falantes não-nativos” em detrimento do uso da língua de “falantes nativos”, por mais que o uso de alguns padrões linguísticos entre ambos possam se distanciar. Para Tribble (2017), a dicotomia “falante nativo” e “falante não-nativo” tem pouco a lhe dizer em termos de escrita de textos acadêmicos. Em consonância com essa premissa, acreditamos que ser falante de uma dada L1 não significa ser proficiente, tampouco garante que se possa escrever textos acadêmicos. Pelas razões aqui elencadas, nesta tese, não adotamos os termos “falante nativo” e “falante não-nativo”, o que quer dizer que eles só serão mencionados quando da referência aos autores que utilizam essa terminologia.

A importância dos estudos com *corpus* de aprendizes é enfatizada por diversos

²⁵ No original: “(...) electronic collections of authentic texts produced by L2 learners, can help to reveal those difficulties and to understand the differences between learner production and the features that characterize native-like language use” (COTOS, 2014, p.203).

pesquisadores (NESSELHAUF, 2004; GRANGER, 2002), principalmente pelo fato de os *corpora* revelarem as dificuldades dos aprendizes de uma determinada língua, o que nos permite analisar a sua produção linguística, a fim de preparar materiais didáticos adequados e intervenções pedagógicas mais eficientes. Há estudos que se interessam em mostrar as descrições da língua para fins didático-pedagógicos.

Allen (2009), ao estudar o uso dos pacotes lexicais no ALESS²⁶, em um *corpus* de aprendizes de inglês, percebeu que há uma diferença considerável entre os pacotes lexicais utilizados na escrita dos estudantes quando comparados à escrita de falantes de inglês como L1. Tal fato corrobora com outras pesquisas (ALTENBERG; GRANGER, 2001; HYLAND; MILTON, 1997) cujos resultados reconhecem que a produção escrita de aprendizes de inglês apresenta características que muito se diferem das utilizadas por falantes nativos. Essas diferenças se concentram, por exemplo, na frequência de uso de itens lexicais e nas combinações fraseológicas.

Com efeito, os estudos com *corpora* de aprendizes fornecem evidências sobre a língua produzida por estes, o que pode permitir a identificação e análise do que é típico e atípico em determinados contextos, além de possibilitar a realização de inferências sobre sua interlíngua, aspecto sobre o qual discutimos na próxima seção.

A interlíngua de aprendizes de inglês e a Análise Contrastiva de Interlíngua (ACI)

Esta seção apresenta o fenômeno linguístico da interlíngua. Nosso propósito centra-se em: a) explicar o conceito de interlíngua; b) apresentar e discutir os principais aspectos da interlíngua; c) traçar um panorama acerca das pesquisas sobre interlíngua, observando como os aspectos da interlíngua se apresentam na escrita acadêmica de aprendizes de inglês. Para alcançar tais propósitos, será necessário apresentar e discutir as propriedades e classificações de interlíngua propostas por pesquisadores da área, bem como daqueles que se interessam pela relação entre a interlíngua e o estudo com *corpora* de aprendizes de inglês.

O termo “interlíngua”, cunhado por Selinker (1972), refere-se ao sistema linguístico de um aprendiz durante o processo de aquisição de uma língua adicional, isto é, a interlíngua é “[...] um sistema linguístico separado, baseado em *output* observável, resultante da tentativa de produção da norma da língua-alvo pelo aprendiz”

²⁶ Active Learning of English for Science Students, o ALESS, é um *corpus* composto por textos escritos em língua adicional, na Universidade de Tóquio.

(SELINKER, 1972, p.214). Esse conceito é atribuído em grande parte à Abordagem da Análise de Erro ou Error Analysis Approach (CORDER, 1967), que afirmava ser o erro um instrumento significativo para o acesso aos processos e estratégias que o aprendiz utiliza na sua produção linguística e que enfatizava a análise cuidadosa dos erros como mais produtiva do que sua mera contagem.

Os estudos de Selinker (1972) enfatizavam as diferenças e semelhanças entre a aquisição de L2 por adultos e crianças, discutindo as explicações mentalistas (Gramática Universal) e cognitivistas para o uso da língua. O autor direciona a hipótese de interlíngua a aprendizes adultos, afirmando que as crianças podiam recorrer à sua gramática universal para aprender uma L2. Nesse ínterim, na presente tese, nos detemos apenas na aquisição de L2 por adultos.

Selinker (1972) acredita que a definição de interlíngua por ele defendida só pode se dar a partir de cinco processos cognitivos que são parte da estrutura psicológica do aprendiz. São eles: a transferência linguística, na qual o aprendiz projeta características da L1 na L2; a super generalização das regras da L2, isto é, o aprendiz aplica as regras aprendidas da L2 para todos os contextos; transferência por treinamento, na qual o aprendiz tende a formular regras que não se aplicam à L2, levado pelo modo como certas regras linguísticas lhe foram apresentadas; a simplificação, quando o aprendiz, por exemplo, usa apenas um tempo verbal nas suas construções e, por fim, as estratégias comunicativas, através das quais o aprendiz omite itens gramaticais por considerá-los redundantes.

Nas investigações na área de ensino/aprendizagem de língua adicional, a descrição da língua do aprendiz desempenha um papel essencial (GASS; SELINKER, 2008) no sentido em que, através do estudo da interlíngua, é possível analisar particularidades importantes dos fenômenos linguísticos na produção dos aprendizes, bem como compreender esse sistema autônomo de regras cuja gramática não pode ser descrita em termos de erros e desvios da norma da L2.

Nesse contexto, a interlíngua, por assumir-se uma língua natural, como um sistema de regras linguísticas, pode ser considerada dinâmica, sistemática e variável (SAVILLE TROIKE, 2006), uma vez que estruturas adquiridas gradualmente vão se integrando ao conhecimento do aprendiz. Dessa forma, a interlíngua pode ser entendida como “[...] as gramáticas intermediárias construídas pelos aprendizes no seu percurso

até a língua alvo” (McLAUGHLIN, 1987, p.60, tradução nossa)²⁷, constituindo-se como gramáticas instáveis e em constante mudança.

Assim, defendemos que a interlíngua se configura como um sistema linguístico produzido pelo aprendiz no processo de aprendizagem de uma língua adicional, que se respalda também no que o aprendiz já sabe da língua-alvo.

Dada a sua importância, o trabalho seminal de Selinker (1972) tem servido para motivar pesquisas (TARONE, 2014) que se interessam por compreender a complexa seara que constitui a aquisição de L2 e, conseqüentemente, muitos estudos têm se debruçado sobre o estudo da interlíngua de aprendizes. Embora existam poucos estudos com *corpora* de aprendizes para tentar descrever características particulares de interlíngua, há algumas investigações sobre interlíngua de falantes não-nativos de inglês em registros escritos que demonstram que uma possível explicação para o não-uso ou sobreuso de itens linguísticos deva-se à transferência da língua nativa, o que acaba interferindo no uso dessa língua pelo aprendiz. Descrevemos a seguir alguns desses estudos.

Mourssi (2013) conduziu um estudo que investigou a influência da língua nativa no uso gramatical de segunda língua com árabes aprendizes de inglês, no qual foram coletados textos escritos em três estágios. A análise quantitativa forneceu evidências empíricas da influência da língua árabe na aquisição de itens linguísticos em inglês, principalmente quanto ao uso das formas do passado simples em inglês. Esse estudo coincide com o realizado por Granger e Rayson (1998), no qual características estilísticas de textos em inglês, produzidos por aprendizes, foram analisados longitudinalmente e, de igual modo, foram evidenciados padrões significativos de sobreuso e subuso de categorias linguísticas.

Outro estudo buscou investigar a interlíngua de aprendizes e foi realizado por Paquot (2013). A autora conduziu uma pesquisa para verificar os efeitos da transferência da L1 de aprendizes de inglês no uso de pacotes lexicais. Ela enfocou nas sequências recorrentes de três palavras, que incluía um verbo lexical constante no ICLE, e foi comparado a nove outros subcorpora de aprendizes de inglês, constantes no ICLE. De acordo com a autora, os resultados demonstraram que a L1 desempenha um papel ativo na aquisição de uma língua adicional. Foi verificado que as manifestações da

²⁷ No original: “[...] the interim grammars constructed by the second-language learners on their way to the target language” (McLAUGHLIN, 1987, p. 60).

influência da L1 no uso dos pacotes lexicais pelo aprendiz são calcadas em várias propriedades do francês, L1 dos aprendizes, incluindo o uso colocacional, os padrões léxico-gramaticais, a função das convenções do discurso e a frequência de uso.

Do ponto de vista histórico, a análise contrastiva passou por um processo de mudança ao longo do tempo. Nas décadas de 1940 e 1950, os linguistas interessados no ensino de línguas evidenciaram o papel da análise contrastiva cuja preocupação era identificar características da língua a ser aprendida. Lado (1957) acreditava que “[...] a transferência da L1 na língua estrangeira podia ajudar os aprendizes a produzirem estruturas gramaticais e lexicais que se desviam da norma da língua alvo” (LADO, 1957, p.58). Por seu turno, Corder (1973) caracterizava a análise contrastiva a partir da ênfase na significância dos erros, com o objetivo de “[...] não somente mapear os erros, mas representar o nível de proficiência do aprendiz” (HASSELGARD; JOHANSSON, 2011, p. 34, tradução nossa)²⁸.

Com o passar dos anos, a análise contrastiva ganha uma nova roupagem metodológica, deixando de envolver somente a comparação entre duas línguas e passando a comparar variedades da mesma língua, principalmente com o início da exploração de *corpora* de aprendizes. Surge, assim, a Contrastive Interlanguage Analysis (CIA), em português, Análise Constrastiva de Interlíngua (GRANGER, 1996). Essa abordagem, considerada neste estudo, busca comparar variedades de língua nativa e a interlíngua de aprendizes e desempenha papel importante na descoberta de características específicas de língua de aprendizes. De acordo com Granger (2015), as pesquisas em torno do viés metodológico da ACI tiveram como desafio inicial a criação de métodos que dessem conta de ajudar os pesquisadores a descobrirem as características distintivas da língua do aprendiz, possibilitando-os a identificação do sobreuso, subuso e das suas inadequações do inglês (HASSELGARD; JOHANSSON, 2011). Dessa forma, dois tipos de comparação eram considerados importantes: uma comparação com a língua nativa (no caso de aprendizagem de segunda língua ou língua estrangeira) e uma comparação de uma produção do aprendiz com produções de outros aprendizes, principalmente com aprendizes de L1 diferentes (GRANGER, 2015). Assim, podemos dizer que a análise contrastiva pode ser utilizada como metodologia de análise tanto para comparar qualitativa e quantitativamente duas línguas distintas quanto para comparar variedades da interlíngua.

²⁸ No original: “[...] not only mapping the errors, but representing the learner’s level of proficiency” (HASSELGARD & JOHANSSON, 2011, p. 34).

A ACI tornou-se cada vez mais comum nos estudos sobre aquisição de línguas e nos estudos com *corpus* de aprendizes e, desse modo, tem permitido a descoberta da frequência e do contexto linguístico de características da interlíngua (GRANGER, 2009). De acordo com essa nova análise baseada em *corpora*, as interferências da língua nativa para a língua-alvo não são concebidas de forma negativa, mas são entendidas como mecanismos cognitivos acionados pela mente humana para a aquisição de uma língua. Desse modo, os erros ou desvios gramaticais “[...] tendem a se diferenciar em tipo e frequência dependendo da L1 dos aprendizes. Uma questão natural levantada busca entender em que medida a interferência da L1 desempenha um papel na produção do aprendiz” (ALTENBERG, 2011, p. 14, tradução nossa). Neste estudo, o tipo de comparação que nos interessa consiste na contraposição de uma língua nativa (o inglês) com a interlíngua dos estudantes, com o objetivo de estabelecer características distintivas de uma interlíngua particular, o que nos possibilita ter acesso ao seu uso linguístico. A maioria dos estudos de interlíngua é baseado em *corpora* de aprendizes e envolve a ACI, permitindo, assim, que estudos comparativos entre *corpora* de falantes da língua como L1 e de aprendizes, ou entre aprendizes de diferentes L1 ou, até mesmo, entre falantes de distintas variedades do inglês possam ser realizados.

De fato, as pesquisas que se interessam pela investigação das características da interlíngua de aprendizes não focalizam apenas no que se distancia do uso atestado em *corpora* de aprendizes, mas também no “[...] o que o aprendiz fez certo? e o mais importante, “qual a natureza da aprendizagem da língua e como melhor podemos ajudar os aprendizes no seu caminho à língua-alvo?” (JOHANSSON, 2008, p.11, tradução nossa)²⁹.

De acordo com Hasselgreen (1994), os aprendizes apresentam forte tendência a usarem palavras mais conhecidas, “[...] palavras com as quais se sintam mais seguros” (HASSELGREEN, 1994, p.237, tradução nossa)³⁰, ou seja, seus “ursos de pelúcia lexicais”, aos quais a autora denomina de “*lexical teddy bears*”. A meu ver, tradicionalmente, o termo deva-se ao fato de os ursos de pelúcia serem vistos como um objeto a ser guardado e cuidado com carinho, geralmente deixados em um espaço bastante individual que, a qualquer hora, podemos dele usufruir. De igual modo, os *lexical teddy bears* devem funcionar, ou seja, nós os colocamos num espaço só nosso,

²⁹ No original: “What does the learner get right? and, most important, What is the nature of language learning? and How can we best help learners on their way to the target language?”.

³⁰ “Words they feel safe with” (HASSELGREEN, 1994, p. 237).

no qual nos sentimos seguros e, por isso, os utilizamos.

É importante enfatizar que nosso estudo se insere em uma perspectiva descritiva e, assim, não normativa da língua, uma vez que nossa preocupação é prover uma análise sobre o uso de características gramaticais, investigando a forma com que os aprendizes usam os verbos acadêmicos nas suas produções escritas. Para tanto, selecionamos os verbos mais frequentes disponíveis na Academic Vocabulary List (AVL) (GARDNER; DAVIES, 2013). Tomamos aqui o conceito de colocação como “[...] a ocorrência de duas ou mais palavras, com um curto intervalo entre elas, em um texto. A medida usual de proximidade é de no máximo quatro palavras” (SINCLAIR, 1991, p.170, tradução nossa)³¹. Com base nesse pressuposto, neste estudo, consideraremos as colocações formadas por duas palavras para analisar em que medida advérbios *-ly* coocorrem com os verbos do inglês acadêmico escrito.

A seguir, serão apresentados alguns relevantes estudos que buscaram investigar e comparar usos linguísticos na produção de aprendizes e na produção de falantes de inglês como L1.

Dentre alguns estudos que verificaram que inadequações na produção de aprendizes de inglês podem ser fruto do pouco conhecimento dos colocados mais frequentes com certos verbos está o realizado por Dutra e Silero (2010), no qual se analisou o uso adequado ou não adequado de agrupamentos de palavras com “verbo + *for*” e “substantivo + *for*” em redações argumentativas, produzidas por estudantes universitários brasileiros. As autoras discutem aspectos da interlíngua desses estudantes, demonstrando como “[...] os erros de aprendizes de inglês, aparentemente restritos ao uso da preposição *for*, podem ser tratados, considerando os colocados³², na abordagem baseada em *corpus*” (DUTRA; SILERO, 2010, p. 912). Elas também enfatizam a importância da abordagem empírica, na qual a língua é analisada, *a priori*, observando a frequência e os motivos que levam os aprendizes a escolherem certos itens lexicais, ao invés de outros, em contextos específicos de uso.

Bertoli e Shepherd (2015) também encontraram diferenças ao investigarem quadrigramas-chave (ex: *on the other hand*) em um *corpus* escrito de aprendizes de inglês, estudantes de um curso de Letras, e contrastaram os resultados com o BAWE.

³¹No original: “Collocation is the occurrence of two or more words within a short space of each other in a text. The usual measure of proximity is a maximum of four words intervening” (SINCLAIR, 1991, p. 170).

³²O termo “colocados” foi usado pela primeira vez por Firth (1957) para se referir às palavras que ocorrem juntas com outras palavras.

As autoras verificaram que há uso em demasia de quadrigramas como *on the other hand*, ou seja, os aprendizes preferiram o uso de colocações com as quais são mais familiarizados, bem como preferindo combinações que não foram encontradas no BAWE, *corpus* de referência do estudo das autoras.

Xiao (2007), em um estudo contrastivo em torno das construções passivas no Chinese Learner English Corpus (CLEC) e no Louvain Corpus of Native English Speakers (LOCNESS), identificou que as construções passivas são mais frequentes no LOCNESS do que no CLEC, por conta dos chineses evitarem o uso das construções passivas em textos formais escritos. Em outras palavras, há uma predileção de uso de construções passivas no LOCNESS, *corpus* de falantes nativos de inglês, principalmente as do tipo curtas (sem o uso da preposição *by*), a fim de evitar a menção ao agente, o que não acontece na produção escrita dos aprendizes chineses. Assim, o uso de tais construções é mais evidente na interlíngua dos aprendizes de inglês do que nos textos escritos por falantes nativos, o que quer dizer que constitui uma característica distintiva que pode relacionar-se com o modo como o aprendiz utiliza a construção passiva na sua L1.

Ao voltarmos nossa atenção para o CorIFA, *corpus* sobre o qual nos debruçamos para a análise nesta pesquisa, percebemos que é formado por textos produzidos por estudantes no contexto de aulas de Inglês para Fins Acadêmicos³³. Trata-se de uma abordagem de ensino-aprendizagem da língua cujo componente principal consiste no estudo de habilidades (JORDAN, 2012) específicas que atendam às necessidades do aprendiz em um dado contexto. Diante disso, é importante destacar o papel dos gêneros e das tarefas de Inglês para Fins Acadêmicos, no intuito de auxiliar o estudante a habituar-se com o contexto acadêmico (SILVA, 1990), onde almeja-se que atue com sucesso.

Com base nisso, defendemos que, ao trabalhar com *corpus* de grandes quantidades de dados, estudos contrastivos são necessários, pois podem nos revelar diferenças na produção de falantes de L1 e de língua adicional, isto é, sua variação sistemática, além de ajudarem a desvelar características inerentes do próprio gênero, da

³³Em inglês, English for Academic Purposes (EAP). Constitui uma ramificação do termo guarda-chuva English for Specific Purposes (ESP), Inglês para Fins Específicos, que “[...] se concentra no estudo e ensino de variedades especializadas do inglês, na maioria das vezes, voltadas para falantes não nativos em contextos acadêmicos e profissionais avançados” (BAWARSHI; REIFF, 2013, p. 60). Mais detalhes, na subseção 2.5.1.

escrita acadêmica e do discurso acadêmico escrito de um determinado grupo de aprendizes e, desse modo, auxiliar na produção de materiais de ensino significativos para os aprendizes.

Na próxima seção, abordamos o discurso acadêmico escrito, trazendo algumas de suas características e entendendo sua importância para a compreensão do uso da língua em contextos acadêmicos.

O discurso acadêmico escrito

A pesquisa em torno do discurso acadêmico tem aumentado muito desde os últimos 30 anos, principalmente entre os linguistas aplicados interessados nas características que diferenciam o registro acadêmico de outros registros escritos e orais (BIBER; GRAY, 2016). Muitos desses pesquisadores têm se debruçado na descrição dos usos distintivos de características linguísticas em registros acadêmicos escritos, em diferentes níveis.

A partir de um estudo com o Longman Grammar of Spoken and Written Corpus (LGSWC), com mais de 5 milhões de palavras, através do qual foi elaborada a Longman Grammar of Spoken and Written English (LGSWE), Biber et al. (1999) fornecem uma descrição de quatro registros (oral, ficção, notícias e discurso acadêmico escrito) e seus padrões de uso mais evidentes, dentre os quais, destacamos a relevância do discurso acadêmico para os propósitos da nossa pesquisa. O subcorpus do discurso acadêmico escrito constitui-se de 2.65 milhões de palavras (livros acadêmicos) e 2.68 milhões de palavras (artigos acadêmicos), provenientes de uma variedade de áreas (Direito, Biologia, Química, Educação, Sociologia).

Biber e Gray (2016), ao descreverem o perfil das características distintivas da escrita acadêmica, a partir do estudo de Biber et al. (1999), afirmam que, dentre os resultados deste, alguns merecem destaque:

- (1) As três classes de palavras mais frequentes do discurso acadêmico escrito são os substantivos, adjetivos e preposições;
- (2) As categorias verbais mais típicas são o verbo copular *be*, as formas passivas (*be + made/given/taken/used*) e verbos derivados, como o de prefixo *re-* e de sufixo *-ize* (*reabsorb, itemize*) e os verbos lexicais, mais comuns que em outros registros: verbos de ação (*use, produce, provide, apply, form, obtain e reduce*);

verbos mentais (*consider, assume, determine*); verbos causativos/de ocorrência ou de existência (*follow, allow, require, include, involve, contain, exist, indicate, represent*); verbos preposicionais específicos (*depend on, consist of, be related to, occur in, result in, be based on*);

(3) Os advérbios e locuções adverbiais são geralmente mais frequentes nos registros orais, contudo há alguns especialmente comuns como, por exemplo, os advérbios de ligação, advérbios amplificadores (*extremely* e *highly*) e os advérbios de grau/extensão (*relatively, fairly* e *slightly*).

Dentre as características descritas por Biber e Gray (2016), nesta pesquisa, vale lembrar que nos interessam os verbos lexicais, que serão abordados mais adiante neste capítulo.

Por diversas vezes, utilizamos o termo “discurso acadêmico escrito” que, num sentido genérico, pode ser considerado o uso que mais tipicamente ocorre nos textos acadêmicos escritos. Quando se falava em escrita acadêmica, por muito tempo, algumas crenças se arraigaram em torno do tema. Biber e Gray (2016) enfatizam que muitos resultados de pesquisa, particularmente, estudos baseados em *corpus*, têm auxiliado a desmitificar algumas delas, tais como: a) a ideia da utilização de uma gramática elaborada no discurso acadêmico (a exemplo, o emprego de orações subordinadas); b) a noção de universalidade das características lexicais e gramaticais. Sobre a primeira, os autores afirmam que: “[...] o discurso acadêmico é apontado como o registro mais propenso a estar em conformidade com as regras do inglês padrão, em parte, por conta das regras serem explicitamente ensinadas em contextos educacionais” (BIBER; GRAY, 2016, p.19, tradução nossa)³⁴. Em relação à segunda crença mencionada, Hyland e Tse (2007), ao abordar questões ligadas ao vocabulário acadêmico, nos informam que não se deve inferir os discursos da academia como algo unitário, dada a variedade de características específicas de determinadas áreas.

Na subseção que se segue, dada a importância da escrita acadêmica, julgamos interessante traçar um panorama acerca da sua relação com as abordagens para o ensino de inglês ao longo dos anos.

³⁴ No original: “Academic prose is singled out as the register most likely to conform to the rules of standard English, in part because those rules are overtly taught in educational contexts” (BIBER; GRAY, 2016, p. 19).

Escrita acadêmica e abordagens para o ensino de inglês

Academic writing is also important because of its role in university education: it is the primary register that students must control for academic success. There is no doubt that the challenges of learning to read and write academic prose is one of the major hurdles of advanced education (BIBER; GRAY, 2016, p.67)³⁵.

O debate sobre escrita acadêmica tem sido intenso recentemente, dada a sua importância no ensino superior e por ser uma das principais ações com as quais os estudantes precisam saber lidar para terem sucesso na academia (BIBER; GRAY, 2016). A escrita acadêmica tem sido pauta de discussão em muitas instituições no Brasil e no exterior. Um dos motivos dessa movimentação relaciona-se à crescente produção acadêmico-científica e os resultados de estudos que apontam a dificuldade dos estudantes com a escrita de textos nos contextos universitários.

Na escrita acadêmica em inglês, também sobressaltam dificuldades, tanto no que diz respeito a seus falantes como L1 (WINGATE, 2015) quanto aos aprendizes como língua adicional, como podemos perceber nas pesquisas conduzidas por Hyland (2009), entre outros, reveladas a partir de estudos baseados em *corpus*.

Considerando-se o variado espectro de trabalhos sobre a escrita acadêmica, em suas mais diversas abordagens, e levando em conta o contexto universitário, diríamos que quatro abordagens sobre esse tema se dedicam a explicá-la. São elas: 1. Habilidades; 2. Processo; 3. Gênero; 4. Letramentos Acadêmicos. Sintetizamos na Tabela 1 abaixo as principais características que, de um ponto de vista genérico, norteiam tais abordagens.

Tabela 1—Principais características das abordagens de ensino de inglês relacionadas à escrita acadêmica

1. HABILIDADES		2. PROCESSO		3. GÊNERO		4. LETRAMENTOS ACADÊMICOS	
Visão de escrita como habilidade	Foco na gramática e características retóricas	Visão de escrita como resultado de processos cognitivos e criativos	Foco no planejamento e crescimento pessoal	Visão de texto como evento comunicativo (SWALES, 1990)	Foco no propósito comunicativo	Visão de escrita como prática social	Foco no contexto, cultura e gênero

³⁵ No original: “A escrita acadêmica é também importante por conta do seu papel no ensino superior: é o primeiro registro que os estudantes têm que controlar para o sucesso acadêmico. Não há dúvidas de que os desafios em aprender a ler e escrever discurso acadêmico é um dos principais obstáculos da educação superior” (BIBER; GRAY, 2016, p. 67).

Objetivo: escrita do inglês padrão.	Forte vinculação à tradição dos Estudos de Composição	Objetivo: escrita reveladora de autoexpressão e criatividade	White; Arndt (1991)	Objetivo: inserção dos estudantes em uma comunidade discursiva	Destaque para a teoria de análise de gêneros (SWALES,1990)	Vinculada à teoria dos letramentos no ensino superior e aos discursos e gêneros institucionalizados	Pilares teóricos: LEA; STREET, 1998
-------------------------------------	---	--	---------------------	--	--	---	-------------------------------------

Fonte: Elaborada pela autora, 2017.

Vemos que na abordagem com foco nas habilidades prepondera uma visão de escrita, com forte vínculo com os Estudos de Composição, oriunda das práticas instrucionais norte-americanas, relacionadas à tradição da Retórica. Aqui, há uma preocupação em demasia com a gramática e as características retóricas, no sentido de levar o estudante a escrever conforme o inglês padrão e a utilizar a estrutura de cinco parágrafos nas redações argumentativas. Essa abordagem, ainda muito recorrente nos cursos de escrita acadêmica, tem sido rechaçada por alguns pesquisadores, principalmente os voltados para a abordagem dos Letramentos Acadêmicos (LEA; STREET, 1998) sob a alegação de caracterizar uma visão simplista de escrita acadêmica, através da qual há uma espécie de remediação do conhecimento linguístico dos aprendizes, dada a sua natureza prescritiva.

A abordagem de processo apresenta uma visão de escrita como expressão pessoal e cognitiva, na qual importam mais o estilo e o conteúdo, ao invés da gramática (WINGATE, 2015). Um dos gêneros mais típicos é a narrativa pessoal, geralmente motivada por tópicos sugeridos pelos professores e não relevantes para a área de atuação acadêmica ou profissional dos envolvidos, o que acaba gerando críticas à abordagem. De igual modo, por não haver uma teoria que a subsidie, há críticas quanto à quais critérios são considerados decisivos para uma boa escrita.

A abordagem de gênero, para a qual destacamos o trabalho pioneiro de John Swales (1990), inaugura a noção de texto como evento comunicativo. Dessa forma, o gênero tem um propósito comunicativo, na medida em que suas características são determinadas pelo contexto social ao qual é pertencente. O produtor do texto, nessa perspectiva, escreve para atender a certas demandas comunicativas e sociais de uma dada comunidade discursiva.

Já a abordagem de Letramentos Acadêmicos, que tem como trabalho seminal o estudo de Lea e Street (1998), consiste em uma abordagem de prática social e crítica, assim como a de gêneros, porém, com alguns pontos de divergência, principalmente no que se refere ao ensino e à pesquisa. Enquanto a abordagem de gêneros apresenta

modelos pedagógicos, a abordagem de Letramentos Acadêmicos apenas tem se debruçado por pesquisas, sem, no entanto, existir sistematização de proposta para o ensino da escrita, além de não preocupar-se com o ensino da gramática da língua, mas sim com a comunicação do significado.

Defendendo a abordagem do Letramento Acadêmico (no singular), em uma de suas obras voltadas para a caracterização e discussão sobre essa abordagem, Wingate (2015) discute como a diversidade social, que se instaurou com o aumento do acesso de estudantes às universidades nas últimas décadas, principalmente, com o número cada vez maior de estudantes internacionais nas universidades britânicas (que constitui seu lugar de fala), se tornou uma questão de extrema relevância a ser considerada nas pesquisas em escrita acadêmica. Segundo Wingate (2015), essas dificuldades relatadas em seu estudo são muito mais de ordem epistemológica do que, propriamente, linguística, haja vista que os estudantes adentram a universidade e logo se deparam com uma variedade de práticas e demandas com as quais precisam lidar e que, muitas vezes, são divergentes das suas experiências anteriores (LEA; STREET, 1998).

Em outras palavras, diríamos que os problemas com a produção e recepção de textos ultrapassam as fronteiras da língua, mas são motivadas pela não familiaridade dos alunos com o discurso acadêmico. Nesse ínterim, Wingate (2015) problematiza que tais dificuldades com a escrita acadêmica entre estudantes das universidades britânicas se estendem também aos próprios falantes de inglês como L1, muitas vezes, minimamente assistidos nesse quesito, já que muitas universidades ofertam cursos de Inglês para Fins Acadêmicos apenas para os estudantes internacionais. Um ponto importante também apontado pela autora relaciona-se à natureza desses cursos que, segundo ela, tendem a ser genéricos e não elaborados com vistas a atender as especificidades de gênero e convenções das áreas, mas sim focar na gramática e nas características retóricas, o que pode ser um mero resgate da abordagem com foco em habilidades.

Considerando que, nesta pesquisa, utilizamos para fins de análise textos produzidos no contexto de aulas de Inglês para Fins Acadêmicos, dedicamos a subseção a seguir para descrever e discutir sobre tal abordagem.

O Inglês para Fins Acadêmicos

O Inglês para Fins Acadêmicos, conhecido, em inglês, como English for Academic Purposes (EAP), diz respeito a uma abordagem de ensino que, desde a

década de 1980, tornou-se um novo campo no ensino de inglês em universidades e outras instituições acadêmicas. O EAP surgiu como uma das ramificações do English for Specific Purposes (ESP) ou Inglês para Fins Específicos e, hoje, é um dos temas de maior interesse por pesquisadores em todo o mundo, tornando-se crucial para se redefinir os rumos do ensino e da pesquisa do inglês no ensino superior (HYLAND; HAMP-LYONS, 2002).

A importância da abordagem, conforme Hyland (2006) bem nos lembra, supera a elaboração e montagem de cursos pré-direcionados a grupos específicos de estudantes, à análise de necessidades e ao desenvolvimento de materiais de suporte para o ensino. As demandas de comunicação acadêmica vão muito além da ênfase nas questões linguístico-gramaticais, o que leva o autor a reconhecer que os estudantes universitários hoje (e aqui também se inclui os falantes de inglês como L1) “[...] têm que assumir novos papéis e envolver-se com conhecimento de novas formas quando entram na universidade” (HYLAND, 2006, p. 2, tradução nossa)³⁶.

Segundo Tribble (2015), existem duas principais abordagens de ensino de EAP: a abordagem fundamentada na tradição do Inglês para Fins Específicos, fortemente vinculada a Halliday, Macintosh e Strevens (1996) e, depois, mais aprofundada por Swales (1990), e a abordagem denominada “Composition Studies”, ou Estudos da Composição, como já tratamos na seção anterior. A primeira interessa-se pela noção de comunidade discursiva e é caracterizada pela ênfase na análise de vários exemplares de texto e na análise de *corpus* para a identificação e compreensão dos padrões léxico-gramaticais em textos dos diversos gêneros, dos quais Tribble (2015) destaca os trabalhos de Hyland (2000) e, mais recentemente, Nesi e Gardner (2012). Em contrapartida, na abordagem de Estudos da Composição, há uma preocupação com os modos retóricos (ilustração, exemplificação, comparação etc.) e com a primazia do gênero redação argumentativa (*essay*).

Assim, ao relatar sobre sua experiência docente com o EAP, Tribble (2017) reforça esse panorama ao afirmar que, inicialmente, o ensino de EAP, na Inglaterra, foi marcado fortemente por dois diferentes paradigmas, a saber: a análise de gênero (SWALES, 1990), na qual já havia a preocupação com a escrita de gêneros com base numa compreensão crítica dos contextos de produção e recepção (TRIBBLE, 2017) e a escrita processual, conhecida como Process Writing (WHITE; ARNDT, 1991), na qual,

³⁶No original: “[...] have to take on new roles and engage with knowledge in new ways when they enter university” (HYLAND, 2006, p. 2).

ao invés da ênfase nos textos e contextos, o ensino era baseado em práticas retóricas muito comuns, na época, nos Estados Unidos. O autor alega que, naquele tempo, a tradição da escrita de redações argumentativas valorizava a ideia da criatividade e regulava as convenções retóricas das áreas.

Outra questão que tem mobilizado discussões em relação à abordagem de EAP diz respeito à natividade, como já discutimos na seção 1.2. Retomando Tribble (2017), o que importa em um texto usado como exemplo em aulas de EAP não é o *status* da L1 do falante, mas sim se é adequado para fins pedagógicos e se pode auxiliar na aprendizagem dos estudantes. Acrescentamos a isso o fato de o BAWE, *corpus* que gerou o subcorpus de referência que utilizamos nesta pesquisa, apresentar somente textos de estudantes proficientes, bem-sucedidos na escrita, sendo que 30% deles não têm o inglês como L1 (TRIBBLE, 2015).

Hyland e Hamp-Lyons (2002, p.3, tradução nossa)³⁷ são categóricos ao afirmarem que:

O Inglês para Fins Acadêmicos refere-se à pesquisa e ensino da língua que foca nas necessidades e práticas comunicativas específicas de grupos particulares em contextos acadêmicos. Isso significa ensino com base numa compreensão das demandas cognitivas, sociais e linguísticas de áreas acadêmicas específicas. Isso leva os professores para além de preparar os aprendizes para estudar inglês para desenvolver novas formas de letramento: munir os estudantes com habilidades comunicativas a fim de participarem de contextos acadêmicos e culturais específicos.

Com base nesse pressuposto, percebemos que um curso de Inglês para Fins Acadêmicos precisa levar em consideração as necessidades de comunicação acadêmica de grupos específicos de estudantes. Para tanto, podemos acrescentar a isso que, a partir das práticas escritas já recorrentes nas diversas áreas, o(s) gênero(s) a ser(em) trabalhado(s) é(são) definido(s) e constitui(em) uma questão crucial em relação à escrita de textos em contexto acadêmico. Os autores ainda acreditam na necessidade de os estudos em Inglês para Fins Acadêmicos desmitificarem a noção da universalidade das convenções acadêmicas, que desconsideram as particularidades de cada área, o que pode provocar nos aprendizes a falsa ideia de que há padrões pré-estabelecidos pela abordagem, aplicável para quaisquer áreas. A noção de comunidade discursiva, nesse

³⁷No original: “English for Academic Purposes refers to language research and instruction that focuses on the specific communicative needs and practices of particular groups in academic contexts. It means grounding instruction in an understanding of the cognitive, social and linguistic demands of specific academic disciplines. This takes practitioners beyond preparing learners for study in English to developing new kinds of literacy: equipping students with the communicative skills to participate in particular academic and cultural contexts” (HYLAND; HAMP-LYONS, 2002, p. 3).

caso, tem se tornado um princípio do EAP, na medida em que considera os gêneros e convenções comunicativas das áreas, a fim de que os aprendizes possam ter efetiva participação em suas respectivas comunidades discursivas.

O gênero acadêmico “redação argumentativa”

Há estudos que trabalham com a noção de gênero sob diferentes perspectivas. Bakhtin (2000) reconhece os gêneros como enunciados relacionados às convenções, através dos quais as pessoas se comunicam. Há aqueles que consideram os gêneros como convenções discursivas (BHATIA, 2004) e socialmente contextualizadas, que acreditam que eles precisam atender às necessidades de um campo comunicativo. Há ainda os estudos que compreendem gênero como “[...] entidades orientadas para objetivos ou propósitos” (ASKEHAVE; SWALES, 2009, p. 221).

Assim, a concepção de gênero assumida nesta pesquisa toma o gênero como “[...] uma classe de eventos comunicativos cujos membros compartilham certo conjunto de propósitos comunicativos” (SWALES, 1990, p. 58). A partir disso, levamos em consideração a proposta de Swales (1990), na qual os gêneros são motivados por propósitos comunicativos e, como tal, acreditamos que desempenham relevante papel no ensino-aprendizagem de línguas.

Há nos contextos acadêmicos o trabalho com os mais diversos gêneros que podem variar de um curso a outro. Hyland e Tse (2007) argumentam que “[...] todas as disciplinas moldam palavras para seus próprios usos”, o que aponta para o uso das palavras e seus colocados, a depender do contexto em que circulam. Para comprovar tal afirmação, os autores discutem o uso do substantivo *process* que, segundo eles, é mais facilmente encontrado em textos na área das Ciências e Engenharias, do que na área de Ciências Sociais. Ainda, o verbo *analyse* tende a ter sentidos diferentes nessas três áreas. Essa constatação aponta para o fato de que, mesmo sendo consideradas palavras comumente usadas nos contextos acadêmicos, elas podem não ser típicas de determinadas áreas e, conseqüentemente, também aparecerão com menos frequência nos textos produzidos por estudantes daquela(s) área(s).

Hyland (2009), ao investigar o discurso acadêmico em inglês, traça um panorama dos principais gêneros que permeiam a academia e, dentre eles, destaca os gêneros argumentativos. Para o autor, os gêneros argumentativos, mais especificamente, as redações argumentativas, são um dos mais requisitados nas universidades,

especialmente, nos cursos de graduação, e são produzidos na medida em que os estudantes se engajam nas práticas acadêmicas e precisam socializar seu conhecimento. Para Hyland (2009, p. 130, tradução nossa)³⁸:

[...] essencialmente, a redação argumentativa envolve a apresentação de um argumento escrito para defender ou explicar uma posição, tipicamente pautada em fontes bibliográficas, ao invés de pesquisa que o estudante, por si próprio, tenha conduzido.

O autor cita Hinkel (2002) que percebeu que, em redações argumentativas de falantes de outras sete línguas que não o inglês, a escrita se assemelhava às narrativas de cunho pessoal, com o uso de primeira pessoa e a ênfase no passado como tempo verbal. Esse fato, a nosso ver, pode caracterizar certa inexperiência desses estudantes que ainda não estão familiarizados com os gêneros exigidos nas universidades, já que, durante sua vida escolar, outros gêneros lhes eram mais comuns. Nas redações argumentativas, os estudantes precisam mostrar a evidência de um argumento, avaliá-lo e sustentar firmemente seu argumento. Sua natureza pode variar significativamente, mas o argumento permanece como a principal característica para um bom desempenho na escrita (WINGATE, 2012).

De acordo com Nesi e Gardner (2012, p.97, tradução nossa)³⁹, esse gênero se estrutura em três partes: “Introdução: apresentação do tópico, definição de termos-chaves, explanação sobre a estrutura, uma sequência lógica de argumentos e uma conclusão”. A partir de um estudo com os textos que compõem o BAWE, as autoras perceberam que muitos estudantes constroem seus argumentos reconhecendo-se como principiantes na sua determinada área e, portanto, muitas vezes, não os discutem criticamente. Ainda, as autoras puderam reconhecer que há diferenças significativas na escrita da redação argumentativa, a depender da área a qual o estudante se vincula. Elas citam como exemplo o fato de as redações argumentativas produzidas por estudantes do curso de Sociologia apresentarem uma estrutura mais desenvolvida na introdução, que permite que o leitor seja logo direcionado à estrutura do texto, além de mais objetividade na exposição dos argumentos.

Ainda sobre o que se refere à argumentação, as autoras afirmam que:

[...] no cerne da argumentação acadêmica, os argumentos devem ser

³⁸ No original: “Essentially, the „essay“ involves the presentation of a written argument to defend or explain a position, typically drawing on library sources rather than research that the student himself or herself has conducted” (HYLAND, 2009, p. 130).

³⁹ No original: “Introduction: introducing topic, defining key terms, explaining structure, a logical sequence of arguments, a conclusion” (NESI; GARDNER, 2012, p. 97).

sustentados pela evidência e relacionados às teorias das áreas. Isso significa que os estudantes precisam desenvolver uma compreensão sobre o que se tem como evidência, e quais justificativas podem ser usadas para construir um argumento (NESI; GARDNER, 2012, p.91-92, tradução nossa)⁴⁰.

Assim, o ensino de gêneros deve se voltar de forma a atender às necessidades comunicativas dos estudantes, oferecendo-os condições para o uso eficaz da língua, uma vez que eles apresentam dificuldades em usar a escrita em seus contextos particulares de aprendizagem (LEA; STREET, 1998). Concordamos com Ramos (2004, p.116) ao argumentar que:

[...] uma proposta de aplicação de gêneros, diferentemente de outras escolhas de ações instrucionais mais consagradas, como, por exemplo, aquelas que trabalham com o foco exclusivamente no texto, propicia ao professor desenvolver um trabalho em que o uso de textos e questões de conhecimento sistêmico são trabalhados não mais de forma estanque, mas observando-se suas condições e situação de produção.

Há muitos estudos sobre escrita acadêmica e análise de *corpora* cujo foco se volta para a investigação de aspectos gramaticais específicos. Nessa direção, no estudo realizado por Almeida (2014, p.77) em relação às colocações formadas por verbos muito frequentes do inglês, tais como *have* e *make*, foi confirmado que três fatores podem intensificar as dificuldades dos aprendizes, principalmente no que se refere à fraseologia: fatores interlinguais relacionados à transferência negativa da L1; fatores intralinguais, referentes à falta de conhecimento na L2, principalmente de restrições linguísticas; e fatores referentes ao ensino inadequado da L2.

Diante desse panorama, este trabalho se insere nessa corrente que objetiva se debruçar sobre o estudo dos padrões de uso, mais especificamente, dos verbos do inglês acadêmico escrito, entendendo que esses são verbos usados amplamente nos registros acadêmicos. Como mencionado anteriormente, o CorIFA é composto, em sua maior parte, por redações argumentativas produzidas por estudantes universitários brasileiros, matriculados em disciplinas de Inglês para Fins Acadêmicos (IFA). No ensino de Inglês para Fins Acadêmicos, predominam o trabalho com determinados gêneros textuais, tais como artigos, ensaios, redações argumentativas e resumos de artigos. As análises dos gêneros próprios desse contexto têm se tornado cada vez mais profundas ao descrever não somente os traços linguísticos, mas também os propósitos e efeitos comunicativos das variedades da língua (BAWARSHI; REIFF, 2013 [2010]) que, segundo Swales

⁴⁰ No original: “[...] at the heart of the academic argumentation claims have to be supported by evidence and related to theories of the disciplines. This means that students need to develop a discipline-based understanding of what counts as evidence, and which warrants can be used to build an argument” (NESI; GARDNER, 2012, p. 91-92).

(1990), se voltam para a avaliação dos propósitos retóricos e para a explicação de escolhas sintáticas e lexicais.

Elaboradas para auxiliar os aprendizes de Inglês para Fins Acadêmicos, muitas listas baseadas em *corpus*, contendo itens lexicais mais frequentemente usados no inglês acadêmico escrito, têm sido desenvolvidas. Na subseção a seguir, abordamos algumas delas, dentre as quais destacamos a AVL, que constitui nossa lista de referência no presente trabalho.

As listas de vocabulário acadêmico: a Academic Word List (AWL), a New Academic Word List (NAWL) e a New Academic Vocabulary List (AVL)

Conforme Hartshorn e Hart (2016), a necessidade de se elaborar listas de vocabulário acadêmico que pudessem ser utilizadas para fins pedagógicos remonta às décadas de 1940 e 1950, com o trabalho pioneiro de Thorndike e Lorge (1944) e West (1953), que acabou provocando, a partir da década de 1970, um interesse cada vez maior por parte de outros pesquisadores, em listas baseadas em critérios de frequência de uso, a partir de textos provenientes de diferentes áreas. Sendo assim, na década de 1980, Xue e Nation (1984), através da University Word List, reuniram as listas produzidas até então. Contudo, esta foi rotulada como insuficiente por pesquisadores, tais como Avril Coxhead, que decide então criar outra, formada por um *corpus*, segundo ela, mais robusto, e que pudesse ser destinado a alunos que estivessem se preparando para ingressar na universidade, já que, como professora universitária, alegava conhecer as dificuldades de seus estudantes em relação ao vocabulário.

Assim, surge a Academic Word List (AWL) (COXHEAD, 2000), lista de vocabulário acadêmico, que tem sido amplamente utilizada por pesquisadores, professores e aprendizes em todo o mundo. Elaborada na Nova Zelândia no ano 2000, o *corpus* no qual a AWL se baseia é composto por 3.5 milhões de palavras, originadas de textos, tais como artigos jornalísticos e livros didáticos universitários dos anos 1990, perfazendo um total de mais de 400 textos acadêmicos escritos, de quatro grandes áreas: Artes, Comércio, Direito e Ciência. Para a elaboração da lista, a autora fez a contagem da frequência e uso de diferentes palavras e incluiu na lista apenas aquelas que apareciam pelo menos 100 vezes no *corpus* e pelo menos 10 vezes em cada uma das quatro áreas citadas. Desse modo, a AWL foi organizada levando em consideração as famílias de palavras, uma vez que a lista contém 570 famílias de palavras encontradas

frequentemente em textos acadêmicos escritos e agrupadas em 10 sublistas que apresentam a frequência da palavra. Assim, a primeira sublista contém palavras com maior frequência e a última sublista, as palavras de menor frequência.

Apesar de sua utilidade, a lista apresenta algumas limitações, dentre as quais podemos elencar: a) o fato de abranger vocabulário somente da década de 1990, o que pode inviabilizar a consulta do que é mais típico na língua recentemente; b) a AWL contém palavras oriundas de apenas quatro áreas, desconsiderando outras áreas acadêmicas; c) as entradas lematizadas não são organizadas por ordem de frequência, o que pode não ajudar o usuário da lista a ter acesso às palavras mais usadas.

Alguns anos depois da criação da AWL, entre 2009 e 2013, Browne, Culligan e Phillips elaboraram o que chamam de New Academic Word List (NAWL)⁴¹, formada por 963 palavras, provenientes das 288 milhões de palavras disponíveis no Cambridge Academic Corpus. Esse *corpus* reúne textos de periódicos e livros acadêmicos, bem como de redações de estudantes tanto do Reino Unido quanto dos Estados Unidos, de uma diversidade de áreas e temas. Os autores da NAWL se preocuparam em: a) considerar os lemas e as famílias de palavras; b) utilizar um *corpus* contemporâneo com grandes quantidades de textos das áreas acadêmicas; c) organizar as entradas lematizadas por ordem de frequência, o que facilita sua consulta. Um ponto interessante que merece destaque quanto à NAWL deve-se ao fato de termos realizado a contagem da frequência das entradas lematizadas dos verbos, o que resultou na identificação de 232. Pensamos em utilizá-la como lista de referência deste estudo, contudo, verificamos que a lista não era apropriada para nosso estudo, visto que, dos 232 verbos, encontramos ocorrências apenas de *facilitate*, *obtain*, *interact*, *affirm*, *publish*, *cure* e *punish*, em número muito baixo, o que inviabilizaria os propósitos da presente pesquisa que utilizou 20 como valor de ponto de corte de frequência absoluta no nosso *corpus* de estudo.

Por sua vez, a AVL⁴² (GARDNER; DAVIES, 2013) constitui uma lista de vocabulário acadêmico que contém 120 milhões de palavras, extraídas de textos publicados de 1990 a 2015, constante de 3.000 lemas que ocorrem nos nove domínios acadêmicos (História, Educação, Ciências Sociais, Direito, Humanidades, Filosofia, Ciência e Tecnologia, Medicina e Administração) do Contemporary Corpus of American English⁴³ (COCA). Segundo Davies (2016), a seção acadêmica do COCA, com

⁴¹ A New AWL pode ser encontrada no *site*: <http://www.newacademicwordlist.org/>.

⁴² Lista de vocabulário acadêmico. Disponível em: <<http://www.academicvocabulary.info/>>. Acesso em: 20 dez. 2016.

⁴³ Disponível em: <<http://corpus.byu.edu/coca/>>. Acesso em: 20 dez. 2016.

103.421.981, da qual é gestada a AVL, possui quase 100 diferentes revistas acadêmicas, selecionadas para cobrir todas as áreas do conhecimento, o que reitera a importância da representatividade desse *corpus*. Gardner e Davies (2013) se preocuparam em: a) considerar os lemas (no total foram gerados 3.015), em detrimento das famílias de palavras, como na AWL; b) utilizar um *corpus* contemporâneo com grandes quantidades de textos que incluem áreas acadêmicas que perfazem nove domínios acadêmicos; c) considerar os lemas das palavras que ocorriam com 20% da frequência esperada em pelo menos sete dos nove domínios.

Hartshorn e Hart(2016, p.72, tradução nossa)⁴⁴, em um estudo comparativo acerca da AWL e da AVL, nos relata que:

[...] a AVL é baseada em um *corpus* muito maior de textos acadêmicos e isso pode incluir uma gama maior de áreas acadêmicas. Além dessas observações, talvez as diferenças mais significativas sejam sobre o uso de lemas pela AVL, ao invés de famílias de palavras e na sua sequência de itens de 1 a 3.015, ao invés de agrupar os membros das famílias de palavras em sublistas e depois, organizá-las alfabeticamente em sublistas.

Retomando as informações aqui apresentadas acerca das listas, consideramos que a AVL, devido à sua extensão e por cobrir muitos domínios acadêmicos, pode melhor orientar os estudantes que precisam de auxílio no que diz respeito ao vocabulário mais frequentemente utilizado no inglês acadêmico escrito, bem como pode fornecer informações importantes que possibilitem o desenvolvimento de estudos pelos pesquisadores da área, mesmo seu vocabulário não sendo separado por área. Por essa razão, a AVL constitui a lista de vocabulário acadêmico que decidimos utilizar como referência na nossa pesquisa.

Sobre a questão da não separação de vocabulário por área, nos reportamos à Hyland e Tse (2007) ao questionarem o termo “vocabulário acadêmico”. Os autores investigaram se, de fato, existe um vocabulário que possa dar conta de todas as áreas acadêmicas. Segundo eles, muitos estudos, tais como o de Coxhead (2000) e Nation(1990), têm se preocupado em organizar uma lista de palavras frequentes no registro acadêmico em inglês, com a firme ideia de que os aprendizes estão em busca do desenvolvimento de seu repertório lexical acadêmico. Talvez essa pressuposição possa

⁴⁴ No original: “[...] the AVL is based on a much larger corpus of academic texts and that it may include a broader range of academic disciplines. Beyond these observations, perhaps the most striking differences are in the AVL’s use of lemmas rather than word families and in its sequencing of items from 1 to 3015 rather than grouping word family members into sublists and then ordering them alphabetically within the sublists” (HARTSHORN; HART, 2016, p. 72).

explicar a criação dessas listas sobre as quais nos dispusemos a explicitar nesta seção (AWL, NAWL e a AVL). Entretanto, apesar de as listas, de certa forma, contribuírem para delimitar o escopo dos termos de maior frequência em determinados *corpora*, elas não são representativas de áreas específicas por não estarem separadas por área. O fato de as listas serem provenientes do registro acadêmico não garante que uma determinada palavra seja fortemente usada em todas as áreas, igualmente. Sendo essa a nossa posição, nesta pesquisa, não concordamos que haja um determinado número de palavras mais comuns para todas as áreas do conhecimento. Em contrapartida, acreditamos que cada área requer diferentes normas e práticas, o que faz com que encontremos características linguísticas distintas entre elas. Vale ressaltar que o CorIFA, nosso *corpus* de estudo, é ainda pequeno para separarmos os textos escritos pelos estudantes de acordo com sua disciplina de origem (seu curso). Desse modo, o *corpus* precisa crescer para podermos fazer isso de maneira relevante.

Resultados que confirmam essa prerrogativa foram encontrados por Hyland e Tse (2007). Para exemplificar, tomemos a família da palavra *analyze* que, de acordo com os autores, apresenta algumas diferenças de uso entre as áreas: a) ocorre mais frequentemente como substantivo nas Ciências Sociais; b) a forma *analytical* é seis vezes mais usada na área de Engenharia. Eles também pontuam que há diferenças de cunho semântico, a exemplo da forma verbal. Na área de Engenharia é mais usada para referir-se a “métodos que compõem alguma substância”, como vemos nos exemplos 1 e 2 (exemplo 1: “The proteins were separated by electrophoresis on 10% polyacrylamide SDS gels and analyzed by autoradiography”; exemplo 2: “We used a variety of methods to analyse fungal spore load, volatiles and toxins”). Já nas Ciências Sociais, a ideia de análise está vinculada ao “pleno cuidado com algo” (exemplo 3: “The major objective of this report is to analyze developments in political sociology over the last half century”; exemplo 4: “Following this, the results were analyzed to determine buyer behaviour”). Esses resultados, como enfatizam os autores, acabam desmitificando a falsa impressão da existência de uniformidade entre as áreas com relação ao vocabulário acadêmico.

De posse desses argumentos, concordamos com Hyland e Tse (2007) quanto à importância de listas que possam definir um repertório lexical específico das diferentes áreas e, desse modo, ser possível melhor auxiliar professores e estudantes universitários no que concerne à escrita acadêmica em inglês, uma vez que reconhecemos “[...] a importância dos ambientes contextuais que refletem diferentes práticas e normas das

áreas” (HYLAND; TSE, 2007, p.251, tradução nossa)⁴⁵. Assim, como bem pontuam os autores, uma das formas de ajudar os estudantes universitários é prover-lhes condições para a compreensão das características dos discursos de suas áreas e, dentre elas, inclui-se o vocabulário.

Na seção que se segue, caracterizamos os verbos lexicais e mostramos sua classificação, conforme Biber et al. (1999).

Verbos lexicais: caracterização e classificação

Alguns pesquisadores têm se dedicado aos estudos sobre os aspectos gramaticais das línguas e têm delineado uma caracterização para a classe gramatical dos verbos. Nesta pesquisa, tomamos como objetos de análise o que chamamos de “verbos do inglês acadêmico escrito”, que constituem verbos lexicais.

Na LGSWE encontramos a distribuição e função das diferentes classes gramaticais nesses quatro registros, dentre elas, a classe de verbos, dividida pelos autores em três principais: verbos lexicais (*run, eat*), verbos primários (*be, have e do*) e verbos modais (*can, will, might*). De acordo com os autores, tal distinção é realizada por conta de suas funções como verbos principais e auxiliares. Os autores conceituam os verbos lexicais, aqueles que nos interessam nesse estudo, da seguinte maneira:

Os verbos lexicais incluem uma classe aberta de palavras que somente funcionam como verbos principais; os três primeiros verbos primários podem funcionar tanto como verbos principais ou verbos auxiliares; e os verbos modais podem funcionar somente como verbos auxiliares (BIBER et al, 1999, p. 358, tradução nossa)⁴⁶.

De acordo com os autores, os verbos lexicais são a classe de verbo mais comum no LSWEC, ocorrendo 100.000 vezes por milhões de palavras (mais que uma palavra em 10). Eles acreditam que “[...] a frequência muito mais alta de verbos lexicais reflete o fato de os verbos principais ocorrerem frequentemente sem qualquer auxiliar nas formas do presente e nas orações subordinadas” (BIBER et al., 1999, p. 359, tradução nossa)⁴⁷. Por outro lado, chegaram à conclusão de que, no discurso acadêmico, há mais

⁴⁵ No original: “[...] the importance of contextual environments which reflect different disciplinary practices and norms” (HYLAND; TSE, 2007, p. 251).

⁴⁶ No original: “Lexical verbs comprise an open class of words that function only as main verbs; the three primary verbs can function as either main verbs or auxiliary verbs; and modal verbs can function only as auxiliary verbs” (BIBER et al., 1999, p. 358).

⁴⁷ No original: “The much higher frequency of lexical verbs reflects the fact that main verbs often occur without any auxiliary in simple tense forms, and in nonfinite clauses” (BIBER et al., 1999, p. 360).

usos de verbos auxiliares primários que marquem, por exemplo, a voz passiva (ex: *they are produced*)⁴⁸ e menos uso de verbos lexicais do tipo *know*, *think* e *mean* que, por sua vez, são mais comuns no discurso oral. Biber et al (1999) nos mostram que os verbos lexicais apresentam características de natureza morfológica, sintática e semântica. Quanto aos elementos morfológicos, os autores afirmam que, geralmente, possuem uma estrutura complexa, tais como os verbos derivados (ex: *itemize* e *soften*) e multipalavras (ex: *bring up* e *rely on*), e podem variar conforme o tempo, aspecto e voz. Quanto às características sintáticas, os verbos lexicais podem ser verbos principais ou verbos frasais (ex: *has written* e *will be writing*). As características semânticas dos verbos lexicais referem-se a processos, ações e estados.

Biber et al. (1999, p. 360) classificam os verbos quanto aos seus domínios semânticos e afirmam que muitos têm mais de um significado, podendo fazer parte de vários domínios semânticos. Nesse caso, o verbo é listado na categoria cujo uso é mais típico. Sendo assim, os verbos lexicais são classificados em: a) verbos de ação; b) verbos mentais; c) verbos causativos; d) verbos de simples ocorrência; e) verbos de existência ou relacionamento; f) verbos aspectuais.

Para fins de melhor visualizarmos alguns exemplos de cada um dos tipos de verbos acima mencionados, organizamos o Tabela 2 a seguir:

Tabela 2 – Classificação e caracterização dos verbos lexicais (BIBER et al, 1999)

Verbos lexicais (BIBER et al, 1999)	
a. Verbos de ação	ex: <i>bring, buy, carry, leave, open, run, work</i>
b. Verbos de comunicação	ex: <i>ask, announce, call, discuss, explain, write</i>
c. Verbos mentais	ex: <i>think, love, know, want, examine, decide, discover</i>
d. Verbos causativos	ex: <i>allow, cause, enable, force, help, require, permit</i>
e. Verbos de simples ocorrência	ex: <i>become, happen, develop, increase, occur</i>
f. Verbos de existência ou relacionamento	ex: <i>exist, live, include, involve, represent, seem</i>
g. Verbos aspectuais	ex: <i>begin, continue, finish, keep, start, stop</i>

Fonte: Elaborada pela autora, 2017.

Os verbos de ação referem-se a ações ou eventos que se realizam a depender da escolha do agente, podendo ser transitivos (ex: *In many of these jobs, women are*

⁴⁸ Todos os exemplos fornecidos, nessa subseção, foram retirados da LGSWE (BIBER et al., 1999), Capítulo 5.

working with women only), ou seja, que requerem objeto, e intransitivos (ex: *They ran, on rubbery legs, through an open gate*), ou seja, não inclui objeto.

O que Biber et al. (1999) chamam de “verbos de comunicação” são tidos como uma subcategoria dos verbos de ação e, como o próprio nome já denota, referem-se aos verbos que se vinculam a tarefas comunicativas (ex: *He might find it impossible to write in the tone or theme he first took up*).

Já os verbos mentais são aqueles cujos sentidos se relacionam a: ações cognitivas, como é o caso de *think* e *know*; emocionais, como *love* e *want*; perceptivas, como *see* e *taste*. Muitos desses verbos, para Biber et al. (1999), descrevem atividades cognitivas, apresentando um significado mais dinâmico, como é o caso de *examine* e *decide*. Verbos que descrevem estados cognitivos, como *believe*, *know*, *remember* e *understand*, bem como os que descrevem estados emocionais, tais como, *enjoy*, *fear*, *feel*, *hate*, *like*, *love*, *prefer*, *suspect* e *want*, são tidos como mais estáveis.

Os verbos causativos, também chamados pelos autores de “verbos de facilitação”, referem-se àqueles que trazem em seu bojo a ideia de um novo estado de coisas. Eles podem ocorrer com um objeto direto ou no complemento da oração, como no exemplo advindo do registro acadêmico: *This information enables the formulation of precise questions*. Vemos, nesse caso, que o verbo *enable* é usado para denotar que o substantivo *information* facilita, permite, provoca “a formulação de questões precisas”.

Utilizados para relatar acontecimentos não intencionais, os verbos de simples ocorrência caracterizam-se pelo fato de o sujeito desempenhar um determinado papel semântico. Um dos exemplos provenientes do registro acadêmico do LGSWC demonstra isso: *The term „yeature“ has occurred many times in this chapter*.

Os verbos de existência ou relacionamento têm como função primordial relatar um estado entre entidades (BIBER et al., 1999). Os mais típicos são verbos, conhecidos como “copulares”, tais como *be*, *seem* e *appear*, usados para ligar um adjetivo ou complemento nominal a um sujeito. Ex: *All these uses seem natural and serviceable*.

Por fim, os verbos aspectuais, como *begin*, *continue*, *finish*, *keep* e *stop*, são usados para se referir ao estágio de progresso de alguma determinada ação, como se pode perceber no exemplo do registro acadêmico do LGSWC: *The conventional turboprop will certainly continue to dominate the market for smaller aircraft*.

Outro ponto importante que merece ser mencionado diz respeito ao gênero, termo que é adotado neste trabalho para se referir à produção escrita do estudante com um propósito comunicativo e social, ao qual Biber et al. (1999) preferem chamar de

“registro”. Ele também pode ser fator determinante para que encontremos certos verbos em um determinado gênero com mais frequência.

No estudo que empreenderam com o LGSWC, Biber et al. (1999) encontraram evidências de que 50% dos verbos que consideram comuns (que ocorrem pelo menos 50 vezes por milhões de palavras) são do tipo “verbos de ação”. A classe de verbos é mais típica nos registros orais que nos registros escritos, talvez por conta dos falantes amplamente se apoiarem em dois domínios semânticos nos registros orais, ou seja, os verbos de ação e verbos mentais.

No que diz respeito à propriedade gramatical, importa-nos sua morfologia e sintaxe; numa perspectiva morfológica, consideramos os verbos como classes possuidoras de radical e morfemas de sufixo, ao passo que, numa perspectiva sintática, nos valem do sujeito e predicado – morfemas funcionais que indicam modo, tempo e pessoa. Também funcionando como propriedade gramatical, há a flexão, que consiste na propriedade mais distintiva dos verbos, já que eles apresentam um determinado número de formas flexionais que são permitidas ou exigidas nas mais variadas construções gramaticais (HUDDLESTON; PULLUM, 2002).

Para Biber (2006), é importante considerar a marcação de tempo, aspecto e voz ao se analisar textos acadêmicos. Reportando-se ao estudo de Biber et al. (1999), o autor afirma que há uma preferência dos aprendizes pelo tempo passado (*past tense*), inclusive diferenças significativas entre o uso nas mais diversas áreas. Nos livros didáticos das áreas de Educação e Humanidades, por exemplo, 30% dos verbos se colocam no tempo passado, enquanto nas áreas de Engenharia e Ciências Naturais, apenas 10%. Os autores interpretam que esses resultados são ocasionados pelo fato da necessidade dos aprendizes de relatarem eventos passados e narrarem episódios históricos, nos textos da área de Humanidades.

Quanto ao aspecto verbal, Biber (2006) chama a atenção para o uso da forma simples e contínua. A primeira é largamente usada em ambos os registros, todavia, a forma contínua é mais frequente que o aspecto perfeito. Este último se configura como muito comum nos registros orais, com destaque para as aulas em laboratório, uma vez que os aprendizes descrevem ações que estão sendo observadas naquele dado momento. Em oposição, o autor afirma que, nos registros escritos, a forma contínua quase nunca ocorre, o que corrobora com Hinkel (2013), ao enfatizar que são formas praticamente raras nos registros acadêmicos. Há a preferência pelo aspecto perfeito, a partir do qual os eventos passados são descritos, de modo a evidenciar suas consequências até o

presente momento.

O uso das vozes ativa e passiva é também contraposto por Biber (2006). Segundo ele, a voz passiva predomina nos registros escritos e a voz ativa nos registros orais, sendo a preferência, por este último, correspondente a 95%. Defendendo um posicionamento similar, Hinkel (2013) também alerta para o fato de a voz passiva ser determinada por convenções do discurso acadêmico e, assim, ser muito comum na escrita acadêmica, principalmente nos textos das áreas de Ciências Naturais e Engenharia.

1.7.1 Os verbos lexicais no discurso acadêmico

No discurso acadêmico, há muitas diferenças quanto ao uso dos verbos lexicais, se comparadas a outros registros.

Segundo Biber et al. (1999), os verbos mais típicos do discurso acadêmico encontrados no LGSWC são pertencentes aos domínios semânticos de “ação” e “existência”, seguidos dos verbos mentais. Os de menor frequência, no registro, constituem os verbos de comunicação, causativos, de ocorrência e aspectuais. Os autores alertam que os verbos *say*, *think* e *know* são mais comumente usados em construções, de igual modo, do tipo *that-clauses*. Além disso, os verbos encontrados no discurso acadêmico podem ser fruto dos tópicos sobre os quais haja a necessidade de fala e de escrita, como bem declaram:

A gama de verbos encontrados nos domínios semânticos reflete os tipos de tópicos que falantes e escritores mais comumente discutem. Embora existam diferenças importantes entre os registros, o inglês geralmente faz mais distinções (como representado por diferentes verbos) entre verbos de ação que em qualquer outro domínio (BIBER et al., 1999, p.465, tradução nossa)⁴⁹.

Biber et al. (1999) alegam que no discurso acadêmico é pouco comum o uso dos verbos mentais, já que nesses textos quase nunca os pensamentos individuais são reportados. Já os verbos de comunicação, como *show*, *ensure* e *indicate*, apresentam frequência moderada e, normalmente, aparecem para retratar a posição de quem escreve, sem necessariamente fazer alusão a pensamentos ou sentimentos:

⁴⁹ No original: “The range of verbs found across semantic domains reflects the kinds of topics that speakers and writers most commonly discuss. Although there are important differences across registers, English generally makes more distinctions (as represented by different verbs) among activity verbs than in any other domain” (BIBER et al., 1999, p. 465).

1. *Other models show that this state of affairs may have been reached only a few hundred million years after Earth's formation.*

2. *Measured values at present indicate that the chain modulus of polyethylene is certainly higher than 280 Gpa.*

Segundo eles, os verbos de comunicação podem ainda apresentar:

a. Alto grau de certeza, como vemos nos usos dos verbos *show*, *prove* e *demonstrate*:

1. *Empirical data **show** that similar processes can be guided quite differently by users with different views on the purposes of communication.*

b. Menor grau de certeza, como em *suggest*, *indicate* e *postulate*:

1. *Information from the planetary probes **indicates** that all the terrestrial planets have undergone differentiation.*

Para os autores, “[...] há categorias específicas de verbos que são típicos do discurso acadêmico” (BIBER et al., 1999, p.78, tradução nossa)⁵⁰. Desse modo, entendemos que essas categorias às quais eles se referem são os chamados “verbos lexicais” que acabamos de descrever. Em Biber et al. (1999), não há menção ao que vimos chamando nesta pesquisa de “verbos do inglês acadêmico escrito”, embora acreditamos que eles possam ser incluídos na categoria de verbos lexicais. Por essa razão, defendemos que estudá-los seja importante, no sentido de que podemos ter subsídios que melhor caracterizem seu uso e, dessa forma, contribuirmos para preencher lacunas, principalmente no que diz respeito aos textos produzidos em inglês e que circulam nos mais variados contextos acadêmicos.

O que intitulamos de “verbos do inglês acadêmico escrito”, nesta pesquisa, originou-se do termo “EAP verbs” (verbos do Inglês para Fins Acadêmicos), destacado por Granger e Paquot (2009, p.210, tradução nossa) em um estudo, no qual são propostas categorias e a importância das colocações que se formam com esses verbos. Sobre isso, as autoras acreditam que “[...] os aprendizes de inglês como língua estrangeira usam com muito pouca frequência a maioria dos verbos do Inglês para Fins Acadêmicos, que são verbos como: incluir (*include*), relatar (*report*) ou relacionar

⁵⁰No original: “[...] there are specific verb categories that are typical of academic prose” (BIBER et al., 1999, p. 78).

(*relate*), que expressam funções retóricas no cerne da escrita acadêmica”⁵¹. Assim, as autoras pontuam que:

O conhecimento insuficiente dos verbos que são tipicamente usados no discurso acadêmico escrito é uma séria deficiência para os aprendizes já que os impede de expressar seus pensamentos em todas as suas nuances e treiná-los em seu estilo esperado (GRANGER; PAQUOT, 2009, p. 194, tradução nossa)⁵².

Esse pressuposto colocado pelas autoras, acerca do pouco conhecimento dos estudantes universitários com o uso desses verbos que julgamos essencial à escrita acadêmica, nos reporta para a necessidade de compreender o papel do discurso acadêmico e da escrita acadêmica nos contextos de ensino superior.

Biber e Gray (2016) afirmam que existem razões muito específicas para que haja, de sua parte, interesse no estudo do discurso acadêmico. Segundo eles, além de a escrita acadêmica ser, inegavelmente, inerente e importante para o contexto em que estão inseridos, os autores enfatizam seu papel de extrema importância no ensino superior, na medida em que os estudantes precisam ter conhecimento dos elementos que subjazem a escrita nesse contexto. Outra razão que justifica a relevância da escrita acadêmica diz respeito à diferença linguística em relação ao discurso oral, bem como a outros registros, o que a torna, a seu ver, uma tarefa desafiadora para os estudantes. O desafio ao qual se referem reside justamente nos aspectos distintivos presentes nesses registros como, por exemplo, o propósito informacional e o público, características “[...] associadas a padrões específicos de variação linguística” (BIBER; GRAY, 2016, p.70, tradução nossa)⁵³. Isso nos chama a atenção para o fato de a escrita acadêmica ser direcionada a um público específico, motivada por propósitos que buscam atender a demandas acadêmicas também diversas e, como resultado disso, os textos apresentam variações quanto às escolhas linguísticas.

No estudo conduzido por Biber et al. (1999), os verbos, geralmente, são menos comuns no discurso acadêmico se comparados a outros registros. Todavia, há verbos que são típicos desse registro, conforme Granger e Paquot (2009) já sinalizaram. Biber e

⁵¹ No original: “EFL learners significantly underuse the majority of „academic verbs”, that is, verbs like: include, report or relate, that express rhetorical functions at the heart of academic writing” (GRANGER; PAQUOT, 2009, p. 210).

⁵² No original: “Insufficient knowledge of verbs that are typically used in academic written discourse is a serious handicap for learners as it prevents them from expressing their thoughts in all their nuances and couching them in the expected style” (GRANGER; PAQUOT, 2009, p. 194).

⁵³ No original: “[...] associated with specific patterns of linguistic variation” (BIBER; GRAY, 2016, p. 70).

Gray (2016) chegaram à conclusão de que a classe gramatical mais comum no discurso acadêmico é a dos substantivos, perfazendo aproximadamente 300 mil/1 milhão de palavras. Apesar dessa constatação, Biber et al. (1999) nos informam que os verbos de alta frequência, tais como *get*, *give*, *take* e *make*, não são tão comuns na escrita acadêmica, o que nos leva a entender que pode haver, nesse discurso, a predominância de verbos lexicais que, por sinal, constituem a categoria de verbos mais comuns no LGSWC.

Em nosso estudo, temos como objetivo principal a investigação sobre a frequência e o uso do que chamamos de “verbos do inglês acadêmico escrito”, verbos estes considerados lexicais. Decidimos por analisar o uso desses verbos a partir das suas colocações formadas pelo padrão “advérbio-ly+ verbo”. Essa decisão metodológica se deu por conta de termos percebido colocações inadequadas nos textos dos estudantes do IFA na disciplina que ministramos em 2015. Outro fator importante se refere ao escopo dos *corpora* que são aqui contrastados, apesar de ambos possuírem textos produzidos por estudantes em contexto acadêmico. O CorIFA inclui textos que, apesar de terem passado por avaliação do professor, não foram rankiados em *high score* ou *low score*, sendo que, no BAWE, apenas os textos com nota mais alta (*high score*) estão incluídos no *corpus*. Sendo assim, julgamos relevante analisar as colocações com determinado tipo de advérbio. É sobre essa classe de palavras que abordamos na seção que se segue.

Os advérbios no discurso acadêmico

Assim como os substantivos, adjetivos e verbos, os advérbios também combinam características morfológicas, sintáticas e semânticas. Os advérbios que nos interessam neste estudo para compor as colocações são os advérbios *-ly*. Para Biber et al. (1999, p.541, tradução nossa)⁵⁴, “[...] advérbios formados por sufixo *-ly* são, em sua maioria, derivados de adjetivos. No discurso acadêmico, estes advérbios são usados para uma variedade de fins”.

Biber et al. (1999) citam a terminação *-ly* (ex: *clearly*) como traço morfológico dos advérbios, apesar de outros não possuírem terminação (ex: *however*, *now*). Quanto às características morfológicas, os autores sinalizam o uso de advérbio como modificadores em locuções adjetivas e adverbiais. Em relação às características

⁵⁴ No original: “[...] adverbs formed with the *-ly* suffix are mostly derived from adjectives. In academic prose, these adverbs are used for a variety of purposes” (BIBER et al., 1999, p. 541).

semânticas, na função de modificadores, Biber et al (1999, p.65, tradução nossa)⁵⁵ acreditam que os advérbios expressam:

- Circunstâncias sob as quais uma ação, processo ou estado acontece, como em: *she writes well, they leave tomorrow;*
- A posição do falante e do escritor em relação à informação contida no restante da oração: *surely you don't believe this it is apparently wrong;*
- Conexão entre as orações: *nevertheless it wasn't true, I'm not sure though.*

De acordo com Biber et al. (1999), os advérbios são comuns nos quatro registros escolhidos para seu estudo: registro oral, ficção, notícias e discurso acadêmico. No entanto, os autores afirmam que eles ocorrem mais vezes na ficção se comparados às notícias e ao discurso acadêmico. Segundo os autores, os advérbios do tipo simples, tais como *well, too, rather, quite, soon* e *here*, e os formados pelo sufixo *-ly*, especialmente os advérbios amplificadores, são muito frequentes nos quatro registros. No discurso acadêmico, quando comparado, por exemplo, ao registro oral, há extremas diferenças de uso. No primeiro, há maior número de ocorrências de advérbios do tipo simples (mais de 60%), ao passo que, no discurso acadêmico, predominam as formas adverbiais de sufixo *-ly* (em torno de 55%). Essas últimas compõem as colocações que nos importam nesta pesquisa: advérbio *-ly*+ verbo. Esses advérbios, em sua grande maioria, são formados por sufixação (BIBER et al., 1999), derivados de adjetivos (ex: *clear* –*clearly*). Os outros advérbios que apresentam uma forma derivada de adjetivos que já possuem esse sufixo, como é o caso de *weekly* e *fatherly*, não constituem nosso objeto de interesse.

Biber et al. (1999) se ativeram a explicar diferentes classes de palavras, dentre elas, os advérbios. Os autores afirmam que os advérbios *-ly* constituem o tipo de advérbio que mais se destaca no registro acadêmico. De acordo com seu estudo, que tem nos servido de base para grande parte da costura teórico-analítica desta pesquisa, no registro acadêmico houve um percentual de 55% de uso de advérbios comuns, com sufixo *-ly*, utilizados nos textos para cumprir diferentes propósitos. Vejamos alguns deles:

⁵⁵ No original: “[...] circumstances under which an action, process, or state takes place, as in: *she writes well; they leave tomorrow; the speaker's or writer's attitude towards the information contained in the rest of the clause: surely you don't believe this, it is apparently wrong; the connection between clauses: nevertheless it wasn't true, I'm not sure though*” (BIBER et al., 1999, p. 65).

1. Marcar a posição do falante/escritor em relação a uma proposição na oração:

Os autores afirmam que, no registro acadêmico, muitos advérbios, tais como *generally*, *possibly*, *probably*, *certainly* e *obviously*, são usados para delimitar a posição/opinião do falante/escritor em relação a uma proposição, anteriormente tratada na oração: “*Obviously this is an overstatement*”.

2. Descrever processos, condições ou instruções:

Ex1: *The solubility decreases **rapidly** as the temperature falls.*

Ex2: *The passage is almost **entirely** in third-person narration.*

Ex3: *Examine footnotes **carefully**.*

3. Atuar como modificadores

Ex1: *Acclimatization is **relatively** unimportant.*

Ex2: *The paper by Donchin et al. is a **particularly** good methodological critique of research in this area.*

Os autores chamam a atenção para o fato de modificadores como *relatively* e *particularly*, encontrados nos exemplos acima, se distinguirem de modificadores mais tipicamente da fala, tais como *pretty* e *rather*.

4. Expressar precisão ou estimativas

Ex1: *After **exactly** five years the tree that had grown up weighed 169 pounds and about three ounces.*

Ex2: ***Approximately** 60 percent of the community are of Polish and Russian ancestry.*

Segundo Biber et al.(1999), alguns advérbios podem ter significados de natureza literal ou metafórica, a depender do contexto no qual são utilizados, a exemplo de *perfectly*. Em “*perfectly arranged*”, o significado do advérbio nos remete à ideia de “*in a perfect manner*”(de maneira perfeita), ao passo que em “*perfectly normal*”, ele passa a ter um sentido metafórico.

São sete principais categorias semânticas dos advérbios, trazidas à tona por Biber et al. (1999). São elas: a) lugar; b) tempo; c) modo; d) extensão; e) adição/restrição; f) posicionamento; g) ligação. Demonstramo-nas na Tabela3, abaixo:

Tabela 3 – As sete principais categorias semânticas de advérbios

	Objetivos	Exemplos
1. Lugar	demonstrar posição ou distância	<i>there, backward, far</i>
2. Tempo	mostrar posição, frequência ou duração	<i>then, recently, always, already</i>
3. Modo	manifestar como uma ação é desempenhada	<i>automatically, quickly, carefully</i>
4. Extensão	descrever a escala de uma determinada característica	
4.1 Amplificadores/intensificadores	aumentar a intensidade da escala	<i>more, very, so, too, extremely, totally</i>
4.2 Diminuidores/atenuadores	reduzir a intensidade da escala	<i>slightly, somewhat, quite</i>
5. Adição/restrição	mostrar um item adicionado a outro	<i>too, also, especially, only</i>
6. Posicionamento	demarcar a posição do falante/escritor	
6.1 Posicionamento epistêmico	* mostrar níveis de certeza ou dúvida * comentar sobre a realidade de uma proposição * mostrar a evidência de uma proposição * mostrar limitações sobre uma proposição * transmitir imprecisão	<i>probably, definitely, apparently, roughly</i>
6.2 Posicionamento atitudinal	expressar a posição em relação a uma proposição	<i>surprisingly, curiously, unfortunately</i>
6.3 Posicionamento de estilo	fazer um comentário sobre o modo de falar adotado por um falante	<i>honestly, frankly, simply</i>
7. Ligação	fazer ligações semânticas e contribuir para a coesão discursiva	
7.1 Enumeração e adição	enumerar e adicionar informação ao discurso	<i>firstly, secondly, thirdly</i>
7.2 Resumo	resumir algo que foi recém mencionado	<i>altogether, overall</i>
7.3 Aposição	prover informação extra	<i>namely, i.e</i>
7.4 Resultado/Inferência	explicar o resultado de algo	<i>therefore, thus</i>
7.5 Contraste/Concessão	contrastar informações	<i>though, alternatively, however</i>
7.6 Transição	mostrar a relação entre as frases	<i>incidentally</i>

Fonte: baseado em Biber et al. (1999). Elaborado pela autora, 2017.

Dentre as categorias semânticas elencadas no quadro, os autores enfatizam que as de maior destaque no discurso acadêmico são as que compõem a categoria de ligação e a categoria de extensão (amplificadores/intensificadores, diminuidores/atenuadores e adição/restrrição). A primeira categoria tem como objetivo fazer ligações semânticas para auxiliar na coesão do texto. Sobre esse aspecto, os autores revelam que a alta frequência de advérbios deste tipo, no *corpus* analisado, “[...] reflete a importância, no discurso acadêmico, em marcar as conexões entre ideias e mostrar explicitamente o desenvolvimento dos argumentos lógicos” (BIBER et al., 1999, p. 562, tradução nossa)⁵⁶. Os advérbios da categoria de extensão, usados para descrever o grau de intensidade de uma dada característica, são muito frequentes, de igual modo, no discurso acadêmico. Os autores afirmam que os advérbios amplificadores/intensificadores, tais como *entirely*, *extremely*, *fully*, *highly* e *strongly* são muito frequentes no discurso acadêmico, como pode ser ilustrado nos exemplos abaixo retirados de Biber et al. (1999, p.565):

1. *But the publication of Darwin's theory of biological evolution gave an **entirely**[new]stimulus to the elaboration of organic theories.*
2. *Indeed it is **extremely** [difficult] to establish any truly satisfactory system of defining the limits of these functions.*
3. *The office slave boasts a waist and slender figure, while the navy rolls along, a**fully** [developed] man of bone and muscle.*
4. *The **highly** [complex] process of adjustment to infection is determined by many variables.*
5. *When both calmodulin and GTP are present, a **strongly** [biphasic] curve is produced with a peak activity at 1.1 x 10⁴ m.*

Segundo sinalizam Biber et al. (1999), os amplificadores em destaque nos exemplos acima funcionam, nessas construções, de forma a manifestar os graus de intensidade, dificuldade, desenvolvimento, complexidade e bifasicidade, respectivamente. Ainda, comentam sobre o uso dos amplificadores como *very*, *so* e *too* que também são muito frequentes no registro oral. Diante desses exemplos, constatamos que, nos exemplos 1 (*entirely*), 4 (*highly*) e 5 (*strongly*), os amplificadores se

⁵⁶No original: “[...] reflects the importance in academic prose of marking the connections between ideas and explicitly showing the development of logical arguments” (BIBER et al., 1999, p. 562).

posicionam na condição de “advérbio + substantivo”, ao passo que em 2 (*extremely*) e 3 (*fully*), estão na posição “advérbio + adjetivo”, porém nenhum dos amplificadores se encontra na condição colocacional de “advérbio + verbo”, padrão ao qual nos dedicamos a investigar.

Ainda sobre os exemplos acima, outro ponto também destacado pelos autores nos informa que, no discurso acadêmico, muitas colocações objetivam estabelecer uma medida de comparação entre dois itens, como é o caso do uso do advérbio “*relatively*”.

1. *However, the morphology is still **relatively** [simple].*
2. *In any **relatively** [stable] society, there exists an equilibrium between the mode of production, the social relations, and the "superstructure".*
3. *Amphiboles occur characteristically in the plutonic rocks and in general are **relatively** [unimportant] minerals of the volcanic rocks.*

O posicionamento do falante/escritor pode se realizar através de determinadas classes de palavras, dentre as quais destacamos os verbos (HYLAND, 2002) e advérbios (BIBER; FINEGAN, 1988; BIBER et al., 1999). Para Biber (1988, p.204, tradução nossa)⁵⁷, o falante/escritor posiciona-se, no discurso acadêmico, a fim de manifestar-se abertamente e expressar “[...] atitudes, sentimentos, julgamentos, ou compromisso referente a mensagem”. Em miúdos, podemos afirmar que quem escreve tende a posicionar-se em relação a alguma proposição e, para tanto, utiliza expressões linguísticas, dentre as quais podemos citar a classe dos verbos e advérbios. Os advérbios de posicionamento são muito comuns no discurso acadêmico, principalmente nos registros orais. São subclassificados por Biber (2006) em: advérbios epistêmicos (advérbios de certeza e advérbios de significância), advérbios atitudinais e advérbios de estilo. Os advérbios epistêmicos são o tipo de advérbio de posicionamento mais comum no discurso acadêmico. Eles são usados pelo falante para expressar o *status* de uma informação, trazendo, para tanto, uma indicação de certeza (ex: *certainly, definitely, undoubtedly, inevitably, obviously*), asseverando a manifestação da verdade do que é dito/escrito. Já os advérbios de significância (ex: *apparently, evidently, possibly, probably, roughly*) são usados para fazer juízo de valor sobre alguma proposição. Os

⁵⁷ No original: “[...] attitudes, feelings, judgments, or commitment concerning the message” (BIBER, 1988, p. 204).

advérbios atitudinais (ex: *amazingly, curiously, sadly, importantly, unfortunately, hopefully*), importantes na comunicação acadêmica oral, principalmente nas falas em sala de aula, por sua vez, referem-se às atitudes do falante/escritor frente a proposições. Os advérbios de estilo (ex: *honestly, mainly, technically, typically, usually, primarily*) que, de acordo com Biber (2006), constituem os advérbios de posicionamento mais frequentes nos registros acadêmicos escritos por ele analisados (livros didáticos e material de aulas), são usados para fazer comentários acerca de uma informação apresentada.

Os advérbios *-ly*, combinados aos verbos do inglês acadêmico escrito, formam o tipo de colocação que nos instiga à pesquisa. Na seção que se segue, apresentamos as contribuições de estudos que se propuseram a definir, explicar, identificar e verificar o uso fraseológico das colocações no inglês acadêmico escrito.

As colocações e a fraseologia no inglês acadêmico escrito

The choice of one word conditions the choice of the next, and of the next again. The item and the environment are ultimately not separable, or certainly not separable by present techniques (SINCLAIR, 2004, p. 19).

As pesquisas em torno das colocações tiveram início com o trabalho do britânico John Firth, que incluiu o conceito de colocação à linguística moderna. Apesar do tempo em que são estudadas, não há um consenso sobre a definição do termo “colocações” (XIAO, 2015). Contudo, sabemos que ele pode ser atribuído a duas diferentes perspectivas: uma delas, trazida à tona por Halliday e Hasan (1976), definem colocações como associações entre palavras, também chamadas de “colocações por coerência” (*coherence collocation*). A outra perspectiva, de natureza fraseológica, mais intimamente ligada à Linguística de Corpus, traz o conceito de “colocação por vizinhança” (*neighbourhood collocation*) e define-a como uma sequência de palavras que coocorrem mais frequentemente umas com as outras. Nesta pesquisa, nos ancoramos nessa última perspectiva, uma vez que entendemos e concordamos que as palavras não são meras combinações aleatórias.

Além disso, concordamos que:

[...] não somente o desenvolvimento da Linguística de Corpus tem facilitado grandemente a análise colocacional, mas a pesquisa sobre colocações baseadas em *corpus*, nas últimas décadas, tem revelado também uma diversidade de resultados interessantes que, até então, eram amplamente

negligenciados em descrições de língua não baseadas em *corpus* (XIAO, 2015, p.115, tradução nossa)⁵⁸.

Xiao (2015) nos informa que a perspectiva fraseológica de colocação tem sido investigada sob diferentes terminologias: “*lexical bundle*” e “*word cluster*”, na Linguística de Corpus, “*mutiword unit*” e “*n-gram*” no processamento de linguagem natural e, ainda, “expressão formulaica”. Um dos grandes pesquisadores do tema, John Sinclair, um dos primeiros pesquisadores a explicar, através de dados baseados em *corpus*, no modo como as palavras coocorrem em padrões específicos, atesta que as colocações incluem combinações binárias com uma organização sintática, podendo consistir em elementos gramaticais que, juntos, integram o significado. As colocações do tipo lexicais podem ser de seis diferentes tipos: substantivo + substantivo, verbo + substantivo (*reach an agreement*), advérbio + verbo (*drink greedily*), adjetivo + substantivo (*heavy smoker*), advérbio + adjetivo (*blissfully happy*) e verbo + frase preposicional.

De acordo com Sinclair (1991), as colocações podem ser explicadas a partir de dois princípios de interpretação que tentam elucidar o modo como os significados são construídos e a necessidade de serem levados em conta, ao observar e descrever a língua. O primeiro deles é intitulado “*open-choice principle*” ou “princípio da livre escolha”, resultante do que ele chama de “um grande número de escolhas complexas”. Esse princípio parte da ideia que as palavras podem ser decompostas item a item. Por outro lado, o “*idiom principle*”, ou “princípio idiomático”, de natureza fraseológica, pressupõe que “[...] o usuário da língua tem acesso a uma grande quantidade de frases semi pré-construídas que constituem escolhas únicas, mesmo podendo ser analisadas em segmentos” (SINCLAIR, 1987, p.319, tradução nossa)⁵⁹. Sobre esse princípio, o autor utiliza o exemplo da colocação “*of course*”, em que *of* não denota uma preposição tal qual encontrada nas gramáticas tradicionais e *course* não se refere a um substantivo contável, uma vez que o seu significado não está na propriedade da palavra, mas no âmbito da oração. O princípio idiomático, segundo o autor, tem sido objeto de estudo de muitos pesquisadores na área da linguagem, mas sua importância ainda tem sido bastante relegada. Contudo, o autor acredita que este deva ser o princípio a ser utilizado

⁵⁸ No original: “[...] not only has the development of corpus linguistics greatly facilitated collocation analysis but corpus-based collocation research over the past decades has also revealed a range of interesting findings that have hitherto been largely overlooked in previous non corpus-based language descriptions” (XIAO, 2015, p. 115).

⁵⁹ No original: “[...] a language user has available to him or her a large number of semipreconstructed phrases that constitute single choices, even though they might appear to be analysable into segments” (SINCLAIR, p.1987, p. 319).

pelos linguistas ao analisar textos.

Quando se fala em colocações é comum pensarmos no fenômeno da coocorrência. Scott e Tribble (2006) nos lembram que a coocorrência nem sempre pode significar uma estreita relação entre dois itens. Por vezes, ela apenas informa que, com frequência, aqueles elementos podem aparecer juntos em certos textos. Os autores, por assim dizer, afirmam que nem sempre há simetria entre o nódulo e o colocado e, metaforicamente, exemplificam que a relação entre as palavras se assemelha à um relacionamento entre um casal, no qual, enquanto uma das partes morre de amores, o outro não corresponde ao afeto com a mesma intensidade.

Vários estudos têm se debruçado na identificação e análise de colocações e padrões léxico-gramaticais, definidos como regularidades lexicais, gramaticais ou semânticas que ocorrem próximas a palavra em análise (DAYRELL, 2011). Segundo Dayrell (2011), são considerados padrões léxico-gramaticais as repetições lexicais em volta do lema, as diferentes formas de um mesmo lema e a recorrência de um grupo semântico específico. Como bem coloca Stubbs (1996, p.36, tradução nossa)⁶⁰, “[...] não há limite entre o léxico e a gramática: léxico e gramática são interdependentes”. Em outras palavras, podemos dizer que essa afirmação retrata o perfil de uma abordagem para o estudo e análise da língua que não polariza léxico e gramática, mas busca a unidade entre ambos, como defendemos também nesta pesquisa.

Conforme atesta Granger e Paquot (2009, p.151, tradução nossa)⁶¹:

[...] cada verbo do Inglês para Fins Acadêmicos tem suas próprias companhias léxico-gramaticais, a saber, sujeitos (este estudo mostra que; a evidência sugere que; esses resultados sugerem que), objetos (apoia a visão/hipótese que; fornece evidência/informação) e advérbios (difere significativamente; varia consideravelmente/amplamente; se aplica igualmente; intimamente relacionado; largamente usado; geralmente aceito) e tende a aparecer em estruturas rotinizadas [...].

Granger (1998), ao conduzir um estudo sobre as colocações na escrita acadêmica de aprendizes de inglês, verificou comparativamente o uso dos advérbios amplificadores que terminavam com o sufixo *-ly* e funcionavam como modificadores no subcorpus de aprendizes franceses de inglês no ICLE e no *corpus* de redações argumentativas da

⁶⁰ No original: “There is no boundary between lexis and grammar: lexis and grammar are interdependent” (STUBBS, 1996, p. 36).

⁶¹ No original: “[...] each EAP verb has its own preferred lexico-grammatical company, namely, subjects (this study shows that; the evidence suggests that; these results suggest that), objects (supports the view/hypothesis that...; provide evidence/information) and adverbs (differ significantly; vary considerably/widely; apply equally; closely related; widely used; generally accepted) and tend to appear in routinized structures [...]” (GRANGER; PAQUOT, 2009, p. 151).

University of Louvain. Os resultados do estudo comprovaram: a) o uso de muitos advérbios modificadores nos textos dos aprendizes franceses, embora muitos fizessem parte do que ela chamou de “combinações criativas”, como em *easily controlled*; b) os advérbios amplificadores utilizados no *corpus* de estudo tinham uma correspondência direta com a L1 dos aprendizes, o francês; c) os aprendizes sobreusam *very*, em detrimento de outros advérbios; d) subuso estatisticamente significativo de amplificadores do subcorpus do ICLE, tanto em número de *types* quanto em número de *tokens*; e) sobreuso de *completely* e *totally* e subuso de *highly* pelos aprendizes franceses de inglês. A autora atribuiu as diferenças à dificuldade com o uso do vocabulário e às questões intralinguais entre a L1 dos aprendizes e o inglês, uma vez que havia muitas combinações inapropriadas, inventadas e escritas pelos aprendizes.

Com base nessas pressuposições teóricas e, entendendo que o conhecimento colocacional é de extrema relevância para o desenvolvimento de habilidades na L1 e na língua adicional (ALTENBERG; GRANGER, 2001; PARTINGTON, 1998), nos dedicamos a descrição e discussão das colocações, formadas por “advérbio *-ly* + verbo”, com os verbos do inglês acadêmico escrito do CorIFA. Compreendemos que o verbo atua como base da combinação, enquanto o advérbio é o colocado, numa relação de dependência sintática, em que se pode reconhecer os diferentes domínios semânticos dos advérbios, a partir de seu uso.

Além disso, Hunston (2002) traz o conceito de fraseologia como determinante para a compreensão do significado das palavras, bem como retoma o princípio da livre escolha e o princípio idiomático, postulados por Sinclair (1991), como vimos discutindo. Para ela, não há distinção entre padrão e significado e entre léxico e gramática. A fraseologia consiste em mais que um conjunto de frases superpostas, o que implica na não ocorrência aleatória de palavras na língua. Assim como Sinclair (1991), Hunston (2002) acredita que esta se constitua o cerne da descrição de línguas e deixa um questionamento que direciona as discussões vindouras: “A fraseologia pode, de fato, ser tomada como a base do que o aprendiz precisa saber ou a fraseologia é simplesmente um adjunto importante para a gramática?” (HUNSTON, 2002, p.138, tradução nossa)⁶².

Nesse sentido, ver a língua a partir de uma perspectiva fraseológica significa rejeitar a visão tradicional de léxico e gramática como entidades separadas, conforme

⁶²No original: “Can phraseology indeed be taken to the basis of what a learner needs to know, or is phraseology simply an important adjunct to grammar?” (HUNSTON, 2002, p.138).

Hunston (2002) discute. Para ela, as distinções em torno do léxico e da gramática se pautam em três vieses: as palavras gramaticais são mais frequentes que as palavras lexicais; as palavras gramaticais estão mais relacionadas a aspectos paradigmáticos, ao passo que as lexicais se relacionam com os sintagmáticos; as palavras gramaticais não têm significado próprio, mas as lexicais, sim.

A seguir, apresentamos os procedimentos metodológicos utilizados para a obtenção dos dados desta pesquisa.

CAPÍTULO 2

PROCEDIMENTOS METODOLÓGICOS

Studies of use are concerned with actual practice, and the extent to which linguistic patterns are common or rare, rather than focusing exclusively on potential grammaticality. As such, adequate investigations of language use must be empirical, analyzing the functions and distribution of linguistic features in natural discourse contexts.

(BIBER, 2000, p.)

O presente estudo tem como objetivo principal analisar a frequência e o uso do que intitulamos Verbos do Inglês Acadêmico Escrito (VIAEs) do subcorpus do CorIFA. Além desse, outros quatro objetivos orientam a presente investigação. O primeiro diz respeito a identificar os verbos mais frequentes no subcorpus do CorIFA, através da geração de uma lista de palavras. O segundo objetivo consiste em identificar quais deles são encontrados na Academic Vocabulary List (AVL) como mais frequentes. O terceiro objetivo é descrever o uso dos VIAEs no subcorpus do CorIFA. O quarto objetivo é comparar os resultados do subcorpus do CorIFA com um subcorpus compilado do British Academic Written English (BAWE), ancorado na Análise Contrastiva de Interlíngua. O quinto objetivo consiste em identificar as colocações formadas por “advérbio-ly + verbo” com os VIAEs mais frequentes e contrastar com as disponíveis no subcorpus do BAWE.

Neste capítulo, dedicamo-nos à apresentação e exemplificação das escolhas metodológicas utilizadas para esta pesquisa. Em um primeiro momento, traçamos o perfil dos aprendizes cujos textos formam o subcorpus do CorIFA. Depois, apresentamos e descrevemos de modo sistemático os *corpora* utilizados no estudo, bem como relatamos a extração de nossas unidades de análise, ou seja, os verbos do inglês acadêmico escrito, para a utilização na pesquisa. Mais adiante, relatamos sobre o AntConc 3.4.3 (ANTHONY, 2011), ferramenta computacional usada para a seleção e extração de dados, bem como abordamos os procedimentos utilizados para a análise.

Perfil dos aprendizes

Os estudantes das disciplinas de Inglês para Fins Acadêmicos (IFA) são provenientes dos mais diversos cursos de graduação e pós-graduação da Universidade

Federal de Minas Gerais (UFMG). Para a realização da matrícula, os estudantes que se candidatam ao IFA fazem um teste de nivelamento ou apresentam resultados de testes internacionais, entre eles o TOEFL iTP⁶³ e iBT, Cambridge English: First (FCE) ou Cambridge English: Advanced (CAE).

O contexto acadêmico no qual os estudantes do IFA estão inseridos tem se preocupado em atender as demandas linguísticas dos aprendizes, levando em consideração o caráter heterogêneo das turmas e primando pelo ensino da língua para os mais diversos propósitos acadêmicos, desde a apresentação oral de trabalhos até a produção escrita de variados gêneros acadêmicos.

Como já explicitada na introdução desta tese, tivemos a experiência docente com uma turma do Inglês para Fins Acadêmicos II (nível B1, de acordo com o Quadro Comum Europeu de Referência para Línguas⁶⁴) no segundo semestre de 2015. Na ocasião, a turma era composta por 25 estudantes pertencentes aos mais diversos cursos de graduação e pós-graduação da UFMG. Percebemos que os estudantes se engajavam bem nas atividades e expressavam o desejo de aperfeiçoar sua proficiência linguística para, por exemplo, apresentar trabalhos científicos nos eventos de suas respectivas áreas, bem como escrever textos em inglês para publicação.

As disciplinas do IFA constam no histórico escolar do estudante por poder fazer parte das disciplinas de formação livre. Caso concorde que as versões do seu texto façam parte do CorIFA, o estudante assina um Termo de Consentimento (Anexo 1) e, em seguida, envia para o endereço eletrônico da administração do *corpus*. Assim, somente os textos acompanhados da assinatura do referido termo são incluídos no CorIFA.

Caracterização e compilação do *corpus* de estudo

A seguir, apresentamos o *corpus* que nos permitiu selecionar o subcorpus utilizado como base para esta pesquisa.

⁶³ O TOEFL iTP é um exame de proficiência que permite avaliar as habilidades de estudantes em língua inglesa e tem sido utilizado em mais de 50 países, inclusive no Brasil. Com o intuito de avaliar o conhecimento linguístico de estudantes universitários brasileiros, o Ministério da Educação, através do Programa Idiomas sem Fronteiras, tem disponibilizado o exame nas universidades públicas e institutos federais de educação como forma de prover oportunidades de acesso de estudantes a programas de intercâmbio acadêmico em universidades estrangeiras. Mais informações sobre o exame estão disponíveis em: https://www.ets.org/toefl_itp.

⁶⁴ A distribuição das disciplinas do Inglês para Fins Acadêmicos se aporta nos níveis estabelecidos no Quadro Comum Europeu de Referência para Línguas (European Framework of Reference for Languages). Na seção 3.2.1, apresentamos tal distribuição.

O subcorpus do CorIFA

[...] learner corpora need to be assembled on the basis of very strict design criteria and a wide range of variables should ideally be recorded for each learner production (GRANGER, 2009, p.17).

O subcorpus do CorIFA é composto de 638 redações argumentativas, no qual encontramos 12.671 *word types* (tipos) e 253.659 *tokens* (itens).

No quadro a seguir, apresentamos o número de textos de cada gênero, compilados nos anos de 2013, 2015 e 2016, a partir dos quais geramos dados que estão dispostos na seção 3.6 deste capítulo.

Quadro 1 – Estrutura do subcorpus do CorIFA

GÊNERO	ANO E NÚMERO DE TEXTOS			TOTAL DE TEXTOS	NÍVEL DOS ALUNOS
	2013	2015	2016		
Redação argumentativa	404	97	137	638	B1 e B2

Fonte: Elaborado pela autora, 2017.

Todas as disciplinas do Inglês para Fins Acadêmicos são presenciais e possuem carga horária de 60h/aula, tendo como objetivo o aprimoramento do conhecimento de inglês dos estudantes para a participação nas diversas práticas acadêmicas (eventos internacionais e publicações), bem como em programas de intercâmbio.

Cada disciplina do IFA demanda a produção escrita de um determinado gênero textual, considerando as necessidades de escrita dos estudantes matriculados. Sendo assim, a partir do ano de 2016, os gêneros textuais solicitados para cada turma estão dispostos na Tabela 4, a seguir:

Tabela 4 – Gêneros textuais solicitados para os estudantes do IFA

IFA I	IFA II	IFA III	IFA IV	IFA V
Resumo	Resumo de artigo/resumo	Redação argumentativa	Revisão de literatura	Artigo de pesquisa
Carta de apresentação	redação argumentativa	_____	_____	_____

Fonte: Elaborada pela autora, 2017.

Todas essas disciplinas do IFA, como mencionado anteriormente, são ofertadas

pela Faculdade de Letras da UFMG e, baseadas em cinco níveis, de intermediário a intermediário-avançado que, elaboradas de acordo com o Quadro Comum Europeu de Referência para Línguas, correspondem respectivamente a: B1 (IFA I e IFA II), B2 (IFA III e IFA IV) e C1 (IFA V). O subcorpus do CorIFA, portanto, é composto, até o momento da escrita deste trabalho, por 253.659 palavras. Esperamos obter, ao final da compilação do CorIFA, no mínimo 200 mil palavras em cada nível dos cursos ofertados.

A maior parte das tarefas que originaram os textos que compõem o CorIFA foram feitas pelos estudantes em casa, exceto as tarefas de 2015/2 que foram feitas em laboratório, por fazerem parte dos *mock tests*⁶⁵. Foi escolhido um teste de cada nível (B1 e B2) dentre os disponíveis na página do Corpus do Inglês sem Fronteiras⁶⁶ (CorIsF) para ser aplicado aos estudantes matriculados nas disciplinas do IFA, com o objetivo de verificar a proficiência e o desempenho dos estudantes na escrita de textos com determinado limite de tempo.

No caso das tarefas realizadas em laboratório, houve restrição de tempo para a escrita do texto. Quanto às tarefas realizadas em casa, os estudantes não tinham limite de tempo, apenas prazo previamente determinado para entregar a versão final. No início da oferta da disciplina de Inglês para Fins Acadêmicos, era necessário os estudantes enviarem aos professores a primeira versão de suas redações. Sendo assim, algumas foram editadas, uma vez que passaram pelo crivo do professor.

Na próxima subseção, descrevemos as instruções das tarefas que deram origem a boa parte das redações argumentativas que compõem o *corpus*, a fim de demonstrar um panorama geral das tarefas que propiciaram a produção escrita dos gêneros do CorIFA, pelos aprendizes.

Dados sobre as instruções das tarefas do subcorpus do CorIFA

⁶⁵ Os *mock tests*, que aqui nos referimos, são testes temáticos, de habilidades integradas que contém as três seções: *reading*, *listening* e *writing*. Tais testes apresentam vídeos, textos e imagens e são organizados por nível de proficiência (seguindo o Quadro Comum Europeu de Referência para Línguas). Além disso, são corrigidos automaticamente e estão disponíveis no *site* do CorIsF (sites.google.com/sites/corpusisf) apenas para os professores de instituições de ensino superior parceiras que tenham interesse em aplicá-los com seus respectivos estudantes.

⁶⁶ O CorIsF é um *corpus* em compilação que agrega textos produzidos por alunos do Programa Inglês sem Fronteiras. Esses alunos são de cursos de todas as áreas do conhecimento, sendo ofertados gratuitamente, mas não fazem parte da grade curricular. O Programa, do MEC e CAPES, iniciou em 2012 e em 2014 se tornou Idiomas sem Fronteiras. Em 2017, a oferta de cursos presenciais de várias línguas, entre elas, francês, italiano, alemão e espanhol, passaram a ser oferecidos em várias universidades brasileiras.

As propostas e tópicos de escrita das tarefas⁶⁷ as quais os aprendizes são solicitados a escrever, a nosso ver, podem certamente influenciar nas suas escolhas linguísticas e, conseqüentemente, no uso de repertório lexical dos textos que compõem o *corpus*.

Tendo em vista tal pressuposto, apresentamos a seguir o número mínimo de palavras e as instruções de boa parte das tarefas do subcorpus do CorIFA, em ordem cronológica, acreditando que eles podem ter influenciado nas escolhas lexicais dos aprendizes e, mais especificamente aqui, no uso dos verbos do inglês acadêmico escrito, que serão analisados no próximo capítulo desta tese.

Apresentamos abaixo o Quadro 2, no qual podemos encontrar as informações referentes ao número de textos, gêneros e nível dos aprendizes que compõem o subcorpus do CorIFA.

Quadro 2 – Estrutura do CorIFA: número de textos, gênero e nível dos alunos

2013.1	Número de textos	Gênero	Nível
	271	Redação argumentativa	B1
	102	Redação argumentativa	B2
2013.2	Número de textos	Gênero	Nível
	805	Carta de apresentação	B1
	276	Carta de apresentação	B2
	317	Redação argumentativa	B2

⁶⁷ Muitos títulos das tarefas propostas para os estudantes do IFA no ano de 2013 e que compõem o CorIFA são sugestões do ICLE, tais como: 1. Crime does not pay; 2. The prison system is outdated. No civilised society should punish its criminals: it should rehabilitate them.; 3. Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.; 4. A man/woman's financial reward should be commensurate with their contribution to the society they live in.; 5. The role of censorship in Western society.; 6. Marx once said that religion was the opium of the masses. If he was alive at the end of the 20th century, he would replace religion with television.; 7. All armies should consist entirely of professional soldiers : there is no value in a system of military service.; 8. The Gulf War has shown us that it is still a great thing to fight for one's country.; 9. Feminists have done more harm to the cause of women than good.; 10. In his novel Animal Farm, George Orwell wrote "All men are equal: but some are more equal than others". How true is this today?; 11. In the words of the old song "Money is the root of all evil"; 12. Europe.; 13. In the 19th century, Victor Hugo said: "How sad it is to think that nature is calling out but humanity refuses to pay heed. "Do you think it is still true nowadays ?; 14. Some people say that in our modern world, dominated by science technology and industrialisation, there is no longer a place for dreaming and imagination. What is your opinion ?". Todos esses títulos estão disponíveis no sítio eletrônico: <https://uclouvain.be/en/research-institutes/ilc/cecl/corpus-collection-guidelines.html>.

2015.2	Número de textos	Gênero	Nível
	42	Redação argumentativa	B1
	55	Redação argumentativa	B2
2016.1	Número de textos	Gênero	Nível
	28	Redação argumentativa	B2
	28	Redação argumentativa	B2
	65	Resumo de artigo	B1
	43	Resumo de artigo	B1
2016.2	Número de textos	Gênero	Nível
	54	Redação argumentativa	B2
	27	Redação argumentativa	B2
	32	Resumo de artigo	B1
	21	Resumo de artigo	B1

Fonte: Elaborado pela autora, 2017.

Mais especificamente no Quadro3 e no Tabela 5, que mostramos a seguir, podemos encontrar informações quanto às instruções das tarefas dos textos produzidos em 2013 e que fazem parte do subcorpus do CorIFA.

Quadro 3– Número mínimo de palavras das tarefas dos textos do subcorpus do CorIFA

	Number of words					
	No less than...					
	<i>Writing 1</i>	<i>Writing 2</i>	<i>Essay 1</i>	<i>Quiz 1</i>	<i>Quiz 2</i>	<i>Final</i>
IFA I	200	250	300	200	250	300
IFA II	200	250	300	200	250	300
IFA III	250	300	350	250	300	350
IFA IV	250	300	350	250	300	350
IFA V	350	400	450	350	400	450

Fonte: Arquivo de tarefas do IFA, 2013.

Na Tabela 5 encontramos as instruções das tarefas dos textos no primeiro semestre do ano de 2013 e que compõem o CorIFA:

Tabela 5– Instruções das tarefas dos textos do subcorpus do CorIFA

Writing 1	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>You want to apply for a foreign University to take an undergraduate course. Write an essay to convince them of your academic skills and of the relevance of the study you will develop.</i></p>
Writing 2	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>Is it more important for a government to spend money on public transportation (buses, subways, etc.) or on roads?</i></p> <p><i>Use specific reasons and examples to support your opinion.</i></p>
Essay 1	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. What is your opinion?</i></p> <p><i>Use specific reasons and examples to support your opinion.</i></p>
Quiz 1	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>Most university degrees are theoretical and do not prepare students for the real world. They are therefore of very little value.</i></p> <p><i>Do you agree or disagree? Use specific reasons and examples to support your position.</i></p>
Quiz 2	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>"For a business, it is acceptable to do anything to make a profit."</i></p> <p><i>Explain why you agree or disagree with this observation, supporting your ideas.</i></p>
Final Exam	<p>Write an essay of at least <u>XXX words</u>.</p> <p><i>In the 19th century, Victor Hugo said: "How sad it is to think that nature is calling out but humanity refuses to pay heed*". Do you think it is still true nowadays?</i></p> <p><i>*pay attention</i></p>

Fonte: Arquivo de tarefas do IFA, 2013.

No Quadro 4, é evidenciado o número de palavras para cada produção escrita, a depender do nível dos aprendizes. Quanto aos textos, temos as seguintes informações: *Writing 1* (doravante W1) e *Writing 2* (doravante W2), os *essays* e *quiz* são referentes à uma redação cujo tema o aprendiz pode escolher dentre os listados um pouco abaixo no

quadro.

Na Tabela 5 percebemos que as tarefas de produção escrita são todas do tipo independente, ou seja, não há recursos textuais e vídeos que sirvam como ponto de partida para a escrita, o que pode colaborar para que o aprendiz utilize de seu repertório lexical sem, necessariamente, ter um suporte que o direcione na escrita.

O W1 apresenta uma proposta de escrita que muito se assemelha com o gênero carta de apresentação (*statement of purpose*), embora não seja assim intitulada na listagem de tarefas a serem produzidas. O aprendiz precisa, dessa forma, escrever um texto endereçado a uma universidade a qual tenha interesse em estudar e, para tanto, há a necessidade de convencer e argumentar acerca de suas habilidades acadêmicas e da relevância do estudo que se propõe a desenvolver.

De certo, os aprendizes nas tarefas aqui mostradas eram convocados a expressar e argumentar suas opiniões e ideias, tendo como ponto de partida uma afirmação ou questionamento, o que lhes dava a oportunidade de refletir e, desse modo, escrever sobre o tópico em questão.

Ainda, outros temas foram sugeridos pelos professores:

Tabela 6– Instruções para textos do IFA– 2013.2

1	Write an essay of at least <u>xxx words</u> . <i>Is it more important to take risks or to plan carefully if you want to be successful?</i> <i>Use specific reasons and examples to support your opinion</i>
2	Write an essay of at least <u>xxx words</u> . <i>When people immigrate, should they adapt to the new country or keep their old customs?</i> <i>Use specific reasons and examples to support your opinion.</i>
3	Write an essay of at least <u>xxx words</u> . <i>Discuss the contributions of artists vs. those of scientists.</i>
4	Write an essay of at least <u>xxx words</u> . <i>Technology causes cultures to become more similar.</i> <i>Do you agree or disagree with this statement? Use specific reasons and examples to support your opinion.</i>
5	Write an essay of at least <u>xxx words</u> . <i>The prison system is outdated. No civilized society should punish its criminals: it</i>

	<p><i>should rehabilitate them.</i></p> <p><i>Do you agree or disagree with this statement? Use specific reasons and examples to support your opinion.</i></p>
6	<p><i>Write an essay of at least xxx words.</i></p> <p><i>Why do you think it is that people's lives are longer now?</i></p> <p><i>Use specific reasons and examples to support your opinion.</i></p>

Fonte: Arquivo de tarefas do IFA, 2013.

É possível perceber que os tipos de tarefas propostas para os aprendizes nas Tabelas 5 e 6 apresentadas anteriormente são todas do tipo tarefa independente. A tarefa do tipo independente envolve somente a escrita e vem acompanhada de uma simples proposta de redação; a tarefa do tipo integrada envolve as habilidades de leitura, compreensão oral e escrita (HIGGINS, 2011) que, geralmente, vêm acompanhadas de um texto escrito, vídeo, tabela, um gráfico ou um infográfico. Tarefas como essa muito se assemelham, por exemplo, com os da seção de escrita do TOEFL iBT, na qual há um trecho para leitura que relata sobre determinado assunto e, em seguida, os estudantes assistem a um vídeo e precisam produzir a redação, de forma a resumir e avaliar o que leram e compreenderam. Ainda, as tarefas do tipo independentes requerem dos aprendizes conhecimento prévio sobre o tópico do texto, ao passo que as tarefas do tipo integradas conduzem os aprendizes a escreverem com base nas informações do texto escrito e do vídeo (WOLFERSBERGER; COOMBE, 2015).

Compilação do CorIFA

A compilação de um *corpus* é definida através da tomada de determinadas decisões (SINCLAIR, 1987) e critérios (DUTRA et al., 2016; GRANGER, 1998), estabelecidos antes mesmo da coleta como, por exemplo, o gênero, o tipo e o tamanho da tarefa executada, nível de proficiência do aprendiz, idade e curso de graduação ou pós-graduação, que constituem os metadados do *corpus*.

O processo de compilação do CorIFA seguiu algumas etapas. Inicialmente, fizemos a busca dos textos no formato *.txt*⁶⁸, coletados no ano de 2013 e que já haviam passado por uma limpeza, ação importante para a preparação do *corpus* para o

⁶⁸ Extensão de arquivo de texto que abre com qualquer editor de texto.

processamento computacional. Em seguida, removemos todos os textos que não possuíam o termo de consentimento assinado pelos estudantes.

Após essa ação, tornamos anônimos os aprendizes e verificamos nas coletas quais os textos pertenciam à categoria independente ou integrada. Em seguida, decidimos criar convenções para nomear os arquivos que contemplassem informações contextuais sobre o texto (REPPEN, 2010). Assim, elaboramos um código de cada texto, contendo caracteres alfa-numéricos, e o padronizamos para todos os textos, no qual constam as seguintes informações: o nome do *corpus*, a abreviação da universidade⁶⁹, o nível do aluno, o tipo de tarefa (independente ou integrada⁷⁰), o gênero, informação referente a a versão 1 ou a versão final, ou seja, se foi editado ou não, o mês e ano no qual o texto foi coletado, o número da produção escrita do aprendiz e, por fim, o número do aluno. O exemplo a seguir demonstra como os textos têm sido codificados:

Tabela 7– Exemplo de código de texto do CorIFA e suas especificações

<CorIFA-UFMG-A2.Int.AEss.E.mar2013.0001.0002>

CorIFA	Nome do <i>corpus</i>
UFMG	Nome da universidade
A2	Nível de proficiência (segundo o Quadro Comum Europeu de Referência para as Línguas)
Int	Tipo de tarefa (Ind = Independente; Int = Integrada)
AEss	Gênero textual (Let = <i>Letter</i> ; SOP = <i>Statement of Purpose</i> ; Abs = <i>Abstract</i> ; Sum = <i>Summary</i> ; Ema = <i>Email</i> ; AEss = <i>Argumentative Essay</i> ; DEss = <i>Descriptive Essay</i> ; NEss = <i>Narrative Essay</i>)
E	Texto editado (E = <i>Edited</i> ; NE = <i>Non-edited</i>)
mar2013	Mês e ano da coleta
0001	Número do texto no <i>corpus</i>

⁶⁹ Há um desejo e interesse de que outras universidades públicas em diferentes estados brasileiros possam colaborar na composição e expansão do CorIFA, com o objetivo de se obter mais dados sobre a interlíngua dos aprendizes universitários brasileiros.

⁷⁰ Chamamos de “tarefa independente” aquelas nas quais o aprendiz produz o texto usando seu conhecimento e experiência de mundo. Por sua vez, a tarefa integrada requer do aprendiz a interpretação de um gráfico, diagrama, tabela ou imagem para então, produzir seu texto (DUTRA; GOMIDE, 2015).

0001	Número do aluno no <i>corpus</i> (cada aluno recebe um número na ordem em que a redação está organizada)
------	--

Fonte: Elaborada pela autora, 2017.

Embora esse código de registro pareça extenso, houve a necessidade de fornecer informações específicas, a fim de facilitar o trabalho dos pesquisadores quando da utilização dos textos. Caso um pesquisador tenha interesse, por exemplo, em investigar determinados padrões léxico-gramaticais apenas nos textos de alunos do nível A2, cujas tarefas foram independentes, ele já os visualiza de imediato, seleciona-os e realiza seu estudo.

A partir de 2015, a proposta para a coleta dos textos no CorIFA não mais tem sido realizada como nos anos anteriores, no qual o professor recebia os textos do estudante e os encaminhava para o grupo de colaboradores do CorIFA. No segundo semestre de 2015, foi criado um formulário⁷¹, via Google Docs⁷², que passou a ser utilizado pelos professores do IFA para a coleta de textos dos seus respectivos estudantes.

A fim de organizarmos as informações referentes ao processo de coleta e aos gêneros requisitados para cada nível do IFA para compor o CorIFA, bem como, diante da necessidade de disponibilizar um formulário, no qual o estudante envia suas produções escritas ao *corpus*, criamos um *website*⁷³, via Google Sites⁷⁴. A seguir, a tela inicial do *website* do CorIFA:

⁷¹ Disponível em: <https://docs.google.com/forms/d/1nnRMNCdnbTzbcBRHOSknuNwhRsxh47ukdh0nYHEtLyw/edit>.

⁷² Pacote de aplicativos do Google, através do qual é possível criar, editar, visualizar e compartilhar documentos.

⁷³ No *website* <https://sites.google.com/site/corpusifa/home> é possível encontrar, dentre outras informações, detalhes sobre a coleta de dados para o CorIFA e informações sobre os gêneros requisitados para cada nível do IFA. No entanto, os textos que compõem o *corpus* ainda não estão disponibilizados para consulta pública.

⁷⁴ O Google Sites é um serviço ofertado pelo Google gratuitamente para a edição de *websites*.

Figura 1– Tela inicial do *website* do CorIFA



Fonte: Disponível em: <<https://sites.google.com/site/corpusifa/>>. Acesso em: 18 jun. 2017.

No *website*, é possível encontrar informações referentes à coleta dos textos, aos gêneros coletados para compor o *corpus*, ao formulário que precisa ser preenchido pelo estudante para o envio e autorização de uso do seu texto, além dos nomes que compõem a equipe organizadora do *corpus*.

2.3 Organização dos dados do subcorpus de estudo

Neste estudo, como será discutido mais adiante, utilizamos como lista de referência a nova Academic Vocabulary List (AVL), uma lista de vocabulário acadêmico em inglês, composta de 3.000 lemas que ocorrem em todos os domínios acadêmicos do COCA. O COCA consiste em um *corpus* com mais de 520 milhões de palavras, contendo textos de 1990 a 2015, escritos por falantes nativos e publicados nos Estados Unidos. Além disso, o *corpus* também possui textos orais, tais como entrevistas e programas de TV, embora não tenhamos acesso aos arquivos em áudio.

Começamos, pois, primeiramente, a identificar os verbos mais frequentes no subcorpus do CorIFA, através da geração da *word list* (lista de palavras), com o

AntConc. Feito isso, verificamos quais palavras podem funcionar como verbos, listamos, contando o número de ocorrências de suas formas lematizadas. Encontramos um total de 426 verbos lexicais (ver Apêndice I). Em seguida, checamos quais desses verbos aparecem na nova AVL (GARDNER; DAVIES, 2013), lista que retrata os 500 lemas mais frequentes do COCA. Como na AVL não há indicação da categoria gramatical à qual as palavras pertencem, contamos os lemas na lista, que se colocam na categoria “verbos” e encontramos um total de 116 verbos. Sendo assim, verificamos que dos 116 verbos que constam na AVL, 114 foram encontrados no subcorpus do CorIFA, contudo, apenas 28 deles foram mais frequentes, considerando o corte de frequência de 20 palavras por 100.000(*improve, provide, develop, increase, adopt, base, produce, relate, contribute, include, influence, achieve, depend, compare, facilitate, require, involve, exist, acquire, control, affect, observe, occur, apply, promote, reduce, conclude, tend*) e, em suas formas lematizadas, constam na nova AVL. Após esse procedimento, geramos as linhas de concordância com os lemas de cada verbo com vistas a obtermos sua frequência absoluta e a frequência normalizada.

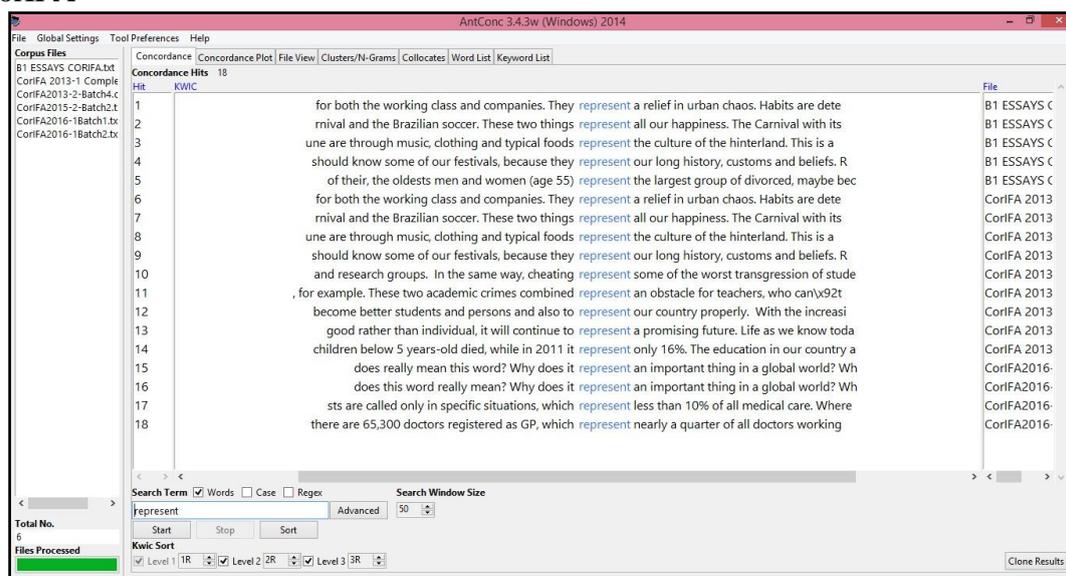
2.3.1 O programa AntConc 3.4.3: um dispositivo de extração de dados

Através do programa AntConc 3.4.3⁷⁵ é possível extrair dados a partir de suas sete ferramentas, a saber: Concordance; Concordance Plot; File View; Clusters/N-Grams; Collocates; Word List e Keyword List. A presente pesquisa utilizou-se da ferramenta Concordance para coletar as ocorrências com os verbos do discurso acadêmico escrito no CorIFA, bem como da ferramenta Collocates, para gerar a lista de colocados. Além disso, buscamos as colocações mais frequentes, utilizando a ferramenta N-Grams.

A título de exemplificação da utilização do *software*, demonstramos na Figura 2 que segue as linhas de concordância geradas, através da ferramenta Concordance, com o uso do lema *represent*, no CorIFA.

⁷⁵ Disponível gratuitamente em: <<http://www.antlab.sci.waseda.ac.jp/software.html>>. Acesso em: 25 maio 2017.

Figura 2– Tela do AntConc 3.4.3 com as linhas de concordância do verbo *represent* no CorIFA



Fonte: AntConc, ANTHONY, 2011.

A palavra de busca, também chamada de “nódulo”, nesse exemplo, aparece na cor azul, no centro da linha de concordância e, por essa razão, pode ser visualizada e identificada mais facilmente. Na presente investigação, a ferramenta Concordance foi utilizada para gerar as linhas de concordância com cada um dos verbos identificados no *corpus* como sendo os mais frequentes. Buscamos todos os lemas de todos esses verbos e, no caso de verbos irregulares, sua forma no passado e particípio passado foram também verificadas.

Distribuição dos verbos lexicais mais frequentes no subcorpus do CorIFA

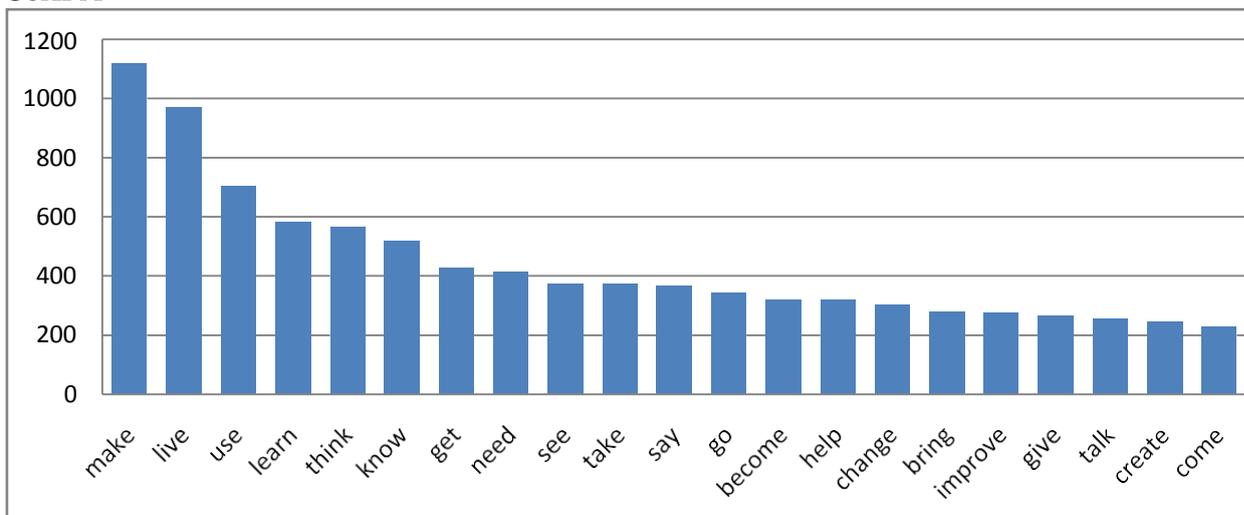
Para gerarmos as linhas de concordância do subcorpus do CorIFA, utilizamos o AntConc. Identificamos, assim, 12.761 *types* e 253.659 *tokens*. Os *types* são referentes às formas das palavras que ocorrem no *corpus*, enquanto que os *tokens* se referem às ocorrências das palavras no *corpus*.

Convém ressaltar que escolhemos para a presente pesquisa a classificação proposta por Biber et al. (1999), na qual os verbos são divididos em: verbos lexicais, verbos primários e verbos modais. Os verbos lexicais, por sua vez, são subdivididos em: verbos de ação, verbos de comunicação, verbos causativos, verbos de simples ocorrência, verbos de existência ou relacionamento e verbos aspectuais, conforme já apresentamos no Capítulo 1. A justificativa para tal escolha se dá por conta da

apresentação de uma classificação de verbos relacionada a um domínio semântico, que abrange várias funções.

No quadro (ver Apêndice I) é possível ver a distribuição de todos os 426 verbos lexicais, constituindo-se os tipos de verbos mais comuns encontrados no subcorpus do CorIFA, com frequência acima de duas ocorrências, incluindo os de natureza polissêmica que foram manualmente contados, a fim de se observar seu uso. A coluna 1 mostra os verbos em frequência decrescente, gerada a partir da lista de palavras; a coluna 2 mostra a classe verbal; a coluna 3 mostra o domínio semântico e a coluna 4 nos fornece o número de ocorrências dos verbos em todas as suas formas lematizadas. Para a obtenção desses números, utilizamos o etiquetador Constituent Likelihood Automatic Word-tagging System (CLAWS 7) (GARDSIDE; SMITH, 1997) e fizemos a contagem das ocorrências dos verbos mais frequentes, a partir do qual listamos no Gráfico 1 abaixo:

Gráfico 1– Número de ocorrências dos verbos lexicais mais frequentes no subcorpus do CorIFA



Fonte: Elaborado pela autora, 2017.

Notamos que um dos verbos lexicais de maior frequência encontrado na lista de palavras é de natureza cognata: *use* (178 ocorrências). Por conta do caráter polissêmico de muitas dessas palavras encontradas no subcorpus do CorIFA (*live*, *need*, *help*, *change*, *talk*), verificamos todas as suas ocorrências através da geração das linhas de concordância e consideramos aqui apenas as ocorrências nas quais funcionam como verbos.

Após gerarmos essa lista de verbos, verificamos que 114 deles estão presentes na AVL, apresentando-se com número de ocorrências que varia entre 01 e 277. Através da verificação de todas as linhas de concordância com tais palavras no *corpus*, contamos o número de ocorrências nas quais funcionam como verbos nas suas formas lematizadas.

Para obtermos os dados concernentes aos VIAES mais frequentes identificados no subcorpus do CorIFA, procedemos às seguintes etapas: geração da lista de palavras para a contagem do número de ocorrências dos verbos. Em seguida, obtivemos sua frequência normalizada, ação essencial quando se trabalha com *corpora* de diferentes tamanhos (BIBER; CONRAD; REPPEN, 1998). Assim, acreditamos ser mais apropriado converter cada frequência em um valor por 100 mil palavras; então, dividimos o número de ocorrências de cada palavra pelo número de palavras do *corpus* (subcorpus do CorIFA = 253.659 palavras) e, depois, multiplicamos por 100.000, considerando o mesmo ponto de corte de frequência de 20 palavras por 100.000 (CORTES, 2008). Sendo assim, os verbos do subcorpus do CorIFA a serem tomados como objeto de estudo, nesta pesquisa, são aqueles cuja frequência normalizada possui valor igual ou maior que 20 e serão mostrados no Capítulo 3.

A seguir, serão apresentados o *corpus* que originou a compilação do nosso subcorpus de referência, os instrumentos para tratar os dados e os procedimentos de análise utilizados para atender aos objetivos do presente estudo.

O British Academic Written English Corpus (BAWE)

O British Academic Written English Corpus (BAWE) consiste em um *corpus* formado por 2.761 textos, coletados por um grupo de pesquisadores, como parte de um projeto intitulado *An Investigation of Genres of Assessed Writing in British Higher Education*, financiado pelo Conselho de Pesquisas Econômicas e Sociais, do Reino Unido, entre os anos 2004 e 2007.

Os textos são provenientes de alunos universitários (da graduação e do mestrado), sendo que 30% deles, como já mencionamos, são de alunos cuja L1 não é o inglês, que foram avaliados como proficientes (notas de, no mínimo, 60%), de quatro universidades na Inglaterra: Oxford, Brookes, Reading e Warwick. Segundo Nesi e Gardner (2012), pesquisadoras envolvidas na compilação do *corpus*, a escolha de tais textos se justifica por estarem em conformidade com as expectativas dos seus respectivos departamentos institucionais, já que apresentam diferentes sistema de notas.

Um dos principais objetivos da compilação do BAWese centra na comparação da escrita acadêmica nas diferentes áreas, a partir da identificação de uma classe de eventos comunicativos.

Na Figura 3, encontramos a interface do BAWE:

Figura 3– Interface do BAWE



Fonte: Disponível em: <<http://www.sketchengine.com>>. Acesso em: 5 maio 2017.

O *corpus* contém 6.968.089 palavras, anotadas por parte do discurso (POS⁷⁶) e categoria semântica, a partir do CLAWS 7. Os textos foram classificados em 13 diferentes famílias de gêneros⁷⁷ (NESI;GARDNER, 2012), conforme podemos ver no Quadro 4abaixo:

Quadro 4– Número de textos em cada família de gênero no BAWE

Famílias de Gêneros	Número de textos
Estudo de caso	194
Crítica	322
Especificação de projeto (da concepção)	93
Escrita de compreensão ⁷⁸	35
Redação argumentativa	1.238

⁷⁶ Em inglês, a abreviação de *parts-of-speech*, ou seja, partes do discurso. Refere-se a um tipo de anotação, como forma de distinção entre as palavras, por exemplo, que tem a mesma escrita, porém significados e pronúncias diferentes (LEECH, 2004).

⁷⁷ As 13 famílias de gêneros identificadas no BAWE e ordenadas, sequencialmente, de acordo com o Quadro 8, são: *case study*, *critique*, *design specification*, *empathy writing*, *essay*, *exercise*, *explanation*, *literature survey*, *methodology recount*, *narrative recount*, *problem question*, *proposal* e *research proposal* (GARDNER; NESI, 2012, tradução nossa).

⁷⁸ Nossa tradução para a família de gênero, originalmente, intitulada *Empathy Writing* que, diz respeito às atividades elaboradas e direcionadas ao público não acadêmico com a função de tornar compreensível determinado conhecimento.

Exercício	114
Explicação	214
Revisão de literatura	35
Metodologia	362
Texto narrativo	75
Questão-problema	40
Proposta	76
Relato de pesquisa	61
Total	2.859

Fonte: Elaborado pela autora, 2017.

Mesmo não sendo esse nosso objetivo aqui, acreditamos ser importante descrever o modelo teórico-analítico utilizado para a classificação das famílias de gênero. Conforme Nesi e Gardner (2012), a classificação em família de gêneros se pauta, em grande parte, na Escola de Sydney, que bebe dos fundamentos da Linguística Sistêmico-Funcional (LSF), na medida em que busca identificar gêneros nos textos de estudantes do ensino fundamental e médio. Além disso, essa classificação recebe influência da concepção de gêneros acadêmicos trazida à tona por Swales (1990), bem como dos estudos em torno da abordagem dos Letramentos Acadêmicos, proposta por Lea e Street (1998). Nesi e Gardner (2012) utilizaram também o modelo de análise multidimensional elaborado por Biber (1998) para auxiliar na caracterização das famílias de gênero. Sobre isso, as autoras comentam que 67 características linguísticas foram etiquetadas no *corpus* e agrupadas em 16 categorias gramaticais, dentre as quais podemos citar: marcadores de tempo e aspecto, advérbios de lugar e tempo, pronomes, formas nominais, formas passivas, verbos modais e classes verbais especializadas.

Os textos que formam as 13 famílias de gênero foram coletados de 35 disciplinas, pertencentes a quatro grandes áreas do conhecimento (Artes e Humanidades, Ciências Sociais, Ciências Naturais e Ciências Físicas), a partir das quais as autoras encontraram diferenças que refletem as especificidades contextuais de cada área. Para esta pesquisa, nos interessamos redações argumentativas cujo número de textos constitui 1.238, conforme pode ser visto na distribuição por área, estabelecida da seguinte forma:

Quadro 5– Distribuição de redações argumentativas por área

	Artes e Humanidades	Ciências Naturais	Ciências Físicas	Ciências Sociais	Total
<i>Essay</i>	602	127	65	444	1.238

Fonte: Adaptado do BAWE corpus Manual(Heuboeck; Holmes; Nesi, 2008).

Vemos que as áreas que lideram em número de redações argumentativas no BAWE são Artes e Humanidades e as Ciências Sociais, o que demonstra que, nessas áreas, os estudantes tendem a utilizar os recursos linguísticos para sustentar sua argumentação, já que esta é uma característica típica dessas áreas. Por outro lado, as Ciências Físicas apresentam o menor número, apenas 65, o que parece convergir com o fato de, na área de exatas, ser mais comum os estudantes escreverem textos pertencentes a outras famílias de gênero.

Para Nesi e Gardner (2012, p.94, tradução nossa)⁷⁹, as redações argumentativas são gêneros de cunho avaliativo que têm como finalidade “[...] demonstrar ou desenvolver a habilidade de raciocinar independentemente e construir um argumento coerente”. No BAWE, as redações argumentativas são advindas de todas as áreas, com maior predominância na área de Artes e Humanidades e nas Ciências Sociais, o que parece justificar o comentário das autoras de que nos cursos de graduação de História, Inglês e Filosofia, os estudantes são, em sua maior parte, avaliados através desse gênero. As autoras, ao identificarem diferentes movimentos retóricos nas redações argumentativas que compõem o *corpus*, aos quais Swales (1990) intitula de *rhetorical moves*, subdividem as redações argumentativas em seis subgêneros, a saber: exposição, discussão, desafio, fatorial, consequencial e comentário. Todos eles têm em comum o mesmo objetivo, ou seja, desenvolver um argumento, contudo, se diferem apenas no modo como os argumentos são configurados.

Outro aspecto apontado por Nesi e Gardner (2012), ao analisarem as redações argumentativas, diz respeito à sua estrutura. Elas perceberam que muitas são organizadas em três parágrafos, o que parece ser característica da tradição de escrita com base em cinco parágrafos (introdução, três argumentos diferentes e conclusão), conforme dita a abordagem dos Estudos de Composição, já discutida no capítulo anterior.

⁷⁹No original: “Essays is to demonstrate/develop the ability to reason independently and construct a coherent argument”(NESI; GARDNER, 2012, p. 94).

Como veremos mais adiante na análise, os usos linguísticos que encontramos nas redações argumentativas do BAWE podem auxiliar os estudantes menos proficientes à compreensão de padrões recorrentes na escrita acadêmica, ao tempo em que possibilitam aos pesquisadores o acesso ao uso de vocabulário e padrões léxico-gramaticais dos textos dos estudantes, apesar de passados pelo crivo avaliativo.

A AVL e o BAWE

A Academic Vocabulary List (AVL) (GARDNER; DAVIES, 2013), lista de vocabulário acadêmico, que constitui nossa lista de referência neste estudo, nos fornece as 500 palavras mais frequentes da seção acadêmica do Contemporary Corpus of American English (COCA), *corpus* disponível *on-line* e de acesso gratuito, composto por mais de 520 milhões de palavras. Os 185 mil textos nele disponibilizados são de natureza diversa: oral, ficcional, textos de revistas, jornais e periódicos acadêmicos.

O interesse desse estudo pela investigação do uso de verbos em gêneros acadêmicos, a fim de identificá-los na AVL se deu por conta da lista ser baseada no COCA. A seção acadêmica do COCA contém 103.421.981 palavras *corpus* que possui 103.421.981 palavras, o que julgamos ser um número bastante representativo.

Por outro lado, nossa escolha pelo BAWE, a fim de compararmos os resultados de uso dos verbos do inglês acadêmico escrito do subcorpus do CorIFA, foi motivado por quatro principais razões:

- A plataforma de busca do BAWE nos permite buscar usos de determinados itens lexicais, em diferentes registros;
- As ferramentas disponíveis na plataforma possibilitam a extração de dados importantes para o estudo e compreensão de certos fenômenos linguísticos;
- O BAWE possui um número bastante representativo de 6.968.089 palavras, organizadas em famílias de gêneros, a partir das quais pudemos selecionar apenas as redações argumentativas, com número de palavras, de igual modo, representativo, perfazendo um total de 1.238 textos, o que possibilitou a construção de um subcorpus, a fim de tornar nossa comparação mais equilibrada;

Para compilarmos o nosso subcorpus, composto pelas redações argumentativas do BAWE, adotamos alguns procedimentos que serão descritos aqui.

Uma vez que, no sítio eletrônico em que o BAWE está hospedado, havia 10 redações argumentativas em cada página, decidimos criar *scripts*⁸¹ utilizando o *software* estatístico R⁸², no intuito de obter o número de *tokens* necessários para o *corpus* por meio da amostragem dos textos disponíveis, que caracterizam o subcorpus analisado.

Como mencionado na subseção anterior, o número de palavras do BAWE (6.968.089 palavras) diverge consideravelmente do número de palavras do subcorpus do CorIFA (253.659 palavras) que analisamos neste estudo. Sendo assim, foi necessário balancear o *corpus*, considerando o dobro do número de palavras (CORTES, 2004) do CorIFA, isto é, aproximadamente 500 mil *tokens*. Em outras palavras, foi feita uma amostragem do BAWE, para que não houvesse discrepância quanto às análises empreendidas.

Além disso, apesar de o BAWE contemplar 13 famílias de gêneros, os textos não estão organizados por gênero. Dessa forma, ao iniciar o processo de compilação do subcorpus, foi necessário fazer a contagem do número mínimo e número máximo de *tokens* para servir como critério para a escolha dos textos. Para chegar à metodologia da nossa pesquisa, realizamos um teste piloto a partir do qual obtivemos os seguintes resultados:

Quadro 6 – Quadro descritivo do teste piloto para a contagem de *tokens*

Valor mínimo <i>tokens</i>– textos	335
Valor máximo <i>tokens</i>– textos	76907
Soma –<i>tokens</i> dos textos	127774
Soma –<i>tokens</i> selecionados	159295
Média –<i>tokens</i> por texto	29040
Desvio padrão	23460,35
Índice de confiança 95%	6931,96306
Intervalo de confiança	22107,67 – 35971

⁸¹ Os *scripts* usados nesta pesquisa foram criados por Araújo e Castro (2017). Os principais *scripts* que utilizamos estão disponibilizados em: https://github.com/rodrigoacastro/GUEDES-A.-S._2018.

⁸² O R é uma linguagem de programação de código livre capaz de compilar e executar grandes quantidades de dados. Disponibiliza uma grande variedade de técnicas estatísticas e gráficas e, por essa razão, é amplamente utilizado entre analistas de dados e pesquisadores de modo geral.

Mínimo	22107,67
Máximo	35971
Margem mínima	500
Margem máxima	5000

Fonte: Elaborado pela autora, 2017.

No teste piloto, adotamos o critério estatístico de cálculo de intervalo de confiança⁸³, uma vez que não há o mesmo número de *tokens* em todos os textos e isso, conseqüentemente, poderia influenciar na configuração e análise dos dados. Isso se deve ao fato de que é necessário um determinado número de *tokens* por texto para o balanceamento do *corpus*. Assim, entendemos que isso reforça a impossibilidade de usar uma amostragem aleatória, visto que há certas restrições nas escolhas dos textos, como o tipo de texto do *corpus* de estudo e os interesses da presente pesquisa.

Com base no teste piloto, foi adotado um dos critérios utilizados na compilação do BAWE, já que os textos dos estudantes que compõem o *corpus* possuem no mínimo, 500, e, no máximo, 5.000 palavras. Assim, utilizamos o critério de 500 a 5.000 *tokens*, com base na própria compilação do *corpus*-fonte, visto que alguns textos estavam fora dessa faixa.

O Quadro 7, a seguir, mostra o número de textos que compõem o nosso subcorpus do BAWE:

Quadro 7– Quadro descritivo para a contagem de *tokens*

Valor Mínimo <i>tokens</i>– textos	376
Valor Máximo <i>tokens</i>– textos	12426
Soma –<i>tokens</i> dos textos	574671
Soma –<i>tokens</i> selecionados	509007
Média –<i>tokens</i> por texto	2262
Desvio padrão	1338,42
Índice de confiança 95%	185,49
Intervalo de confiança	2076,45 - 2447,43
Mínimo	500
Máximo	5000
Margem mínima	500
Margem máxima	5000

⁸³ O intervalo de confiança é um critério estatístico que estima e calcula o grau de confiança correspondente a uma margem de erro.

Fonte: Elaborado pela autora, 2017.

A tabela descritiva do subcorpus compilado consta de 239 textos, que totalizaram 509.007 *tokens*.

Diante disso, seguimos as etapas:

1. Primeiro, todas as redações argumentativas do BAWE em formato .txt foram exportadas no R;
2. Foi realizada a limpeza dos textos retirando as etiquetas em parênteses angulares (*tags*);
3. No R, foi feita a amostragem de 260 textos no formato .txt, com base em *scripts* elaborados para esse propósito, dentre o total de 1.238 textos que perfazem o conjunto de redações argumentativas do BAWE;
4. Para cada texto, foi verificado o total de *tokens* utilizando o *software* AntConc 3.4.3;
5. Com base no número mínimo e máximo de *tokens* definidos (500 e 5.000, respectivamente), foram selecionados os textos necessários para chegarmos a aproximadamente 500.000, alcançando o total de 509.007, que corresponde à soma de *tokens* selecionados para compor o subcorpus, a partir dos quais foram separados os textos da análise;
6. Compilado o subcorpus, assim como realizado no estudo piloto, a fim de compor a descrição do *corpus*, foram calculadas medidas estatísticas, como a média de *tokens* por texto (compilados em sua integralidade), cujo valor é 2.262;
7. Por fim, os textos foram importados para o AntConc para a continuidade da análise.

O tamanho final do subcorpus de análise foi de 509.007 (de palavras), tendo em vista que esse número equivale a duas vezes o tamanho do subcorpus do CorIFA.

A lista de referência

Instrumentos

Como já mencionado previamente, a manipulação de dados deste estudo será

realizada com o AntConc 3.4.3, uma vez que, através dele, é possível verificar a frequência dos verbos do inglês acadêmico escrito e suas colocações mais frequentes, por meio da análise das linhas de concordância.

Procedimentos de análise

Como comentamos inicialmente neste capítulo, um dos objetivos desta pesquisa consiste em identificar e analisar os verbos do inglês acadêmico escrito (VIAEs) e seu uso no subcorpus do CorIFA, contrastando com os usos no BAWE, mais especificamente, com as colocações formadas por advérbios *-ly* + verbo. Assim, análises quantitativas e qualitativas serão realizadas no subcorpus de estudo, bem como no subcorpus de referência, a fim de identificarmos as diferenças e semelhanças no uso e nos padrões de uso do inglês, presentes na escrita acadêmica dos estudantes universitários.

Para tanto, alguns critérios serão considerados, a começar pela necessidade de se fazer um ponto de corte (CORTES, 2008) de frequência do tipo conservador para a delimitação dos VIAEs a serem analisados e comparados aos disponíveis na nova AVL (GARDNER; DAVIES, 2013), já que a partir desse recorte, reconhecemos a viabilidade e exequidade da investigação e alcance dos objetivos pretendidos. A opção pelo ponto de corte se justifica pela necessidade de se gerar uma lista de VIAEs usando as ferramentas de *corpus*, com base na frequência, para determinar o número a ser incluído na análise. Com base em Cortes (2008), decidimos analisar os VIAEs encontrados na frequência normalizada de 20 vezes por 100 mil palavras.

Desse modo, com base nos propósitos delineados para este estudo, a análise fará o seguinte percurso: no primeiro momento, identificação e análise dos VIAEs; no segundo momento, os VIAEs obtidos serão comparados aos disponíveis na AVL; em seguida, serão identificadas e analisadas as colocações formadas por advérbios *-ly*+ verbo no subcorpus do CorIFA, e serão comparadas às colocações constantes no subcorpus do BAWE. A análise será feita entendendo que a comparação entre essas combinações de palavras também podem revelar características da interlíngua dos aprendizes. Assim, o cálculo aqui utilizado para o ponto de corte de frequência pode ser definido pela fórmula “tamanho do *corpus*/100.000”(CORTES, 2008).

A partir dessas decisões metodológicas, buscamos analisar o uso que os estudantes universitários de Inglês para Fins Acadêmicos fazem desses verbos nas suas

produções escritas, bem como das colocações do tipo “advérbio *-ly*+ verbo” presentes nesses textos.

A fim de identificar tais colocações, com o AntConc 3.4.3, geramos as linhas de concordância com todos os lemas dos VIAEs. Para tanto, observamos os advérbios que colocam os VIAEs (nódulo) em até duas posições à esquerda e até duas posições à direita. O nosso objetivo, portanto, se concentra em fazer uma descrição do uso dos VIAEs e o comportamento colocacional com os advérbios de sufixo *-ly*, verificando, desse modo, seu uso fraseológico.

Como utilizamos o etiquetador CLAWS 7, de início, ficamos em dúvida sobre quais *tags* (etiquetas) usar, a fim de listarmos os advérbios terminados com o sufixo *-ly*, pois há 15 delas disponíveis no etiquetador e a etiqueta que poderia ser útil aos nossos objetivos, a *RR*, serve para os advérbios de modo geral. Para identificarmos os advérbios *-ly* do subcorpus do CorIFA, utilizando o R, geramos quatro listas com cada um dos arquivos para buscar os termos terminados em *-ly* e sua lista de frequência. Para eliminar os termos das listas que não são advérbios (ex: *early*, *family* e *likely*), mas apresentam essas terminações, utilizamos uma lista de exclusão ou *stopwords* e, assim, fizemos a limpeza dos arquivos e identificamos apenas os advérbios *-ly*.

Com base nos postulados teórico-metodológicos da Análise Contrastiva de Interlíngua, realizaremos uma contraposição do inglês com a interlíngua dos estudantes a partir do subcorpus do CorIFA e procederemos à aplicação de um teste estatístico.

A partir do exposto no capítulo, acreditamos que pode haver pouco uso dos verbos do inglês acadêmico escrito no nosso *corpus* de aprendiz, cuja distribuição de frequência foi apresentada neste capítulo, talvez em decorrência da dificuldade dos aprendizes na utilização de vocabulário recorrente do discurso acadêmico. Além disso, esse fato corrobora com os resultados de pesquisas que nos informam sobre os padrões de uso linguístico encontrados em *corpora* de aprendizes de inglês como língua adicional, como o conduzido por Egbert et al (2016):

Resultados de pesquisa com *corpus* informam nossa compreensão de como palavras e características gramaticais são realmente usadas. Esses padrões de uso da língua geralmente, não são encontrados em livros didáticos tradicionais, então os achados das pesquisas com corpus podem desempenhar papel importante no ensino de línguas (EGBERT et al., 2016, p.119, tradução nossa)⁸⁴.

⁸⁴ No original: “Findings from corpus research inform our understanding of how words and grammatical features are actually used. These patterns of language use are usually not addressed in traditional language textbooks, so the findings from corpus research can play an important role in language teaching[...]” (EGBERT et al., 2016, p. 119).

Como já mencionado anteriormente, além da análise dos verbos do inglês acadêmico escrito, essa pesquisa se propõe a analisar o uso das colocações “advérbio - *ly*+ verbo” com cada um dos VIAEs, além de gerarmos sua lista de colocados. Através da ferramenta Collocates no AntConc, podemos obter a lista de colocados com cada verbo acadêmico investigado neste estudo. De posse disso, mostramos suas ocorrências, discutimos e analisamos o seu uso. Utilizamos como critério as duas palavras à esquerda e duas palavras à direita. Descrevemos a seguir sobre esses dois tipos de colocados (à esquerda e à direita) e mostramos as categorias que foram usadas para classificar as ocorrências.

São considerados colocados aqui, os itens lexicais que se colocam dentro de um dimensão de duas palavras à esquerda (posições L1 e L2) a duas palavras à direita (posições R1 e R2) do nóculo, com uma frequência mínima de duas ocorrências. Através da lista de colocados é possível verificar se algum item lexical se destaca em termos de frequência ou força de associação.

Descrição do teste estatístico para extração de dados: o *log-likelihood* (LL)

Os estudos linguísticos, mais particularmente, os estudos em Linguística de Corpus, têm se apossado amplamente nos testes estatísticos para a análise de dados de modo mais acurado, tendo em vista seu alto grau de comparabilidade, objetividade, replicabilidade e precisão (GRIES, 2013), levando-nos a entender que “[...] o uso de quantificação na linguística de *corpus* vai além da simples contagem” (McENERY; WILSON, 2001, p.81, tradução nossa)⁸⁵.

A simples quantificação da frequência de dados pode nos levar a conclusões equivocadas. Por conta disso, utilizamos técnicas estatísticas chamadas de “testes de significância estatística” para verificarmos se hipóteses sobre as diferenças observadas entre dados de frequência são provavelmente verdadeiras ou provavelmente falsas.

Na presente investigação, utilizamos o teste estatístico *log-likelihood*, por compreendermos que ele é adequado para atender aos objetivos que delineamos para este estudo, ou seja, pode nos fornecer dados importantes, a partir de *corpora* com número menor de palavras, como é o caso do CorIFA, que apresenta alto grau de

⁸⁵ No original: “The use of quantification in corpus linguistics typically goes well beyond simple counting” (McENERY; WILSON, 2001, p. 81).

confiabilidade e permite a comparação de significância entre fenômenos da língua (DUNNING, 1983).

O *log-likelihood* consiste em um texto paramétrico, que leva em consideração a frequência absoluta dos itens lexicais, utilizado para extrair palavras que são típicas de um determinado *corpus*, como também para mensurar o nível de significância estatística de uma palavra específica. O *log-likelihood*, assim, compara a frequência de itens lexicais em mais de um *corpus* e tem por objetivo validar o nível de significância estatística de uma palavra. Para que a frequência de determinado item lexical seja considerada estatisticamente significativa, o LL da palavra precisa ser maior que o valor de 3.84 ($p < 0.05$), ($p = \text{valor}$), como definido por McEnery e Hardie (2012), e que constitui o *cut-off point* estatístico que adotamos aqui.

O LL de algum item lexical é calculado tomando como base o que consta na Figura 8:

Figura 5 – Exemplo de distribuição de dados para obtenção do LL

	Corpus 1	Corpus 2	Total
Frequency of word	a	b	a+b
Frequency of other words	c-a	d-b	c+d-a-b
Total	c	d	c+d

Fonte: Disponível em: <<http://ucrel.lancs.ac.uk/llwizard.html>>.

O valor „c“ é referente ao número de palavras no *corpus* 1, ao passo que o valor „d“ corresponde ao número de palavras no *corpus* 2. Os valores „a“ e „b“ são denominados de “valores observados”.

A título de exemplificação, a Figura 2 abaixo, que consiste em um recorte do cálculo do LL elaborado por pesquisadores da Universidade de Lancaster e disponibilizada⁸⁶ *on-line* gratuitamente, nos mostra como os dados são distribuídos para a obtenção do valor de LL:

⁸⁶ Disponível em: <<http://ucrel.lancs.ac.uk/llwizard.html>>. Acesso em: 10 maio 2016.

Figura 6– Distribuição de dados para obtenção do LL

	Corpus 1	Corpus 2
Frequency of word	<input type="text"/>	<input type="text"/>
Corpus size	<input type="text"/>	<input type="text"/>
<input type="button" value="Calculate"/>		<input type="button" value="Clear form"/>

Fonte: Disponível em: <<http://ucrel.lancs.ac.uk/llwizard.html>>.

Para calcularmos o valor de LL de um item lexical, consideramos: na primeira coluna, a frequência da palavra (*frequency of word*) no *corpus* de estudo e, logo abaixo, o número de palavras desse *corpus* (*corpus size*); na coluna 2, registramos a frequência absoluta do item lexical no *corpus* de referência e, em seguida, o seu número de palavras.

Para obtermos, por exemplo, o valor de LL do verbo *acquire*, no subcorpus do CorIFA, procedemos da seguinte maneira:

Figura 7– Distribuição de dados para obtenção do LL – *acquire*

	Corpus 1	Corpus 2
Frequency of word	73	92
Corpus size	253659	509007
<input type="button" value="Calculate"/>		<input type="button" value="Clear form"/>

Fonte: Disponível em: <<http://ucrel.lancs.ac.uk/llwizard.html>>.

Na Figura 7, a coluna referente ao *corpus* 1 corresponde à frequência absoluta (73) do verbo *acquire*, no subcorpus do CorIFA; mais abaixo, é registrado o número de palavras do *corpus* (253.659); na coluna referente ao *corpus* 2, o número 92 corresponde à frequência absoluta do mesmo item lexical no subcorpus do BAWE e, logo abaixo desse dado, encontramos o número de palavras do subcorpusdo BAWE, ou seja, 509.007 (número de palavras utilizado aqui para o cálculo do LL de cada verbo). Para chegarmos ao LL do verbo *acquire*, após o preenchimento do quadro, basta inserir as informações que precisamos (apenas números) e depois clicar em *calculate*.

O resultado de LL desse verbo, bem como de todos os outros verbos e o resultados das combinações dos verbos do inglês acadêmico escrito com os advérbios -*ly*, a serem analisados nessa pesquisa, estão dispostos no capítulo a seguir.

CAPÍTULO 3

RESULTADOS E ANÁLISE DE DADOS

No descomeço era o verbo.
 Só depois é que veio o delírio do verbo.
 O delírio do verbo estava no começo, lá onde a
 criança diz: Eu escuto a cor dos passarinhos.
 A criança não sabe que o verbo escutar não funciona
 para cor, mas para som.
 Então se a criança muda a função de um verbo, ele
 delira.
 E pois.
 Em poesia que é voz de poeta, que é a voz de fazer
 nascimentos –
 O verbo tem que pegar delírio.

(*Manoel de Barros, 2011*)

Introdução

Neste capítulo, apresentamos os resultados da pesquisa quanto à frequência e uso dos verbos do inglês acadêmico escrito, encontrados no subcorpus do CorIFA e disponíveis na AVL. Para tanto, inicialmente, apresentamos um quadro no qual constam todos os verbos, de modo geral, acompanhados de sua frequência e classificação quanto ao domínio semântico, à luz de Biber et al. (1999). Em seguida, apresentamos o quadro com os verbos do inglês acadêmico escrito mais frequentes no subcorpus do CorIFA, segundo o *cut-off point* de 20 X 100.000 palavras, considerando sua frequência normalizada. De posse disso, identificamos sua classificação semântica e descrevemos o comportamento dos VIAEs a serem analisados na pesquisa. Ainda, com o intuito de identificarmos quais verbos apresentam diferenças significativas de frequência entre o subcorpus do CorIFA e o subcorpus do BAWE, aplicamos o teste estatístico *log-likelihood* (LL). Por fim, trazemos as considerações finais do capítulo e as limitações deste estudo.

Resultados gerais

Aqui apresentamos os resultados referentes à frequência normalizada dos verbos do inglês acadêmico escrito mais frequentes no subcorpus do CorIFA e presentes na Academic Vocabulary List (AVL). Em seguida, abordamos sobre o domínio semântico

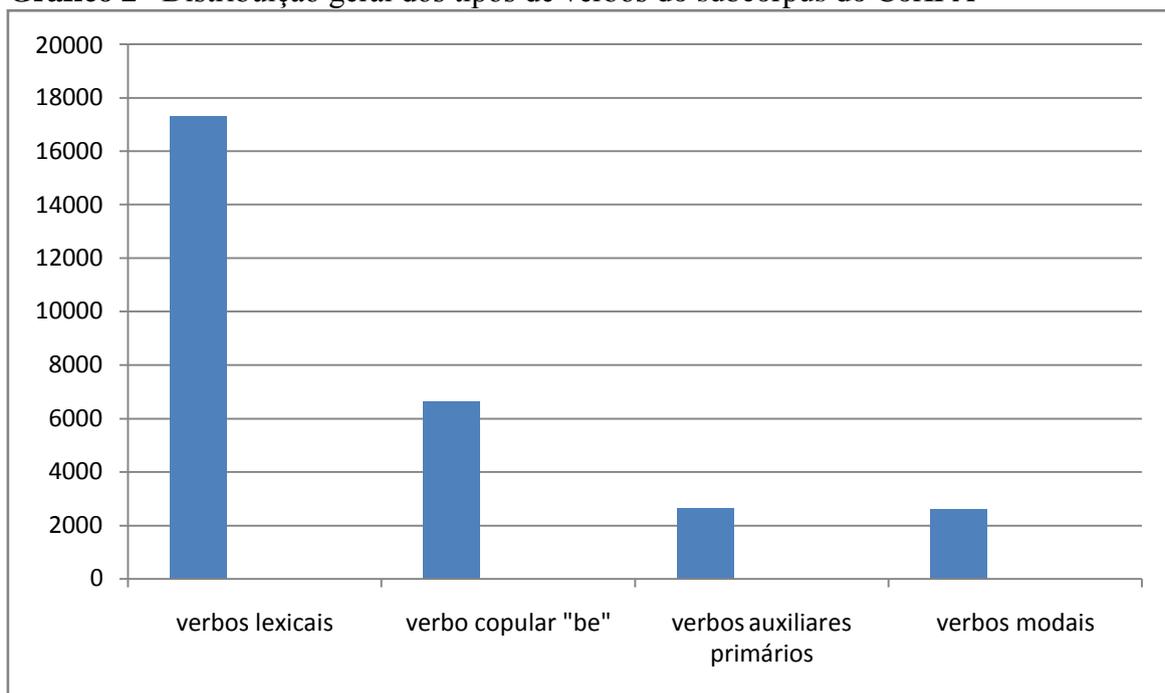
desses verbos, discutindo e contrastando os padrões de uso encontrados no subcorpus do CorIFA e no subcorpus do BAWE.

Frequência geral

Como mencionado no capítulo anterior, o critério utilizado para a escolha dos verbos do inglês acadêmico escrito para análise, nesta pesquisa, consistiu na geração da lista de verbos mais frequentes do subcorpus do CorIFA, com ponto de corte de frequência de 20 X 100.000 palavras e posterior verificação na nova Academic Vocabulary List(AVL), na qual Gardner e Davies (2013) elencaram as 500 palavras mais frequentes dos nove domínios acadêmicos do COCA. Entendemos que, como na AVL, Gardner e Davies (2013) não identificam as POS, tais como, substantivos, verbos, adjetivos, advérbios, ou seja, não há uma especificação da classe gramatical da palavra. Consideramos as entradas lematizadas e fizemos a contagem dos verbos considerando todos os seus lemas, o que nos oportuniza uma avaliação mais acurada da forma, das funções e dos significados das palavras (GARDNER; DAVIES, 2013).

Encontramos 426 verbos lexicais no subcorpus do CorIFA, o que julgamos ser uma alta incidência que, segundo Biber et al. (1999), pode ser ocasionada por conta de serem verbos principais que ocorrem sem o uso de verbos auxiliares nas suas formas mais simples. Podemos perceber, a partir do quadro do Apêndice I, que a forma lematizada “verbo + *ing*” tem menos ocorrências em relação aos outros lemas, exceto em *cheating* (135 ocorrências) e *looking* (50 ocorrências). Parece-nos que a forma lematizada que mais se destaca é a forma infinitiva e o presente simples, precedidas de primeira ou segunda pessoa do singular e terceira pessoa do plural. Outro aspecto observado nos dados do subcorpus diz respeito ao pouco uso da forma lematizada do particípio passado dos verbos, fenômeno que corrobora com o estudo de Chafe e Danielewicz (1987) em que, apenas poucos verbos, tais como *acquire*, *allow*, *cause* e *connect* apresentam uso mais recorrente da forma do particípio passado se comparado aos outros lemas, o que pode ser decorrente do uso frequente de construções passivas.

A partir da necessidade de melhor caracterizar os tipos de verbos gerados pela *wordlist* do CorIFA, apresentamos o Gráfico 2, abaixo:

Gráfico 2– Distribuição geral dos tipos de verbos do subcorpus do CorIFA

Fonte: Elaborado pela autora, 2017.

O Gráfico 2 mostra a distribuição dos verbos do subcorpus do CorIFA. Os verbos lexicais são os mais frequentes e, a partir deles, extraímos os verbos do inglês acadêmico escrito que serão analisados neste estudo. Em virtude disso, na subseção a seguir, disponibilizamos a distribuição desses verbos lexicais conforme seu domínio semântico.

Distribuição de domínios semânticos no CorIFA e no BAWE

O Quadro 8 fornece o número total de verbos lexicais (426) constantes no subcorpus do CorIFA e no subcorpus do BAWE (990), distribuídos em sete domínios semânticos:

Quadro 8– Frequência normalizada de verbos lexicais comuns de cada domínio semântico (sem as formas lematizadas) baseado em Biber et al. (1999)

Domínio semântico	Número de verbos lexicais no subcorpus do CorIFA
Ação (e.g <i>open, buy</i>)	357
Mental (e.g <i>think, know</i>)	26

Comunicação (e.g <i>say, write</i>)	13
Existência (e.g <i>stay, exist</i>) ou relacionamento (e.g <i>involve, include</i>)	10
Simple ocorrência (e.g <i>become, develop</i>)	7
Facilitação ou Causa (e.g <i>enable, require</i>)	7
Aspectual (e.g <i>start, continue</i>)	6

Fonte: Elaborado pela autora, 2017.

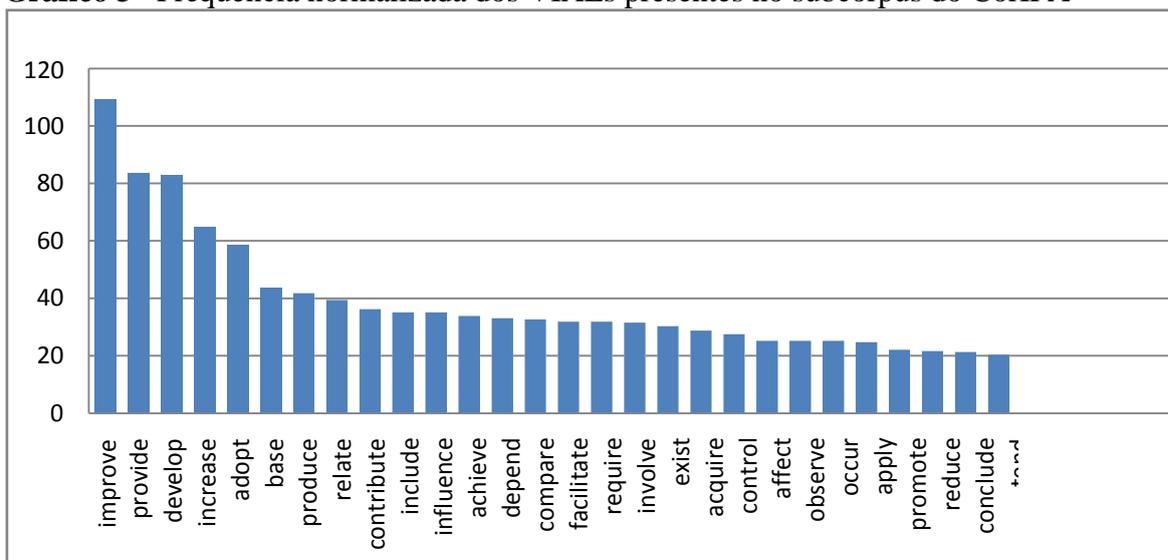
Os verbos mais comuns no subcorpus do CorIFA são os verbos de ação, com número de ocorrências bem maior que a soma de todos os verbos constantes nos outros domínios. Por outro lado, os verbos mentais, de comunicação, de existência, de ocorrência, causa e de aspecto têm frequência muito baixa.

Para Biber et al. (1999), a gama de verbos encontrada no seu *corpus* de estudo “[...] reflete os tipos de tópicos que os falantes e escritores mais comumente discutem” (BIBER et al., 1999, p.365, tradução nossa)⁸⁷. Para os autores, por exemplo, verbos como *ask, tell, write speak*, considerados “verbos de comunicação”, são usados para retratar atividades executadas, ao tempo em que verbos como *find, consider, know, need*, dentre outros, além de expressar ações, se referem também a sentimentos abstratos, ideias e opiniões. Mais especificamente no registro acadêmico, os verbos de existência são muito comuns, assim como os verbos de ação. Além disso, a referência a ações de naturezas físicas, mentais e de comunicação são pouco usadas nesse registro, o que acaba implicando na maior frequência de verbos causativos e de ocorrência em relação aos outros registros. A variedade de verbos que encontramos no subcorpus do CorIFA, dessa forma, pode ser originária dos tipos de tarefa que os aprendizes realizaram, como já comentamos na subseção 2.2.1.

Frequência normalizada dos VIAEs mais frequentes no subcorpus do CorIFA e do BAWÉ

Os VIAEs, a serem tomados como objeto de estudo, estão representados quanto à sua frequência normalizada no Gráfico 3 abaixo:

⁸⁷ No original: “[...] reflects the kinds of topics that speakers and writers most commonly discuss” (BIBER et al., 1999, p. 365).

Gráfico 3– Frequência normalizada dos VIAEs presentes no subcorpus do CorIFA

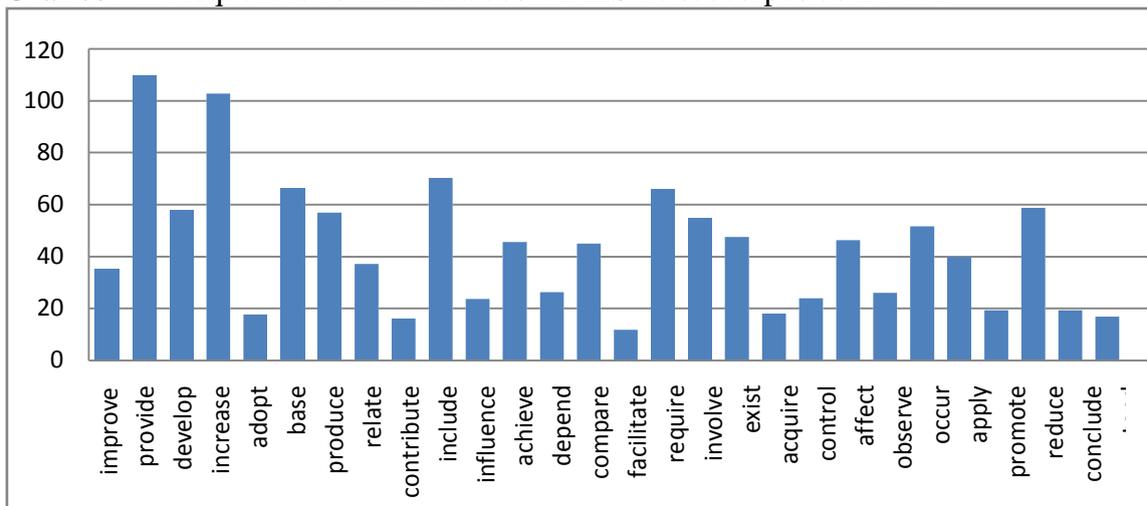
Fonte:Elaborado pela autora, 2017.

Segundo Biber et al. (1999, p.8, tradução nossa)⁸⁸, “[...] um dos mais importantes usos de uma investigação baseada em *corpus* é fornecer informações sobre sua frequência de uso”.

Os verbos do inglês acadêmico escrito de maior frequência normalizada no subcorpus do CorIFA foram: *improve* (109,2), *provide* (83,57), *develop* (82,78), classificados quanto ao domínio semântico, como verbo de facilitação/causa, mental e de simples ocorrência, respectivamente.

Os mesmos procedimentos adotados para o cálculo da frequência normalizada dos verbos do inglês acadêmico escrito no subcorpus do CorIFA foram também utilizados para obtermos a frequência normalizada dos VIAEs no subcorpus do BAWE. Dessa maneira, calculamos: (frequência absoluta (número de ocorrências) da palavra/509.007 X 100.000). Fizemos o seguinte cálculo: (frequência absoluta (número de ocorrências da palavra)/509.007 X 100.000). Os resultados estão dispostos no Gráfico 4 a seguir:

⁸⁸ No original: “[...] one of the most important uses of corpus-based investigation is to provide information about frequency of use” (BIBER et al.,1999, p. 8).

Gráfico 4 – Frequência normalizada dos VIAES no subcorpus do BAWE

Fonte: Elaborado pela autora, 2017.

Dentre os verbos do inglês acadêmico escrito de maior frequência normalizada no subcorpus do BAWE, podemos destacar: *provide* (109,62) e *increase* (102,74), classificados quanto ao domínio semântico, como verbo mental e de simples ocorrência, respectivamente. Por outro lado, sete VIAEs apresentam frequência normalizada abaixo de 20 ocorrências: *acquire* (18,07), *adopt* (17,68), *conclude* (19,25), *contribute* (16,1), *facilitate* (11,78), *promote* (19,25) e *tend* (16,89).

O Quadro 9 nos mostra os resultados da frequência absoluta e normalizada dos VIAEs mais frequentes em ambos os subcorpora, obtidos após a geração de WordList no AntConc:

Quadro 9– Frequência absoluta e frequência normalizada dos verbos do inglês acadêmico escrito mais frequentes no subcorpus do CorIFA e do BAWE

VIAE	Frequência absoluta subcorpus do CorIFA	Frequência normalizada (por 100.000 palavras)	VIAE	Frequência absoluta subcorpus do BAWE	Frequência normalizada (por 100.000 palavras)
<i>improve</i>	277	109,2	<i>improve</i>	179	35,16
<i>provide</i>	212	83,57	<i>provide</i>	558	109,62
<i>develop</i>	210	82,78	<i>develop</i>	295	57,95
<i>increase</i>	164	64,65	<i>increase</i>	523	102,74
<i>adopt</i>	148	58,34	<i>adopt</i>	90	17,68
<i>base</i>	111	43,75	<i>base</i>	337	66,2
<i>produce</i>	106	41,78	<i>produce</i>	290	56,97
<i>relate</i>	100	39,42	<i>relate</i>	189	37,13
<i>contribute</i>	92	36,26	<i>contribute</i>	85	16,1
<i>include</i>	89	35,08	<i>include</i>	358	70,33

<i>influence</i>	89	35,08
<i>achieve</i>	86	33,9
<i>depend</i>	84	33,11
<i>compare</i>	83	32,72
<i>facilitate</i>	81	31,93
<i>require</i>	81	31,93
<i>involve</i>	80	31,53
<i>exist</i>	77	30,35
<i>acquire</i>	73	28,77
<i>control</i>	70	27,59
<i>affect</i>	64	25,23
<i>observe</i>	64	25,23
<i>occur</i>	64	25,23
<i>apply</i>	63	24,83
<i>promote</i>	56	22,07
<i>reduce</i>	55	21,68
<i>conclude</i>	54	21,28
<i>tend</i>	52	20,49

<i>influence</i>	121	23,77
<i>achieve</i>	232	45,57
<i>depend</i>	134	26,32
<i>compare</i>	229	44,98
<i>facilitate</i>	60	11,78
<i>require</i>	336	66,01
<i>involve</i>	279	54,81
<i>exist</i>	242	47,54
<i>acquire</i>	92	18,07
<i>control</i>	122	23,96
<i>affect</i>	236	46,36
<i>observe</i>	133	26,12
<i>occur</i>	263	51,66
<i>apply</i>	203	39,88
<i>promote</i>	98	19,25
<i>reduce</i>	299	58,74
<i>conclude</i>	98	19,25
<i>tend</i>	86	16,89

Fonte: Elaborado pela autora, 2017.

No quadro à esquerda estão listados os 28 VIAEs mais frequentes no subcorpus do CorIFA, considerando o ponto de corte de frequência aqui adotado (aqueles cuja frequência normalizada seja igual ou superior a 20). A partir dela, buscamos a frequência dos VIAEs no subcorpus do BAWE (quadro à direita). Os dados parecem indicar que são verbos amplamente usados por aprendizes de inglês, em suas produções escritas. Contudo, há a questão do tópico/assunto do texto que pode certamente influenciar no uso do vocabulário e, conseqüentemente, na escolha dos verbos, como já tínhamos apontado anteriormente. Ademais, verificamos que os aprendizes usam verbos de vários domínios semânticos nos seus textos escritos, embora com algumas diferenças, conforme mostraremos mais adiante.

Há diferenças quanto ao uso dos VIAEs entre os subcorpora, dentre as quais destacamos: a) a alta frequência de *improve* no subcorpus do CorIFA (109,2) em relação ao subcorpus do BAWE (35,16); b) a alta frequência de *develop* (82,78) no subcorpus do CorIFA comparado ao subcorpus do BAWE (57,95); c) a alta frequência de *adopt* (58,34) no subcorpus do CorIFA em oposição ao subcorpus do BAWE (17,68); d) a frequência de *increase* no subcorpus do CorIFA quase 50% (64,65) menor quando comparado à frequência no subcorpus do BAWE (102,74); e) verbos como

include, occur, reduce e *require*, que apresentam número de ocorrências duas vezes maior no subcorpus do BAWE comparados ao subcorpus do CorIFA.

Características compartilhadas entre os subcorpora do CorIFA e do BAWE

Vemos que os verbos que lideram o *ranking* de frequência no subcorpus do CorIFA são: *improve, provide, develop, increase* e *adopt*. Comparando os resultados do subcorpus do BAWE com os resultados obtidos a partir do subcorpus do CorIFA, percebemos que há três VIAES considerados mais frequentes em ambos os subcorpora: *develop, increase* e *provide*. Vejamos algumas linhas de concordância com os verbos tais verbos do subcorpus do CorIFA:

(1) *Every time we **develop** new technologies, we give to our creativity an entire new world of possibilities.*

(2) *Furthermore, the dance have **increased** its range of movements and have improved them as well.*

(3) ***Providing** scholarship requires a lot of investments and expenses from the government.*

No exemplo 1, o verbo *develop*, cujo lema representa o tempo verbal presente, é usado para se referir ao substantivo *technologies*, formando, assim, uma colocação verbo-nominal. Há 96 ocorrências desse lema, sendo que, em 54 delas, está precedido de partícula marcadora de infinitivo. Além desse aspecto, encontramos linhas de concordância em que o lema se combina aos verbos modais *can, must* e *will*.

Evidenciamos, no segundo exemplo, o lema *increased*, na condição de particípio passado, usado para compor a forma do presente perfeito, forma verbal que se apresenta também na segunda oração “*have improved them as well*”. Em todo o subcorpus do CorIFA, há 36 ocorrências do lema *increased*, porém apenas sete ocorrências nas quais é usado na referida forma verbal, restando para as demais ocorrências a função de verbo no passado.

O terceiro exemplo nos mostra o lema *providing* sendo usado no início da oração, em uma frase com *-ing* na posição de sujeito. Das 18 ocorrências do lema, quatro são precedidas da preposição *by* e apenas um compõe a forma verbal do presente contínuo: “*It has improved as time is passing and it is **providing** to us a new way of life[...]*”.

Olhando mais especificamente para os lemas do verbo *develop*, no subcorpus do CorIFA, encontramos:

- (1) *When we go abroad we **develop** a different point of view about the world and people.*
- (2) *The new thinking is what **develops** our world, with the research, new projects, and this can happen only when we can dream and imagine.*
- (3) *Thirdly, for a kid that is **developing** its mind, learn a foreign language is a big stimulation.*
- (4) *All these aspects were only possible because our society had **developed** technologically to attend our public needs.*

Sobre o exemplo 1, vemos que o aprendiz usa o verbo *develop* com o sentido de “desenvolver”, referindo-se a viajar para o exterior como um elemento que leva as pessoas ao “desenvolvimento” de um ponto de vista diferente em relação ao mundo e as pessoas.

Já no exemplo 2, o lema *develops* coloca com pronome (*our*) e substantivo (*world*) na posição direita. Parece-nos que no exemplo 2, bem como no exemplo 3, o aprendiz usa *develop the world* e *develop the mind*, com o sentido de “desenvolver o mundo”, “desenvolver a mente”. Combinações como essas nos chamam a atenção para o fato da provável junção de palavras ao verbo de forma aleatória, sem se atentar para seu uso colocacional.

No exemplo 4, por seu turno, o lema *developed*, utilizado numa construção no passado perfeito, coloca, na posição esquerda, com o advérbio *-ly* “*technologically*”.

No subcorpus do BAWE, encontramos em muitas linhas o lema *develop* associado aos verbos modais, tais como *can* e *will*, bem como na sua forma infinitiva (precedido de preposição *to*). Vejamos algumas linhas de concordância com os lemas do verbo *develop*:

- (1) *“These linguists have then been able to **develop** her points and back them up with evidence however many of the approaches still have their flaws”.*
- (2) *“The Gothic is a genre which reflects and **develops** women's fantasies”.*
- (3) *“Several solutions have been identified to assist organisations in **developing***

training schemes for the successful adaptation of families and spouses”.

(4) “*Most of the frameworks and models are **developed** on the basis of large company*”.

Das quatro linhas que apresentamos acima, vemos que os usos do VIAE *develop* se diferem dos usos encontrados nos exemplos anteriormente mostrados, oriundos do subcorpus do CorIFA. No exemplo 1, o lema *develop* foi utilizado com a ideia de desenvolver pontos/questões/perspectivas; nos exemplos 2, 3 e 4, o verbo parece ser sinônimo de “criar/construir/elaborar”.

No exemplo 3, convém ressaltar que das 65 ocorrências do lema *developing* no subcorpus do BAWE, notamos que, em apenas duas, há o uso da forma verbal presente contínuo. A predominância, na maior parte das linhas, é do uso do lema precedido das preposições *by*, *in* e *of*.

Já o uso mais frequente do lema *developed* encontrado no subcorpus do BAWE diz respeito às construções passivas, tais como *are developed*, *were developed*, *would be developed* *has been developed*, no presente perfeito (*have/has developed*) e passado perfeito (*had developed*).

Características não compartilhadas entre os subcorpora do CorIFA e do BAWE

Como ressaltamos no Capítulo 1, na subseção 1.6.2, a AVL possui 500 entradas lematizadas, das quais 116 são verbos. Apenas 28 deles, aos quais intitulamos de VIAEs, são encontrados no subcorpus do CorIFA, considerando o ponto de corte de frequência de 20 x 100.000 palavras. Como já enfatizamos, os VIAEs de maior frequência normalizada encontrados em tal subcorpus foram: *improve* (109,20), *develop* (82,78), *provide* (83,57), *increase* (64,65) e *adopt* (58,34). Os verbos *occur*, *affect* e *observe* apresentam a mesma frequência normalizada: 25,23. Encontramos baixa frequência normalizada em *apply* (24,83), *reduce* (21,68), *conclude* (21,28) e *tend* (20,49). Isso indica o pouco uso desses VIAEs pelos aprendizes, o que pode ser reconhecido como uma evidência da dificuldade em usar alguns verbos lexicais (GRANGER; PAQUOT, 2009), gama lexical limitada e, até mesmo, o desconhecimento de quais verbos lexicais são mais recorrentes no discurso acadêmico escrito em inglês. Tais aspectos podem estar ligados ao nível de proficiência linguística dos aprendizes, ao pouco *input* ao longo de sua aprendizagem de inglês, atrelado ao desconhecimento de elementos próprios do

discurso acadêmico em inglês, como a escolha dos verbos lexicais.

No subcorpus do BAWE, por sua vez, os VIAEs de maior frequência normalizada são *provide* (109,62) e *increase* (102,74), classificados como verbo de ação e verbo de simples ocorrência, respectivamente. Nesse sentido, parece haver no BAWE uma tendência dos aprendizes em relatar ações e/ou eventos, bem como relatar eventos.

Retomando os resultados que aqui informamos acerca do subcorpus do CorIFA, podemos afirmar que os textos que compõem o subcorpus do CorIFA são todos provenientes de tarefas independentes (*independent tasks*), o que nos leva a crer que os aprendizes não tinham outros textos com os quais se apoiar para produzirem suas redações argumentativas. Outro ponto a ser considerado diz respeito às diferenças com relação a frequência de uso dos verbos. No subcorpus do CorIFA somente sete verbos estão entre a frequência normalizada entre 40 e 110 (*improve, provide, develop, increase, adopt, base, produce*), enquanto no BAWE, há 15 verbos de frequência normalizada entre 40 e 110 (*provide, develop, increase, base, produce, include, achieve, compare, require, involve, exist, affect, occur e reduce*).

Esses resultados podem nos reportar ao processo de ensino e aprendizagem de vocabulário acadêmico, mais particularmente, de verbos típicos do discurso acadêmico. Os 28 verbos identificados como verbos do inglês acadêmico escrito no CorIFA indica que são comumente usados no discurso acadêmico. Dentre eles, observamos que 23 deles são cognatos (*acquire, adopt, affect, apply, base, compare, conclude, contribute, control, depend, exist, facilitate, include, influence, involve, observe, occur, produce, promote, reduce, relate, require etend*), ou seja, são de origem latina. Talvez o uso desses verbos nos remeta a uma explicação interlingual por serem cognatos (apresentam semelhanças na escrita e no significado quando comparados ao português).

Ainda, os dados que encontramos sobre os VIAES identificados no CorIFA, nos levam a crer que os aprendizes parecem optar pela utilização de itens lexicais com os quais são mais familiarizados, ou seja, usam “[...] as palavras com as quais se sentem seguros” (HASSELGREN, 1994, p.237, tradução nossa)⁸⁹, também conhecidas como “*lexical teddy bears*”, como a própria autora denomina.

Hasselgren (1994) chegou a esse termo quando da realização de uma pesquisa com noruegueses universitários, aprendizes avançados de inglês, que estavam em seu primeiro ano de estudo e outros no sexto ano. O objetivo da pesquisadora era investigar como tais aprendizes usavam vocabulário em inglês nas suas produções escritas,

⁸⁹ No original: “[...] the words they feel safe with” (HASSELGREN, 1994, p. 237).

observando suas escolhas lexicais e seus efeitos no discurso, além de verificar em que medida usavam as colocações. Os resultados dessa pesquisa apontaram para o fato de os aprendizes utilizarem largamente o vocabulário que lhes era familiar, escolhendo frases e palavras que muito se assemelhavam com as da sua L1.

O gênero que os aprendizes escrevem também pode influenciar no uso que fazem do repertório lexical. Convém, desse modo, fazer uma breve retomada sobre os gêneros que permeiam a AVL, o CorIFA e o subcorpus do CorIFA, por nós compilado. A AVL, como já comentado anteriormente, consiste em uma lista que se baseia nas seções acadêmicas do COCA, originada a partir da publicação de artigos acadêmicos. Por sua vez, o CorIFA, apesar de conter outros gêneros (e.g cartas de apresentação), é, em sua maioria, formado por redações argumentativas. Por fim, o subcorpus do CorIFA que criamos contém apenas redações argumentativas. Essa nossa escolha foi intencional, uma vez que constitui um dos gêneros mais comuns no ensino superior.

Gilquin e Paquot (2007), ao realizarem um estudo que dava conta de analisar as funções retóricas utilizadas por aprendizes avançados de inglês como língua adicional, constantes em 14 subcorpora que compõem o ICLE (China, Tchecoslováquia, Dinamarca, Finlândia, França, Alemanha, Itália, Japão, Noruega, Polônia, Rússia, Espanha, Suécia e Turquia), verificaram que há entre eles uma tendência quanto ao uso de características comuns na fala, ao escreverem textos acadêmicos.

Além disso, esses resultados corroboram com a premissa de Granger e Paquot (2009) de que:

[...] os aprendizes de inglês como língua estrangeira subusam significativamente a maioria dos verbos do inglês para fins acadêmicos, que são verbos como: incluir (*include*), relatar (*report*) ou relacionar (*relate*), que expressam funções retóricas no cerne da escrita acadêmica (GRANGER; PAQUOT, 2009, p.210, tradução nossa)⁹⁰.

Essa observação de Granger e Paquot (2009) pode nos ajudar a refletir sobre a baixa frequência normalizada dos VIAEs (os que não atingiram o ponto de corte de frequência de 20 x 100.000 palavras) no subcorpus do CorIFA. A nosso ver, esses dados revelam características importantes sobre as escolhas linguísticas dos aprendizes. Isso posto, é válida uma apresentação e discussão sobre o panorama da distribuição da frequência dos VIAEs: aqueles que os aprendizes brasileiros usam com pouca frequência e aqueles que eles não utilizam, mas estão presentes na AVL. Vejamos:

⁹⁰ No original: “EFL learners significantly underuse the majority of „academic verbs”, that is, verbs like: include, report or relate, that express rhetorical functions at the heart of academic writing” (GRANGER; PAQUOT, 2009, p. 210).

Quadro 10– Distribuição da frequência dos VIAEs no subcorpus do CorIFA (f<20)

VIAES no subcorpus do CorIFA com frequência normalizada (f<20)		
VIAEs	Número de ocorrências	Frequência normalizada
1. <i>recognize</i>	49	19,31
2. <i>obtain</i>	48	18,92
3. <i>support</i>	48	18,92
4. <i>represent</i>	45	17,74
5. <i>present</i>	44	17,34
6. <i>focus</i>	43	16,95
7. <i>note</i>	42	16,55
8. <i>discuss</i>	41	16,16
9. <i>generate</i>	37	14,58
10. <i>define</i>	36	14,19
11. <i>result</i>	36	14,19
12. <i>analyze</i>	34	13,40
13. <i>describe</i>	34	13,40
14. <i>experience</i>	34	13,40
15. <i>establish</i>	34	13,40
16. <i>publish</i>	34	13,40
17. <i>associate</i>	28	11,03
18. <i>maintain</i>	28	11,03
19. <i>reflect</i>	28	11,03
20. <i>transform</i>	28	11,03
21. <i>limit</i>	26	10,24
22. <i>permit</i>	26	10,24
23. <i>ensure</i>	25	9,85
24. <i>regard</i>	25	9,85
25. <i>construct</i>	24	9,46
26. <i>emerge</i>	23	9,06
27. <i>explore</i>	23	9,06
28. <i>link</i>	23	9,06
29. <i>demonstrate</i>	22	8,67

30. <i>suggest</i>	22	8,67
31. <i>enable</i>	21	8,27
32. <i>seek</i>	21	8,27
33. <i>design</i>	20	7,88
34. <i>perform</i>	20	7,88
35. <i>argue</i>	19	7,49
36. <i>evaluate</i>	19	7,49
37. <i>form</i>	19	7,49
38. <i>participate</i>	19	7,49
39. <i>consist</i>	18	7,09
40. <i>encourage</i>	18	7,09
41. <i>report</i>	18	7,09
42. <i>vary</i>	18	7,09
43. <i>determine</i>	17	6,70
44. <i>emphasize</i>	16	6,30
45. <i>indicate</i>	16	6,30
46. <i>perceive</i>	16	6,30
47. <i>assume</i>	15	5,91
48. <i>organize</i>	15	5,91
49. <i>propose</i>	15	5,91
50. <i>test</i>	15	5,91
51. <i>conduct</i>	13	5,12
52. <i>select</i>	13	5,12
53. <i>address</i>	12	4,73
54. <i>enhance</i>	12	4,73
55. <i>measure</i>	12	4,73
56. <i>lack</i>	12	4,73
57. <i>incorporate</i>	12	4,73
58. <i>cite</i>	11	4,33
59. <i>express</i>	11	4,33
60. <i>contain</i>	10	3,94
61. <i>reveal</i>	10	3,94
62. <i>state</i>	10	3,94

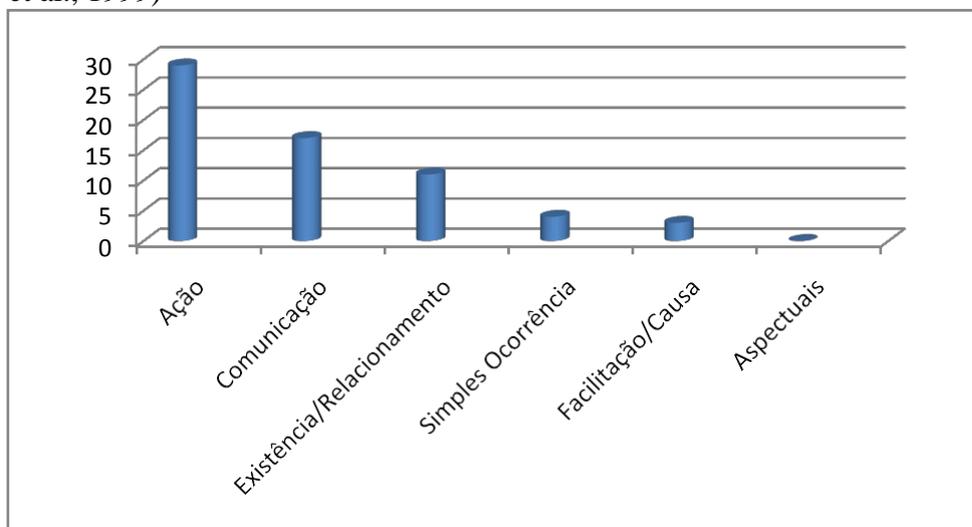
63. <i>illustrate</i>	9	3,54
64. <i>inform</i>	9	3,54
65. <i>refer</i>	9	3,54
66. <i>rely</i>	7	2,75
67. <i>review</i>	7	2,75
68. <i>constitute</i>	6	2,36
69. <i>differ</i>	6	2,36
70. <i>extend</i>	6	2,36
71. <i>predict</i>	6	2,36
72. <i>yield</i>	6	2,36
73. <i>arise</i>	5	1,97
74. <i>display</i>	5	1,97
75. <i>derive</i>	5	1,97
76. <i>imply</i>	5	1,97
77. <i>locate</i>	5	1,97
78. <i>characterize</i>	4	1,57
79. <i>engage</i>	4	1,57
80. <i>estimate</i>	4	1,57
81. <i>interpret</i>	4	1,57
82. <i>account</i>	3	1,18
83. <i>view</i>	3	1,18
84. <i>assess</i>	2	0,78
85. <i>examine</i>	2	0,78
86. <i>attempt</i>	1	0,39
87. <i>employ</i>	0	0
88. <i>range</i>	0	0

Fonte: Elaborado pela autora, 2017.

Dos 88 VIAEs dispostos no Quadro 10, dois deles não apresentam ocorrências no subcorpusdo CorIFA: *employ* e *range*; nove apresentam número de ocorrências igual ou inferior a cinco e 63 são cognatos. A baixa frequência normalizada de todos esses verbos listados no quadro nos mostra que há pouco uso, nas redações argumentativas dos aprendizes, de verbos que costumam ocorrer no discurso acadêmico escrito. O

Gráfico 5 mostra a distribuição dos VIAEs pouco usados pelos aprendizes brasileiros, levando em conta seu domínio semântico:

Gráfico 5 – Distribuição dos VIAEs ($f < 20$) no CorIFA por domínio semântico (BIBER et al., 1999)



Fonte: Elaborado pela autora, 2017.

Dentre os verbos listados com frequência normalizada menor que 20 palavras por 100.000, encontramos ocorrências de apenas um verbo de existência (*reveal*) que, para os autores, são os mais comuns no discurso acadêmico (além do verbo copular *be*). Em contrapartida, há ocorrências de verbos de relacionamento, tais como *contain* (3,94) e *represent* (17,74). Os VIAEs de menor frequência normalizada dentre os dispostos no Gráfico 5 foram os verbos de simples ocorrência, facilitação (ex: *enable* e *permit*) ou causa (ex: *ensure*) e os aspectuais. Sobre esse aspecto, o estudo de Biber et al. (1999), ao analisar a distribuição dos verbos com base em seu domínio semântico, obteve como resultado o fato de os verbos causativos, de ocorrência e aspectuais apresentarem também frequência muito baixa em todos os registros, porém, no que diz respeito ao registro acadêmico, a frequência desses verbos foi um pouco maior em relação aos demais registros.

Verificando o uso desses verbos no subcorpus do CorIFA, observamos o seguinte:

a) Quanto aos verbos de ação *engage*, *extend* e *account*:

- O primeiro é usado em apenas quatro linhas, porém em nenhuma delas com a ideia de engajar-se ou envolver-se, mas sim com a conotação de

estabelecer relações com alguém: “[...]they prefer to stay connected to the internet rather than to **engage** with other persons;

- O verbo *extend* é encontrado apenas com dois lemas no subcorpus (*extend* e *extended*), sendo este último, com seis ocorrências, utilizado no particípio passado em construções passivas e no presente perfeito, como em: “The development of drugs has **extended** the life of humans [...]”;
- O verbo *account*, encontrado em apenas quatro linhas, aparece em duas delas em construções na forma passiva, como vemos na seguinte linha: “Thus, the productivity gains must also be **accounted** jointly with the loss of workday”.

Acreditamos que pode haver uma relação da baixa frequência desses verbos com seus respectivos domínios semânticos, principalmente, se lembrarmos que os verbos de maior frequência normalizada no CorIFA, conforme mostramos na seção 3.3, são, justamente, os classificados como verbos de ação e comunicação. Entendemos, assim, que no subcorpus do CorIFA predominam verbos de ação e comunicação e estes são utilizados sob uma perspectiva polarizada: tanto em maior frequência quanto em frequência muito baixa.

Resultados do teste *log-likelihood* com os verbos do inglês acadêmico escrito do CorIFA

Como já afirmamos no Capítulo 2, as questões de análise de dados têm recebido atenção recentemente e as medidas estatísticas têm sido muito utilizadas para interpretar dados quantitativos. Diante disso, nessa subseção, apresentamos uma discussão sobre os resultados do teste *log-likelihood* aplicado aos verbos do inglês acadêmico escrito mais frequentes do CorIFA, com o intuito de verificar se as diferenças na frequência dos verbos do inglês acadêmico escrito no CorIFA são estatisticamente significativas e para compreendermos se tais diferenças são aleatórias ou não. Os dados constantes no quadro que se segue, de acordo com McEnery e Hardie (2012), são referentes a:

- O1: frequência observada no *corpus1* (nesse caso, o subcorpus do CorIFA);
- O2: frequência observada no *corpus2* (nesse caso, o subcorpus do BAWE);
- % 1 e % 2: frequências relativas nos textos;

- + indica sobreuso no *corpus1* em relação ao *corpus 2*;
- - indica subuso no *corpus1* em relação ao *corpus 2*;

Listamos, no quadro abaixo, os verbos do inglês acadêmico escrito mais frequentes do CorIFA a serem analisados nesta pesquisa, em ordem alfabética, com seus respectivos valores de LL:

Quadro 11– Valor de *log-likelihood* e indicação de sobreuso ou subuso⁹¹

Verbo	Valor de <i>log-likelihood</i> (LL)	Sobreuso	Subuso
<i>achieve</i>	5,73		X
<i>acquire</i>	8,58	X	
<i>adopt</i>	82,97	X	
<i>affect</i>	20,76		X
<i>apply</i>	11,65		X
<i>base</i>	15,28		X
<i>compare</i>	6,47		X
<i>conclude</i>	0,35	X	
<i>contribute</i>	26,20	X	
<i>control</i>	0,87	X	
<i>depend</i>	2,67	X	
<i>develop</i>	15,22	X	
<i>exist</i>	12,64		X
<i>facilitate</i>	34,53	X	
<i>improve</i>	143,70	X	
<i>include</i>	39,22		X
<i>increase</i>	28,88		X
<i>influence</i>	7,58	X	
<i>involve</i>	20,88		X
<i>observe</i>	0,05		X

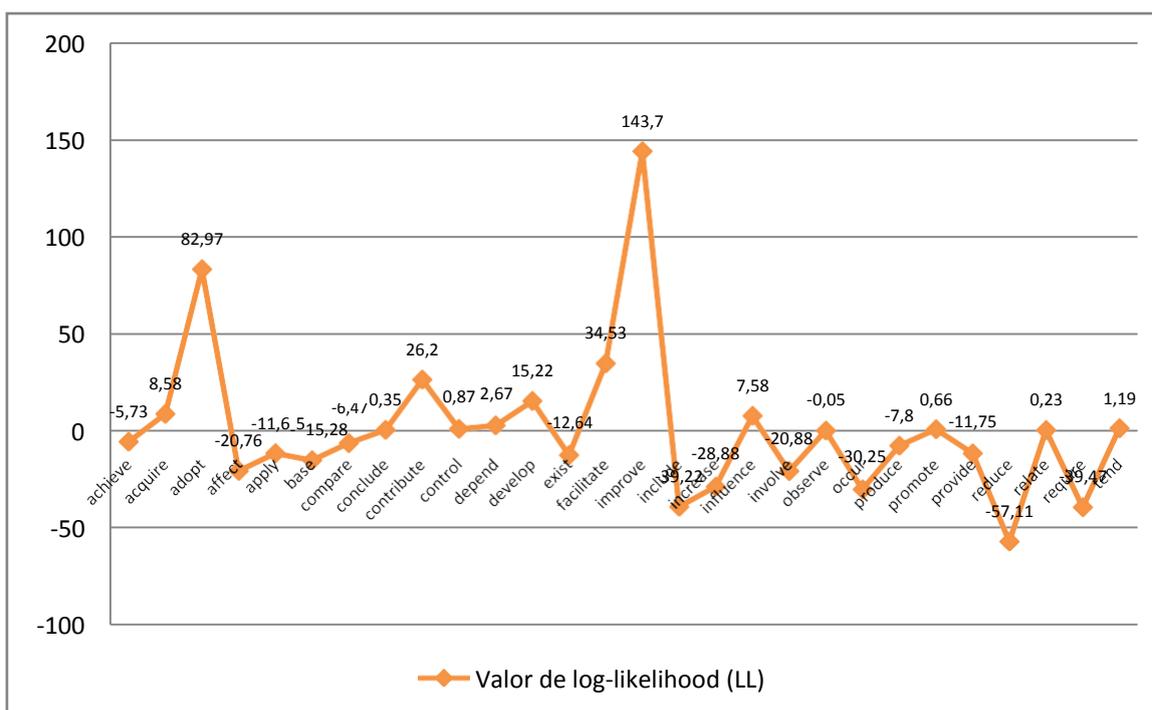
⁹¹ Os termos “sobreuso” e “subuso”, mencionados por McEnery e Hardie (2012), são usados, nesta pesquisa, como forma de se estabelecer um parâmetro de quais verbos podemos considerar significativos do ponto de vista estatístico, nos *corpora* de aprendizes em investigação, na comparação com o *corpus* de referência. Quando tratamos neste texto sobre os referidos termos, estamos nos pautando em McEnery e Hardie (2012) que os relacionam, respectivamente, à alta e baixa frequência de determinado item lexical.

<i>occur</i>	30,25		X
<i>produce</i>	7,80		X
<i>promote</i>	0,66	X	
<i>provide</i>	11,75		X
<i>reduce</i>	57,11		X
<i>relate</i>	0,23	X	
<i>require</i>	39,47		X
<i>tend</i>	1,19	X	

Fonte: Elaborado pela autora, 2017.

A distribuição dos valores de LL desses verbos está mais bem representada no Gráfico 6:

Gráfico 6 – Distribuição dos valores de log-likelihood (LL) dos VIAEs mais frequentes no subcorpus do CorIFA



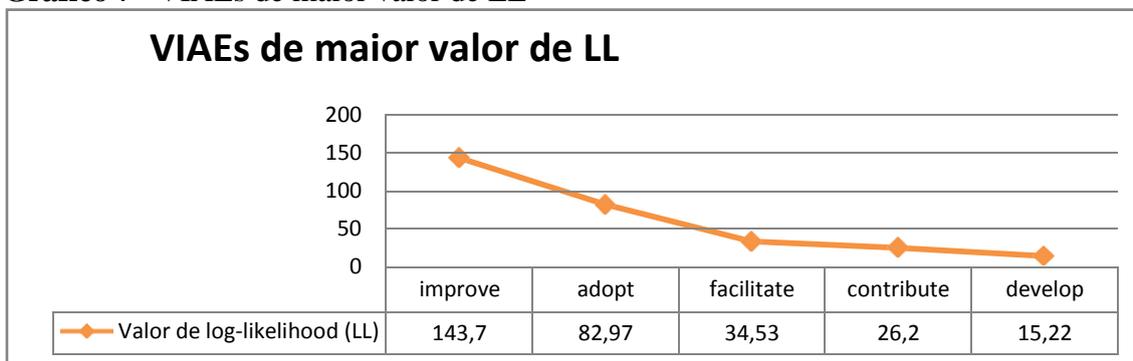
Fonte: Elaborado pela autora, 2017.

A partir do Gráfico 6, podemos observar que alguns VIAEs apresentam valor de LL bem maiores que outros.

Conforme apresentado no Gráfico 7 abaixo, os verbos de maior valor de LL são: *improve* (143,70), *adopt* (82,97), *facilitate* (34,53), *contribute* (26,2) e *develop* (15,22), sendo que *improve* constituiu verbo com o maior número de ocorrências no *corpus*, ou

seja, 277.

Gráfico 7– VIAEs de maior valor de LL

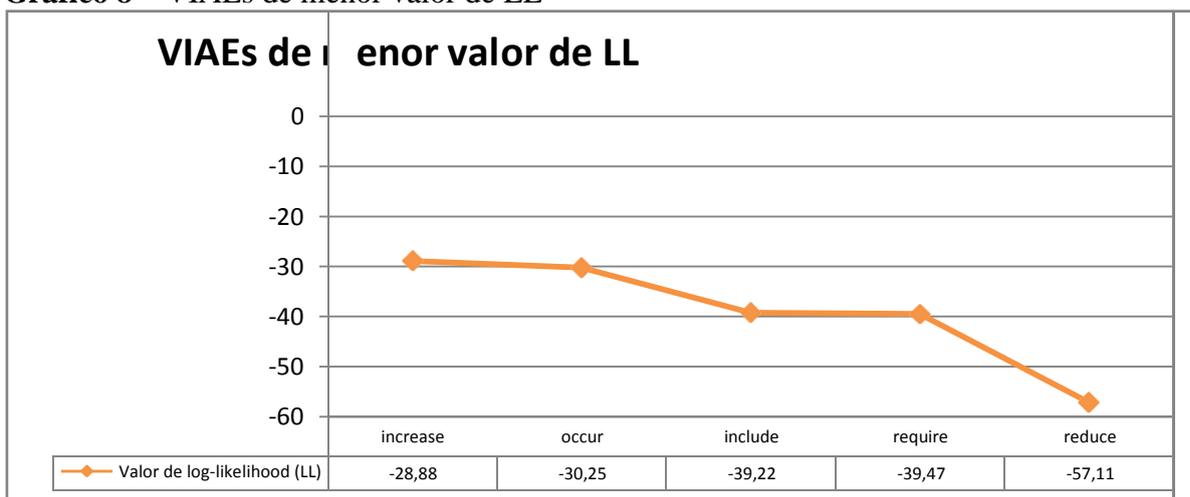


Fonte: Elaborado pela autora, 2017.

Dessa forma, podemos afirmar que esses verbos apresentam nível de significância estatística alto. Os valores de LL dos verbos *control*, *depend* e *relate* são inferiores a 3,84, o que implica em serem considerados estatisticamente não significativos.

Ademais, 14 dos 26 verbos do gráfico apresentam valores de LL indicativos de subuso, conforme aponta McEnery e Hardie (2012). No Gráfico 8 abaixo, encontramos os cinco VIAEs de menor valor de LL:

Gráfico 8 – VIAEs de menor valor de LL



Fonte: Elaborado pela autora, 2017.

Os verbos de mais baixo valor de LL são *require* (-39,47), *include* (-39,22) e *occur* (-30,25). Segundo os autores, não há valores de LL negativos, no entanto, eles

preferem o uso dos sinais de “+” e “-” como forma de indicar sobreuso e subuso no *corpus* 1 em relação ao *corpus*2, respectivamente. Assim, de posse desse argumento, podemos dizer que os verbos aos quais mencionamos (*achieve*, *base*, *exist*, *include*, *increase*, *involve*, *observe*, *produce*, *provide* e *require*) apresentam subuso no subcorpus do CorIFA em relação ao subcorpusdo BAWE (que constitui, aqui, nosso *corpus*2).

Um dos valores que destacamos diz respeito ao verbo *require*. No CorIFA, sua frequência absoluta é de 81 ocorrências e no subcorpusdo BAWE, o número de ocorrências foi bastante alto: 336. Todavia, seu valor de LL, apesar de ser bem maior que 3,84, ou seja, 39,47, ainda possui indicativo de subuso. Atribuímos esse resultado possivelmente ao grande número de ocorrências no nosso subcorpus de referência. Entendemos que quanto maior a diferença no número de ocorrências, maior diferença encontramos no seu uso em ambos os subcorpora e, conseqüentemente, maior também será seu valor de LL. Podemos citar como exemploo verbo *adopt* que ocorreu 58,34 vezes no subcorpus do CorIFA e 17,68 vezes no subcorpus do BAWE. O valor de LL entre os dois é bem alta, representando 82,97, o que nos indica que a diferença quanto ao uso desse VIAE entre os subcorpora é significativa. Com o verbo *achieve*, notamos também algumas diferenças. O VIAE apresenta maior frequência normalizada no subcorpus do BAWE (45,57) em relação ao subcorpus do CorIFA (33,9), tendo como valor de LL -5,73, o que consideramos valor bastante baixo, mesmo com a diferença não tão grande entre as frequências normalizadas.

Síntese dos resultados

No Quadro12 abaixo, apresentamosos dados relativos aos VIAEs no subcorpus do CorIFA e no subcorpus do BAWE, bem como seu valor de LL, a fim de melhor visualizarmos e destacarmos os resultados obtidos, até então, nesta pesquisa.

Quadro 2 – Síntese dos resultados dos VIAEs – subcorpus do CorIFA e subcorpus do BAWE

VIAEs	Frequência absoluta de <i>tokens</i> no subcorpus do CorIFA	Frequência normalizada subcorpusdo CorIFA	Frequência absoluta de <i>tokens</i> no subcorpus do BAWE	Frequência normalizada subcorpus do BAWE	<i>Log-likelihood</i> (LL)
<i>achieve</i>	86	33,9	232	45,57	-5,73
<i>acquire</i>	73	28,77	92	18,07	+ 8,58
<i>adopt</i>	148	58,34	90	17,68	+ 82,97

<i>affect</i>	64	25,23	236	46,36	-20,76
<i>apply</i>	63	24,83	203	39,88	-11,65
<i>base</i>	11	43,75	337	66,20	-15,28
<i>compare</i>	83	32,72	229	44,98	-6,47
<i>conclude</i>	54	21,28	98	19,25	+0,35
<i>contribute</i>	92	36,26	85	16,10	+26,20
<i>control</i>	70	27,59	122	23,96	+0,87
<i>depend</i>	84	33,11	134	26,32	+2,67
<i>develop</i>	210	82,78	295	57,95	+15,22
<i>exist</i>	77	30,35	242	47,54	-12,64
<i>facilitate</i>	81	31,93	60	11,78	+ 34,53
<i>improve</i>	277	109,20	179	35,16	+143,70
<i>include</i>	89	35,08	358	70,33	-39,22
<i>increase</i>	164	64,65	523	102,74	-28,88
<i>influence</i>	89	35,08	121	23,77	+7,58
<i>involve</i>	80	31,53	279	54,81	-28,88
<i>observe</i>	64	25,23	133	26,16	-0,05
<i>occur</i>	64	25,23	263	51,66	-30,25
<i>produce</i>	106	41,78	290	56,97	-7,80
<i>promote</i>	56	22,07	98	19,25	+0,66
<i>provide</i>	212	83,57	558	109,62	-11,75
<i>reduce</i>	55	21,68	299	58,74	-57,11
<i>relate</i>	100	39,42	189	37,13	+0,23
<i>require</i>	81	31,93	336	66,01	-39,47
<i>tend</i>	52	20,49	86	16,89	+1,19

Fonte: Elaborado pela autora, 2017.

Há alguns dados disponíveis no quadro anterior que merecem destaque, dentre os quais podemos citar: o verbo *improve*, de maior frequência normalizada no subcorpus do CorIFA, é menos frequente no subcorpus do BAWE e apresenta o maior valor de LL, ou seja, + 143,70, o que representa uma diferença estatisticamente significativa. Outro VIAE de valor de LL alto é o verbo *adopt*(+82,97) que, como comentamos na seção 3.4, tem uma diferença bem evidente quanto à frequência normalizada no subcorpus do CorIFA (58,34) em relação ao subcorpus do BAWE (17,68). O VIAE *facilitate* é bem mais frequente no subcorpus do CorIFA quando comparado ao subcorpus do BAWE e apresenta valor de LL alto (+34,53), o que indica que a diferença desse VIAE entre ambos os subcorpora é estatisticamente significativa.

A seguir, descrevemos e analisamos o comportamento das colocações dos verbos do inglês acadêmico escrito do CorIFA.

As colocações no subcorpus do CorIFA e do BAWE: descrição das colocações formadas por advérbios *-ly* do tipo maximizadores

O próximo passo consistiu na identificação e descrição linguística das colocações do advérbio *-ly* com os verbos do inglês acadêmico escrito, utilizados pelos

aprendizes no subcorpus do CorIFA, estabelecendo comparações com o utilizado no subcorpus do BAWE. Entendemos que, além da frequência, o contexto de uso no qual os verbos foram utilizados, isto é, seu contexto lexical, é de suma importância para o significado da palavra e seu papel gramatical (ELLIS, 2008).

Para dar conta da descrição dessas colocações, num primeiro momento, identificamos as características lexicais e sintáticas dos verbos e, em seguida, buscamos seus padrões de uso nas linhas de concordância. Considerando que os advérbios exercem o papel de modificadores dos verbos, decidimos pela descrição das colocações formadas por “advérbio *-ly* + verbo”, uma vez que, a nosso ver, com a classe de advérbios, poderíamos encontrar menor variação lexical, se comparado a classe dos substantivos ou adjetivos. Além disso, concordamos com Granger (1998) ao nos revelar que esse tipo de colocação “[...] constitui uma categoria de colocação particularmente rica, envolvendo um complexo jogo de restrições semânticas, lexicais e estilísticas e cobrindo o espectro do todo colocacional, que varia a partir de colocabilidade restrita” (GRANGER, 1998, p.2, tradução nossa)⁹².

Foi realizado um refinamento para chegarmos aos advérbios que aqui nos interessam (advérbios *-ly*), como já descrevemos no Capítulo 2. Foram encontrados 332 *word types* e 373 *word tokens*, o que significa que houve uma grande variedade de advérbios no subcorpus do CorIFA, porém com frequência muito baixa, principalmente como colocados de VIAEs. Vinte e dois dos advérbios-*ly* identificados possuem equivalentes na sua L1, o português, o que sugere que os aprendizes usam advérbios apoiando-se em adjetivos cognatos que lhes são mais conhecidos.

Como vimos no Capítulo 1, os advérbios *-ly* aparecem em sete diferentes categorias semânticas, de acordo com a classificação de Biber et al. (1999). Observemos na Tabela 8 os advérbios *-ly* comuns ao subcorpus do CorIFA e ao subcorpus do BAWE e sua respectiva classificação semântica:

Tabela 8 – Advérbios *-ly* comuns ao subcorpus do CorIFA e ao subcorpus do BAWE

Advérbios <i>-ly</i>	Classificação semântica (BIBER et al., 1999)
<i>actually</i>	posicionamento epistêmico
<i>adversely</i>	contraste

⁹²No original: “[...] constitute a particularly rich category of collocation, involving as they do a complex interplay of semantic, lexical and stylistic restrictions and covering the whole collocational spectrum, ranging from restricted collocability” (GRANGER, 1998, p.2).

<i>directly</i>	modo
<i>especially</i>	restrição
<i>highly</i>	intensificador
<i>largely</i>	intensificador
<i>mainly</i>	posicionamento epistêmico
<i>negatively</i>	posicionamento epistêmico
<i>only</i>	restrição
<i>really</i>	posicionamento epistêmico
<i>severely</i>	posicionamento epistêmico
<i>significantly</i>	posicionamento epistêmico
<i>strongly</i>	intensificador
<i>totally</i>	maximizador/amplificador
<i>usually</i>	tempo

Fonte: Elaborada pela autora, 2017.

Os advérbios de posicionamento epistêmico⁹³ se destacam dentre os advérbios comuns a ambos os subcorpora. A partir deles, os aprendizes colocam-se avaliativamente nos textos. Biber (2006), em um estudo no qual foram contrastadas expressões de posicionamento utilizadas em registros acadêmicos, dentre eles, os advérbios de posicionamento, afirma que “[...] em alguns casos os falantes e escritores, em registros acadêmicos, parecem estar mais preocupados com as expressões de posicionamento que com a comunicação de „fatos“” (BIBER, 2006, p.87, tradução nossa)⁹⁴. Para o autor, o posicionamento é aspecto crucial quando se trata de textos acadêmicos. Desse modo, os advérbios de posicionamento “[...] focam no valor de verdade de uma proposição, comentando sobre fatores tais como certeza, realidade, fontes, limitações e precisão da proposição” (BIBER et al., 1999, p.764, tradução nossa)⁹⁵. Em relação a isso, acreditamos que, a depender da seção do texto acadêmico, como, por exemplo, um artigo, expressões de posicionamento podem aparecer em maior quantidade.

⁹³ Devido ao escopo do trabalho, a questão da classificação semântica dos advérbios de posicionamento epistêmico será abordada em um trabalho posterior.

⁹⁴ No original: “[...] in some cases speakers and writers in university registers seem more concerned with the expression of stance than with the communication of „facts“” (BIBER, 2006, p. 87).

⁹⁵ No original: “[...] focus on the truth value of a proposition, commenting on factors such as certainty, reality, sources, limitations, and precision of the proposition” (BIBER et al., 1999, p. 764).

No Quadro 13, encontramos a frequência de uso dos advérbios de posicionamento epistêmico, comuns a ambos os subcorpora.

Quadro 13– Número de ocorrências dos advérbios de posicionamento epistêmico

Advérbio de posicionamento epistêmico	Número de ocorrências no subcorpus do CorIFA	Número de ocorrências no subcorpus do BAWE
<i>actually</i>	84	88
<i>mainly</i>	72	66
<i>negatively</i>	4	10
<i>really</i>	189	60
<i>severely</i>	7	11
<i>significantly</i>	12	76

Fonte: Elaborado pela autora, 2017.

Os advérbios *actually*, *mainly*, *negatively* e *severely* apresentam frequência bem semelhante entre ambos os subcorpora. Chamaram a nossa atenção o número grande de ocorrências de *really* no subcorpus do CorIFA quando comparado ao subcorpus do BAWE. Por sua vez, *significantly* apresenta seis vezes o número de ocorrências a mais no subcorpus do BAWE em relação ao subcorpus do CorIFA.

Analisando a tabela disponível no Apêndice II, notamos que a preferência de uso dos estudantes, no subcorpus do BAWE, é de colocações com advérbios do tipo amplificadores (tais como *highly*, *greatly* e *vastly*), advérbios de restrição (como *only* e *particularly*) e de posicionamento epistêmico, que manifesta nível de dúvida (a exemplo de *possibly*) e que expressa nível de certeza diante da proposição (a exemplo de *clearly*).

Ao olharmos os advérbios de grau do tipo maximizadores e intensificadores (cuja função é indicar a intensidade ou grau de uma ação), nos subcorpora como um todo, os que colocam com a categoria verbos, de forma genérica, e depois, mais especificamente, os que colocam com os VIAES, são os demonstrados na Quadro 14:

Quadro14 – Distribuição dos advérbios (número de ocorrências) do tipo maximizadores e seus colocados verbais

Advérbios de grau:	Subcorpus do CorIFA	Subcorpus do BAWE

maximizadores						
	Verbos em geral	Verbos lexicais	VIAEs	Verbos em geral	Verbos lexicais	VIAEs
<i>absolutely</i>	9	3	3	11	4	3
<i>perfectly</i>	8	7	0	17	6	1
<i>definitely</i>	30	2	0	10	7	1
<i>totally</i>	23	15	3	11	7	1
<i>completely</i>	47	7	1	69	29	2

Fonte: Elaborado pela autora, 2017.

Notamos que os advérbios *-ly* são amplamente utilizados em ambos os subcorpora, com verbos (de modo geral), com destaque para *completely*, *definitely* e *totally*, os de maior número de ocorrências, sendo que os dois últimos aparecem em maior número no subcorpus do CorIFA em comparação ao subcorpus do BAWE. Quando refinamos a busca para o funcionamento dos advérbios dispostos no Quadro 14 como colocados de verbos lexicais, é notório o pouco uso dando os primeiros sinais nos dois subcorpora e, mais fortemente, no subcorpus do CorIFA. Como vemos no Quadro 14, houve uma diminuição no uso dos advérbios do tipo maximizadores, à medida que refinamos nossa busca com verbos. Uma redução ainda mais drástica pode ser percebida quanto ao uso desses advérbios como colocados de VIAEs. Há uma redução bastante perceptível, principalmente quanto ao uso de *completely*, que passa de 47 para sete ocorrências como colocados de verbos lexicais e somente uma ocorrência como colocado de VIAE, no subcorpus do CorIFA. No subcorpus do BAWE, o advérbio passa de 69 para 29 com verbos lexicais e apenas duas ocorrências como colocado de VIAE.

No subcorpus do CorIFA, a frequência é bastante baixa, tanto que *perfectly* e *definitely* não apresentam ocorrências com os VIAEs. De igual modo, no subcorpus do BAWE, o número de ocorrências é também muito reduzido. O subuso dos advérbios do tipo maximizadores como colocados dos VIAEs, nos dois subcorpora, pode ser indício de que tais advérbios são mais comumente usados pelos estudantes para modificar a classe de adjetivos ou outros advérbios, em detrimento dos verbos de natureza lexical. Além disso, os aprendizes podem, talvez, ter buscado evitar o uso desses maximizadores (uma vez que funcionam para remeter a uma condição hiperbólica) a fim de não

potencializar determinadas afirmações no seu mais alto grau. Vejamos algumas ocorrências desses advérbios maximizadores, com VIAEs, no subcorpus do CorIFA:

(1) *Technological innovation **absolutely depends** on dreaming and innovative thoughts.*

(2) *In addition is interesting that the measurement that they have to do **totally involves** math and its principles.*

(3) *The human beings are **completely involved** with different technologies.*

Nas linhas de concordância acima estão negritados os advérbios maximizadores amplificadores que representam a expressão do mais alto grau de uma determinada proposição. As colocações formadas a partir de tais advérbios (*absolutely depends*, *totally involves* e *completely involved*) podem ter sido escolhidas pelos aprendizes por conta de possuírem correspondentes na sua L1. Ao checarmos as mesmas colocações, até mesmo com outros lemas dos mesmos verbos no subcorpus do BAWE, não encontramos ocorrências, o que pode confirmar que os aprendizes usam poucas colocações comuns no discurso dos falantes de inglês como L1, preferindo combinações de palavras mais atípicas.

Observemos algumas linhas de concordância com esses advérbios maximizadores com VIAEs no BAWE:

(1) *Since, nowadays, there is enormous use of agrochemicals in modern agriculture, the use of environmentally friendly alternatives to chemical pesticides are **absolutely required**.*

(2) *Team's main decision was the development of a new product ('Tiger'), **totally based** on future sales' forecasts.*

(3) *[...]however, his attempts to check his emotions and control his furor are often compromised, and the vision of Rome as a civilising force where emotional control is key to peace is proved impossible to **achieve completely**.*

A partir das linhas de concordância apresentadas acima vemos que os advérbios maximizadores *absolutely*, *totally* e *completely* são usados no subcorpus do BAWE com VIAEs diferentes dos encontrados no subcorpus do CorIFA. Nos exemplos 1 e 2,

encontramos as colocações *absolutely required* e *totally based*, respectivamente, cujos advérbios foram utilizados na posição esquerda dos VIAEs.

No exemplo 1, a colocação é usada numa construção na forma passiva (*be* + advérbio + particípio passado do verbo), ao passo que no exemplo 2, a colocação *totally based* não faz parte de uma conjugação verbal. Quanto ao exemplo 3, podemos afirmar que encontramos 69 ocorrências do advérbio *completely* no subcorpus do BAWE em comparação a 47 ocorrências no subcorpus do CorIFA. Esses dados nos informam que, apesar de sua frequência ser um pouco alta em ambos os subcorpora, quando se trata de seu uso com VIAEs, a situação é bem diferente: apenas uma ocorrência no subcorpus do CorIFA e duas ocorrências no subcorpus do BAWE.

Diante disso, percebemos que, de modo geral, os aprendizes usam vários advérbios *-ly* na combinação com os VIAEs, contudo, no subcorpus do BAWE, a diversidade é, notoriamente, bem maior. Percebemos que os lemas que formam o presente simples e o particípio dos VIAEs predominam nessas combinações; há poucas ocorrências de colocações cujo lema esteja no gerúndio.

Os lemas mais frequentes com os advérbios maximizadores no subcorpus do CorIFA foram os referentes ao presente simples e ao particípio passado. O advérbio *absolutely*, de frequência muito baixa no subcorpus do CorIFA (apenas nove ocorrências), se combina em três linhas de concordância com verbos no presente simples. *Perfectly* apresenta frequência muito baixa (oito ocorrências), porém em duas linhas aparece sendo usado com verbo no particípio passado. Verificando no subcorpus do BAWE, vemos que *perfectly* coloca com três verbos no presente simples (como no exemplo 1 abaixo), bem como com um verbo no particípio passado (exemplo 2 abaixo).

(1) *Carby-Hall outlines **perfectly** the significant change of policy of statute interpretation in British courts [...].*

(2) *This follows from the fact that with rational expectations, policy in period is **perfectly** anticipated [...].*

Biber et al. (1999), ao nos chamarem a atenção sobre a importância do contexto na semântica dos advérbios, nos lembram que o significado dos advérbios pode variar a depender do contexto de uso, já que podem apresentar significados literais e metafóricos. Os autores citam como exemplos a colocação *perfectly arranged*, segundo eles, usada com o sentido literal (“de maneira perfeita”), e as combinações *perfectly*

normal e *perfectly safe*, usadas num sentido metafórico como sinônimo de *completely*. Ainda, os autores mencionam que os advérbios de modo “[...] às vezes perdem seu valor semântico literal quando são usados como amplificadores” (BIBER et al., 1999, p.555, tradução nossa)⁹⁶. Assim, citam o exemplo “*And Carl was perfectly awful*”, para demonstrar que o advérbio *perfectly*, nesse caso, é usado com o sentido metafórico de “*absolutely*”. Com base nisso, podemos afirmar que, no exemplo 1, *perfectly* evidencia um sentido literal ao se colocar com o verbo lexical *outlines*, ao passo que no exemplo 2, o advérbio apresenta sentido mais metafórico.

Já *totally* apresenta 23 ocorrências em todo o subcorpus do CorIFA. O advérbio não aparece em frases no gerúndio, apenas na forma presente (seis ocorrências) e no particípio passado (nove ocorrências). Das 47 ocorrências do advérbio *completely* no subcorpus, 40 têm adjetivos como colocados, tais como *different*, *infeasible* e *theoretical*. Das sete ocorrências em que o advérbio coloca com verbos lexicais, todos estavam no particípio passado. Eis alguns exemplos:

(1) [...] *her cell phone and her car are essential to make it through the day and that she is **completely** satisfied.*

(2) *The human beings are **completely** involved with different technologies.*

A seguir, são apresentados os resultados da frequência das colocações de advérbios *-ly* com *adopt*, *develop*, *improve*, *increase* e *provide*, os cinco verbos de maior frequência normalizada do subcorpus do CorIFA.

Resultados –colocações com *adopt*, *develop*, *improve*, *increase* e *provide*

Como já afirmamos de início, além de descrevermos a frequência e o uso dos verbos do inglês acadêmico escrito (VIAEs), por uma questão de decisão metodológica, nesta pesquisa, nos atemos a descrever e analisar as colocações (advérbio *-ly* + verbo) utilizadas pelos aprendizes com os cinco verbos mais frequentes (*adopt*, *develop*, *improve*, *increase* e *provide*) identificados no subcorpus do CorIFA, nosso subcorpus de estudo. Elegemos, assim, cinco advérbios *-ly* que mais frequentemente coocorrem com

⁹⁶No original: “[...] manner adverbials sometimes lose their literal semantic value when they are used as amplifiers” (BIBER et al., 1999, p. 555).

eles no referido subcorpus, considerando seu posicionamento sintático e os dispusemos nos quadros que se seguem.

É importante ressaltar que houve uma tentativa de se realizar uma análise quantitativa das colocações a partir do teste estatístico Mutual Information (MI), através do qual seria possível medir a força de associação entre os verbos e os advérbios *-ly* que o acompanham. No entanto, quando o fizemos, obtivemos número abaixo de três. Para o MI ser considerado significativo, seria necessário estar acima desse número. No sítio eletrônico do COCA (<https://corpus.byu.edu/mutualinformation.asp>), Mark Davies pontua que os resultados abaixo de três são estranhos e podem ser ocasionados pelo baixo número de ocorrências.

A seguir, mostramos a frequência dos advérbios *-ly* encontrados como colocados de *adopt* nos subcorpus do CorIFA e no subcorpus do BAWE. Iniciemos com a frequência para o verbo *adopt*, na posição esquerda:

Quadro 15 – Frequência para verbo *adopt*

ADV	Subcorpus CorIFA	Frequência à esquerda	Subcorpus BAWE	Frequência à esquerda
<i>Advérbio 1</i>	<i>certainly</i>	0	<i>eventually</i>	1
<i>Advérbio 2</i>	-	0	<i>newly</i>	1
<i>Advérbio 3</i>	-	0	<i>formally</i>	1
<i>Advérbio 4</i>	-	0	<i>recently</i>	1
<i>Advérbio 5</i>	-	0	<i>previously</i>	0

Fonte: Elaborado pela autora, 2017.

Vemos que o verbo *adopt* não coocorre, na posição esquerda, com nenhum advérbio *-ly* no subcorpus do CorIFA. Já no subcorpus do BAWE, apesar da baixíssima frequência, encontramos na posição esquerda, os seguintes advérbios *-ly* coocorrentes com *adopt*: *eventually*, *newly*, *formally*, *recently* e *previously*.

Na posição direita, temos:

Quadro 16 – Frequência para verbo *adopt*

ADV	Subcorpus CorIFA	Frequência à direita	Subcorpus BAWE	Frequência à direita
<i>Advérbio 1</i>	<i>certainly</i>	2	<i>eventually</i>	0
<i>Advérbio 2</i>	-	-	<i>newly</i>	0
<i>Advérbio 3</i>	-	-	<i>formally</i>	0
<i>Advérbio 4</i>	-	-	<i>recently</i>	0
<i>Advérbio 5</i>	-		<i>previously</i>	1

Fonte: Elaborado pela autora, 2017.

Já na posição direita, o verbo *adopt* coloca *comcertainly*, advérbio de posicionamento epistêmico que permite ao escritor avaliar a proposição na oração (BIBER et al, 1999) que, nesse caso, é expressa pelo nível de certeza da proposição, como podemos ver na seguinte linha de concordância:

(1) *The technology has developed with the human being, who **certainly depends**of the technology for their own develops.*

Na colocação *certainly depends*, identificamos um desvio colocacional (NESSELHAUF, 2005) quanto à escolha da preposição *of*. A inadequação quanto ao uso pode ser decorrente da falta de conhecimento colocacional do aprendiz, uma vez que o verbo *depend* requer a companhia da preposição *on*, ao invés de *of*. Podemos atribuir esse uso inadequado ao fato de, no português, usarmos a colocação “depende de”, em que a preposição pode ter sido vista como correspondente a *of*, como em “*two bottles of red wine*”, em português, duas garrafas **de** vinho tinto.

No BAWE, encontramos apenas *previously* como colocado na posição direita de *adopt*, ambos usados na seguinte linha de concordância:

(1) *The ATER policy has a more holistic approach, of equity and social inclusion, instead of diffusion of “technological packages” as **adopted previously**.*

O falante de L1 provavelmente utilizou o advérbio de modo *previously* como forma de referir-se à algo anteriormente mencionado no texto. Nesse caso, a difusão dos pacotes tecnológicos.

Encontramos também frequência muito baixa dos advérbios-lyna posição esquerda, como no Quadro 17:

Quadro 17 – Frequência para verbo *develop*

ADV	Subcorpus CorIFA	Frequência à esquerda	Subcorpus BAWE	Frequência à esquerda
<i>Advérbio 1</i>	<i>responsibly</i>	0	<i>fully</i>	0
<i>Advérbio 2</i>	<i>especially</i>	1	<i>highly</i>	0
<i>Advérbio 3</i>	<i>technologically</i>	0	<i>continuously</i>	1
<i>Advérbio 4</i>	<i>quickly</i>	0	<i>only</i>	1
<i>Advérbio 5</i>	<i>only</i>	1	<i>properly</i>	0

Fonte: Elaborado pela autora, 2017.

Um dos verbos de maior frequência normalizada no subcorpus do CorIFA, *develop* coloca com poucos advérbios *-ly* tanto no subcorpus do CorIFA quanto no subcorpus do BAWE. Identificamos na posição esquerda apenas dois: *especially* e *only*, ambos semanticamente classificados como advérbios restritivos. Estes têm como objetivo enfatizar uma parte da proposição, direcionando a ela um valor de verdade.

Vejamos as linhas de concordância nas quais ocorrem:

(1) *It's useful in some occasions **especially developing** new technology.*

(2) *Penicillin was **only developed** because Alexander Fleming dreamed about a new and effective drug.*

Na linha 1, quando verificamos no subcorpus a sentença que antecede a exposta, verificamos que o aprendiz se refere à importância e limitação do conhecimento acadêmico. Para ele, o conhecimento acadêmico pode ser útil em algumas situações e, “em especial”, para propiciar o desenvolvimento de tecnologia. Notamos que há a necessidade de uso preposicional entre *especially* e *developing*, porém esta foi suprimida pelo aprendiz. Já na linha 2, o aprendiz restringe o desenvolvimento da penicilina ao fato de Alexander Fleming ter sonhado com essa nova e efetiva droga, como se fora essa situação, não houvesse possibilidade da droga ser elaborada.

No subcorpus do BAWE, na posição esquerda, há o uso de *continuously* e *only*, classificados semanticamente como advérbio de modo e advérbio de restrição, respectivamente:

(1) *Further, my need to **continuously develop** as well as my personality, education and the Portuguese culture [...].*

(2) *Furthermore a child will **only develop** their writing skills after many years of practice and this runs somewhat parallel to the development of their speech and vocabulary.*

No exemplo 1, a colocação em destaque é usada numa sentença em que há o uso de primeira pessoa do singular, aspecto que comumente não encontramos nos textos acadêmicos escritos. O advérbio *-ly* utilizado, classificado semanticamente como advérbio de modo, sucede a preposição *to*, numa oração que não requer

objeto. Verificando a sentença, parece se tratar de texto escrito por um não falante de inglês como L1, podendo fazer parte dos 30% dos textos do BAWE que não foram produzidos por falantes de inglês como L1. Essa nossa hipótese se deve pela menção à cultura portuguesa como sendo a sua cultura de origem.

Como no exemplo 1 (*further*), coincidentemente, no exemplo 2, encontramos outra conjunção aditiva (*furthermore*) iniciando a sentença. Nesse caso, a colocação é usada numa oração transitiva direta, na qual *only* sucede o verbo modal *will*.

Em relação à posição direita, encontramos diferentes usos de advérbios, como pode ser visualizado no Quadro 18:

Quadro 18– Frequência para verbo *develop*

ADV	Subcorpus CorIFA	Frequência a direita	Subcorpus BAWE	Frequência a direita
<i>Advérbio 1</i>	<i>responsibly</i>	2	<i>fully</i>	2
<i>Advérbio 2</i>	<i>especially</i>	0	<i>highly</i>	2
<i>Advérbio 3</i>	<i>technologically</i>	2	<i>continuously</i>	0
<i>Advérbio 4</i>	<i>quickly</i>	1	<i>only</i>	0
<i>Advérbio 5</i>	<i>only</i>	0	<i>properly</i>	1

Fonte: Elaborado pela autora, 2017.

No subcorpus do CorIFA, *responsibly*, *technologically* e *quickly*, classificados como advérbios de modo, retratam como uma determinada ação é desempenhada (BIBER et al, 1999). Mostramos a seguir uma linha constando uma colocação em que um desses advérbios foi usado:

(1) *Since the invention of the first vapor machine, technology has **developed quickly** and has brought a lot of facilities to our lives.*

Empregada numa oração no tempo verbal presente perfeito, a colocação em destaque na sentença acima é usada para se referir ao modo como a tecnologia tem se desenvolvido desde um período específico de tempo. Convém lembrar que, como já pontuamos anteriormente, o substantivo *technology* é um dos colocados mais frequentes do verbo *develop* no subcorpus do CorIFA, talvez por conta dos temas das tarefas escritas ou por ser o substantivo muito utilizado em português combinado a esse verbo, quando fazemos referência ao “desenvolvimento de tecnologia”.

Vejamos o comportamento colocacional dos advérbios *-ly* com o verbo *improve*,

a começar pelo Quadro 19:

Quadro 19– Frequência para verbo *improve*

ADV	Subcorpus CorIFA	Frequência à esquerda	Subcorpus BAWE	Frequência à esquerda
<i>Advérbio 1</i>	<i>surely</i>	2	<i>greatly</i>	2
<i>Advérbio 2</i>	<i>significantly</i>	0	<i>significantly</i>	2
<i>Advérbio 3</i>	<i>daily</i>	0	<i>specifically</i>	1
<i>Advérbio 4</i>	<i>highly</i>	2	<i>actually</i>	1
<i>Advérbio 5</i>	<i>certainly</i>	0	<i>only</i>	1

Fonte: Elaborado pela autora, 2017.

No subcorpus do CorIFA, o verbo *improve* é o de maior frequência normalizada (109,2), classificado por Biber et al. (1999) quanto ao seu domínio semântico como verbo de simples ocorrência (haja vista relata ações nãovolitivas, isto é, que não acontecem espontaneamente). Com uma alta recorrência (95 ocorrências) de uso em construções com infinitivo (*to improve*), seus colocados mais frequentes são: *the* (181), *and* (127), *to* (126), *of* (93), *in* (65), *technology* (43), *our* (36) e *quality* (35). Há apenas quatro ocorrências de colocações do tipo “advérbio *-ly* + verbo” no subcorpus, o que sugere que os aprendizes fazem pouco uso de advérbios desse tipo combinados aos lemas do verbo *improve*. Observemos as ocorrências de *surely* e *significantly*, encontradas na posição esquerda:

(1) *The concept of sustainable development, although it started later, is a advance in technology and will **surely improve** the world.*

(2) *As a result, the implementation of that technology would **improve significantly** the electoral process around the world.*

No exemplo 1, vemos o advérbio *surely*, classificado como advérbio de posicionamento, na categoria semântica de posição epistêmica. O advérbio evidencia a realidade de uma proposição, marcando o nível de certeza de acontecimento de um dado evento ou situação, na medida em que o aprendiz coloca um posicionamento avaliativo em relação ao sujeito da oração que, nesse caso, é o “conceito de desenvolvimento sustentável” e, dessa forma, se posiciona frente a proposição. A esse respeito, Hyland (2008, p.12, tradução nossa)⁹⁷ cita os advérbios de natureza epistêmica como uma das

⁹⁷ No original: “[...] stance is impersonally expressed” (HYLAND, 2008, p. 12).

formas a partir das quais a “posição é expressa impessoalmente”.

No exemplo 2, o advérbio *significantly*, diferentemente dos outros exemplos que já mostramos até aqui, aparece como colocado na posição direita do nóculo. Notamos que o advérbio pode sugerir duas categorias semânticas, dentre as cunhadas por Biber et al. (1999): a primeira seria a de posição de opinião, em que o aprendiz expressa sua avaliação positiva sob a proposição, enquanto a outra seria a de amplificador/intensificador, na qual se percebe um aumento da intensidade da escala, na medida em que há uma construção condicional; o aprendiz se refere a algo que melhoraria consideravelmente “o processo eleitoral por todo o mundo”, se uma outra ação acontecesse, nesse caso, “a implementação daquela tecnologia”.

Verificando outros advérbios *-ly* em colocações com o verbo *improve*, encontramos:

(3) *I don't believe that people who spend much of their time studying and trying to get a prominent academic life would do such thing as going to another country with their government's money, only to have fun and not **actually improve** his academic and professional skills*

(4) *Based on these arguments, I think that all this technology in the world today has **really improved** the way that we live.*

(5) *Non-electronic products can be **technologically improved** too.*

Percebemos que, no exemplo 3, há o uso do pronome de primeira pessoa do singular (*I don't believe*), o que sugere uma intenção do aprendiz em mostrar, mais explicitamente, a evidência da proposição. Halliday (2004, p.613-625) chama as construções do tipo *I believe, I think*, que carregam funções semelhantes aos adjuntos modais, de “metáforas de modalidade”. Parece haver entre os aprendizes, não falantes de inglês como L1, uma tendência de uso dessa construção típica do registro oral, em sua escrita, assim como outros estudos já confirmaram (ALTENBERG, 1998; HERRIMAN; ARONSSON, 2009). Como o advérbio *surely*, mostrado no exemplo 1, *actually* também é classificado como advérbio de posicionamento e é usado para marcar uma posição epistêmica por parte do aprendiz. Na classificação de Biber et al. (1999), *actually* se configura como um advérbio de posicionamento, no qual o falante/escritor faz referência ao valor de verdade da proposição, como se quisesse reforçar seus aspectos de certeza e realidade. Para os autores, “[...] os falantes usam advérbios de posicionamento para se referir a seus julgamentos e atitudes, denunciar a natureza

fatídica do que está sendo dito e para marcar exatamente como eles significam seus enunciados” (BIBER et al, 1999, p.766-767, tradução nossa)⁹⁸. Embora ainda seja prematuro afirmar, parece que os advérbios de posicionamento podem auxiliar no entendimento de como os aprendizes, nas redações argumentativas, se colocam avaliativamente na argumentação, já que o papel desses advérbios consiste em expressar a avaliação do falante/escritor em relação à proposição central (BIBER, 2006).

Por sua vez, no exemplo 4, assim como no exemplo 2, há o uso do pronome de primeira pessoa do singular “I”, seguido de *think*, verbo que, segundo Biber et al. (1999), não é comum no registro acadêmico, mas sim no registro oral. O advérbio *really*, similarmente aos advérbios *surely* e *actually* (exemplos 1 e 2, respectivamente), é classificado como advérbio de posicionamento (BIBER, 2006), ou seja, é usado pelo aprendiz com uma construção no presente perfeito simples, no intuito de evidenciar uma realidade da qual concorda e, por assim dizer, insere-se na categoria semântica de posição epistêmica.

No exemplo 5, encontramos o advérbio de modo “*technologically*” que, a nosso ver, constitui uma forma de expressar “*in a technological way*”. A colocação *technologically improved* faz parte de uma construção passiva, a partir do uso do verbo modal *can*, na qual há uma tentativa do aprendiz de manifestar a forma como uma ação pode ser desempenhada.

Diante desses dados, vemos que o número de ocorrências das colocações formadas por “advérbio -ly+ verbo”, com o verbo *improve*, foi bastante baixa. No subcorpus do BAWE, encontramos 13 ocorrências, dentre as quais mostraremos algumas a seguir:

(1) *The primary aim is to **specifically improve** the health of the least well off.*

(2) *Oral corticosteroids: a single systematic review of short-term (2-4 weeks treatment) randomised controlled trials has shown when compared to placebo, oral corticosteroids **significantly improve** lung function.*

(3) *Systematic review evidence has indicated that inhalation of β_2 agonists over a period of 1 week - 6 months **significantly improves** FEV 1 and symptoms of dyspnoea when compared to placebo 1.*

⁹⁸ No original: “Speakers use stance adverbials to convey their judgments and attitudes, to claim the factual nature of what they are saying, and to mark exactly how they mean their utterances to be understood” (BIBER et al., 1999, p. 766-767).

(4) *Kidde has improved significantly over the year, generating more sales and becoming more profitable.*

A colocação *specifically improve*, no exemplo 1, é usada para descrever um objetivo que se deseja alcançar, já que está antecedida da construção “*The primary aim is to*”. O advérbio *specifically*, categorizado semanticamente como advérbio de restrição, é utilizado para modificar um elemento específico da oração (que nesse caso é o lema *improve*).

Nos exemplos 2, 3 e 4, encontramos o advérbio *significantly*, como colocado do verbo *improve*. Os exemplos 2 e 3 parecem ser provenientes da área das Ciências Naturais, uma vez que abordam questões ligadas à saúde. No exemplo 2, o aprendiz se reporta ao uso oral de corticóides que, segundo ele, pode ajudar consideravelmente o pulmão a exercer sua função. O exemplo 3 parece retratar um experimento a respeito de uma determinada substância que, após um certo período de uso (*over a period of 1 week - 6 months*), provocou melhorias na saúde de alguém. Em ambas as linhas, na função de amplificador/intensificador, o advérbio junto ao verbo demonstra a eficácia e efeito positivo de um tratamento, no decorrer de um período de tempo específico (*2-4 weeks*). No exemplo 4, a categoria semântica do advérbio é a mesma encontrada nos exemplos 2 e 3, ou seja, é um advérbio de modo, porém, o contexto em que a colocação é produzida é diferente. Essa linha de concordância deve ter sido, provavelmente, originada de um texto da área de Administração, Economia ou Finanças, já que no campo semântico, há referência às palavras *sales* (vendas) e *profitable* (lucrável). Um ponto importante acerca desse exemplo nos aponta para a delimitação do período de tempo (*over the year*), assim como nos exemplos 2 e 3, o que sugere que o uso dessa colocação, no discurso acadêmico, pode estar diretamente ligado à referência de tempo.

Ainda demos conta de que parece haver no subcorpusdo BAWÉ a mesma tendência que encontramos no subcorpus do CorIFA quanto ao lema com o maior número de ocorrências antecedido de advérbio: *improve*. O lema *improve* possui 188 ocorrências, ao passo que as ocorrências para os outros lemas foram: *improves* (13), *improving* (25) e *improved* (51).

Por ser um verbo ergativo, *improve* pode exercer a função de verbo transitivo ou intransitivo. Nos casos do subcorpus do CorIFA, notamos que, das 277 ocorrências do VIAE *improve*, poucas fazem parte de orações intransitivas, o que significa dizer que há uma preferência entre os aprendizes ao uso do verbo com orações transitivas, isto é, que

requerem objeto, como podemos perceber nos exemplos a seguir:

(1) *A simple example is to do exercises to **improve** our health.*

(2) *Technology has **improved** our methods of interacting with one another.*

(3) [...] *we have an amazing group of instruments to **improve** our knowledge.*

(4) *Technology has the goal of facilitate and **improve** any process and create new products using science.*

Já entre o subcorpusdo BAWE, encontramos também poucas ocorrências de orações intransitivas, apenas nove, considerando seu número de palavras (duas vezes maior que o subcorpus do CorIFA), sendo que, em quatro linhas de concordância, as orações intransitivas fazem parte de construções na voz passiva, como em: *can be improved* e *could be improved*, fenômeno linguístico que não encontramos no subcorpus do CorIFA. As linhas de 1 a 4 trazem tais orações:

(1) *Recommendations will be made on how performance **can be improved**.*

(2) *Rather, by relying in some policy rules, economic performance **can be improved**.*

(3) *The accuracy of the experiments **could be improved** in several ways.*

(4) *The experimental method **could be improved** by melting and forming the compound in a material with a much higher melting point [...].*

Contrastando os resultados dos usos de *improve* na voz ativa e passiva no subcorpus do BAWE, já que no subcorpus do CorIFA não houve ocorrências de construções passivas com o verbo, percebemos que, nas linhas 1 e 2, as construções passivas são posicionadas sintaticamente no final da sentença, ao passo que nas linhas 3 e 4, elas são acompanhadas de preposição (*in* e *by*). Outra observação importante diz respeito aos substantivos aos quais as construções se referem. Nas linhas 1 e 2, a forma passiva *can be improved* faz referência ao substantivo *performance*, enquanto nas linhas 3 e 4, *could be improved* diz respeito aos experimentos e/ou métodos que “poderiam ser

melhorados/implementados”.

Verificando ambos os subcorpora, reconhecemos, também, que a pouca variação lexical quanto ao uso dos advérbios com o verbo *improve* pode ser um indício de que os aprendizes brasileiros usam poucos advérbios antecedendo-o, apesar de percebermos que, no subcorpus do BAWE, há uma variedade um pouco maior de advérbios; o único comum aos dois subcorpora é *actually*. Outra questão que merece ser pontuada diz respeito ao uso do pronome de primeira pessoa do singular “I”, como vimos em dois exemplos no subcorpus do CorIFA.

Vejamos outras linhas de concordância do subcorpus do BAWE:

(5) *To **increasingly improve** model C, we could think about other variables that may also affect the outcome of exam performance.*

(6) *The effects of price discriminating policies are fairly ambiguous, but it has been argued that they can **actually improve** efficiency levels.*

Na linha 5, como colocado do VIAE *improve*, tem-se o advérbio intensificador *increasingly*, usado para marcar o grau de aumento de uma dada característica. Já na linha 6, com o uso do advérbio modalizador epistêmico *actually*, há uma tentativa do aprendiz em qualificar um certo estado de coisas, na medida em que manifesta certeza.

As linhas 7 e 8 trazem as colocações *greatly improve* e *greatly improved*. O advérbio *greatly*, classificado como advérbio de modo, descreve a ação.

(7) *Also having more than one receiver would **greatly improve** the accuracy of the robot.*

(8) *For most animals stereotyping can be **greatly improved** if their environment improves, suggesting they retain the capacity for normal behaviour (Tudge, 1991).*

Nossa pesquisa revela que *develop* e *improve* são os verbos de maior frequência normalizada no subcorpus do CorIFA e no subcorpus do BAWE. A partir dos quadros apresentados, interessante enfatizar que são também os verbos que aparecem com mais advérbios nos subcorpora, o que nos leva a crer que tanto os aprendizes brasileiros quanto os falantes de inglês como L1 usam largamente esses dois verbos, bem como utilizam com frequência as colocações lexicais – “advérbio -ly + verbo” em suas

redações argumentativas.

Vejam os a frequência de increase no Quadro 20:

Quadro 20 – Frequência para verbo *increase*

ADV	Subcorpus CorIFA	Frequência à direita	Subcorpus BAWE	Frequência à direita
<i>Advérbio 1</i>	-	0	<i>dramatically</i>	4
<i>Advérbio 2</i>	-	0	<i>steadily</i>	3
<i>Advérbio 3</i>	-	0	<i>greatly</i>	3
<i>Advérbio 4</i>	-	0	<i>slightly</i>	3
<i>Advérbio 5</i>	-	0	<i>only</i>	1

Fonte: Elaborado pela autora, 2017.

Podemos perceber que os advérbios *-ly* combinados ao verbo *increase* no subcorpus do BAWE aparecem na posição direita, ao passo que, no subcorpus do CorIFA, nenhum advérbio *-ly* ocorre com o referido verbo.

Vejam um dos usos de *dramatically* no subcorpus do BAWE:

(1) *Although significant acquisitions have been made, Gearing has not **increased dramatically** and remains in relatively low level.*

O exemplo mostra a combinação *increased dramatically* sendo usada numa construção passiva, posicionada sintaticamente, após uma oração que expressa uma ideia de concessão, a partir do uso da conjunção concessiva (*although*). O advérbio *dramatically*, classificado semanticamente como advérbio de modo, foi usado pelo aprendiz proficiente para se referir a um substantivo próprio (*Gearing*) que, segundo ele, não “sofreu aumento drástico”. Vemos que a combinação parece expressar o sentido de “aumentou significativamente”.

No subcorpus do BAWE, há diversas linhas de concordância com o advérbio *slightly*, semanticamente classificado como advérbio de grau/extensão do tipo diminuidor (BIBER et al., 1999), contendo números, porcentagens e fórmulas, como nas linhas de concordância 1 e 2 abaixo, como também constando a referência a taxa (rate), tal qual o exemplo 3 abaixo. Tais linhas estão dispostas a seguir:

(1) [...] *dividends have only **increased slightly** inline with sales (a 6% increase from 2002 to 2003).*

(2) *This value **increases slightly** in 2003.*

(3) *The flow rate of water was **increased slightly** so that transitions could be established and the minimum fluidisation velocity of the particles could be determined.*

Já na posição esquerda, no subcorpus do CorIFA, encontramos duas ocorrências de combinações de *increase* com advérbios *-ly*: *significantly increased* e *immensely increased*, conforme vemos no lado esquerdo do quadro abaixo:

Quadro 21– Frequência para verbo *increase*

ADV	Subcorpus CorIFA	Frequência à esquerda	Subcorpus BAWE	Frequência à esquerda
<i>Advérbio 1</i>	<i>significantly</i>	1	<i>mutually</i>	1
<i>Advérbio 2</i>	<i>immensely</i>	1	<i>greatly</i>	1
<i>Advérbio 3</i>	-	0	<i>gradually</i>	1
<i>Advérbio 4</i>	-	0	-	0
<i>Advérbio 5</i>	-	0	-	0

Fonte:Elaborado pela autora, 2017.

No CorIFA, duas combinações de advérbios *-ly* com *increase* foram encontradas. Vejamos:

(1) *Nowadays the life expectation has been **significantly increased** [...].*

(2) *Since then, the variety and quality of devices has **immensely increased** in a way that it is hard to understand from where we came.*

Nas linhas 1 e 2 acima, vemos que as combinações são usadas em construções passivas. Na linha 1, o verbo *increase* está relacionado à “expectativa de vida” que, segundo o aprendiz, tem “aumentado significativamente”. Na linha 2, há o uso do advérbio *immensely*, cujo uso apenas encontramos duas ocorrências no subcorpus do BAWE combinado com o lema *benefited*, mas apenas três ocorrências, na combinação com adjetivos (e.g *uncompassionate, popular, strong* e *pathetic*), o que pode nos indicar que o uso do advérbio não é tão comum ao subcorpus, tampouco combinado a verbos.

No subcorpus do BAWE, destacamos as ocorrências:

(1) *This light is said to be linearly polarized because the electric field vector traces out*

a straight line from the maximum amplitude to the minimum amplitude as the two components **mutually increase and decrease**.

(2) *Women remained subordinate and failed to **greatly increase** their influence within the world of governance [...].*

Na linha 1, o advérbio de modo *mutually* é combinado ao VIAE *increase* para fazer referência a componentes que crescem e diminuem concomitantemente, numa escala de máxima e mínima amplitude. Acreditamos que a linha pode ser advinda de um texto escrito por um aprendiz proficiente, pertencente a algum curso da área de Exatas e, talvez, o tempo verbal presente seja justificado por isso, uma vez que pode ser comum nessa área o uso de descrições técnicas utilizando-se do tempo verbal presente.

Na linha 2, *greatly*, como advérbio intensificador, opera como elemento que potencializa o verbo. Nesse caso, é feita uma referência a permanente subordinação e falha feminina para aumentar sua influência no mundo da governança.

Assim, percebemos que o uso de *increase* com os advérbios *-ly* acontece em menor frequência no subcorpus do CorIFA quando comparado ao subcorpus do BAWE. No subcorpus do BAWE, do ponto de vista do posicionamento sintático, os advérbios *-ly* são mais frequentes à direita do VIAE *increase*.

Outro VIAE listado entre os cinco de maior frequência no CorIFA está *provide*.

Quadro 22– Frequência para verbo *provide*

ADV	Subcorpus CorIFA	Frequência à esquerda	Subcorpus BAWE	Frequência à esquerda
Advérbio 1	<i>consequently</i>	2	<i>only</i>	4
Advérbio 2	-	0	<i>effectively</i>	0
Advérbio 3	-	0	<i>mainly</i>	1
Advérbio 4	-	0	<i>simply</i>	1
Advérbio 5	-	0	<i>really</i>	1

Fonte: Elaborado pela autora, 2017.

No subcorpus do CorIFA, há apenas uma ocorrência de um advérbio *-ly* que ocorre na posição esquerda a ele: *consequently*. Vejamos como acontece essa combinação:

(1) *Westcott (2012) claims that the resistance training reverse muscle loss caused by aging, facilitate physical function, recharge resting metabolism and **consequently provide** a reducing of the body fat.*

É possível perceber que a combinação *consequently provide* é usada para referir-se à “perda muscular”, aspecto abordado pelo citado teórico (Westcott). Segundo ele, tal perda traz alguns benefícios e, como “consequência disso”, aponta a redução da gordura corpórea. Por seu turno, no subcorpus do BAWE, não encontramos ocorrências dessa combinação, mas na posição direita dois advérbios *-ly* (*effectively* e *only*) se combinam ao VIAE, como nos mostra o quadro:

Quadro 23 – Frequência para verbo *provide*

ADV	Subcorpus CorIFA	Frequência à direita	Subcorpus BAWE	Frequência à direita
Advérbio 1	<i>consequently</i>	0	<i>effectively</i>	2
Advérbio 2	-	0	<i>only</i>	2
Advérbio 3	-	0	<i>mainly</i>	0
Advérbio 4	-	0	<i>simply</i>	0
Advérbio 5	-	0	<i>really</i>	0

Fonte: Elaborado pela autora, 2017.

Vejamos na linha abaixo um dos usos de *effectively* combinado à *provide*:

(1) [...] *McDonagh's purpose seems to have been to **provide effectively** what constituted as an artistic blueprint of the original performance [...]*.

A combinação *provide effectively* é antecedida de uma construção na qual pensamos que o aprendiz proficiente tentava usar o tempo verbal presente perfeito, mas, ao invés disso, usa uma construção no infinitivo. O advérbio *effectively*, classificado como advérbio de modo, “[...] expressa informação sobre como uma ação é desempenhada” (BIBER et al., 1999, p.553, tradução nossa)⁹⁹. Nesse caso, a ação desempenhada diz respeito a uma impressão de que o objetivo de alguém parece ter sido o de fornecer algo, de maneira efetiva.

Diante dos dados que apresentamos nesta seção, podemos afirmar que os advérbios *-ly* que se combinam a *adopt*, *develop*, *improve*, *increase* e *provide* parecem ser mais frequentes na posição esquerda no subcorpus do CorIFA e na posição direita no subcorpus do BAWE. Esses resultados nos ajudam a entender que pode haver uma preferência dos estudantes universitários proficientes em escrita acadêmica a usar os advérbios *-ly* na posição direita, quando combinados a verbos lexicais.

⁹⁹ No original: “[...] express information about how an action is performed” (BIBER et al, 1999, p.553).

Os advérbios e sua classificação semântica

O advérbio que mais coloca com os VIAES, no subcorpus do BAWE, constitui o advérbio de restrição *only*, usado com 17 VIAEs (*achieve, acquire, affect, apply, contribute, depend, develop, exist, improve, include, increase, involve, observe, occur, produce, provide e require*). Esse tipo de advérbio, segundo Biber et al. (1999), é tipicamente utilizado pelos falantes para restringir uma dada informação/posição.

Os advérbios *-ly* colocam em menor frequência no subcorpus do CorIFA, bem como no subcorpus do BAWE, com os verbos *achieve, acquire e adopt*. Já com *affect*, no subcorpus do CorIFA, há 10 colocações, ao passo que há maior variedade de advérbios *-ly* na combinação com os VIAEs, no subcorpus do BAWE, representando o dobro de ocorrências, em relação ao subcorpus de estudo. Deles, há três em comum entre os dois subcorpora: *adversely, directly e mainly*, classificados por Biber et al. (1999) como advérbio de contraste (no qual informações são contrapostas), advérbio de modo (no qual descreve-se a forma como dada situação acontece) e advérbio de posicionamento epistêmico (nesse caso, usado para mostrar certas limitações em relação à proposição), respectivamente.

As colocações formadas por “advérbio *-ly* + *apply*” são de diversas categorias semânticas no subcorpus do BAWE. Eis alguns exemplos:

(1) *Through this approach, judges can either choose between two conflicting definitions, as used in Adler v. George [1964] or **simply apply** the legislation[...].*

(2) *Expedia **successfully applied** a combination of sound, graphics and videos to offer customer a virtual tour of hotel room in their website.*

(3) *This defence could be **potentially applied** to a 'site manager' case.*

(4) *A common criticism of Liberalism hinges on the view that Liberalism's emphasis on individual rights and freedoms is a conceptual idea which cannot be **universally applied** to all cultures and corners of the world.*

Em todos os quatro exemplos, a posição sintática do advérbio é de antecessor do verbo. Nos exemplos 1 e 2, o verbo *apply* funciona como verbo transitivo e é usado no sentido de “aplicar algo”, o que requer um objeto direto. Mais especificamente, no

exemplo 1, o uso da colocação *simply apply*, formada pelo advérbio de restrição *simply*+ VIAE, está atrelado à ideia de aplicação de leis, reforçada pelo uso dos substantivos *judges* e *legislation*. No exemplo 2, a colocação *successfully applied*, formada pelo advérbio de modo *successfully* + VIAE, faz referência ao sucesso na combinação de elementos como som e vídeos, usada para compor um determinado *site* de oferta de hotéis.

Já nos exemplos 3 e 4, nas colocações *potentially applied* e *universally applied*, ambas acompanhadas da preposição *to* e usadas em construções passivas, o verbo *apply* funciona como verbo intransitivo.

O lema *applying* não aparece em nenhuma colocação com advérbios *-ly* no subcorpus do BAWE. Por seu turno, o lema *applies* (com frequência de 20 ocorrências no subcorpus) apresenta apenas uma colocação com advérbio *-ly*, sendo usado para se referir a um conceito que não se aplica diante de uma situação pontual:

(1) *The concept of validity only **really applies** when there is a single testable 'truth' [...].*

No subcorpus do CorIFA encontramos apenas uma ocorrência com *especially*, considerado advérbio de restrição na classificação semântica de Biber et al. (1999). No exemplo abaixo, o advérbio apresenta desvio ortográfico (já que sua grafia correta é *especially*), provavelmente por conta de o aprendiz conhecer o adjetivo *special* e, assim, atribuir a ele a derivação do advérbio.

(1) *A continuous increment of government policies for environment education, **specialy applied** to children, may absolutely produce a better scenario in a long run.*

Quando buscamos, nos dois subcorpora, a distribuição dos advérbios do tipo intensificadores, encontramos:

Quadro 24– Distribuição dos advérbios do tipo intensificadores

Advérbios de grau: intensificadores	Subcorpus do CorIFA			Subcorpus do BAWE		
	Colocados Verbos em geral	Colocados Verbos lexicais	Colocados VIAEs	Colocados Verbos em Geral	Colocados Verbos lexicais	Colocados VIAEs

<i>highly</i>	17	9	0	111	43	4
<i>deeply</i>	5	3	0	7	6	2
<i>strongly</i>	17	17	4	36	26	2
<i>considerably</i>	4	4	2	18	7	1
<i>largely</i>	11	8	2	71	47	6
<i>extremely</i>	51	5	1	73	13	0

Fonte: Elaborado pela autora, 2017.

Como mostrado no Quadro 24, os aprendizes usam com pouca frequência os advérbios *considerably*, *deeply* e *largely*. Os advérbios *highly* e *deeply* apresentam baixa frequência, colocando em menor número com os verbos lexicais, ao passo que não colocam com nenhum VIAE. Em contrapartida, no subcorpus do BAWE, há amplo uso de *highly* como colocado de outras classes de palavras (adjetivos e outros advérbios), contudo, combinam apenas com quatro VIAEs: *affect*, *base*, *depend* e *develop*, formando as combinações *highly affected*, *highly based*, *highly depends* e *highly developed*.

(1) *It should be noted that the calculation of the particular ratio is **highly affected** by the industry trends in inventory management.*

(2) *Knowing that the concept is **highly based** on experimentation, it seems sufficient to have basic knowledge of bakery production, to develop a new product.*

(3) *This essay analysed the merits and disadvantages of two stabilisation mechanisms used by Central Banks, fixed exchange rates and inflation targeting, concluding that it is difficult to determine the superiority of one mechanism over the other as their success **highly depends** on the nation's economic state prior to its implementation.*

(4) *Therefore, the archaeology of the Viking Age, weather settlements or ships, is highly developed in Denmark, Norway and Sweden.*

Por sua vez, *deeply* é o maximizador de menor frequência no subcorpus do BAWE, com sete ocorrências, incluindo, duas combinações com VIAEs (*deeply discussed* e *deeply influence*), conforme mostrado nos exemplos a seguir:

(1) *When one issue is deeply discussed, has the social acceptance and support, another arises.*

(2) *However, the industry structure is not static and the actions of the different players will deeply influence the evolution of industry.*

O uso de *strongly* no subcorpus do CorIFA nos chama a atenção para o fato de que, em todas as ocorrências do advérbio, este coloca com verbos lexicais, sendo que dois deles são VIAEs (*strongly affected* e *strongly influence*). No subcorpus do BAWE, a frequência absoluta do advérbio é de 36 e, em 26 ocorrências, coloca com verbos lexicais, incluindo três VIAEs: *affect*, *influence* e *suggest*, formando as colocações *strongly affected*, *strongly influence*, *strongly suggest* e *strongly suggesting*, conforme mostramos nos exemplos abaixo:

(1) *Thus the flow of a fluid is **strongly affected** by its viscosity.*

(2) *It has long been recognised that emotions exert powerful effects on memory, and that what we feel can **strongly influence** what we remember.*

(3) *Evidence in figure 6, **strongly suggests** that Wnt2b and Wnt8c signalling is mediated via catenin activation in the fore- and hindlimb respectively.*

(4) *Together with other lexical choices **strongly suggesting** falsehood[...].*

Já *largely*, também classificado como advérbio maximizador, apresenta frequência muito baixa no subcorpus do CorIFA, enquanto no subcorpus do BAWE há uso mais frequente, com 47 combinações com verbos lexicais e seis combinações com VIAEs (*largely depends*, *largely depend*, *largely influenced*, *depended largely*, *focuses largely*, *contributed largely*). Observemo-nos:

(1) *[...]the success of the implementation the new rights to information and consultation **largely depends** on the awareness employees have of their existence and the ways to set procedures in motion.*

(2) *It seems evident then that the successful implementation of the ICE regulations would **largely depend** on the sufficient expertise of the employee representatives balloted by their fellow workers.*

(3) *Grassland species structure and composition is also **largely influenced** by activities of herbivores.*

(4) *Rather, the type of policies, whether helpful or repressive, **depended largely** on the category of the poor the policy was directed at.*

(5) *Through out the writing off this essay the impression gained from the media studied has been one that **focuses largely** on the negative aspects [...].*

(6) *Since we still know relatively little about it, the discovery of Graveney boat **contributed largely** to the understanding of the craft of shipbuilding.*

Interessante notar que nas linhas de concordância 1, 2 e 3, o advérbio *largely* é usado como colocado dos VIAEs *depend* e *influence*, se encontra na posição sintática à esquerda dos verbos. Esse dado pode nos informar sobre a provável preferência, no subcorpus do BAWE, do uso do advérbio *largely* na posição esquerda ao verbo. Quando checamos o referido subcorpus, vemos que, das 71 ocorrências do advérbio, em 32 delas, *largely* aparece na posição esquerda dos verbos lexicais, fato que talvez possa justificar seu uso nas linhas de 1 a 3. Ainda, os verbos *depend* (linhas 1 e 2) e *influence* (linha 3) são acompanhados de preposições (*on* e *by*). As colocações *largely depends* e *largely depend* parecem atribuir ideia de sucesso na implementação de algo, como pode ser comprovado pelas expressões: *the success of the implementation* (linha 1) e *the successful implementation* (linha 2). Na linha 3, a colocação *largely influenced* encontra em uma construção passiva, na qual há o sentido de “sofrer a influência de”. Nesse caso, “as atividades de herbívoros” sofrem influência “da estrutura e composição de espécies de plantas”.

Nas linhas de 4 a 6, observamos a posição contrária do advérbio, ou seja, aparece à direita dos VIAES *depend*, *focus* e *contribute*. Na linha 4, a colocação *depended largely* é utilizada para se referir ao tipo de medida à qual algo/alguém se baseava/se apoiava amplamente. Já na linha 5, a colocação *focuses largely* se reporta à impressão obtida a partir de uma determinada mídia, na qual aspectos negativos são colocados em evidência. Na linha 6, é atribuída à descoberta do barco Graveney a compreensão da embarcação da construção naval.

Possivelmente, a variedade de palavras com as quais os aprendizes combinaram *highly*, *strongly*, *largely* e *extremely* sugere seu amplo uso com adjetivos e, até mesmo, outros advérbios. Tomando como exemplo *extremely*, notamos que, em ambos os subcorpora, tal advérbio coloca frequentemente com os adjetivos *necessary* e *important* (subcorpus do CorIFA), *difficult*, *important* e *different* (subcorpus do BAWE). Dessa

forma, parece haver entre os estudantes uma preferência pelo uso desse advérbio com outras classes de palavras em detrimento dos verbos.

Vejamos os exemplos 1 e 2 extraídos do subcorpus do CorIFA:

(1) *The cost-benefit is an extremely important factor in any project, because for any company as there is less spending greater is their profits.*

(2) *The exchange of knowledge is extremely important and necessary to be improved, but today it is more restrained.*

E os exemplos 3, 4 e 5 extraídos do subcorpus do BAWE:

(3) *Thus it is extremely difficult to prove the triumph of any particular economic theory.*

(4) *Genetic analysis of each of the island rail populations is extremely important.*

(5) *Yet structurally they are extremely different.*

As linhas de concordância 1 e 2 abaixo mostram as colocações *highly praised* e *strongly agree*, combinações de advérbios do tipo amplificadores com verbos lexicais. A colocação *strongly agree* sugere uma tentativa do aprendiz de defender um posicionamento pessoal, por conta do uso do pronome pessoal de primeira pessoa “I”. Além disso, quando verificamos no subcorpus do BAWE, não encontramos ocorrências de tal colocação.

(1) *Friendship and hospitality are **highly** praised traits in our society.*

(2) *In conclusion I **strongly** agree that technology has made a world a better place to live.*

Nas linhas 3 e 4, os aprendizes combinaram *considerably* e *largely* a lemas do VIAE *contribute* e, na linha 5, a colocação *extremely required*. Vejamos:

(3) *In summary, technological advancement contributes **considerably** to an improvement in our quality of lives.*

(4) *This new technology has **largely** contributed for the decentralization of information.*

(5) *In the case of using the exact same words, putting them in quotation marks are **extremely** required.*

Como mostrado acima, nas linhas 3 e 4, as colocações trazem em seu bojo advérbios intensificadores (*considerably* e *largely*), ao passo que, na linha 5, encontramos o advérbio maximizador *extremely*, amplamente usado pelos estudantes no subcorpus do BAWE, mas que apresenta frequência baixa no subcorpus do CorIFA.

Algumas construções¹⁰⁰ nos chamaram a atenção durante a observação do subcorpus do CorIFA. As “combinações criativas” ou “combinações nãocongruentes”, a exemplo de *badly injured*, *finely detailed* and *widely held*, encontradas por Granger (1998) no ICLE, também nos chamam a atenção na escrita dos aprendizes no subcorpus do CorIFA e, provavelmente, podem soar estranho aos falantes de inglês como L1. São elas: (1) *formulated mathematically*; (2) *fall perpendicularly*; (3) *die needlessly*; (4) *cruelly exploited*; (5) *speak incessantly*. Vejamos as sentenças nas quais essas combinações criativas (em negrito) aparecem no subcorpus do CorIFA:

(1) *Before, for example, Isaac Newton **formulated mathematically** his law of universal gravitation [...].*

(2) *Why do some things always **fall perpendicularly** to the ground?*

(3) *Without these devices, several people **die needlessly**.*

(4) *These categories of animals use to be, respectively, abandoned by their human family and **cruelly exploited** at work.*

(5) *That is justified on the fact that the regular school have monotonous lessons, where the teacher **speak incessantly** for one or two hours [...].*

¹⁰⁰ Além das combinações criativas, mesmo não sendo esse nosso objetivo aqui, achamos curioso o fato da utilização de muitos advérbios *-ly*, no início de orações, como em: (1) ***Obviously**, if a person is welcome in some place [...]*; (2) ***Honestly**, I never thought about this subject before*; (3) ***Undoubtedly**, there are advances which came to [...]*; (4) ***Sincerely**, there is no place to regress*. Nesses exemplos, as palavras em negrito funcionam como adjuntos adverbiais e evidenciam a tentativa do aprendiz em marcar um nível de certeza, demonstrando sua posição na escrita em relação a uma proposição na oração (BIBER et al., 1999).

Não encontramos no subcorpus do BAWE as combinações que destacamos acima; apenas encontramos uma ocorrência dos advérbios *mathematically* e *perpendicularly*, embora somente com este último em uma colocação com verbo lexical (em negrito): “*There is a one refractive index for light entering the crystal that is polarized parallel to the director and another for light **polarized perpendicularly** to the director*”.

Conclusão do capítulo

Os resultados apontaram para uma grande gama de uso de VIAEs (acima do ponto de corte de frequência de 20 palavras por 100.00), ou seja, 23 verbos, no subcorpus do CorIFA, com uma concentração na utilização de sete verbos (acima de 40 palavras por 100,000). Isso nos mostra uma grande diferença em relação ao BAWE (21 VIAEs acima de 20 e 14 VIAEs acima de 40).

Houve preferência pelo uso de verbos, classificados semanticamente como verbos de ação (e.g. *apply* e *provide*). Ainda, os resultados nos informam que há uma preferência dos aprendizes pelo uso de advérbios *-ly*, classificados semanticamente como intensificadores (e.g. *extremely*, *highly* e *strongly*), mas houve baixa ocorrência de advérbios *-ly* coocorrendo com os VIAEs.

Além disso, a força de atração entre os VIAEs e os advérbios não pôde ser testada devido ao baixo número de ocorrências, tais como 0,1,2 (mesmo quando analisados os verbos mais frequentes no CorIFA), o que nos impossibilitou a aplicação do teste estatístico Mutual Information.

CONCLUSÃO

Retomando o verso 4 da epígrafe, “Eu escuto a cor dos passarinhos”, no início desta seção, nos deparamos com uma construção, proferida por uma criança, segundo o poeta, na qual, apesar de sintaticamente apropriada, revela uma combinação fraseológica incomum, se tomada fora do contexto literário. Em outras palavras, podemos dizer que a criança escolhe usar o verbo “escutar”, no tempo verbal presente, sem se dar conta de que tal verbo não coloca com o substantivo “cor”. Esse argumento pode ser justificado ao nos voltarmos para os versos 5 e 6: “A criança não sabe que o verbo escutar não funciona para cor, mas para som”. Nesse momento, o poeta esclarece

que a escolha do verbo “escutar” aconteça por conta do desconhecimento lexical da língua, já que o uso dessa colocação verbo-nominal “escuto a cor” não é típico na língua, apesar de ser possível de realização em textos literários, nos quais as palavras aparecem em sentido figurado e, assim, provocam e evocam outros sentidos.

Escolhemos retratar tais versos da epígrafe justamente porque percebemos que fenômenos linguísticos como este parecem acontecer, como vimos nos resultados desta pesquisa, com muitos verbos lexicais encontrados como mais frequentes no CorIFA. Os aprendizes tendem a fazer escolhas lexicais que se distanciam, muitas vezes, dos usos mais típicos do inglês acadêmico escrito, devido, talvez, à pouca exposição e ao pouco conhecimento das fraseologias da língua.

Como nos lembra Sinclair (1991), o significado de uma palavra bem como seus padrões léxico-gramaticais são interrelacionados, o que nos leva a perceber que o contexto de uso na qual a palavra se insere nos ajuda a compreender as características distintivas entre as escolhas linguísticas dos aprendizes e dos falantes de inglês como L1, como vimos fazendo neste estudo.

A metodologia utilizada na presente pesquisa consistiu na identificação e descrição da frequência e do uso dos Verbos do Inglês Acadêmico Escrito (VIAEs) no subcorpus do CorIFA, tendo como base a Academic Vocabulary List (AVL). Feito isso, nos propusemos a contrastar os resultados referentes a esses verbos com os resultados encontrados no subcorpus do BAWE. Foram identificados 28 VIAEs no subcorpus do CorIFA, dentre os quais elegemos os cinco de maior frequência normalizada nesse mesmo subcorpus, ou seja, *adopt*, *develop*, *improve*, *increase* e *develop* para verificarmos seu uso colocacional com os advérbios *-ly*, tanto na posição esquerda quanto direita.

Na Introdução, apontamos as quatro perguntas de pesquisa norteadoras deste trabalho:

1. Quais verbos são mais frequentes no subcorpus do CorIFA?;
2. Quais desses verbos estão na Academic Vocabulary List(AVL)?
3. Como se dá o uso dos verbos do inglês acadêmico escrito no subcorpus do CorIFA?;
4. Com quais advérbios *-ly*esses verbos coocorrem?

Retomando a primeira delas, notamos que há uma variedade lexical muito grande em relação ao uso de verbos no subcorpus do CorIFA. Encontramos um total de

426 verbos, sendo que os mais frequentes são classificados semanticamente como verbos de ação. Percebemos ainda que há a repetição de muitos desses verbos no subcorpus, o que torna os textos mais repetitivos e parece indicar que a variação lexical é menor.

No que diz respeito à segunda pergunta da pesquisa, podemos afirmar que, considerando o ponto do corte de frequência aqui adotado (20 palavras por 100.000), os verbos mais frequentes no subcorpus do CorIFA foram: *improve, provide, develop, increase, adopt, base, produce, relate, contribute, include, influence, achieve, depend, compare, facilitate, require, involve, exist, acquire, control, affect, observe, occur, apply, promote, reduce, conclude* e *tend*. Como já afirmamos anteriormente, a maior parte desses verbos do inglês acadêmico escrito é cognato e se classifica semanticamente como verbo de ação. Percebemos que a preferência de uso de verbos cognatos pode estar ligado ao pouco conhecimento do aprendiz de outros verbos que são mais tipicamente usados no discurso acadêmico escrito.

A terceira pergunta buscou saber como se dá o uso dos verbos do inglês acadêmico escrito no subcorpus do CorIFA. Os resultados indicaram que os VIAEs são utilizados em muitas construções passivas, sendo que a maioria deles se repete diversas vezes ao longo do texto. Ademais, os lemas de maior predominância foram os utilizados nos tempos verbais presente, passado e particípio passado. Quanto à classificação semântica, os verbos de ação são os mais frequentes dentre os identificados no subcorpus do CorIFA.

Por fim, a última pergunta de pesquisa teve como objetivo conhecer os advérbios *-ly* que colocavam com os verbos do inglês acadêmico escrito anteriormente identificados. Os resultados revelaram que os advérbios *-ly* são largamente utilizados pelos aprendizes, contudo sua frequência de uso com os VIAEs é muito baixa, aspecto que também notamos no subcorpus do BAWE. Apesar disso, os advérbios de conhecimento epistêmico (e.g *really, significantly*) aparecem com mais frequência com tais verbos.

A análise dos dados do subcorpus do CorIFA revelou que os aprendizes brasileiros utilizam muitos verbos do inglês acadêmico escrito em seus textos. Além disso, muitos são os advérbios *-ly* encontrados no subcorpus, contudo, no que se refere à sua combinação com verbos lexicais, a fim de formarem colocações lexicais advérbio-*ly*+ verbo (ou verbo + advérbio), o número de ocorrências é bem menor. Em outras palavras, nossa hipótese perpassava a ideia de que apareceriam colocados (advérbios -

ly) com os VIAEs, mas a frequência foi muito baixa, o que nos motivou a fazer uma discussão qualitativa. De modo geral, os aprendizes parecem conhecer inúmeros verbos do inglês acadêmico escrito, porém seu uso, na maior parte dos casos, apresenta inadequações quanto à escolha dos advérbios *-ly* que melhor combinem com os VIAEs aqui estudados.

Desse modo, nesta pesquisa, buscamos uma melhor compreensão de como os estudantes brasileiros de Inglês para Fins Acadêmicos, mais particularmente, os estudantes da Universidade Federal de Minas Gerais, usam as estruturas de verbos do inglês acadêmico escrito com advérbios de terminação *-ly*.

Esperamos com esta pesquisa provocar reflexões no sentido de colaborar para que os estudantes estejam mais preparados com o uso de determinados padrões léxico-gramaticais em inglês ao adentrarem a universidade e se deparem com a escrita de textos acadêmicos, como, por exemplo, as redações argumentativas.

Limitações do estudo

A maior limitação encontrada neste estudo diz respeito ao tamanho dos subcorpora. Como nos limitamos ao gênero redação argumentativa, isso impactou o tamanho do CorIFA e levou-nos a compilar um subcorpus pequeno do BAWE. Isso nos impôs uma restrição em termos de frequência, o que não possibilitou a identificação de colocações e nem tampouco a força de associação entre os VIAEs estudados e os diversos advérbios *-ly*.

Desse modo, considerando os resultados deste estudo, percebemos que há entre os aprendizes de inglês, representados no subcorpus do CorIFA, uma tendência quanto ao uso de verbos mais comuns e que são pouco usados no inglês acadêmico escrito de falantes de inglês como L1. Esses resultados nos instigam a pensar de que maneira o ensino-aprendizagem de inglês em contextos escolares/acadêmicos pode influenciar nas escolhas léxico-gramaticais dos aprendizes e, desse modo, refletirem na sua interlíngua.

Assim, analisar *corpora* de aprendizes nos possibilita não apenas conhecer os itens lexicais e gramaticais, nesse caso, os mais frequentes verbos que intitulamos de VIAEs, bem como suas colocações com advérbios *-ly*, mas acima de tudo, caracterizar e descrever sua interlíngua, através da escrita de um determinado gênero, produzido em contexto acadêmico.

Por conta disso, os resultados desta pesquisa nos fizeram pensar sobre como a

comparação entre *corpus* de aprendizes e *corpus* composto por textos de estudantes cuja L1, em sua grande maioria, é o inglês, pode oferecer importantes contribuições para o ensino, mais particularmente, de inglês como língua adicional, na medida em que evidenciam as dificuldades encontradas pelos aprendizes e que são, muitas vezes, desconhecidas pelos seus professores. Além disso, uma vez identificadas as dificuldades é possível informá-las ao ensino de Inglês para Fins Acadêmicos e, assim, desenvolver materiais ligados a essa área. Nesse sentido, acreditamos na possibilidade de produção de atividades didáticas ou, até mesmo, na criação de uma lista ou material de consulta sobre os verbos com suas colocações mais frequentes, no intuito de auxiliar os estudantes universitários, aprendizes de inglês, nas escolhas lexicais e gramaticais em sua escrita, que sejam mais apropriadas e condizentes com o uso da língua circulante no discurso acadêmico escrito, na tentativa de contemplar diferentes áreas do conhecimento. Desse modo, a lista poderia ser um ponto de partida para auxiliar os aprendizes na compreensão e uso dessas combinações fraseológicas do inglês em suas produções acadêmicas escritas.

Reppen (2010) nos chama a atenção para pesquisas que têm mostrado a importância do ensino de vocabulário nas aulas de Inglês para Fins Acadêmicos, através de diferentes abordagens. Segundo ela, os professores podem usar a Key Word in Context (KWIC), ferramenta que faz uma busca de palavras em um texto ou conjunto de textos, fazendo com que determinadas palavras sejam listadas em colunas. Essa ferramenta constitui importante recurso para uma atividade de escrita, que tenha como objetivo aumentar a variedade de verbos que os aprendizes usam quando escrevem textos acadêmicos. Ainda, a autora defende que os aprendizes, na escrita acadêmica, precisam incorporar ideias e informações de uma gama de fontes e, desse modo, a escolha lexical dos verbos é crucial.

Apesar das restrições quanto ao número de palavras do subcorpusdo CorIFA, bem como do subcorpusdo BAWÉ, ao qual precisamos compilar para que tivéssemos o balanceamento dos dados, a presente pesquisa constitui um passo importante justamente porque aponta vários aspectos que distinguem os *corpora* de aprendizes do *corpus* de falantes de inglês como L1 quanto ao uso dos verbos típicos do inglês acadêmico escrito e suas escolhas lexicais e gramaticais. Ao mostrarmos como esses usos se efetivam na produção escrita dos aprendizes, estamos buscando melhor compreender os textos que produzem e, dessa forma, pensar em como podemos auxiliá-los na tarefa de chamar a sua atenção para conhecer e entender como operam os usos lexicais e gramaticais do

inglês na escrita acadêmica.

Como parte dos textos que compõem nosso subcorpus do CorIFA não estava acompanhada das suas respectivas notas, não foi possível realizar um estudo contrastivo entre os *low-scoring essays* ou, até mesmo, dos *high-scoring essays*. Assim, poderíamos verificar outras características que não foram consideradas nesta pesquisa, bem como nos possibilitaria ver os dados com um olhar mais refinado, no que diz respeito ao uso dos VIAEs e suas colocações com advérbios *-ly*.

Outra limitação que pode nos levar a desenvolver futuros estudos nesse campo diz respeito à própria configuração das disciplinas de Inglês para Fins Acadêmicos e aqui também incluímos o contexto de boa parte das universidades brasileiras, em que as aulas de Inglês para Fins Acadêmicos continuam sendo ofertadas para grupos provenientes de diferentes cursos de graduação e pós-graduação. Como nos lembra Wingate (2015, p.39)¹⁰¹:

[...] o formato mais comum de Inglês para Fins Acadêmicos é de aulas de língua inglesa direcionadas a estudantes falantes não-nativos de uma diversidade de áreas com o objetivo de aperfeiçoar sua precisão gramatical e uso do estilo. O ensino de conteúdo não é relacionado a nenhum contexto acadêmico específico.

Como não há turmas divididas por área, não há um direcionamento que dê conta de atender as especificidades linguísticas e de gêneros que perpassam cada uma em particular. Sobre isso, Gardner (2007) relata que, fazendo um contraponto entre textos das áreas de História e Engenharia, percebe-se que enquanto os primeiros escrevem mais redações argumentativas, entre os estudantes de Engenharia prepondera a escrita de artigos científicos, propostas para projetos profissionais e relatórios que, muitas vezes, não são produzidos individualmente. Em miúdos, o propósito social e as características léxico-gramaticais entre as áreas se diferem significativamente, o que pode certamente nos auxiliar na realização de outras pesquisas que dêem conta de desvelar resultados que contribuam para o melhor desenvolvimento da escrita acadêmica em inglês de aprendizes em contextos universitários.

¹⁰¹ No original: “[...] the most widespread EAP format is that of English language classes delivered to non-native speaker students from a range of disciplines with the aim to improve their grammatical accuracy and use of style. The teaching content is not related to any specific academic context [...]” (WINGATE, 2015, p. 39).

REFERÊNCIAS

ALLEN, D. Lexical Bundles in Learner Writing: An Analysis of Formulaic Language in the ALESS Learner Corpus. *Kamaba Journal of English Education*, v.1, p. 105-127, 2009.

ALMEIDA, V.C. *Investigando colocações em um corpus de aprendiz*. Tese de doutorado, Universidade Federal de Minas Gerais, MG, Brasil, 2014.

ALTENBERG, B. *On the phraseology of spoken English: the evidence of recurrent word-combinations*. In: COWIE, A. P. (Ed.). *Phraseology: theory, analysis and applications* (p. 101-122). Oxford: Oxford University Press, 1998.

ALTENBERG, B.; GRANGER, S. The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, v.22, n.2, p.173-195, 2001.

ANTHONY, L. AntConc (Version 3.2.4) [Computer Software]. Tokyo, Japan: Waseda University. Disponível em: <<http://www.antlab.sci.waseda.ac.jp/>>. 2011. Acesso em: 22 mar. 2016.

ASKEHAVE, I.; SWALES, M. J. *Identificação de gênero e propósito comunicativo: um problema e uma possível solução*. In: BEZERRA, B. G.; BIASI-RODRIGUES, B.; CAVALCANTE, M. M. (Orgs.). *Gêneros e sequências textuais*. Recife: Edupe, 2009, p. 221-247.

BAKHTIN, M. Os gêneros do discurso. In: BAKHTIN, M. *Estética da criação verbal*. 3. ed. São Paulo: Martins Fontes, 2000, p. 277-326.

BARROS, M. *Poesia Completa*. São Paulo: Leya, 2011.

BAWARSHI, A. S.; REIFF, M. J. *Gênero: história, teoria, pesquisa e ensino*. São Paulo: Parábola, 2013 [2010].

BERBER SARDINHA, T. Linguística de corpus: histórico e problemática. *Delta*, v. 16, n. 2, 2000.

BERTOLI, P. P.; SHEPERD, T.M.G. Escrita acadêmica: um estudo exploratório de quadrigramas. *The ESPecialist*, vol. 36, n. 2 (p. 241-262), 2015.

BHATIA, V.K. *Worlds of Written Discourse: A Genre-Based View*. London and New York: Continuum, 2004.

BIBER, D. *Variation across Speech and Writing*. Cambridge: Cambridge University Press, 1988.

BIBER, D. *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins, 2006.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus Linguistics: investigating language*

structure and use. Cambridge: Cambridge University Press, 1998.

BIBER, D.; CONRAD, S. Register, genre and style. Cambridge: Cambridge University Press, 2009.

BIBER, D.; FINEGAN, E. *Adverbial stance types in english*. *Discourse processes*, 11, p. 1-34, 1998.

BIBER, D.; GRAY, B. *Grammatical Complexity in Academic English: Linguistic Change in Writing*. Cambridge: Cambridge University Press, 2016.

BIBER, D.; JOHANSSON, S.; LEECH, G.; CONRAD, S.; FINEGAN, E. *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson, 1999.

CHAFE, W.; DANIELEWICZ, J. Properties of Spoken and Written Language. *Thematic Development in English Texts*, In: GHADESSY, M. (Ed.), *Thematic Development in English Text*. London: Pinter Publisher, 1987, p.83-113.

CHARLES, M.; HUNSTON, S.; PECORARI, D. *Academic Writing: At the Interface of Corpus and Discourse*. New York: Continuum, 2009.

CONRAD, S. What can a corpus tell us about grammar. In: O'KEEFE, A; MCCARTHY, M. *The Routledge Handbook of Corpus Linguistics*. New York: Routledge, 2010, p.227-240.

CORDER, S. P. The Significance of Learners' Errors. *International Review of Applied Linguistics in Language Teaching*, 5, p. 161-170, 1967.

CORTES, V. A Comparative Analysis of Lexical Bundles in Academic History Writing in English and Spanish. *Corpora*, v.3, n.1, p.43-57, 2008.

COTOS, E. Enhancing writing pedagogy with learner corpus data . *ReCALL*, 26, 2014, p. 202-224.

COXHEAD, A. A New Academic Word List. *TESOL Quaterly*, v. 34, n.2, p.213-238, 2000.

DAVIES, M. Google Scholar and COCA-Academic: Two very different approaches to examining academic English. *Journal of English for Academic Purposes*, 12, p. 155-165, 2013.

DAYRELL, C. *Corpora no ensino do inglês acadêmico: padrões léxico-gramaticais em abstracts de pós-graduandos brasileiros*. IN: VIANA, V.; TAGNIN, S. E. O. *Corpora bi Ensino de Línguas Estrangeiras*. São Paulo: Hub Editorial, 2011.

DUNNING, T. *Accurate Methods for the Statistics of Surprise and Coincidence*. *Computational Linguistics*, v.19, n.1, p.61-74, 1993.

DUTRA, D. P.; BERBER SARDINHA, T. *Pacotes Lexicais em Corpora de Aprendizagem*. In: IBANOS, A.M.T.; MOTTIN, L.P.; SARMENTO, S.; BERBER SARDINHA, T.

(Orgs.). *Pesquisas e Perspectivas em Linguística de Corpus*. São Paulo: Mercado de Letras, 2015, p.57-80.

DUTRA, D.P.; GOMIDE, A.R.; OLIVA, K.P. GUEDES, A.S. *Corpus de aprendizes do Inglês sem Fronteiras: caminhos para compreender a interlíngua de alunos universitários brasileiros*. In: SARMENTO, S.; ABREU-E-LIMA, D.M.; FILHO, W.B.M. (Orgs.). *Do Inglês sem Fronteiras ao Idiomas sem Fronteiras. A construção de uma política linguística para a internacionalização*. Belo Horizonte: Editora UFMG, 2016.

DUTRA, D.P.; GOMIDE, A. R. Compilation of a University Learner Corpus. *Brazilian English Language Teaching Journal*. Porto Alegre, December 2015, v.6, p. 21-33.

DUTRA, D.P; SILERO, R.P. Descobertas linguísticas para pesquisadores e aprendizes: a Linguística de Corpus e o ensino de gramática. *Revista Brasileira de Linguística Aplicada*, v.10, n.4, p.909-930, 2010.

EGBERT, J.; STAPLES, S.; BIBER, D. *Corpus Research*. In: BROWN, J.D.; COOMBE, C. *The Cambridge Guide to Research in Language Teaching and Learning*. Cambridge: CUP, 2016.

ELLIS, R. *Second language acquisition & language pedagogy*. Clevedon: Multilingual Matters Ltd, 1992.

ELLIS, R. *Second Language Acquisition*. Oxford: OUP, 1997, p. 31-35.

ELLIS, N.C. *Phraseology*. The periphery and the heart of language. In: MEUNIER, F.; GRANGER, S. (Eds.). *Phraseology in Language Learning and Teaching*. Amsterdam: John Benjamins Publishing, 2008, p.1-13.

FILLMORE, C. J. *The Case for Case Reopened*. In: COLE, P. & SADOCK, J. M. (eds.). *Syntax and Semantics 8. Grammatical Relations*, 1977, p. 59-81.

FILLMORE, C. J. *On the Organization of Semantic Information in the Lexicon*. In: *Papers from the Parasession of the Lexicon*. Chicago: Chicago Linguistic Society, 1978, p. 1-11.

FILLMORE, C. "Corpus Linguistics" or "Computer-Aided Armchair Linguistics". In: STARTVIK, J. *Directions in Corpus Linguistics*, Berlin: Mouton, 1992, p.35-60.

FIRTH, J. R. Modes of meaning. In: *Papers in linguistics 1934-1951*. Oxford: Oxford University Press, 1957, p.190-215.

GARDNER, D.; DAVIES, M. A New Academic Vocabulary List. *Applied Linguistics*, 35(3), p. 305-327, 2013.

GARSIDE, R.; SMITH, N. 1997. *A hybrid grammatical tagger: CLAWS4*. In: GARSIDE, R.; LEECH, G. ; McENERY, A. (Eds.). *Corpus Annotation: Linguistic Information from Computer Text Corpora*, p.102-121. London: Longman, 1997.

GASS, S; SELINKER, L. *Second language acquisition: An introductory course*. 3^a ed. New York: Routledge: 2008.

GILQUIN, G.; PAQUOT, M. *Spoken Features in Learner Academic Writing: Identification, Explanation and Solution*. In: Proceedings of Corpus Linguistics, 2007. Birmingham, UK, 27-30 July, 2007.

GILQUIN, G., GRANGER, S. & PAQUOT, M. Learner corpora: The missing link in EAP pedagogy. *Journal of English for Academic Purposes*, v.6, n.4, p.319-335, 2007.

GRANGER, S. From CA to CIA and back: an integrated approach to computerized bilingual and learner corpora. In: AIJMER, K., ALTENBERG, B. & JOHNSON, M. (Eds.). *Languages in contrast. Text-based cross-linguistic studies*. Lund: Lund University Press, 1996, p.37-51.

GRANGER, S. The computer learner corpus: a versatile new source of data for SLA research. In: GRANGER, S. (ed.) *Learner English on Computer*. London: Longman, 1998, p. 3-18.

GRANGER, S. *A Bird's-eye View of Learner Corpus Research*. In: GRANGER, S.; HUNG, J.; PETCH-TYSON, S. (eds.). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 2002.

GRANGER, S. The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research, *TESOL Quarterly* 37:3, p. 538-546, 2003.

GRANGER, S. The contribution of learner corpora to second language acquisition and foreign language teaching. In: AIJMER, K. (ed.). *Corpora and language teaching*. Amsterdam: John Benjamins, 2009, p.13-32.

GRANGER, S. Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research* 1:1, 2015, p. 7-24.

GRANGER, S., DAGNEAUX, E., MEUNIER, F. & PAQUOT, M. *The International Corpus of Learner English. Version 2*. (Handbook + CD-ROM). Louvain-la-Neuve: Presses universitaires de Louvain, 2009.

GRANGER, S.; PAQUOT, M. *Lexical Verbs in Academic Discourse: A corpus-driven study of learner use*. In: CHARLES, M.; HUNSTON, S.; PECORARI, D. (Eds.). *Academic Writing: At the Interface of Corpus and Discourse*. London, Continuum, 2009, p. 193-214.

GRANGER, S.; RAYSON, P. Automatic Lexical Profiling of Learner Texts. In: GRANGER, S. (Ed.) *Learner English on Computer*. London: Addison Wesley Longman, 1998, p.119-131.

GRIES, T.S. *Statistical tests for the analysis of learner corpus data*. John Benjamins Publishing, Amsterdam, New York: 2013.

GUNDERSON, L. *ESL (ELL) Literacy Instruction: A Guidebook to Theory and Practice*, 2a ed. Routledge, 2009.

HALLIDAY, M. A. K.; HASAN, R. *Cohesion in English*. New York: Longman, 1976.

HALLIDAY, M. A. K.; A. MCINTOSH; P. STREVENS. *The Linguistic Sciences and Language Teaching*. London: Longman, 1996.

HARTSHORN, K. J.; HART, M. J. Comparing the academic word list with the academic vocabulary list: analyses of frequency and performance of English language learners. *The Journal of Language Teaching and Learning*, 2016, p.70-87.

HASSELGÅRD, H., & JOHANSSON, S. *Learner corpora and contrastive interlanguage analysis*, p.33-62. In: MEUNIER, F.; DE COCK, S.; GÄETANELLE, G.; PAQUOT, M. (Eds.). *Taste for Corpora - In honour of Sylviane Granger*. Amsterdam: John Benjamins Publishing Company, 2011.

HASSELGREN, A. *Lexical teddy bears and advanced learners: a study into the ways Norwegian students cope with English vocabulary*. *International Journal of Applied Linguistics*. Volume 4, Issue 2, December 1994, p. 237-258.

HIGGINS, C. (Ed.). *Identity formation in globalizing contexts. Language learning in the new Millennium* (Vol. 1). Berlin, Germany: Walter de Gruyter, 2011.

HINKEL, E. *Second language writers' text: Linguistic and rhetorical features*. Mahwah: Lawrence Erlbaum, 2002.

HINKEL, E. Research findings on teaching grammar for academic writing. *English Teaching*, 68(4), p. 3-21, 2013.

HUDDLESTON, R.; PULLUM, G.K. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press, 2002.

HUNSTON, S. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press, 2002.

HUNSTON, S. Corpora and language teaching: Issues of language description. In: HUNSTON, S. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press, 2009.

HYLAND, K.. *Disciplinary Discourses: Social Interactions in Academic Writing*. London: Longman, 2000.

HYLAND, K. *English for academic purposes: an advanced resource book*. New York/Canada: Routledge, 2006.

HYLAND, K. As can be seen: Lexical Bundles and disciplinary variation. *English for Specific Purposes*. v.27, n.1, p.4-21, 2008.

HYLAND, K. *Academic Discourse*. London, Continuum, 2009.

HYLAND, K.; MILTON, J. Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*. v. 16, n. 2, p. 183-205, 1997.

HYLAND, K.; HAMP-LYONS, L. EAP: issues and directions. *Journal of English for Academic Purposes*, 1, p.1-12, 2002.

HYLAND, K.; TSE, P. Is there an "academic vocabulary"? *TESOL Quarterly*, v. 41, n. 2, p. 235-254, 2007.

JOHANSSON, S. *Some aspects of the development of corpus linguistics in the 1970s and 1980s*. In: LÜDELING, A.; KYTÖ, M. (Ed.). *Corpus linguistics: an international handbook*. Berlin/New York: Walter de Gruyter, 2008.

JORDAN, R.R. *English for Academic Purposes. A guide and resource book for teachers*. Cambridge: Cambridge University Press, 2012.

JUDD, E L.; TAN, L; WALBERG, H. J. *Teaching additional languages*. UNESCO. International Academy of Education. International Bureau of Education, 2003.

LADO, R. *Linguistics across Cultures*. Ann Arbor: University of Michigan Press, 1957.

LEA, M.; STREET, B. Student writing in higher education: an academic literacies approach. *Studies in Higher Education*, v. 23, n.2, p. 157-172, 1998.

LEECH, G. *Developing Linguistic Corpora: a Guide to Good Practice. Adding Linguistic Annotation*. Lancaster: Lancaster University, 2004.

MARTIN, J. *English Text: System and Structure*. Amsterdam: Benjamins, 1992.

MARTINEZ, R.; SCHIMTT, N.A Phrasal Expressions List. *Applied Linguistics*, v. 33, n.3, p. 299-320, 2012.

McENERY, A.; XIAO, Z. *What corpora can offer in language teaching and learning*. In: HINKEL, E.(ed.). *Handbook of Research in Second Language Teaching and Learning (Vol. 2)*. London / New York: Routledge, 2010.

McENERY, T.; HARDIE, A. *Corpus linguistics: method, theory and practice*. Cambridge, New York: Cambridge University Press. 2012.

McENERY, T.; WILSON, A. *Corpus Linguistics*. 2a.ed. Edinburgh: Edinburgh University Press, 2001.

MCLAUGHLIN, B. *Theories of Second Language Learning*. London: Edward Arnold, 1987.

MEUNIER, F. Corpus Linguistics and second/foreign language learning: exploring multiple paths. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v.11, n.2, p.459-477, 2011.

MOURSSI, A. Crosslinguistic Influence of L1 (Arabic) in Acquiring Linguistic Items of L2 (English): An Empirical Study in the Context of Arab Learners of English as Undergraduate Learners. *Theory and Practice in Language Studies*, v. 3, n. 3, p. 397-403, 2013.

NESI, H.; GARDNER, S. *Genres across the Disciplines. Student Writing in Higher Education*. Cambridge: Cambridge University Press, 2012.

NESSSELHAUF, N. *How learner corpus analysis can contribute to language teaching: A study of support verb constructions*. In: ASTON, G.; BERNARDINI, S. & STEWART, D. (Eds.). Amsterdam and Philadelphia: Benjamins, 2004, p.109-124.

NESSSELHAUF, N. *Collocations in a Learner Corpus*. Amsterdam and Philadelphia: Benjamins, 2005.

OSS, D.B. Educação linguística e formação de professores de inglês. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v.13, n.3, p.677-709, 2013.

PAQUOT, M. *Exemplification in learner writing A cross-linguistic perspective*. In: MEUNIER, F.; GRANGER, S. *Phraseology in Foreign Language Learning and Teaching*. Amsterdam and Philadelphia: Benjamins, 2008, p.101-119.

PAQUOT, M. Lexical bundles and L1 transfer effects. *International Journal of Corpus Linguistics*, v. 18, n. 3, p. 391-417, 2013.

PAQUOT, M. Cross-linguistic influence and formulaic language: recurrent word sequences in French learner writing. In: Roberts, L. (ed.). *EUROSLA Yearbook*. Amsterdam and Philadelphia: John Benjamins, 2014, p. 240-261.

PAQUOT, M.; GRANGER, S. Formulaic language in learner corpora. *Annual Review of Applied Linguistics* 32, p. 130-149, 2012.

PARTINGTON, A. *Patterns and Meanings. Using corpora for English language research and teaching*. Amsterdam: John Benjamins, 1998.

RAMOS, R.C.G. Gêneros textuais: uma proposta de aplicação em cursos de Inglês para Fins Específicos. *The ESPECIALIST*, v. 25, n. 2, p.107-129, 2004.

REPPEN, R. *Bringing it into the classroom: Example activities descriptions of corpora in the classroom*. In: *Corpora in the Language Classroom*. Cambridge: Cambridge University Press, 2010, p.61-71.

REPPEN, R. Building a corpus: What are the key considerations? In: O'KEEFFE, Anne; MCCARTHY, Michael. *The Routledge Handbook of Corpus Linguistics*. New York: Routledge, 2010, p. 31-37.

RÖMER, U. *Corpus-research and practice. What help do teachers need and what can we offer?*. In: AIJMER, K. *Corpora and Language Teaching*. Philadelphia: John Benjamins Publishing Company, 2009, p.83-120.

- SAVILLE-TROIKE, M. *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press, 2006.
- SCHLATTER, M.; GARCEZ, P. M. *Educação linguística e aprendizagem de uma língua adicional na escola*. In: Referencias curriculares do Estado do Rio Grande do Sul: linguagens, códigos e suas tecnologias. Porto Alegre: Secretaria de Estado da Educação/Departamento Pedagógico, 2009.
- SCOTT, M.; TRIBBLE, C. *Textual patterns: Key words and corpus analysis in language education*. Vol. 22. John Benjamins Publishing, 2006.
- SELINKER, L. Interlanguage. *International Review of Applied Linguistics*, 10, p.209-231, 1972.
- SILVA, T. *Second language composition instruction: developments, issues and directions in ESL*. In: KROLL, B. (Ed.). *Second Language Writing*. Cambridge: Cambridge University Press, 1990.
- SINCLAIR, J. *Looking up*. An account of the COBUILD project in lexical computing. London & Glasgow: Collins Cobuild, 1987.
- SINCLAIR, J. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.
- SINCLAIR, J. 2005. *Corpus and Text - Basic Principles*. In: WYNNE, M. (Ed.). *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford: Oxbow Books: 1-16.
- STUBBS, M. *Text and corpus analysis: computer-assisted studies of language and culture*. Oxford: Blackwell, 1996.
- SWALES, J. M. *Genre analysis: english in academic and research settings*. Nova Iorque: Cambridge University Press, 1990.
- TARONE, E. On the variability of interlanguage systems. *Applied Linguistics* 4, p.142-163, 1983.
- TARONE, E. *Enduring themes from the interlanguage hypothesis*. In: HAN, Z-H.; TARONE, E. (Eds.). *Interlanguage: 40 years later* (p. 7-26). Amsterdam: John Benjamins, 2014.
- THORNDIKE, E. L.; LORGE, I. *The teacher's word book of 30,000 words*. Oxford, England: Bureau of Publications, Teachers Co, 1944.
- TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. Amsterdam: John Benjamins, 2001.
- TRIBBLE, C. Writing academic English further along the road. What is happening now in EAP writing instruction?. *English Language Teaching Journal*, v. 69, n. 4, p. 442-462, 2015.

TRIBBLE, C. ELFA vs. Genre: A new paradigm war in EAP writing instruction?.
Journal of English for Academic Purposes, v. 25, p. 30-44, 2017.

WAIBEL, B. *Phrasal verbs. German and Italian learners of English compared*.
Saarbrücken: VDM, 2008.

WEST, M. *A general service list of English words*. London: Longman, 1953.

WHITE, R.; ARNDT, V. *Process writing*. Harlow: Longman, 1991.

WINGATE, U., „Argument!“ helping students understand what essay writing is about.
Journal of English for Academic Purposes, v. 11, p. 145-154, 2012.

WINGATE, U. *Academic Literacy and Student Diversity. The Case for Inclusive Practice*. Bristol:Multilingual Matters, 2015.

WOLFERSBERGER, M.; COOMBE, C. *Integrated Reading and Writing Assessment : History, Processes, and Challenges*. In: EVSANS, N.W. *ESL Readers and Writers in Higher Education: understanding challenges, providing support*. Routledge, 2015.

XIAO, Z. What can SLA learn from contrastive corpus linguistics? The case of passive constructions in Chinese learner English. *Indonesian Journal of English Language Teaching* 3(2): p. 1-19, 2007.

XIAO, Z. *Collocation*. In: BIBER, C.; REPPEN, R. *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 2015, p. 106-124.

APÊNDICES

APÊNDICE I – Verbos do CorIFA com frequência mínima de cinco ocorrências (organizado por ordem decrescente de frequência)

Verbo	Classe verbal	Domínio semântico	Número de ocorrências
<i>be</i>	verbo auxiliar primário	verbo copular	6.617 (2.922 <i>is</i> ; 1.165 <i>be</i> ; 1.344 <i>are</i> ; 245 <i>am</i> ; 420 <i>was</i> ; 145 <i>being</i> ; 197 <i>been</i> ; 152 <i>were</i>)
<i>have</i>	verbo auxiliar primário	-----	2.001 (1208 <i>have</i> ; 569 <i>has</i> ; 52 <i>having</i> ; 172 <i>had</i>)
<i>can</i>	verbo modal		978
<i>make</i>	verbo lexical	verbo de ação	1.117 (524 <i>make</i> ; 130 <i>makes</i> ; 115 <i>making</i> ; 348 <i>made</i>)
<i>do</i>	verbo auxiliar primário	-----	646 (452 <i>do</i> ; 61 <i>does</i> ; 77 <i>doing</i> ; 56 <i>did</i>)
<i>learn</i>	verbo lexical	verbo mental	583 (495 <i>learn</i> ; 7 <i>learns</i> ; 22 <i>learning</i> ; 59 <i>learned</i>)
<i>will</i>	verbo modal		494 (0 ocorrências de <i>won't</i>)
<i>live</i>	verbo lexical	verbo de existência	969 (518 <i>live</i> ; 296 <i>lives</i> ; 130 <i>living</i> ; 25 <i>lived</i>)
<i>would</i>	verbo modal		441 (2 ocorrências de <i>wouldn't</i>)
<i>study</i>	verbo lexical	verbo mental	187 (84 <i>study</i> ; 0 <i>studies</i> ; 82 <i>studying</i> ; 18

			<i>studied</i>)
<i>think</i>	verbo lexical	verbo mental	564 (397 <i>think</i> ; 9 <i>thinks</i> ; 132 <i>thinking</i> ; 18 <i>thought</i>)
<i>know</i>	verbo lexical	verbo mental	519 (317 <i>know</i> ; 20 <i>knows</i> ; 51 <i>knowing</i> ; 11 <i>knew</i> ; 120 <i>known</i>)
<i>get</i>	verbo lexical	verbo de ação	426 (317 <i>get</i> ; 24 <i>gets</i> ; 54 <i>getting</i> ; 27 <i>got</i> ; 4 <i>gotten</i>)
<i>could</i>	verbo modal		271
<i>take</i>	verbo lexical	verbo de ação	373 (227 <i>take</i> ; 40 <i>takes</i> ; 55 <i>taking</i> ; 24 <i>took</i> ; 27 <i>taken</i>)
<i>go</i>	verbo lexical	verbo de ação	344 (198 <i>go</i> ; 27 <i>goes</i> ; 105 <i>going</i> ; 14 <i>went</i>)
<i>develop</i>	verbo lexical	verbo de ação	210 (96 <i>develop</i> ; 13 <i>develops</i> ; 32 <i>developing</i> ; 71 <i>developed</i>)
<i>say</i>	verbo lexical	verbo de comunicação	210 (124 <i>say</i> ; 17 <i>says</i> ; 35 <i>saying</i> ; 66 <i>said</i>)
<i>improve</i>	verbo lexical	verbo mental	277 (188 <i>improve</i> ; 13 <i>improves</i> ; 25 <i>improving</i> ; 51 <i>improved</i>)
<i>become</i>	verbo lexical	verbo de única ocorrência	318 (170 <i>become</i> ; 35 <i>becomes</i> ; 31 <i>becoming</i> ; 82 <i>became</i>)
<i>cheat</i>	verbo lexical	verbo de ação	192 (43 <i>cheat</i> ; 11 <i>cheats</i> ; 135 <i>cheating</i> ; 3 <i>cheated</i>)
<i>want</i>	verbo	verbo mental	205

	lexical		(173 <i>want</i> ; 20 <i>wants</i> ; 0 <i>wanting</i> ; 12 <i>wanted</i>)
<i>see</i>	verbo lexical	verbo mental	374 (269 <i>see</i> ; 2 <i>sees</i> ; 22 <i>seeing</i> ; 58 <i>saw</i> ; 23 <i>seen</i>)
<i>should</i>	verbo modal		178
<i>bring</i>	verbo lexical	verbo de ação	280 (92 <i>bring</i> ; 76 <i>brings</i> ; 17 <i>bringing</i> ; 95 <i>brought</i>)
<i>create</i>	verbo lexical	verbo de ação	246 (112 <i>create</i> ; 10 <i>creates</i> ; 37 <i>creating</i> ; 87 <i>created</i>)
<i>come</i>	verbo lexical	verbo de ação	228 (73 <i>come</i> ; 78 <i>comes</i> ; 20 <i>coming</i> ; 57 <i>came</i>)
<i>talk</i>	verbo lexical	verbo de comunicação	254 (182 <i>talk</i> ; 11 <i>talks</i> ; 55 <i>talking</i> ; 6 <i>talked</i>)
<i>give</i>	verbo lexical	verbo de ação	264 (118 <i>give</i> ; 42 <i>gives</i> ; 39 <i>giving</i> ; 16 <i>gave</i> ; 49 <i>given</i>)
<i>believe</i>	verbo lexical	verbo mental	139 (111 <i>believe</i> ; 12 <i>believes</i> ; 3 <i>believing</i> ; 13 <i>believed</i>)
<i>consider</i>	verbo lexical	verbo mental	191 (55 <i>consider</i> ; 1 <i>considers</i> ; 44 <i>considering</i> ; 91 <i>considered</i>)
<i>find</i>	verbo lexical	verbo de ação	184 (129 <i>find</i> ; 2 <i>finds</i> ; 5 <i>finding</i> ; 48 <i>found</i>)
<i>may</i>	verbo modal		126

<i>understand</i>	verbo lexical	verbo mental	143 (121 <i>understand</i> ; 3 <i>understands</i> ; 8 <i>understanding</i> ; 11 <i>understood</i>)
<i>happen</i>	verbo lexical	verbo de simples ocorrência	208 (53 <i>happen</i> ; 100 <i>happens</i> ; 25 <i>happening</i> ; 30 <i>happened</i>)
<i>provide</i>	verbo lexical	verbo de ação	206 (79 <i>provide</i> ; 56 <i>provides</i> ; 18 <i>providing</i> ; 53 <i>provided</i>)
<i>try</i>	verbo lexical	verbo de ação	156 (95 <i>try</i> ; 8 <i>tries</i> ; 48 <i>trying</i> ; 5 <i>tried</i>)
<i>apply</i>	verbo lexical	verbo de ação	63 (21 <i>apply</i> ; 6 <i>applies</i> ; 2 <i>applying</i> ; 34 <i>applied</i>)
<i>travel</i>	verbo lexical	verbo de ação	172 (123 <i>travel</i> ; 11 <i>travels</i> ; 31 <i>traveling</i> ; 7 <i>traveled</i>)
<i>must</i>	verbo modal		96
<i>eat</i>	verbo lexical	verbo de ação	162 (116 <i>eat</i> ; 4 <i>eats</i> ; 34 <i>eating</i> ; 5 <i>ate</i> ; 3 <i>eaten</i>)
<i>spend</i>	verbo lexical	verbo de ação	147 (94 <i>spend</i> ; 3 <i>spends</i> ; 17 <i>spending</i> ; 33 <i>spent</i>)
<i>allow</i>	verbo lexical	verbo de facilitação ou causa	133 (44 <i>allow</i> ; 27 <i>allows</i> ; 16 <i>allowing</i> ; 46 <i>allowed</i>)
<i>adopt</i>	verbo lexical	verbo de ação	148 (97 <i>adopt</i> ; 1 <i>adopts</i> ; 4 <i>adopting</i> ; 46 <i>adopted</i>)
<i>teach</i>	verbo lexical	verbo de ação	85 (57 <i>teach</i> ; 5 <i>teaches</i> ; 0 <i>teaching</i> ; 21

			<i>taught</i>)
<i>keep</i>	verbo lexical	verbo de aspecto	119 (100 <i>keep</i> ; 5 <i>keeps</i> ; 12 <i>keeping</i> ; 2 <i>kept</i>)
<i>imagine</i>	verbo lexical	verbo mental	118 (97 <i>imagine</i> ; 0 <i>imagines</i> ; 7 <i>imagining</i> ; 14 <i>imagined</i>)
<i>write</i>	verbo lexical	verbo de comunicação	63 (36 <i>write</i> ; 0 <i>writes</i> ; 5 <i>writing</i> ; 18 <i>wrote</i> ; 4 <i>written</i>)
<i>contribute</i>	verbo lexical	verbo de ação	92 (51 <i>contribute</i> ; 14 <i>contributes</i> ; 6 <i>contributing</i> ; 21 <i>contributed</i>)
<i>read</i>	verbo lexical	verbo mental	104 (61 <i>read</i> ; 0 <i>reads</i> ; 43 <i>reading</i>)
<i>feel</i>	verbo lexical	verbo mental	137 (84 <i>feel</i> ; 13 <i>feels</i> ; 38 <i>feeling</i> ; 2 <i>felt</i>)
<i>receive</i>	verbo lexical	verbo de ação	121 (80 <i>receive</i> ; 3 <i>receives</i> ; 16 <i>receiving</i> ; 22 <i>received</i>)
<i>choose</i>	verbo lexical	verbo de ação	89 (52 <i>choose</i> ; 2 <i>chooses</i> ; 14 <i>choosing</i> ; 9 <i>chose</i> ; 12 <i>chosen</i>)
<i>gain</i>	verbo lexical	verbo de ação	119 (40 <i>gain</i> ; 6 <i>gains</i> ; 1 <i>gaining</i> ; 72 <i>gained</i>)
<i>relate</i>	verbo lexical	verbo mental	69 (11 <i>relate</i> ; 1 <i>relates</i> ; 2 <i>relating</i> ; 86 <i>related</i>)
<i>meet</i>	verbo lexical	verbo de ação	72 (50 <i>meet</i> ; 6 <i>meets</i> ; 3 <i>meeting</i> ; 13 <i>met</i>)
<i>achieve</i>	verbo	verbo de ação	86

	lexical		(56 <i>achieve</i> ; 0 <i>achieves</i> ; 5 <i>achieving</i> ; 25 <i>achieved</i>)
<i>solve</i>	verbo lexical	verbo de ação	79 (58 <i>solve</i> ; 0 <i>solves</i> ; 9 <i>solving</i> ; 12 <i>solved</i>)
<i>pay</i>	verbo lexical	verbo de ação	113 (95 <i>pay</i> ; 7 <i>pays</i> ; 2 <i>paying</i> ; 9 <i>paid</i>)
<i>deal</i>	verbo lexical	verbo mental	96 (71 <i>deal</i> ; 8 <i>deals</i> ; 16 <i>dealing</i> ; 1 <i>dealt</i>)
<i>realize</i>	verbo lexical	verbo mental	78 (60 <i>realize</i> ; 1 <i>realizes</i> ; 0 <i>realizing</i> ; 17 <i>realized</i>)
<i>produce</i>	verbo lexical	verbo de ação	106 (64 <i>produce</i> ; 9 <i>produces</i> ; 14 <i>producing</i> ; 19 <i>produced</i>)
<i>answer</i>	verbo lexical	verbo de ação	82 (57 <i>answer</i> ; 15 <i>answers</i> ; 4 <i>answering</i> ; 6 <i>answered</i>)
<i>enjoy</i>	verbo lexical	verbo de ação	84 (54 <i>enjoy</i> ; 0 <i>enjoys</i> ; 24 <i>enjoying</i> ; 6 <i>enjoyed</i>)
<i>speak</i>	verbo lexical	verbo de comunicação	134 (67 <i>speak</i> ; 10 <i>speaks</i> ; 45 <i>speaking</i> ; 2 <i>spoke</i> ; 10 <i>spoken</i>)
<i>acquire</i>	verbo lexical	verbo de ação	73 (29 <i>acquire</i> ; 2 <i>acquires</i> ; 1 <i>acquiring</i> ; 41 <i>acquired</i>)
<i>seem</i>	verbo copular	verbo de facilitação ou de causa	81 (27 <i>seem</i> ; 48 <i>seems</i> ; 2 <i>seeming</i> ; 2 <i>seemed</i>)
<i>accept</i>	verbo lexical	verbo de ação	55 (24 <i>accept</i> ; 0 <i>accepts</i> ; 8 <i>accepting</i> ;

			23 accepted)
<i>include</i>	verbo lexical	verbo de existência	89 (19 <i>include</i> ; 21 <i>includes</i> ; 40 <i>including</i> ; 9 <i>included</i>)
<i>treat</i>	verbo lexical	verbo de ação	87 (55 <i>treat</i> ; 5 <i>treats</i> ; 6 <i>treating</i> ; 21 <i>treated</i>)
<i>cause</i>	verbo lexical	verbo de facilitação ou causa	142 (51 <i>cause</i> ; 32 <i>causes</i> ; 12 <i>causing</i> ; 47 <i>caused</i>)
<i>offer</i>	verbo lexical	verbo de ação	77 (42 <i>offer</i> ; 17 <i>offers</i> ; 6 <i>offering</i> ; 12 <i>offered</i>)
<i>require</i>	verbo lexical	verbo de facilitação	81 (17 <i>require</i> ; 32 <i>requires</i> ; 2 <i>requiring</i> ; 30 <i>required</i>)
<i>discover</i>	verbo lexical	verbo mental	71 (45 <i>discover</i> ; 3 <i>discovers</i> ; 1 <i>discovering</i> ; 22 <i>discovered</i>)
<i>base</i>	verbo lexical	verbo mental	94 (2 <i>base</i> ; 0 <i>bases</i> ; 3 <i>basing</i> ; 89 <i>based</i>)
<i>lose</i>	verbo lexical	verbo de ação	65 (27 <i>lose</i> ; 4 <i>loses</i> ; 8 <i>losing</i> ; 26 <i>lost</i>)
<i>share</i>	verbo lexical	verbo de ação	73 (34 <i>share</i> ; 5 <i>shares</i> ; 11 <i>sharing</i> ; 23 <i>shared</i>)
<i>communicate</i>	verbo lexical	verbo de comunicação	87 (75 <i>communicate</i> ; 0 <i>communicates</i> ; 10 <i>communicating</i> ; 2 <i>communicated</i>)
<i>dance</i>	verbo lexical	verbo de ação	34 (13 <i>dance</i> ; 0 <i>dances</i> ; 19 <i>dancing</i> ; 2 <i>danced</i>)

<i>depend</i>	verbo lexical	verbo mental	84 (21 <i>depend</i> ; 46 <i>depends</i> ; 13 <i>depending</i> ; 4 <i>depended</i>)
<i>follow</i>	verbo lexical	verbo de ação	72 (28 <i>follow</i> ; 5 <i>follows</i> ; 18 <i>following</i> ; 21 <i>followed</i>)
<i>agree</i>	verbo lexical	verbo de ação	96 (79 <i>agree</i> ; 0 <i>agrees</i> ; 0 <i>agreeing</i> ; 17 <i>agreed</i>)
<i>facilitate</i>	verbo lexical	verbo de ação	80 (29 <i>facilitate</i> ; 18 <i>facilitates</i> ; 8 <i>facilitating</i> ; 25 <i>facilitated</i>)
<i>put</i>	verbo lexical	verbo de ação	83 (70 <i>put</i> ; 3 <i>puts</i> ; 10 <i>putting</i>)
<i>connect</i>	verbo lexical	verbo de ação	76 (25 <i>connect</i> ; 3 <i>connects</i> ; 1 <i>connecting</i> ; 47 <i>connected</i>)
<i>grow</i>	verbo lexical	verbo de única ocorrência	72 (30 <i>grow</i> ; 1 <i>grows</i> ; 28 <i>growing</i> ; 3 <i>grew</i> ; 10 <i>grown</i>)
<i>involve</i>	verbo lexical	verbo de existência	80 (1 <i>involve</i> ; 34 <i>involves</i> ; 12 <i>involving</i> ; 33 <i>involved</i>)
<i>visit</i>	verbo lexical	verbo de ação	62 (36 <i>visit</i> ; 4 <i>visits</i> ; 13 <i>visiting</i> ; 9 <i>visited</i>)
<i>ask</i>	verbo lexical	verbo de comunicação	62 (37 <i>ask</i> ; 6 <i>asks</i> ; 5 <i>asking</i> ; 14 <i>asked</i>)
<i>focus</i>	verbo lexical	verbo mental	56 (31 <i>focus</i> ; 2 <i>focuses</i> ; 10 <i>focusing</i> ; 13 <i>focused</i>)
<i>invest</i>	verbo lexical	verbo de ação	73 (34 <i>invest</i> ; 4 <i>invests</i> ; 20 <i>investing</i> ; 25

			<i>invested</i>)
<i>reach</i>	verbo lexical	verbo de ação	55 (38 <i>reach</i> ; 2 <i>reaches</i> ; 0 <i>reaching</i> ; 15 <i>reached</i>)
<i>prepare</i>	verbo lexical	verbo de ação	50 (25 <i>prepare</i> ; 5 <i>prepares</i> ; 0 <i>preparing</i> ; 20 <i>prepared</i>)
<i>participate</i>	verbo lexical	verbo de ação	19 (15 <i>participate</i> ; 0 <i>participates</i> ; 2 <i>participating</i> ; 2 <i>participated</i>)
<i>pass</i>	verbo lexical	verbo de ação	72 (48 <i>pass</i> ; 3 <i>passes</i> ; 12 <i>passing</i> ; 9 <i>passed</i>)
<i>notice</i>	verbo lexical	verbo mental	64 (38 <i>notice</i> ; 11 <i>notices</i> ; 1 <i>noticing</i> ; 14 <i>noticed</i>)
<i>continue</i>	verbo lexical	verbo de ação	49 (36 <i>continue</i> ; 9 <i>continues</i> ; 0 <i>continuing</i> ; 4 <i>continued</i>)
<i>expect</i>	verbo lexical	verbo mental	51 (18 <i>expect</i> ; 3 <i>expects</i> ; 2 <i>expecting</i> ; 28 <i>expected</i>)
<i>affect</i>	verbo lexical	verbo mental	64 (29 <i>affect</i> ; 15 <i>affects</i> ; 9 <i>affecting</i> ; 11 <i>affected</i>)
<i>open</i>	verbo lexical	verbo de ação	51 (46 <i>open</i> ; 2 <i>opens</i> ; 2 <i>opening</i> ; 1 <i>opened</i>)
<i>send</i>	verbo lexical	verbo de ação	56 (31 <i>send</i> ; 4 <i>sends</i> ; 15 <i>sending</i> ; 6 <i>sent</i>)
<i>leave</i>	verbo lexical	verbo de ação	53 (23 <i>leave</i> ; 7 <i>leaves</i> ; 17 <i>leaving</i> ; 6

			<i>left</i>
<i>forget</i>	verbo lexical	verbo mental	77 (54 <i>forget</i> ; 2 <i>forgets</i> ; 6 <i>forgetting</i> ; 9 <i>forgot</i> ; 6 <i>forgotten</i>)
<i>mention</i>	verbo lexical	verbo de comunicação	50 (17 <i>mention</i> ; 1 <i>mentions</i> ; 8 <i>mentioning</i> ; 24 <i>mentioned</i>)
<i>move</i>	verbo lexical	verbo de ação	47 (16 <i>move</i> ; 10 <i>moves</i> ; 13 <i>moving</i> ; 8 <i>moved</i>)
<i>build</i>	verbo lexical	verbo de ação	68 (50 <i>build</i> ; 3 <i>builds</i> ; 5 <i>building</i> ; 10 <i>built</i>)
<i>decide</i>	verbo lexical	verbo mental	32 (21 <i>decide</i> ; 6 <i>decides</i> ; 2 <i>deciding</i> ; 3 <i>decided</i>)
<i>dismiss</i>	verbo lexical	verbo de ação	34 (8 <i>dismiss</i> ; 1 <i>dismisses</i> ; 11 <i>dismissing</i> ; 14 <i>dismissed</i>)
<i>regard</i>	verbo lexical	verbo mental	22 (0 <i>regard</i> ; 5 <i>regards</i> ; 15 <i>regarding</i> ; 2 <i>regarded</i>)
<i>compare</i>	verbo lexical	verbo de ação	83 (18 <i>compare</i> ; 4 <i>compares</i> ; 20 <i>comparing</i> ; 41 <i>compared</i>)
<i>conclude</i>	verbo lexical	verbo de aspecto	54 (36 <i>conclude</i> ; 2 <i>concludes</i> ; 13 <i>concluding</i> ; 3 <i>concluded</i>)
<i>decide</i>	verbo lexical	verbo mental	32 (21 <i>decide</i> ; 6 <i>decides</i> ; 2 <i>deciding</i> ; 3 <i>decided</i>)
<i>search</i>	verbo lexical	verbo de ação	26 (18 <i>search</i> ; 0 <i>searches</i> ; 7 <i>searching</i> ;

			1 searched)
<i>obtain</i>	verbo lexical	verbo de ação	48 (27 <i>obtain</i> ; 2 <i>obtains</i> ; 2 <i>obtaining</i> ; 17 <i>obtained</i>)
<i>stay</i>	verbo lexical	verbo de existência	42 (33 <i>stay</i> ; 4 <i>stays</i> ; 4 <i>staying</i> ; 1 <i>stayed</i>)
<i>watch</i>	verbo lexical	verbo de ação	52 (28 <i>watch</i> ; 2 <i>watches</i> ; 20 <i>watching</i> ; 2 <i>watched</i>)
<i>begin</i>	verbo lexical	verbo de aspecto	70 (42 <i>begin</i> ; 11 <i>begins</i> ; 9 <i>beginning</i> ; 8 <i>began</i> ; 0 <i>begun</i>)
<i>catch</i>	verbo lexical	verbo de ação	30 (2 <i>catch</i> ; 0 <i>catches</i> ; 0 <i>catching</i> ; 28 <i>caught</i>)
<i>recognize</i>	verbo lexical	verbo mental	49 (20 <i>recognize</i> ; 2 <i>recognizes</i> ; 0 <i>recognizing</i> ; 27 <i>recognized</i>)
<i>influence</i>	verbo lexical	verbo de ação	50 (23 <i>influence</i> ; 5 <i>influences</i> ; 3 <i>influencing</i> ; 19 <i>influenced</i>)
<i>lead</i>	verbo lexical	verbo de ação	56 (33 <i>lead</i> ; 15 <i>leads</i> ; 8 <i>leading</i> ; 0 <i>leaded</i>)
<i>tend</i>	verbo lexical	verbo mental	52 (45 <i>tend</i> ; 6 <i>tends</i> ; 0 <i>tending</i> ; 1 <i>tended</i>)
<i>appreciate</i>	verbo lexical	verbo mental	21 (12 <i>appreciate</i> ; 0 <i>appreciates</i> ; 2 <i>appreciating</i> ; 7 <i>appreciated</i>)
<i>avoid</i>	verbo lexical	verbo de ação	51 (40 <i>avoid</i> ; 4 <i>avoids</i> ; 7 <i>avoiding</i> ; 0 <i>avoided</i>)

<i>discuss</i>	verbo lexical	verbo de comunicação	41 (21 <i>discuss</i> ; 3 <i>discusses</i> ; 5 <i>discussing</i> ; 12 <i>discussed</i>)
<i>observe</i>	verbo lexical	verbo mental	64 (38 <i>observe</i> ; 0 <i>observes</i> ; 5 <i>observing</i> ; 21 <i>observed</i>)
<i>suffer</i>	verbo lexical	verbo mental	50 (19 <i>suffer</i> ; 4 <i>suffers</i> ; 26 <i>suffering</i> ; 1 <i>suffered</i>)
<i>save</i>	verbo lexical	verbo de ação	49 (28 <i>save</i> ; 5 <i>saves</i> ; 5 <i>saving</i> ; 11 <i>saved</i>)
<i>analyze</i>	verbo lexical	verbo de ação	34 (18 <i>analyze</i> ; 0 <i>analyzes</i> ; 6 <i>analyzing</i> ; 10 <i>analyzed</i>)
<i>arrive</i>	verbo lexical	verbo de ação	46 (26 <i>arrive</i> ; 8 <i>arrives</i> ; 5 <i>arriving</i> ; 7 <i>arrived</i>)
<i>listen</i>	verbo lexical	verbo mental	40 (25 <i>listen</i> ; 0 <i>listens</i> ; 15 <i>listening</i> ; 0 <i>listened</i>)
<i>reduce</i>	verbo lexical	verbo de ação	55 (23 <i>reduce</i> ; 3 <i>reduces</i> ; 10 <i>reducing</i> ; 19 <i>reduced</i>)
<i>represent</i>	verbo lexical	verbo de existência	45 (18 <i>represent</i> ; 17 <i>represents</i> ; 8 <i>representing</i> ; 2 <i>represented</i>)
<i>wait</i>	verbo lexical	verbo de ação	35 (20 <i>wait</i> ; 0 <i>waits</i> ; 13 <i>waiting</i> ; 2 <i>waited</i>)
<i>win</i>	verbo lexical	verbo de ação	37 (6 <i>win</i> ; 1 <i>wins</i> ; 3 <i>winning</i> ; 27 <i>won</i>)
<i>buy</i>	verbo	verbo de ação	46

	lexical		(38 <i>buy</i> ; 2 <i>buys</i> ; 6 <i>buying</i> ; 0 <i>bought</i>)
<i>occur</i>	verbo lexical	verbo de única ocorrência	53 (28 <i>occur</i> ; 24 <i>occurs</i> ; 0 <i>occurring</i> ; 1 <i>occured</i>)
<i>intend</i>	verbo lexical	verbo mental	17 (7 <i>intend</i> ; 7 <i>intends</i> ; 1 <i>intending</i> ; 2 <i>intended</i>)
<i>celebrate</i>	verbo lexical	verbo de ação	42 (30 <i>celebrate</i> ; 1 <i>celebrates</i> ; 7 <i>celebrating</i> ; 4 <i>celebrated</i>)
<i>spread</i>	verbo lexical	verbo de ação	41 (34 <i>spread</i> ; 0 <i>spreads</i> ; 6 <i>spreading</i> ; 1 <i>spreaded</i>)
<i>tell</i>	verbo lexical	verbo de comunicação	28 (15 <i>tell</i> ; 2 <i>tells</i> ; 2 <i>telling</i> ; 9 <i>told</i>)
<i>walk</i>	verbo lexical	verbo de ação	37 (30 <i>walk</i> ; 3 <i>walks</i> ; 4 <i>walking</i> ; 0 <i>walked</i>)
<i>adapt</i>	verbo lexical	verbo de ação	46 (25 <i>adapt</i> ; 0 <i>adapts</i> ; 1 <i>adapting</i> ; 20 <i>adapted</i>)
<i>attend</i>	verbo lexical	verbo de ação	18 (9 <i>attend</i> ; 0 <i>attends</i> ; 7 <i>attending</i> ; 2 <i>attended</i>)
<i>copy</i>	verbo lexical	verbo de ação	24 (14 <i>copy</i> ; 1 <i>copies</i> ; 7 <i>copying</i> ; 2 <i>copied</i>)
<i>finish</i>	verbo lexical	verbo de aspecto	14 (12 <i>finish</i> ; 0 <i>finishes</i> ; 2 <i>finishing</i> ; 0 <i>finished</i>)
<i>prefer</i>	verbo lexical	verbo mental	39 (39 <i>prefer</i> ; 0 <i>prefers</i> ; 0 <i>preferring</i> ; 0 <i>preferred</i>)

<i>return</i>	verbo lexical	verbo de ação	39 (31 <i>return</i> ; 1 <i>returns</i> ; 5 <i>returning</i> ; 2 <i>returned</i> ;
<i>stimulate</i>	verbo lexical	verbo de ação	41 (23 <i>stimulate</i> ; 7 <i>stimulates</i> ; 5 <i>stimulating</i> ; 6 <i>stimulated</i>)
<i>expand</i>	verbo lexical	verbo de ação	25 (20 <i>expand</i> ; 0 <i>expands</i> ; 3 <i>expanding</i> ; 2 <i>expanded</i>)
<i>explain</i>	verbo lexical	verbo de comunicação	33 (23 <i>explain</i> ; 2 <i>explains</i> ; 3 <i>explaining</i> . 5 <i>explained</i>)
<i>might</i>	verbo modal		21
<i>motivate</i>	verbo lexical	verbo mental	25 (2 <i>motivate</i> ; 3 <i>motivates</i> ; 0 <i>motivating</i> ; 20 <i>motivated</i>)
<i>promote</i>	verbo lexical	verbo de ação	56 (23 <i>promote</i> ; 11 <i>promotes</i> ; 8 <i>promoting</i> ; 14 <i>promoted</i>)
<i>surround</i>	verbo lexical	verbo de ação	29 (4 <i>surround</i> ; 5 <i>surrounds</i> ; 4 <i>surrounding</i> ; 16 <i>surrounded</i>)
<i>aim</i>	verbo lexical	verbo mental	34 (8 <i>aim</i> ; 14 <i>aims</i> ; 9 <i>aiming</i> ; 3 <i>aimed</i>)
<i>join</i>	verbo lexical	verbo de ação	16 (12 <i>join</i> ; 0 <i>joins</i> ; 0 <i>joining</i> ; 4 <i>joined</i>)
<i>publish</i>	verbo lexical	verbo de ação	24 (9 <i>publish</i> ; 0 <i>pulishes</i> ; 5 <i>publishing</i> ; 10 <i>published</i>)
<i>remember</i>	verbo lexical	verbo mental	40 (36 <i>remember</i> ; 0 <i>remembers</i> ; 2 <i>remembering</i> ; 2 <i>remembered</i>)

<i>transform</i>	verbo lexical	verbo de única ocorrência	28 (14 <i>transform</i> ; 2 <i>transforms</i> ; 3 <i>transforming</i> ; 9 <i>transformed</i>)
<i>interact</i>	verbo lexical	verbo de ação	33 (31 <i>interact</i> ; 0 <i>interacts</i> ; 0 <i>interacting</i> ; 2 <i>interacted</i>)
<i>permit</i>	verbo lexical	verbo de facilitação	26 (11 <i>permit</i> ; 6 <i>permits</i> ; 1 <i>permitting</i> ; 8 <i>permitted</i>)
<i>prove</i>	verbo lexical	verbo de ação	39 (20 <i>prove</i> ; 2 <i>proves</i> ; 4 <i>proving</i> ; 13 <i>proved</i>)
<i>punish</i>	verbo lexical	verbo de ação	20 (8 <i>punish</i> ; 0 <i>punishes</i> ; 1 <i>punishing</i> ; 11 <i>punished</i>)
<i>worry</i>	verbo lexical	verbo mental	24 (10 <i>worry</i> ; 5 <i>worries</i> ; 2 <i>worrying</i> ; 7 <i>worried</i>)
<i>complete</i>	verbo lexical	verbo de ação	58 (53 <i>complete</i> ; 2 <i>completes</i> ; 0 <i>completing</i> ; 3 <i>completed</i>)
<i>decrease</i>	verbo lexical	verbo de ação	47 (15 <i>decrease</i> ; 6 <i>decreases</i> ; 10 <i>decreasing</i> ; 16 <i>decreased</i>)
<i>define</i>	verbo lexical	verbo de ação	36 (13 <i>define</i> ; 10 <i>defines</i> ; 0 <i>defining</i> ; 13 <i>defined</i>)
<i>prevent</i>	verbo lexical	verbo mental	41 (25 <i>prevent</i> ; 2 <i>prevents</i> ; 4 <i>preventing</i> ; 10 <i>prevented</i>)
<i>reflect</i>	verbo lexical	verbo mental	30 (15 <i>reflect</i> ; 9 <i>reflects</i> ; 3 <i>reflecting</i> ; 3 <i>reflected</i>)

<i>wash</i>	verbo lexical	verbo de ação	32 (18 <i>wash</i> ; 1 <i>washes</i> ; 10 <i>washing</i> ; 3 <i>washed</i>)
<i>describe</i>	verbo lexical	verbo de ação	34 (13 <i>describe</i> ; 17 <i>describes</i> ; 0 <i>describing</i> ; 4 <i>described</i>)
<i>design</i>	verbo lexical	verbo de ação	22 (6 <i>design</i> ; 0 <i>designs</i> ; 2 <i>designing</i> ; 14 <i>designed</i>)
<i>enable</i>	verbo lexical	verbo de facilitação	21 (5 <i>enable</i> ; 3 <i>enables</i> ; 4 <i>enabling</i> ; 9 <i>enabled</i>)
<i>evaluate</i>	verbo lexical	verbo mental	19 (8 <i>evaluate</i> ; 0 <i>evaluates</i> ; 2 <i>evaluating</i> ; 9 <i>evaluated</i>)
<i>organize</i>	verbo lexical	verbo de ação	15 (8 <i>organize</i> ; 2 <i>organizes</i> ; 0 <i>organizing</i> ; 5 <i>organized</i>)
<i>dominate</i>	verbo lexical	verbo de ação	16 (5 <i>dominate</i> ; 3 <i>dominates</i> ; 2 <i>dominating</i> ; 11 <i>dominated</i>)
<i>kill</i>	verbo lexical	verbo de ação	25 (10 <i>kill</i> ; 1 <i>kills</i> ; 7 <i>killing</i> ; 7 <i>killed</i>)
<i>overcome</i>	verbo lexical	verbo mental	20 (20 <i>overcome</i> ; 0 <i>overcomes</i> ; 0 <i>overcoming</i> ; 0 <i>overcame</i>)
<i>relax</i>	verbo lexical	verbo de ação	16 (13 <i>relax</i> ; 3 <i>relaxes</i> ; 0 <i>relaxing</i> ; 5 <i>relaxed</i>)
<i>combine</i>	verbo lexical	verbo de ação	23 (0 <i>combine</i> ; 11 <i>combines</i> ; 0 <i>combining</i> ; 12 <i>combined</i>)
<i>invent</i>	verbo	verbo de ação	22

	lexical		(7 <i>invent</i> ; 0 <i>invents</i> ; 0 <i>inventing</i> ; 15 <i>invented</i>)
<i>maintain</i>	verbo lexical	verbo de ação	23 (21 <i>maintain</i> ; 2 <i>maintains</i> ; 0 <i>maintaining</i> ; 5 <i>maintained</i>)
<i>establish</i>	verbo lexical	verbo mental	34 (11 <i>establish</i> ; 3 <i>establishes</i> ; 0 <i>establishing</i> ; 20 <i>established</i>)
<i>fascinate</i>	verbo lexical	verbo mental	8 (0 <i>fascinate</i> ; 0 <i>fascinates</i> ; 7 <i>fascinating</i> ; 1 <i>fascinated</i>)
<i>inspire</i>	verbo lexical	verbo mental	10 (1 <i>inspire</i> ; 1 <i>inspires</i> ; 2 <i>inspiring</i> ; 6 <i>inspired</i>)
<i>integrate</i>	verbo lexical	verbo mental	24 (6 <i>integrate</i> ; 2 <i>integrates</i> ; 2 <i>integrating</i> ; 14 <i>integrated</i>)
<i>demonstrate</i>	verbo lexical	verbo de ação	22 (11 <i>demonstrate</i> ; 3 <i>demonstrates</i> ; 3 <i>demonstrating</i> ; 5 <i>demonstrated</i>)
<i>explore</i>	verbo lexical	verbo de ação	23 (15 <i>explore</i> ; 0 <i>explores</i> ; 2 <i>exploring</i> ; 6 <i>explored</i>)
<i>remain</i>	verbo lexical	verbo de ação	18 (7 <i>remain</i> ; 7 <i>remains</i> ; 2 <i>remaining</i> ; 2 <i>remained</i>)
<i>suggest</i>	verbo lexical	verbo de ação	22 (13 <i>suggest</i> ; 5 <i>suggests</i> ; 4 <i>suggesting</i> ; 0 <i>suggested</i>)
<i>appear</i>	verbo lexical	verbo de existência	17 (7 <i>appear</i> ; 4 <i>appears</i> ; 6 <i>appearing</i> ; 0 <i>appeared</i>)
<i>carry</i>	verbo	verbo de ação	28

	lexical		(16 <i>carry</i> ; 2 <i>carries</i> ; 3 <i>carrying</i> ; 7 <i>carried</i>)
<i>consist</i>	verbo lexical	verbo de ação	18 (3 <i>consist</i> ; 13 <i>consists</i> ; 1 <i>consisting</i> ; 1 <i>consisted</i>)
<i>introduce</i>	verbo lexical	verbo de ação	24 (5 <i>introduce</i> ; 0 <i>introduces</i> ; 5 <i>introducing</i> ; 14 <i>introduced</i>)
<i>perform</i>	verbo lexical	verbo de ação	20 (6 <i>perform</i> ; 0 <i>performs</i> ; 0 <i>performing</i> ; 14 <i>performed</i>)
<i>wonder</i>	verbo lexical	verbo mental	18 (13 <i>wonder</i> ; 2 <i>wonders</i> ; 1 <i>wondering</i> ; 2 <i>wondered</i>)
<i>approve</i>	verbo lexical	verbo de ação	11 (0 <i>approve</i> ; 0 <i>approves</i> ; 0 <i>approving</i> ; 7 <i>approved</i>)
<i>associate</i>	verbo lexical	verbo mental	28 (1 <i>associate</i> ; 0 <i>associates</i> ; 0 <i>associating</i> ; 27 <i>associated</i>)
<i>defend</i>	verbo lexical	verbo de ação	16 (12 <i>defend</i> ; 2 <i>defends</i> ; 2 <i>defending</i> ; 0 <i>defended</i>)
<i>enrich</i>	verbo lexical	verbo de ação	6 (3 <i>enrich</i> ; 0 <i>enriches</i> ; 3 <i>enriching</i> ; 0 <i>enriched</i>)
<i>extend</i>	verbo lexical	verbo de ação	9 (3 <i>extend</i> ; 0 <i>extends</i> ; 0 <i>extending</i> ; 6 <i>extended</i>)
<i>select</i>	verbo lexical	verbo de ação	13 (6 <i>select</i> ; 0 <i>selects</i> ; 2 <i>selecting</i> ; 5 <i>selected</i>)
<i>survive</i>	verbo	verbo de ação	19

	lexical		(17 <i>survive</i> ; 0 <i>survives</i> ; 0 <i>surviving</i> ; 2 <i>survived</i>)
<i>affirm</i>	verbo lexical	verbo de comunicação	19 (11 <i>affirm</i> ; 1 <i>affirms</i> ; 1 <i>affirming</i> ; 6 <i>affirmed</i>)
<i>determine</i>	verbo lexical	verbo de ação	17 (7 <i>determine</i> ; 0 <i>determines</i> ; 3 <i>determining</i> ; 7 <i>determined</i>)
<i>force</i>	verbo lexical	verbo de facilitação	24 (5 <i>force</i> ; 8 <i>forces</i> ; 0 <i>forcing</i> ; 11 <i>forced</i>)
<i>perceive</i>	verbo lexical	verbo mental	16 (11 <i>perceive</i> ; 1 <i>perceives</i> ; 0 <i>perceiving</i> ; 4 <i>perceived</i>)
<i>suppose</i>	verbo lexical	verbo de ação	15 (2 <i>suppose</i> ; 0 <i>supposes</i> ; 2 <i>supposing</i> ; 11 <i>supposed</i>)
<i>wake</i>	verbo lexical	verbo de ação	20 (16 <i>wake</i> ; 0 <i>wakes</i> ; 4 <i>waking</i> ; 0 <i>woke</i>)
<i>abandon</i>	verbo lexical	verbo de ação	13 (1 <i>abandon</i> ; 0 <i>abandons</i> ; 0 <i>abandoning</i> ; 12 <i>abandoned</i>)
<i>accelerate</i>	verbo lexical	verbo de ação	13 (6 <i>accelerate</i> ; 0 <i>accelerates</i> ; 2 <i>accelerating</i> ; 5 <i>accelerated</i>)
<i>accomplish</i>	verbo lexical	verbo mental	6 (5 <i>accomplish</i> ; 0 <i>accomplishes</i> ; 1 <i>accomplishing</i> ; 0 <i>accomplished</i>)
<i>assimilate</i>	verbo lexical	verbo mental	27 (13 <i>assimilate</i> ; 2 <i>assimilates</i> ; 6 <i>assimilating</i> ; 6 <i>assimilated</i>)
<i>behave</i>	verbo	verbo de ação	14

	lexical		(12 <i>behave</i> ; 0 <i>behaves</i> ; 2 <i>behaving</i> ; 0 <i>behaved</i>)
<i>claim</i>	verbo lexical	verbo de comunicação	25 (12 <i>claim</i> ; 8 <i>claims</i> ; 4 <i>claiming</i> ; 1 <i>claimed</i>)
<i>deserve</i>	verbo lexical	verbo mental	9 (6 <i>deserve</i> ; 1 <i>deserves</i> ; 1 <i>deserving</i> ; 1 <i>deserved</i>)
<i>enhance</i>	verbo lexical	verbo mental	12 (12 <i>enhance</i> ; 0 <i>enhances</i> ; 0 <i>enhancing</i> ; 0 <i>enhanced</i>)
<i>expel</i>	verbo lexical	verbo de ação	10 (3 <i>expel</i> ; 0 <i>expels</i> ; 2 <i>expelling</i> ; 5 <i>expelled</i>)
<i>invite</i>	verbo lexical	verbo de ação	11 (3 <i>invite</i> ; 0 <i>invites</i> ; 1 <i>inviting</i> ; 7 <i>invited</i>)
<i>isolate</i>	verbo lexical	verbo de ação	16 (0 <i>isolate</i> ; 2 <i>isolates</i> ; 0 <i>isolating</i> ; 14 <i>isolated</i>)
<i>modify</i>	verbo lexical	verbo de ação	19 (8 <i>modify</i> ; 0 <i>modifies</i> ; 7 <i>modifying</i> ; 4 <i>modified</i>)
<i>release</i>	verbo lexical	verbo de ação	18 (13 <i>release</i> ; 0 <i>releases</i> ; 0 <i>releasing</i> ; 5 <i>released</i>)
<i>run</i>	verbo lexical	verbo de ação	10 (10 <i>run</i> ; 0 <i>runs</i> ; 0 <i>running</i> ; 0 <i>ran</i>)
<i>satisfy</i>	verbo lexical	verbo de ação	9 (5 <i>satisfy</i> ; 1 <i>satisfies</i> ; 0 <i>satisfying</i> ; 3 <i>satisfied</i>)
<i>serve</i>	verbo lexical	verbo de ação	14 (3 <i>serve</i> ; 0 <i>serves</i> ; 1 <i>serving</i> ; 10

			served)
<i>admire</i>	verbo lexical	verbo mental	6 (4 admire; 0 admires; 0 admiring; 2 admired)
<i>grant</i>	verbo lexical	verbo mental	8 (5 grant; 2 grants; 0 granting; 15 granted)
<i>inform</i>	verbo lexical	verbo de ação	9 (2 inform; 0 informs; 2 informing; 5 informed)
<i>innovate</i>	verbo lexical	verbo de ação	8 (5 innovate; 0 innovates; 3 innovating; 0 innovated)
<i>seek</i>	verbo lexical	verbo de ação	21 (13 seek; 5 seeks; 3 seeking; 0 seeked)
<i>sleep</i>	verbo lexical	verbo de ação	16 (12 sleep; 0 sleeps; 4 sleeping; 0 slept)
<i>absorb</i>	verbo lexical	verbo mental	12 (10 absorb; 0 absorbs; 0 absorbing; 2 absorbed)
<i>assume</i>	verbo lexical	verbo mental	15 (7 assume; 0 assumes; 4 assuming; 4 assumed)
<i>close</i>	verbo lexical	verbo de ação	38 (30 close; 2 closes; 0 closing; 6 closed)
<i>heat</i>	verbo lexical	verbo de ação	5 (4 heat; 0 heats; 1 heating; 0 heated)
<i>ignore</i>	verbo lexical	verbo de ação	15 (9 ignore; 0 ignores; 2 ignoring; 4 ingnored)

<i>impose</i>	verbo lexical	verbo de ação	11 (0 impose; 2 imposes; 0 imposing; 9 imposed)
<i>neglect</i>	verbo lexical	verbo de ação	11 (0 neglect; 0 neglects; 0 neglecting; 11 neglected)
<i>restrict</i>	verbo lexical	verbo de ação	13 (1 restrict; 0 restricts; 0 restricting; 12 restricted)
<i>update</i>	verbo lexical	verbo de ação	19 (12 update; 0 updates; 1 updating; 6 updated)
<i>attach</i>	verbo lexical	verbo de ação	3 (2 attach; 0 attaches; 0 attaching; 1 attached)
<i>complement</i>	verbo lexical	verbo de ação	4 (2 complement; 0 complements; 2 complementing; 0 complemented)
<i>complicate</i>	verbo lexical	verbo mental	4 (0 complicate; 0 complicates; 0 complicating; 4 complicated)
<i>disagree</i>	verbo lexical	verbo de ação	9 (7 disagree; 0 disagrees; 2 disagreeing; 0 disagreed)
<i>locate</i>	verbo lexical	verbo de ação	6 (0 locate; 0 locates; 0 locating; 6 located)
<i>confirm</i>	verbo lexical	verbo de ação	11 (9 confirm; 0 confirms; 0 confirming; 2 confirmed)
<i>eliminate</i>	verbo lexical	verbo de ação	5 (5 eliminate; 0 eliminates; 0 eliminating; 0 eliminated)

APÊNDICE II – Advérbios *-ly* utilizados com os VIAESem ambos os subcorpora

VIAEs	Colocados Advérbios <i>-ly</i>		Frequência Subcorpus CorIFA	Colocados Advérbios <i>-ly</i>		Frequência Subcorpus BAWE
	Esquerda	Direita		Esquerda	Direita	
<i>achieve</i>	<i>usually</i>	----	2	<i>exactly</i> <i>completely</i> <i>easily</i> <i>only</i> <i>successfully</i>	<i>completely</i> <i>mutually</i>	7
<i>acquire</i>	<i>hardly</i>	<i>really</i>	4	<i>only</i> <i>instantly</i>	----	2
<i>adopt</i>	----	<i>certainly</i>	2	<i>eventually</i> <i>formally</i> <i>newly</i> <i>recently</i>	<i>previously</i>	5
<i>affect</i>	<i>positively</i> <i>mainly</i> <i>probably</i> <i>directly</i> <i>adversely</i>	----	10	<i>only</i> <i>significantly</i> <i>indirectly</i> <i>adversely</i> <i>clearly</i> <i>directly</i> <i>greatly</i> <i>heavily</i> <i>highly</i> <i>possibly</i> <i>severely</i> <i>vastly</i> <i>substantially</i> <i>particularly</i> <i>consequently</i> <i>greatly</i> <i>negatively</i> <i>positively</i>	<i>mainly</i>	22
<i>apply</i>	<i>specially</i>	----	2	<i>simply</i> <i>only</i> <i>generally</i> <i>necessarily</i> <i>really</i> <i>unsuccessfully</i> <i>potentially</i> <i>universally</i> <i>selectively</i> <i>effectively</i> <i>inaccurately</i> <i>successfully</i> <i>easily</i>	<i>only</i> <i>heavily</i> <i>broadly</i>	17

				<i>readily</i>		
<i>base</i>	<i>finally</i>	----	2	<i>mostly</i> <i>locally</i> <i>quarterly</i> <i>loosely</i> <i>nationally</i> <i>enormously</i> <i>firmly</i> <i>initially</i> <i>mainly</i> <i>hopefully</i> <i>totally</i> <i>functionally</i> <i>purely</i> <i>highly</i> <i>primarily</i> <i>roughly</i> <i>autonomously</i> <i>psychologically</i> <i>entirely</i>	<i>largely</i> <i>primarily</i>	25
<i>compare</i>	----	----	0	<i>subconsciously</i> <i>directly</i> <i>inevitably</i> <i>mainly</i>	<i>correctly</i> <i>exactly</i>	6
<i>conclude</i>	----	----	0	<i>ultimately</i> <i>finally</i>	<i>firmly</i>	5
<i>contribute</i>	<i>largely</i>	<i>practically</i> <i>considerably</i> <i>strongly</i>	6	<i>respectively</i> <i>inevitably</i> <i>only</i> <i>surely</i> <i>probably</i> <i>clearly</i> <i>slightly</i>	<i>entirely</i> <i>significantly</i> <i>substantially</i> <i>externally</i> <i>heavily</i> <i>directly</i> <i>greatly</i> <i>enormously</i> <i>largely</i>	19
<i>control</i>	<i>chemically</i> <i>carefully</i>	<i>severely</i>	6	<i>overtly</i> <i>effectively</i> <i>specifically</i> <i>properly</i> <i>emotionally</i>	----	6
<i>depend</i>	<i>certainly</i>	<i>totally</i>	4	<i>merely</i> <i>largely</i> <i>consequently</i> <i>highly</i> <i>only</i> <i>especially</i> <i>daily</i>	<i>greatly</i> <i>only</i> <i>entirely</i> <i>positively</i> <i>heavily</i> <i>directly</i> <i>largely</i>	18
<i>develop</i>	----	<i>responsibly</i>	2	<i>continuously</i>	<i>virtually</i>	14

				<i>only mainly actually economically newly fully highly</i>	<i>properly autonomously significantly</i>	
<i>exist</i>	----	----	0	<i>only actually really widely</i>	<i>independently necessarily jurally</i>	17
<i>facilitate</i>	<i>daily</i>	<i>daily</i>	5	<i>actually</i>	----	1
<i>improve</i>	<i>surely incredibly highly completely</i>	<i>significantly certainly actually daily</i>	23	<i>specifically significantly increasingly actually historically greatly only slightly</i>	----	15
<i>include</i>	----	----	0	<i>explicitly definitely normally typically only usually currently generally necessarily</i>	----	10
<i>increase</i>	----	----	0	<i>slightly accordingly mutually greatly gradually initially steadily actually equivalently rapidly continually consequently speedily only substantially markedly artificially dramatically considerably abnormally</i>	<i>sharply daily gradually slightly indefinitely steadily significantly drastically only indefinitely slowly greatly slightly dramatically</i>	55

<i>influence</i>	<i>negatively</i>	----	2	<i>strongly solely deeply positively greatly heavily profoundly largely highly clearly</i>	<i>heavily</i>	17
<i>involve</i>	<i>completely directly totally</i>	----	6	<i>normally usually actively necessarily commonly currently directly mainly increasingly indirectly romantically personally</i>	<i>similarly relatively only predominantly especially poorly</i>	25
<i>observe</i>	----	----	0	<i>immediately clearly only commonly easily</i>	----	5
<i>occur</i>	<i>annually typically usually</i>	----	6	<i>normally only probably rarely typically actually patently</i>	<i>freely usually only seemingly simultaneously differently</i>	15
<i>produce</i>	<i>absolutely</i>	----	2	<i>only automatically annually actually collaboratively successfully</i>	<i>only significantly exceptionally socially slightly highly organically tightly locally entirely globally</i>	20
<i>promote</i>	<i>Only</i>	----	2	<i>incorrectly broadly simultaneously consequently</i>	----	4
<i>provide</i>	<i>consequently</i>	----	2	<i>merely</i>	<i>effectively</i>	17

				<i>only</i> <i>automatically</i> <i>really</i> <i>originally</i> <i>mainly</i> <i>concurrently</i> <i>simply</i> <i>effectively</i> <i>namely</i>	<i>luxuriously</i> <i>only</i> <i>internally</i>	
<i>reduce</i>	----	<i>significantly</i>	3	<i>significantly</i> <i>considerably</i> <i>effectively</i> <i>linearly</i> <i>potentially</i> <i>gradually</i> <i>substantially</i> <i>dramatically</i> <i>greatly</i> <i>visibly</i> <i>largely</i> <i>highly</i> <i>severely</i> <i>successfully</i>	<i>sharply</i>	23
<i>relate</i>	<i>antigenically</i> <i>usually</i> <i>friendly</i>	<i>friendly</i>	6	<i>directly</i> <i>distantly</i> <i>closely</i> <i>ancestrally</i> <i>semantically</i> <i>intimately</i>	<i>negatively</i> <i>positively</i> <i>directly</i>	25
<i>require</i>	----	----	0	<i>only</i> <i>simply</i> <i>consequently</i> <i>generally</i> <i>usually</i> <i>absolutely</i>	<i>only</i> <i>accordingly</i> <i>hourly</i> <i>occasionally</i>	13
<i>tend</i>	----	----	0	<i>simply</i> <i>generally</i>	----	2

Quadro 27 – Análise de Mutual Information dos cinco VIAEs mais frequentes do subcorpus do CorIFA

VIAEs	INFORMAÇÃO MÚTUA
<i>increase</i>	1,0500402
<i>provide</i>	0,950271
<i>improve</i>	0,673012
<i>develop</i>	0,673012
<i>adopt</i>	0,500402

Fonte: Elaborado pela autora, 2017.

ANEXOS

ANEXO 1 – Carta de consentimento livre e esclarecido

CARTA DE CONSENTIMENTO LIVRE E ESCLARECIDO: Para os participantes

Caro(a) Senhor(a)

A coordenação das disciplinas de "Inglês para Fins Acadêmicos" da UFMG conduz pesquisas que tem por objetivo estudar o desenvolvimento das habilidades de leitura, de escrita, de audição e de fala de aprendizes de inglês para fins acadêmicos. Cada projeto de pesquisa está devidamente autorizado pela Câmara de Pesquisa da Faculdade de Letras da UFMG.

A fim de que os projetos possam ser desenvolvidos, é necessária a sua autorização, vez que as pesquisas constarão da coleta das suas redações produzidas enquanto aluno do curso. A sua participação nesta pesquisa é voluntária e não determinará qualquer risco nem tratá desconfortos. Além disso, sua participação é importante para o aumento do conhecimento a respeito dos processos de aquisição e desenvolvimento das quatro habilidades supracitadas por alunos universitários brasileiros, podendo beneficiar outros alunos futuramente na melhoria do ensino de língua inglesa no nível superior.

Informamos que o/a Sr(a). tem a garantia de acesso, em qualquer etapa dos estudos, sobre qualquer esclarecimento de eventuais dúvidas. Se tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com a coordenação do programa (3409-3839) ou com o Comitê de Ética em Pesquisa (CoEP) da Universidade Federal de Minas Gerais, situado na Av. Antônio Carlos, 6627. Unidade Administrativa II - 2º andar - Campus Pampulha, telefone 3409-4592 / 3409-4027.

Também é garantida a liberdade da retirada de consentimento a qualquer momento e deixar de participar do estudo.

Fica também garantido que as informações obtidas serão analisadas em conjunto com as de outras pessoas, não sendo divulgada a identificação de nenhum dos participantes.

O/A Sr(a). tem o direito de ser mantido atualizado sobre os resultados parciais das pesquisas e caso seja solicitado, todas as informações que solicitar lhe serão fornecidas.

Não existirão despesas ou compensações pessoais para o participante em qualquer fase dos estudos. Também não há compensação financeira relacionada à sua participação.

Os participantes das pesquisas comprometem-se a utilizar os dados coletados somente para pesquisa, e os resultados serão veiculados através de artigos científicos, em revistas especializadas e/ou em encontros científicos e congressos, sem nunca tomar possível a sua identificação.

Abaixo se encontra o Termo de Consentimento Livre e Esclarecido, para concordância caso não tenha ficado qualquer dúvida.

Deise Prina Dutra - Coordenadora Geral do IFA/UFMG

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO *

Acredito ter sido suficientemente informado a respeito dos estudos conduzidos pela coordenação do Inglês para Fins Acadêmicos/UFMG. Ficaram claros para mim quais são os propósitos dos estudos, os procedimentos a serem realizados, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro, também, que a minha participação é isenta de despesas e que tenho garantia do acesso aos resultados e de esclarecer minhas dúvidas a qualquer tempo. Concordo voluntariamente em participar e estou ciente de que poderei retirar o meu consentimento a qualquer momento sem penalidade ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

Concordo

Discordo

