

BEVILACQUA, C. R.; MACHADO, M. A. Uma análise da presença de unidades fraseológicas em dicionários bilíngues espanhol/português e sua representatividade no espanhol atual. *ReVEL*, vol. 15, n. 29, 2017. [www.revel.inf.br]

UMA ANÁLISE DA PRESENÇA DE UNIDADES FRASEOLÓGICAS EM DICIONÁRIOS BILÍNGUES ESPANHOL/PORTUGUÊS E SUA REPRESENTATIVIDADE NO ESPANHOL ATUAL

**AN ANALYSIS OF THE PRESENCE OF PHRASEOLOGICAL UNITS IN
SPANISH/PORTUGUESE BILINGUAL DICTIONARIES AND ITS IMPORTANCE
IN CURRENT SPANISH**

Cleci Regina Bevilacqua¹

Manuela Arcos Machado²

cleci.bevilacqua@gmail.com

arcosmanuela@gmail.com

RESUMO: Este trabalho insere-se no âmbito de estudos do Léxico e estabelece a relação entre Fraseologia e Lexicografia, tendo como objeto de estudo as Unidades Fraseológicas (UFs) e sua representação em dicionários bilíngues. Assim, o objetivo deste trabalho é fazer uma análise quantitativa e qualitativa da fraseologia presente em dois dicionários bilíngues – Dicionário Santillana para Estudantes (2009) e Dicionário Escolar Michaelis Espanhol (2006). O referencial teórico que sustenta o trabalho baseia-se, sobretudo, na proposta de Corpas Pastor (1996), que classifica as UFs em três esferas: colocações, locuções e enunciados fraseológicos. Fundamenta-se ainda em pressupostos da Lexicografia (Bugueño, 2013; Welker, 2008; Duran; Xatara, 2007). A coleta das UFs analisadas foi feita no intervalo lematístico da letra C, correspondente a 1% do total de entradas de cada obra. Os procedimentos metodológicos seguiram três etapas: a) análise do *front matter* das obras, a fim de buscar informações sobre a seleção e lematização das UFs; b) análise quantitativa e qualitativa da representação dessas unidades na microestrutura dos dicionários; e c) análise da representatividade de uso e geográfica das UFs coletadas, considerando sua ocorrência no *Corpus del Español del Siglo XXI – CORPES XXI* –, da Real Academia Espanhola (RAE). Entre os resultados obtidos, destacamos a ausência de critérios tanto para a seleção das UFs como para sua lematização em ambos os dicionários.

PALAVRAS-CHAVE: fraseologia; dicionários bilíngues; CORPES XXI; língua espanhola.

¹ Professora do Departamento de Línguas Modernas e do Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul. Coordenadora do Projeto Termisul e do Grupo de Pesquisa Termisul no CNPq.

² Aluna de Mestrado no Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul.

ABSTRACT: The present study fits within the scope of the Lexicon studies and establishes the relation between Phraseology and Lexicography, having the Phraseological Units (PUs) and their representation in bilingual dictionaries as object of study. Thus, this work aims to carry out a qualitative and quantitative analysis of the phraseology found in two bilingual dictionaries - *Diccionario Santillana para Estudiantes* (2009) and *Diccionario Escolar Michaelis Espanhol* (2006). The theoretical framework that supports the study is mostly based on the proposition of Corpas Pastor (1996), who categorizes the PUs into three spheres: collocations, phrases and phraseological statements. It is also based on Lexicography assumptions (Bugueño, 2013, Welker, 2008, and Duran & Xatara, 2007). The collect of the analyzed PUs was done in the lemmatical interval of letter C, corresponding to 1% of the total entries of each work. The methodological procedures followed three steps: a) analysis of the *front matter* of the dictionaries to seek information on the selection and stemming of the PUs; b) quantitative and qualitative analysis of these units' representation in the micro-structure of the dictionaries and c) analysis of the use and geographical representativeness of the collected PUs, considering their occurrence in *Corpus del Español del Siglo XXI – CORPES XXI –*, of *Royal Spanish Academy*. Among the results obtained, we highlight the absence of criteria for both the selection of PUs and their lemmatization in both dictionaries.

KEYWORDS: phraseology; bilingual dictionaries; CORPES XXI; Spanish.

INTRODUÇÃO

O ensino do Espanhol como Língua Estrangeira (ELE), mesmo que difundido no mundo todo, ainda carece de uma oferta variada e acessível de materiais lexicográficos especializados no par de línguas português-espanhol. No que tange à Fraseologia, embora haja um número considerável de dicionários fraseológicos³, constata-se ainda uma carência desse tipo de obra voltada especificamente para aprendizes, sobretudo aprendizes brasileiros dessa língua. Considerando que o foco deste texto é a representação das Unidades Fraseológicas (UFs) em dicionários bilíngues português-espanhol, é importante definir a Fraseologia e as UFs. Entendemos a Fraseologia como a disciplina que se ocupa das combinações lexicais com certo grau de fixação e que tem como objeto de estudo as UFs. Tais unidades, segundo Corpas Pastor (1996:15), são definidas como “combinações lexicais formadas por ao menos duas palavras gráficas, cujo limite superior se situa no nível da oração composta” (tradução nossa)⁴.

Nesse contexto, alguns dicionários bilíngues espanhol/português (Es/Pt) e português/espanhol (Pt/Es), como o *Diccionario Santillana para estudiantes* (2009) e o *Diccionario Escolar Michaelis Espanhol* (2006), aqui analisados, procuram incluir

³ Alguns exemplos são o dicionário *Redes* (2004), o *Diccionario Fraseológico del Español Moderno* (1994) e o *Diccionario de Colocaciones del Español* (DiCE) de acesso *on-line*. Vale ressaltar que esses dicionários, além de serem monolíngues, são de difícil acesso, pois dependem de importação e apresentam um alto custo.

⁴ No original: “[...] combinaciones lexicales formadas por al menos dos palabras gráficas, cuyo límite superior se sitúa en el nivel de la oración compuesta” (Corpas Pastor, 1996:15).

as UFs em seus verbetes. Contudo, ao consultar esses dicionários, o usuário poderá constatar que a fraseologia presente neles muitas vezes é apresentada de forma assistemática, isto é, não há critérios claros para a lematização dessas unidades. Ora aparecem na entrada de uma das palavras que formam a UF, ora na de outra. Por outro lado, muitas unidades lematizadas não se mostram representativas na língua espanhola, apresentando um baixo ou nulo índice de frequência em *corpora* de referência⁵, sendo utilizadas somente em regiões específicas, como veremos em nossa análise. Além disso, pode-se constatar essa assistemática já no *front matter* (introdução) das obras, pois, em geral, elas não oferecem informações quanto à seleção e à lematização desse tipo de unidade.

De forma complementar, há poucos os estudos sobre a inserção das UFs em dicionários em geral e bilíngues em particular, à exceção de reivindicações pontuais sobre a necessidade de introduzir esse tipo de informação no dicionário, como vemos em Hausmann (1990), por exemplo. Dessa forma, buscamos, com o presente trabalho, contribuir para os estudos fraseológicos e para sua interface com a Lexicografia e a Metalexigrafia, assim como mostrar sua importância no processo de aquisição do ELE e o papel dessas obras nesse processo.

Levando em conta as constatações anteriores, o objetivo deste trabalho é fazer uma análise quantitativa e qualitativa da fraseologia presente nos dois dicionários bilíngues já referidos acima, escolhidos por serem amplamente utilizados, tanto nas escolas públicas e privadas, quanto em cursos livres (ver Mattos, 2010:124; Noimann, 2007:13; Fioravanti, 2015:103). Assim, serão analisados intervalos lematizados da letra C correspondentes a 1% do total de entradas de cada dicionário com o intuito de observar: a) a lematização das UFs e sua representatividade numérica em relação ao conjunto de entradas; b) sua representação na microestrutura; e c) sua representatividade numérica e geográfica no espanhol atual, a partir de buscas realizadas no *Corpus del Español del Siglo XXI* (CORPES XXI), da Real Academia Española (RAE).

⁵ Tagnin (2013) define um *corpus* de referência como “*corpus* que serve de termo de comparação para o *corpus* de estudo. Em geral, deve ter três a cinco vezes o tamanho do *corpus* de estudo”. Já segundo a própria RAE, um *corpus* de referência tem como propósito básico servir para obter as características globais que uma língua apresenta em um determinado momento de sua história. No caso do espanhol, um *corpus* de referência deve conter textos de todos os tipos (de diferentes meios, isto é, textos orais e escritos, e de diferentes gêneros) e oriundos de todos os países que constituem o mundo hispânico.

Para dar conta dos objetivos aqui propostos, inicialmente, trazemos o referencial teórico utilizado, principalmente relativo às UFs e aos dicionários bilíngues. Num segundo momento, apresentamos nossa metodologia indicando como foram escolhidos os intervalos lexicais observados, selecionadas as UFs e realizada sua análise qualitativa e quantitativa. Também apresentaremos a descrição dos dicionários analisados e o CORPES XXI, ferramenta utilizada para a verificação da frequência e distribuição geográfica das fraseologias. Por fim, serão apresentados os dados obtidos, seguidos de nossas considerações finais. Buscamos, desse modo, trazer contribuições para os estudos lexicográficos e metalexográficos relacionados à fraseologia e à aquisição de línguas estrangeiras, neste caso, o espanhol.

1 REFERENCIAL TEÓRICO

Nesta seção, apresentamos alguns aspectos relativos à função e à estrutura dos dicionários bilíngues e às características e à tipologia das UFs. Para tanto, nos fundamentamos em autores como Duran e Xatara (2007), Welker (2008), Bugueño (2013) e Corpas Pastor (1996).

1.1 DICIONÁRIOS BILÍNGUES: FUNÇÕES E CARACTERÍSTICAS

A aquisição da competência lexical⁶ tem um papel fundamental no processo de aprendizagem de uma Língua Estrangeira (LE), pois se constitui em um dos elementos que conformam a competência linguística e possibilita, portanto, que o aluno entenda e se expresse adequadamente na língua que está aprendendo.

Sob este ponto de vista, os dicionários bilíngues exercem um papel fundamental na aquisição de uma LE e, por esta razão, podem ser considerados também como dicionários pedagógicos. Tal perspectiva é fundamentada em autores como Welker (2008:22-24) que, ao trazer a citação de Söfgen (1991) indica que os dicionários bilíngues, desde o início de sua história, revelavam uma forte orientação didática, pois a função dessas obras era ajudar os estudantes de línguas estrangeiras (Söfgen, 1991, *apud* Welker, 2008:23).

⁶ Entendemos a competência lexical a partir da proposta de Scaramucci (1995:236) que a define “de forma ampla e abrangente, incorporando os vários níveis de conhecimento da palavra (pragmático-discursivo, lexical, semântico, sintático, morfológico, fonológico), além de uma dimensão cognitiva ou de uso”. A autora considera ainda que essa competência faz parte da competência comunicativa de um aprendiz de língua.

Nessa mesma linha, Duran e Xatara (2007) afirmam que

Há cerca de 30 anos, os dicionários, que sempre foram um importante acessório para o aprendiz de idiomas, começaram a refletir a preocupação de atender adequadamente as necessidades desse usuário. A especialização da Lexicografia para essa finalidade cresceu tanto que passou a ter denominação própria: Lexicografia Pedagógica (LP) e compreende tanto dicionários bilíngues quanto dicionários monolíngues para estrangeiros. (Duran; Xatara, 2007: 204).

As autoras indicam ainda que obras dessa natureza têm a preocupação de, entre outros aspectos, simplificar a busca, oferecer informações claras sobre o uso, alertar para enganos comuns que possam ser cometidos e minimizar a possibilidade de incompreensão ou conclusões ambíguas por parte dos aprendizes.

Como seu próprio nome indica, esse tipo de obra está conformado por duas partes: uma que parte da LE para a língua materna, e outra que parte da língua materna para a LE. No caso de um aprendiz brasileiro de espanhol, a primeira direção tem a função de decodificação (compreensão), ou seja, de auxiliar o aprendiz a compreender o significado e o uso das palavras da LE. A segunda direção tem a função de codificação (produção), isto é, de ajudar o aprendiz a produzir textos na língua que está aprendendo.

Para dar conta das funções mencionadas, tais obras devem apresentar um conjunto de informações que assegurem sua coerência interna. Esse conjunto de informações está conformado pelo *front matter*, macroestrutura, microestrutura, medioestrutura e *back matter*. Discorreremos sobre cada uma dessas estruturas a seguir.

Bugueño (2013:23) define o *front matter* como as páginas introdutórias do dicionário. Para o autor, essa parte do dicionário deve cumprir três funções: em primeiro lugar, deve assinalar o usuário ao qual o dicionário se destina; em segundo, indicar a função do dicionário e, em terceiro, deve servir como um manual de instruções que facilitem a utilização da obra.

A macroestrutura, por sua vez, consiste na ordenação dos signos-lemas (entradas), e deve respeitar um algoritmo de busca, isto é, respeitar uma ordem de inclusão das entradas, como, por exemplo, a ordem alfabética. Além disso, possui duas dimensões complementares: a) a definição macroestrutural quantitativa, que se refere à quantidade de palavras que um dicionário deve lematizar e b) a definição macroestrutural qualitativa, relativa aos tipos de unidades lematizadas (Bugueño, 2013:23).

Por sua vez, a microestrutura é definida como o conjunto de informações oferecidas para o signo-lema. Segundo Bugueño (2013:25), a microestrutura possui uma função predicativa, uma vez que é a responsável por oferecer a definição do verbete. Ademais, apresenta dois comentários básicos: o comentário de forma e o comentário semântico, que podem ser pensados pela concepção saussureana significante-significado, na qual o comentário de forma refere-se ao significante, e o comentário semântico, ao significado. Assim, o comentário de forma oferece informações referentes à ortografia, morfologia, divisão silábica, etc., enquanto que o comentário semântico oferece a definição do signo-lema, sinônimos, exemplos, etc.

Finalmente, a medioestrutura refere-se à rede de remissivas estabelecida entre as distintas entradas, e o *back matter*, às informações apresentadas no final da obra (por exemplo, listas de verbos, bibliografia consultada, etc.).

A partir das afirmações anteriores, no presente trabalho, consideramos os dicionários bilíngues como dicionários pedagógicos pela função que cumprem no processo de aprendizagem de determinada LE. Por terem essa função, pensamos que é fundamental que estejam elaborados a partir de princípios que sustentem a coerência entre as estruturas mencionadas acima e que ofereçam um conjunto de UFs adequado ao seu público e função, indicando inclusive os critérios de seleção e de lematização de ditas unidades. A justificativa para essa inserção deve-se ao fato de elas fazerem parte da competência lexical do aprendiz, conforme mencionado anteriormente.

1.2 A FRASEOLOGIA: CARACTERÍSTICAS, TIPOLOGIA E SUA INSERÇÃO EM DICIONÁRIOS

Corpas Pastor (1996: 20), além de formular a definição de UFs anteriormente apresentada, propõe um conjunto de propriedades que as caracterizam e uma tipologia que permite classificá-las em três grandes grupos. As características propostas pela autora são:

- a) Alta frequência de uso e coocorrência dos elementos que as conformam: refere-se à frequência com que os elementos constituintes de uma UF aparecem combinados, que será superior à frequência da aparição individual dos mesmos elementos na língua (coocorrência).

- b) Institucionalização: está relacionada ao uso, à repetição e à frequência de aparição. Assim, o uso frequente de uma UF, isto é, sua repetição, determina sua convencionalização.
- c) Idiomaticidade: relativa à propriedade semântica que apresentam certas UFs, a partir da qual o significado global das unidades não se deduz pelo significado individual de cada um de seus elementos constitutivos, mas do conjunto dos elementos.
- d) Variação: estabelece que a fixação das UFs é relativa. Dessa forma, muitas unidades podem apresentar certa variação léxica, isto é, podem ter, no discurso, diferentes variantes. Por exemplo, a UF *Todo queda en casa* tem como variante *Todo queda en familia*. Apesar dessa variação, ainda apresentam certo grau de fixação, uma vez que não aceitam como variante o componente *hogar*, por exemplo.
- e) Gradação: indica que as UFs podem apresentar as características anteriores (frequência, institucionalização, idiomaticidade e variação) em diferentes graus.

Para classificá-las, Corpas Pastor (1996) dividiu-as em três esferas:

- a) Colocações: são unidades fixadas somente na norma, isto é, sintagmas gerados pelas regras da língua, mas que, pelo uso, apresentam certo grau de restrição combinatória. São combinações estáveis, pré-fabricadas. Por exemplo, em espanhol *fumador empedernido* (fumante inveterado).
- b) Locuções: são unidades com alta frequência de coaparição e uso e que possuem coesão semântica, isto é, uma impossibilidade de substituir, eliminar, reordenar ou manipular sintaticamente seus elementos constitutivos. São classificadas segundo a função que desempenham na oração em: locuções nominais, adjetivas, adverbiais, verbais, prepositivas e conjuntivas. Por exemplo, *a menudo* (com frequência) é uma locução adverbial.
- c) Enunciados fraseológicos: constituem enunciados e atos de fala por si mesmos, além de estarem fixados na fala e fazerem parte do acervo sociocultural da comunidade falante. Por exemplo, *quien calla otorga* (quem cala consente).

Independentemente da classificação das UFs, é possível constatar que os dicionários bilíngues incluem, em certa medida, essas unidades, embora nem sempre o façam de forma sistemática ou previamente esclarecida ao usuário, conforme apontaremos em nossa análise. Nesse sentido, Corpas Pastor (2004) explica que, embora a lexicografia teórica reconheça a importância de se oferecer informação fraseológica nos dicionários bilíngues, durante muito tempo foram poucos os estudos sobre o assunto, exceto algumas manifestações pontuais sobre a necessidade de introduzir esse tipo de informação nos dicionários.

Dessa forma, verifica-se que uma das deficiências mais comuns dos dicionários bilíngues é, precisamente, o relativo abandono a que estão submetidas as UFs (Corpas Pastor, 2004). Tal fato pode ser comprovado pela falta de sistematicidade em relação à inserção e ao tratamento dessas unidades tanto na sua macro quanto na sua microestrutura. Para a referida autora, a localização da informação fraseológica nos dicionários é um dos temas mais debatidos entre estudiosos da área (Hausmann, 1990, por exemplo), conforme já apontamos. Alguns autores (Zgusta, 1971; Verstraten, 1992 *apud* Corpas Pastor, 2004) defendem sua inclusão na macroestrutura, ou seja, que as UFs sejam inseridas como entradas específicas e independentes. Embora não haja um consenso sobre esse aspecto, é possível observar que a maioria dos dicionários coincide em apresentar a fraseologia no comentário semântico da microestrutura, isto é, na parte do verbete em que se oferece o significado da palavra lematizada, tal como veremos nos dicionários analisados. Como consequência, o usuário, quando precisa, por exemplo, procurar o significado ou o equivalente⁷ de uma expressão em espanhol como *hacer el caldo gordo*, e não encontra orientações no *front matter* sobre onde buscá-la, não sabe se deve partir do verbo *hacer*, ou do substantivo *caldo*.

Um dos pontos para poder definir se uma UF deve constar de uma obra lexicográfica deveria ser a sua ocorrência de uso. No entanto, não é fácil definir um parâmetro para sua inserção, pois sua definição depende das fontes e da metodologia

⁷ Os equivalentes fazem parte do comentário semântico nos dicionários bilíngues. A partir de Bugueño (2013:28), poderíamos considerá-los como uma “paráfrase explanatória sinonímica”, ou seja, se oferece uma ou mais formas correspondentes na língua materna (no sentido língua estrangeira/língua materna) ou na língua estrangeira (sentido língua materna/língua estrangeira). Sobre os equivalentes, Tarp (2008: 58) afirma que: “Nos dicionários bilíngues é inconveniente e estruturalmente complexo colocar, junto aos equivalentes, todos os dados que o aprendiz precisa para satisfazer suas necessidades; por isso, o dicionário bilíngue deve ser concebido como uma *ponte bilíngue* que conduz a outro dicionário baseado na língua estrangeira e no qual os dados possam ser localizados de forma mais manuseável e acessível” (tradução nossa).

prevista pelos lexicógrafos. Nesse sentido, cabe ressaltar que atualmente os dicionários têm sido construídos a partir de *corpus*, o que permitira utilizar os critérios de frequência para a seleção das entradas, incluindo as UFs. No entanto, definir um critério de corte depende de vários fatores, entre eles o tamanho do *corpus* e a finalidade e a função da obra que se quer elaborar.

Como um dos propósitos do presente trabalho é verificar o uso e a representatividade geográfica das UFs selecionadas nos dicionários analisados, trazemos algumas perspectivas sobre esse tema e apresentamos nossa posição sobre este aspecto. Em relação à representatividade de uma palavra ou expressão lematizada, Duran (2004:97) afirma que parece bastante lógico definir a nomenclatura de um dicionário bilíngue pedagógico pela sua frequência em um *corpus* representativo. Em relação às UFs, especificamente, para Haensch (1892) e Wotjiak (2007) citados por Mattos (2010:159) o critério de frequência – embora não seja o único – ajuda na verificação da pertinência de lematização de uma UF. Entretanto, embora se considere relevante o critério de frequência para dicionários bilíngues, não observamos um consenso em relação a uma proposta que determine quão frequente uma palavra (ou expressão) deva ser em um *corpus* representativo para ser lematizada.

Contudo, Biderman (1984), ao tratar de dicionários escolares, oferece um critério de frequência sistematizado para a composição de sua macroestrutura. Considerando que os dicionários escolares por ela descritos se assemelham em determinados aspectos aos dicionários bilíngues aqui analisados, optamos por tomar como referência o critério de frequência proposto pela autora. Pensamos que entre os aspectos que aproximam os dicionários escolares e os bilíngues estão a função pedagógica e o número de entradas. Assim, para Biderman, os dicionários escolares devem estar de acordo com o nível do aprendiz e, portanto, devem conter entre 10.000 e 30.000 verbetes – tamanho semelhante ao dos dicionários bilíngues aqui analisados⁸.

⁸ O edital do PNLD 2012 (Brasil, 2012: 1-2) prevê quatro tipos de dicionários escolares: 1) tipo 1 destinado ao 1º ano do Ensino Fundamental (EF), contém um mínimo de 500 e um máximo de 1.000 verbetes e apresenta proposta lexicográfica adequada às demandas do processo de alfabetização inicial; 2) tipo 2 destinado para estudantes do 2º ao 5º ano do EF, contém de 3.000 a 15.000 verbetes e deve dar conta da consolidação do domínio da escrita e da organização da linguagem; 3) tipo 3 destinado aos alunos do 6º ao 9º ano do EF, contém entre 19.000 e 35.000 entradas e sua proposta lexicográfica deve caracterizar-se como um dicionário padrão de uso escolar; 4) tipo 4 destinado aos estudantes do ensino médio, contém entre 40.000 e 100.000 verbetes e sua proposta lexicográfica deve seguir os parâmetros de um dicionário padrão, mas adequado às demandas do seu público.

Assim, Biderman (1984) propõe que a confecção de um dicionário deva ser baseada num *corpus* representativo (no mínimo, cinco milhões de ocorrências abrangendo diferentes gêneros, tais como o literário, o jornalístico, o técnico e o acadêmico), e que permita estabelecer uma hierarquia de frequência para as entradas. Em relação a esse último aspecto, a autora afirma que, em um dicionário escolar, devem ser incluídas palavras com frequência igual ou superior a 10 ocorrências, de modo a oferecer um vocabulário pertinente e representativo da língua para o usuário que se encontra na posição de aprendiz. Dessa forma, para examinar a representatividade na língua espanhola das UFs obtidas nos intervalos lematizados dos dicionários Michaelis (2006) e Santillana (2009), tomaremos como critério de frequência proposto por Biderman (1984) para a elaboração de um dicionário escolar.

2. METODOLOGIA

A coleta e a análise dos dados foram divididas em três etapas. Na primeira, trazemos as informações relativas à fraseologia encontradas no *front matter* de ambos os dicionários, a fim de demonstrar se apresentam critérios de coleta e lematização das UFs, bem como orientações para sua consulta. Essas informações foram coletadas a partir da leitura dos textos apresentados nessa parte das obras.

A segunda etapa refere-se à coleta das UFs registradas pelos dicionários. Para tanto, selecionamos intervalos lematizados da letra C de ambos os dicionários. Tomamos o equivalente a 1% do número de entradas total de cada dicionário, de modo a gerar resultados numéricos proporcionalmente adequados. Desse modo, levando em conta que o dicionário Santillana (2009) contém 38.000 entradas e o Michaelis (2006), 28.000, os intervalos analisados se referem aos primeiros 380 verbetes desde o início da letra C para o primeiro dicionário, e os primeiros 280 do segundo, no sentido Es/Pt⁹. Nosso objetivo, nessa etapa, foi observar: 1) a existência de sistematicidade na inclusão de UFs – ainda que não informada no *front matter* dos dicionários, para identificar padrões de inserção e 2) a representatividade numérica dessas estruturas em relação ao total de entradas dos dicionários.

Os dados coletados nessa etapa foram registrados em um quadro que contém as seguintes informações: “entrada” – refere-se ao signo-lema em que se encontra lematizada a UF –; “UF lematizada”; “UF em português”, quando existente nessa

⁹No apêndice é possível visualizar todas as UFs analisadas.

língua e “significado da UF” – definição em português da UF. Os dados obtidos na segunda etapa foram registrados em uma tabela que se encontra no apêndice.

Na terceira e última etapa, apresentamos a análise qualitativa e quantitativa das UFs selecionadas. O propósito da análise quantitativa foi obter uma amostra da presença de UFs nos dicionários, no sentido Es/Pt, a partir da comparação do número de UFs identificadas com o número de verbetes de cada obra. Por sua vez, a análise qualitativa foi realizada com 10 UFs e foi dividida em dois momentos: 1) análise da microestrutura dos verbetes selecionados, para compreender como, de fato, ocorre a inserção das UFs e identificar se há um padrão de apresentação de seu equivalente no português e de seu respectivo significado e 2) análise das UFs obtidas por meio dos filtros de frequência do CORPES XXI para poder verificar a representatividade numérica de uso e de distribuição geográfica de cada uma das UFs.

Antes de iniciarmos análise dos dados, apresentaremos brevemente os dicionários analisados e o CORPES XXI, a fim de oferecer um panorama sobre as obras e a ferramenta utilizada na coleta dos dados estatísticos.

2.1 DESCRIÇÃO DOS DICIONÁRIOS MICHAELIS (2006) E SANTILLANA (2009)

A descrição dos dicionários foi feita considerando as informações apresentadas na seção 1.1, relativa à constituição dos dicionários bilíngues.

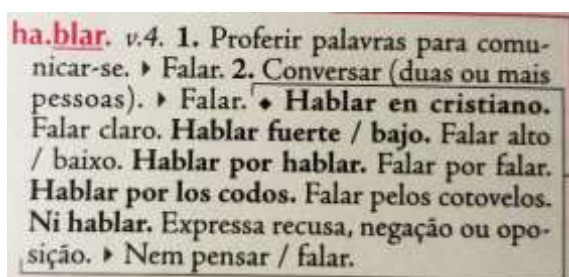
Tanto o dicionário Santillana (2009) quanto o Michaelis (2006) se apresentam como obras voltadas a um usuário estudante brasileiro em processo de aprendizagem de ELE. Portanto, no que concerne à sua função, considerando que os dois dicionários apresentam duas direções (Es/Pt e Pt/Es), podemos dizer que podem, para os aprendizes brasileiros, exercer tanto a função de decodificação ou compreensão do espanhol (Es/Pt) quanto a de codificação ou produção nessa língua (Pt/Es). Também podemos caracterizá-los como dicionários bilíngues de orientação semasiológica, isto é, que partem dos significantes para os significados, oferecendo o signo-lema (entrada) seguido seus equivalentes. Ambos os dicionários também sugerem um critério sincrônico para a inserção de entradas, uma vez que se propõem a registrar formas que correspondem a usos frequentes num recorte atual da língua.

Embora numa análise mais geral os dois dicionários apresentem características comuns, em especial no que diz respeito a sua classificação e função,

cada um contém suas especificidades, inclusive no que tange ao tratamento da fraseologia, conforme veremos na seção 3.1.

O dicionário Santillana (2009) se apresenta como uma publicação voltada a estudantes brasileiros de espanhol que estejam entre os níveis iniciante e intermediário e possui uma macroestrutura de 38.000 entradas. Apresenta um *front matter* de seis páginas, em que se oferecem explicações sobre a estrutura dos verbetes, remissões, exemplos de uso, informações adicionais, etc. Para ilustrar sua microestrutura, trazemos o exemplo da palavra *hablar*.

Figura 1: Exemplo de microestrutura do *Diccionario* Santillana (2009).

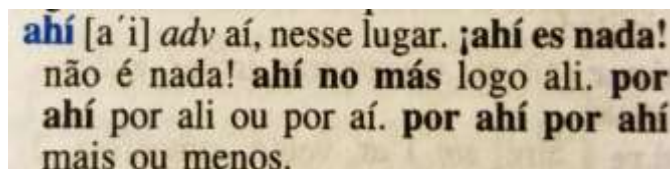


Fonte: Santillana (2009:XI).

Como é possível observar, a obra oferece informação sobre a categoria gramatical da entrada, sua definição em português e seu equivalente, e várias fraseologias com seus equivalentes ou definições em português.

O dicionário Michaelis (2006), por sua vez, possui uma macroestrutura de 28.000 verbetes. Oferece um *front matter* de duas páginas, em que há informações sobre os verbetes no que diz respeito à classe gramatical, transcrição fonética, tradução, etc. Apresentamos o exemplo de sua microestrutura com a entrada da palavra *ahí*, que contém informação sobre a categoria gramatical da entrada, seu equivalente em português, e várias fraseologias com seus respectivos equivalentes.

Figura 2: Exemplo de microestrutura do *Diccionario* Michaelis (2006).



Fonte: Michaelis (2006).

Pelas informações anteriores, parece que o dicionário Santillana (2009) oferece um conjunto maior de entradas e também uma microestrutura e *front matter*

mais completos que o Michaelis (2006). Contudo, essa é uma constatação inicial que deve ser comprovada com estudos mais detalhados futuramente.

2.2 CORPES XXI

O CORPES XXI é o *corpus* de referência mais recente da RAE, cujo propósito é registrar o espanhol atual. Para isso, possui 225 milhões de formas, que ocorrem em textos produzidos entre os anos 2001 a 2012. Tais textos são oriundos de diferentes meios (10% de registros orais e 90% escritos) e de diferentes gêneros textuais. Os textos escritos procedem de livros (40%), publicações de jornais (40%), material de internet (7,5%) e miscelânea¹⁰ (2,5%). Para os textos orais não são indicadas as fontes de coleta. Para ambos os registros, os textos foram classificados em dois grandes gêneros, que se encontram subdivididos por temáticas: ficção (novela, relato, roteiro e teatro) e não ficção (livros científicos e imprensa).

No que respeita à representatividade geográfica do *corpus*, 30% das formas são procedentes de registros da Espanha e 70% de registros da América. O material produzido na América, por sua vez, se classifica em nove zonas linguísticas, a saber: Andina, Antilhas (caribenha), Caribe continental, Chilena, Estados Unidos, México e América Central e Rio da Prata. Como novidade em relação aos demais *corpora* da RAE, o CORPES XXI inclui também textos oriundos da Guiné Equatorial e das Filipinas¹¹.

Para nossa análise, utilizaremos o filtro chamado *estadística*, que oferece, tal qual seu nome sugere, informações estatísticas sobre a frequência absoluta da forma pesquisada, além da distribuição das ocorrências dessas formas nas zonas linguísticas referidas anteriormente. A Figura 3, a seguir, ilustra a consulta da UF *no caber en sí*, na qual é possível observar a frequência absoluta da expressão no *corpus* (34 ocorrências) e a sua distribuição geográfica.

¹⁰A RAE não especifica que tipo de textos de refere à classificação miscelânea.

¹¹Essas informações foram retiradas do dossiê informativo do CORPES XXI oferecido pela própria RAE em seu site oficial, disponível em: http://www.rae.es/sites/default/files/CORPES_Sistema_de_codificacion_12_2015.pdf

Figura 3: Exemplo de consulta de UF.



Fonte: CORPES XXI.

Portanto, nossa análise das UFs por meio das ferramentas do CORPES XXI está orientada por dois critérios: a) a frequência, isto é, o número de registros que a UF pesquisada possui no *corpus* e b) critério de distribuição diatópica, isto é, observar a presença da UF nas diferentes zonas linguísticas do espanhol.

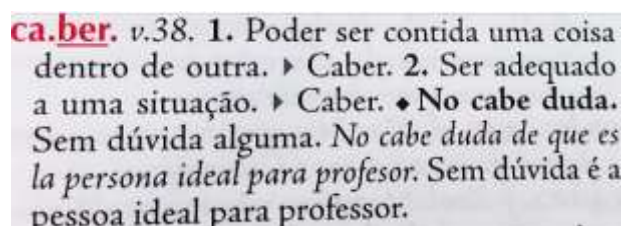
3. DESCRIÇÃO E ANÁLISE DOS DADOS

A análise dos dados obtidos, tal como a metodologia, foi dividida em três momentos: em 3.1 analisamos qualitativamente o *front matter* de cada dicionário em relação ao registro e ao tratamento das UFs na sua microestrutura, com o intuito de observar se há alguma sistematicidade na lematização das UFs e no padrão de informações oferecidas. Em 3.2 analisamos quantitativa e qualitativamente as UFs obtidas intervalos lematizados da letra C e, em 3.3, analisamos uma amostra de 10 UFs de cada dicionário no CORPES XXI, sob uma perspectiva quantitativa, ao observar seus índices de frequência total e sua distribuição geográfica, e qualitativa, tendo por base o critério de frequência proposto por Biderman (1984).

3.1 ANÁLISE DO FRONTMATTER

A análise do *front matter* do *Dicionário Santillana* (2009) permitiu identificar que há um pequeno parágrafo chamado “locuções”, em que consta a seguinte explicação: “incluiram-se a muitos verbetes palavras, locuções, refrãos e expressões idiomáticas derivadas ou relacionadas à cabeça do verbete. Elas são introduzidas pelo símbolo ♦ e podem apresentar definição analítica, versão ao português e/ou exemplo de uso” (Santillana, 2009:XI). A Figura 4 exemplifica a inserção da UF *no cabe duda* no verbete da entrada *caber*, identificada pelo símbolo ♦ e seguida do equivalente no português (*sem dúvida alguma*), de exemplo de uso da UF em espanhol e da tradução do exemplo para o português.

Figura 4: Verbetes *caber*.



ca.ber. v.38. 1. Poder ser contida uma coisa dentro de outra. ▶ Caber. 2. Ser adequado a uma situação. ▶ Caber. ♦ **No cabe duda.** Sem dúvida alguma. *No cabe duda de que es la persona ideal para profesor.* Sem dúvida é a pessoa ideal para professor.

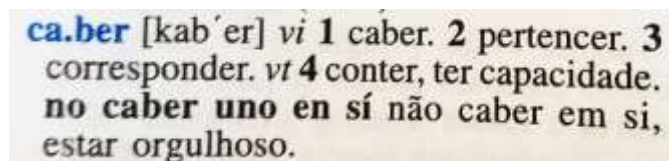
Fonte: Santillana (2009).

Contudo, o dicionário não indica os critérios de seleção e de inserção dessas unidades nos verbetes, isto é, não aponta nem os critérios que justificam sua inclusão na obra, nem a palavra a partir da qual a UF é lematizada. Assim, se o usuário necessita consultar, por exemplo, a expressão *no cabe duda*, não saberá por qual entrada pesquisar, se pelo advérbio de negação *no*, se pelo verbo *caber* ou pelo substantivo *duda*. Nesse caso, optou-se por inseri-la na entrada do verbo *caber*, e não nas entradas de *duda* e *dudar*, que poderiam ser outras possibilidades de inserção da UF. Esse dado poderia levar a pensar que o critério de inserção seria na entrada do verbo, considerado o núcleo da expressão. Contudo, para confirmar essa possibilidade, seria necessário realizar a análise de um conjunto maior de UFs lematizadas para poder comprová-la.

Por sua vez, o dicionário Michaelis (2006) oferece uma pequena parte chamada “expressões”, onde se lê: “expressões usuais são apresentadas em ordem alfabética e destacadas em negrito” (Michaelis, 2006:IX). Utiliza-se a fonte em negrito para indicar que se trata de uma expressão. Da mesma forma que o dicionário anterior, o Michaelis (2006) tampouco esclarece os critérios de seleção e de inserção

das UFs em seus verbetes. A seguir (Figura 5), apresentamos a entrada de *caber*, na qual aparece a UF *no caber en sí* e seu equivalente em português.

Figura 5: Verbetes *caber*.



ca.ber [kab'er] vi **1** caber. **2** pertencer. **3** corresponder. vt **4** conter, ter capacidade. **no caber uno en sí** não caber em si, estar orgulhoso.

Fonte: dicionário Michaelis (2006).

Novamente, podemos observar que a expressão é inserida na entrada do verbo *caber* e vai acompanhada de seu equivalente e sentido em português. A expressão não foi encontrada na entrada de *sí*, o que faz pensar que o critério de inserção da unidade é a forma verbal, que poderia ser considerada seu núcleo. Do mesmo modo que afirmamos para o dicionário Santillana (2009), para comprovar essa informação, posto que não está explicitada no *front matter*, é preciso uma análise mais aprofundada de um conjunto maior de verbetes, o que poderá ser feito em estudos futuros.

Pelas informações encontradas no *front matter*, podemos afirmar que ambos os dicionários não explicitam para o consulente os critérios de seleção das UFs lematizadas nem a forma de sua inserção na microestrutura. Também não indicam as informações que serão dadas sobre as UFs. Apesar dessas constatações, ambos os dicionários têm em comum os seguintes aspectos: incluem UFs na sua microestrutura, destacam as unidades com alguma marca tipográfica – símbolo ou negrito – e inserem seu equivalente em português. O que os difere é que o Michaelis (2006) não inclui exemplos para as UFs nem em espanhol nem em português. Além disso, podem inserir uma mesma UF em entradas diferentes. Por exemplo, a UF *caja de ahorros*, encontra-se na entrada do substantivo *ahorro*, no dicionário Michaelis (2006), mas foi inserida na entrada *caja*, no dicionário Santillana (2009).

3.2 ANÁLISE QUANTITATIVA E QUALITATIVA DAS UF NA MICROESTRUTURA DOS DICIONÁRIOS

No dicionário Santillana (2009), a inserção de UFs foi identificada em 41 dos 380 verbetes analisados, identificando-se um total 73 UFs, pois uma mesma entrada pode conter mais de uma UF. Levando em conta o cálculo proporcional, os resultados sugerem que o dicionário oferece 19% de UFs em relação ao seu número total de

entradas. No dicionário Michaelis (2006), 35 entradas do total das 280 analisadas lematizavam UFs, obtendo-se um total de 58 UFs. Calculando proporcionalmente, os resultados sugerem uma presença de 20% de UFs no dicionário em relação ao número total de entradas. A síntese desses dados é apresentada na Tabela 1, a seguir.

Tabela 1: Total de UFs identificadas em cada obra no intervalo lematizado.

	Entradas analisadas	Total de entradas contendo UFs	Total de UFs identificadas	% em relação às entradas analisadas
Santillana (2009)	380	41	73	19%
Michaelis (2006)	280	35	58	20%

Fonte: As autoras.

No que se refere à análise qualitativa, buscaram-se identificar as seguintes informações: entrada em que aparece a UF, UF lematizada, UF em português e significado da UF em português. O Quadro 1 apresenta uma amostra dos resultados coletados para o dicionário Santillana (2009), e o Quadro 2, para o dicionário Michaelis (2006).

Quadro 1: Informações oferecidas para as UFs – Dicionário Santillana (2009).

Entrada	UF	UF em português	Significado em português da UF
Cada	<i>Cada loco con su tema y cada lobo con su senda</i>	Cada macaco no seu galho	X
Calabaza	<i>Dar calabaza</i>	X	Reprovar em um exame
Caliente	<i>En caliente</i>	No calor dos fatos	No momento, sem deixar para depois.
Calle	<i>Echar a la calle</i>	X	Expulsar alguém do trabalho, de casa, etc. Demitir.
Calma	<i>Después de la tormenta siempre viene la calma.</i>	Depois da tormenta, sempre vem a bonança.	X
Campamento	<i>Levantar el campamento</i>	Levantar acampamento	Desmontar as instalações do acampamento a fim de partir para outro lugar.
Cántaros	<i>Llover cántaros</i>	Chover cântaros	X
Cara	<i>Cara de pascua</i>	X	Risonha e tranquila
Carne	<i>Ser de carne y hueso</i>	Ser de carne e osso	Ser sensível, humano. Ter sentimentos.

Fonte: As autoras.

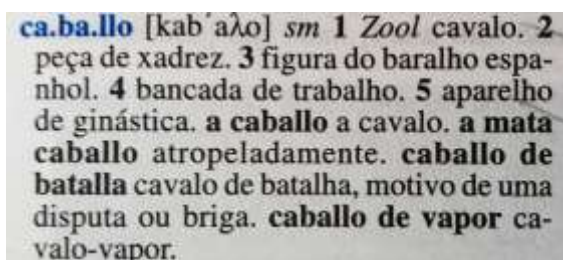
Quadro 2: Informações oferecidas para as UFs – Dicionário Michaelis (2006).

Entrada	UF	UF em português	Significado em português da UF
Caballo	<i>Caballo de batalla</i>	Cavalo de batalha	Motivo de uma disputa ou briga
Cabello	<i>Ponérsele a uno los cabellos en punta</i>	Deixar alguém de cabelo em pé	X
Caber	<i>No caber uno en sí</i>	Não caber em si	Estar orgulhoso
Cable	<i>Echar un cable</i>	X	Ajudar quem está em apuros
Cabra	<i>Cargar las cabras a uno</i>	X	Culpar um inocente
Capote	<i>Decir algo para su capote</i>	Falar com seus botões	X

Fonte: As autoras.

A partir dos dados anteriores, foram identificados três padrões – A, B e C – com relação às informações oferecidas para as UFs na microestrutura. No padrão A, os dicionários oferecem o equivalente da expressão na língua portuguesa, seguida de sua definição. A Figura 6 exemplifica o padrão A no dicionário Michaelis (2006).

Figura 6: Padrão A no verbete *caballo*.

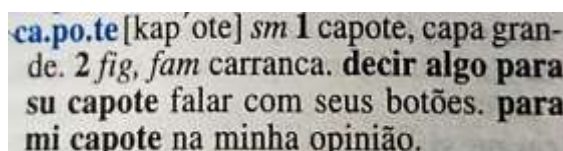


ca.ba.llo [kab'alo] *sm* 1 *Zool* cavalo. 2 peça de xadrez. 3 figura do baralho espanhol. 4 bancada de trabalho. 5 aparelho de ginástica. **a caballo** a cavalo. **a mata caballo** atropeladamente. **caballo de batalla** cavalo de batalha, motivo de uma disputa ou briga. **caballo de vapor** cavalo-vapor.

Fonte: Michaelis (2006).

Na entrada *caballo*, lematiza-se a expressão *caballo de batalla*, seguida de seu equivalente no português: 'cavalo de batalha'. Em seguida, também é oferecido o significado da expressão: "motivo de uma disputa ou briga". Dessa forma, no padrão A, os dicionários lematizam a UF em espanhol, apresentam seu equivalente e sua definição em português.

Figura 7: Padrão B no verbete *capote*.



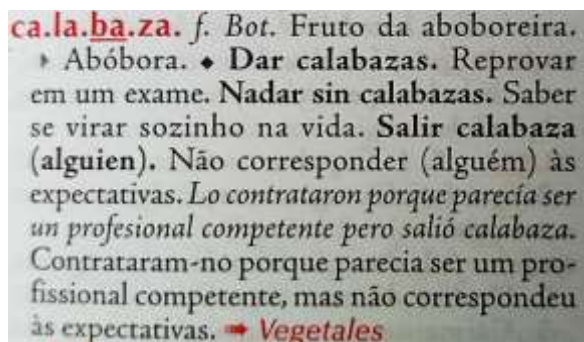
ca.po.te [kap'ote] *sm* 1 capote, capa grande. 2 *fig, fam* carranca. **decir algo para su capote** falar com seus botões. **para mi capote** na minha opinião.

Fonte: Michaelis (2006).

No padrão B (Figura 7), os dicionários lematizam a UF em espanhol e somente a UF equivalente no português, sem oferecer seu significado. Por exemplo, para

decirle algo para su capote, indica-se a expressão ‘falar com seus botões’, mas não apresenta seu significado como no padrão anterior. Parece ser que o dicionário parte do princípio de que o aprendiz sabe o significado de certas expressões na língua portuguesa, mas de outras não.

Figura 8: Padrão C no verbete *calabaza*.



Fonte: Santillana (2009).

Por fim, o padrão C (Figura 8) oferece somente o significado em português, indicando provavelmente a ausência da UF nessa língua. Por exemplo, o dicionário Santillana (2009) lematiza, na entrada *calabaza*, a UF *dar calabaza*, seguida apenas do seu significado: ‘reprovar em uma prova ou em um exame’. O Quadro 3 apresenta um resumo dos padrões de resultados obtidos.

Quadro 3: Padrões de resultados.

	O dicionário oferece:
Padrão A	✓ UF equivalente no português ✓ Significado
Padrão B	✓ UF equivalente no português × Significado
Padrão C	× UF equivalente no português ✓ Significado.

Fonte: As autoras.

A partir dos padrões anteriores, foi possível identificar a seguinte distribuição das UFs analisadas para ambos os dicionários (Tabela 2).

Tabela 2: Total de UFs por padrão em ambos os dicionários.

Padrão	Número de UFs em Santillana (2009)	Número de UFs em Michaelis (2006)
A	8	7
B	32	22
C	33	29
Total de UFs	73	58

Fonte: As autoras.

Uma vez que o dicionário é voltado para aprendizes brasileiros de espanhol que, dependendo do nível de aprendizagem em que se encontrem, podem não conhecer o significado de uma expressão na sua própria língua materna (português), podemos considerar o ‘padrão A’ como sendo o ideal, uma vez que oferece um conjunto maior de informações para o usuário. Em seguida, estaria o ‘padrão C’, pois o consulente, ao utilizar a direção Es/Pt do dicionário, aciona a função de decodificação da LE e, ao ter acesso ao significado da UF em espanhol, pode compreendê-la mais facilmente.. Finalmente, ‘padrão B’, oferece somente a expressão equivalente no português, sem o seu significado. Contudo, a compreensão da UF consultada em espanhol dependerá do conhecimento por parte do aprendiz da UF em sua língua materna¹².

3.3 ANÁLISE QUANTITATIVA E QUALITATIVA DAS UFs COM BASE NO CORPES XXI

A parte final de nossa análise consiste em observar a representatividade no espanhol atual das UFs identificadas nas duas obras. Conforme já explicitado, Biderman (1984) defende que a confecção de um dicionário e a escolha de entradas devem ser baseadas em um *corpus* representativo e indica como corte de frequência para a seleção das entradas uma frequência igual ou superior a 10. A partir dessa proposta e do princípio de que os dicionários bilíngues possuem uma função pedagógica, definimos aplicar para o presente trabalho, esse mesmo corte de frequência como um possível critério de seleção das UF a serem incluídas em um dicionário. Portanto, é esse corte de frequência que guiará a análise realizada a partir dos dados estatísticos do CORPES XXI. Por questões de espaço e pertinência, escolhemos 10 UFs do intervalo lematizado analisado de cada um dos dicionários.

Para demonstrar como o levantamento de dados estatísticos foi realizado no CORPES XXI, retomamos a Figura 4, apresentada na seção 3.1 deste trabalho, na qual podemos ver a busca pela expressão *no caer en sí*. A partir do filtro *estadística*, obtivemos para essa UF uma frequência absoluta de 34 ocorrências. Considerando o critério de corte de frequência estabelecido acima, essa UF mereceria estar registrada

¹² Dessa forma, o consulente aprendiz deve saber previamente o significado de, por exemplo, *cavalovapor, estar com a corda no pescoço, quem cala consente, ficar a zero, tirar a máscara*, etc. Se não souber, o dicionário gera no consulente a necessidade de outra busca a fim de obter o sentido dessas expressões no português.

em um dicionário bilíngue, indicando que se trata de uma unidade de uso frequente na língua e, portanto, pertinente para a aprendizagem.

Além de sua frequência de ocorrência absoluta, buscamos analisar a frequência em relação à distribuição geográfica (diatópica) de uso das UFs. Esse dado é importante de ser observado, pois os dicionários aqui estudados têm a pretensão de oferecer um vocabulário da língua espanhola como um todo e não de regiões geográficas específicas. Assim, o índice de distribuição da frequência por zonas pode auxiliar a identificar as UFs representativas da língua espanhola de um modo geral. No caso de *no caber en sí*, observamos que sua distribuição diatópica é ampla, uma vez que aparece registrada em sete das nove zonas linguísticas da língua espanhola: 12 ocorrências para México e América Central, 8 para Espanha, 5 para a zona Chilena, 5 para o Rio da Prata, 2 para a região Andina, 1 para a Guiné Equatorial, e 1 para as Antilhas. Esse procedimento foi aplicado para cada uma das 10 UFs analisadas para cada dicionário. Nas Tabelas 3 e 4, a seguir, apresentamos os índices de frequência absoluta no CORPES XXI e o número de zonas linguísticas em que as unidades são utilizadas.

Tabela 3: UFs e seus índices coletados utilizando-se o filtro estatístico do CORPES XXI.

Santillana (2009)		
UF	Frequência Absoluta	Zonas registradas
<i>Estar como una cabra</i>	61	9
<i>La cabra siempre tira al monte</i>	7	2
<i>Cada loco con su tema y cada lobo con su senda</i>	0	0
<i>Cada oveja con su pareja</i>	8	3
<i>Nadar sin calabaza</i>	0	0
<i>Hacer el caldo gordo</i>	10	4
<i>En calzas prietas</i>	0	0
<i>Cambiar de chaqueta</i>	2	1
<i>Levantar el campamento</i>	4	2
<i>Dar caña</i>	6	1

Tabela 4: UFs e seus índices coletados utilizando-se o filtro estatístico do CORPES XXI.

Michaelis (2006)		
UF	Frequência Absoluta	Zonas registradas
<i>No estar en sus cabales</i>	131	8
<i>Estar pendiente de un cabello</i>	0	0
<i>No caber en si</i>	34	7
<i>Alzar cabeza</i>	0	0
<i>Echar un cable</i>	8	1
<i>De cabo a rabo</i>	154	9
<i>Cargar las cabras a uno</i>	0	0
<i>Hasta las cachas</i>	14	5
<i>Caiga quien caiga</i>	80	8
<i>Dar una de cal y otra de arena</i>	35	7

Podemos constatar que há UFs com alto índice de frequência absoluta em diferentes zonas geográficas, como, por exemplo, *estar como una cabra* (SANTILLANA, 2009), com uma frequência absoluta de 61 e também apresentando ocorrências nas nove zonas linguísticas. O mesmo ocorre com *de cabo a rabo* (MICHAELIS, 2006), com uma frequência absoluta de 154 ocorrências e presença nas 9 zonas linguísticas da língua espanhola. No entanto, chama a atenção o registro nulo de algumas UFs, como, por exemplo, as expressões *nadar sin calabaza* (SANTILLANA, 2009) e *cargar las cabras* (MICHAELIS, 2006).

A fim de sistematizar os resultados apresentados nas Tabelas 3 e 4 e de analisá-los sob o índice de frequência por nós estabelecidos, apresentamos as Tabelas 5 e 6 (Santillana e Michaelis, respectivamente) com os resultados de frequência absoluta, os quais revelam quantas das UFs pesquisadas no *corpus* possuem frequência igual a zero, frequência menor que dez, ou frequência igual ou maior que dez.

Tabela 5: Frequência absoluta das UFs a partir do corte de frequência maior ou igual a 10.

Resultados Santillana (2009)		
UFs com frequência = 0	UFs com frequência < 10	UFs com frequência ≥ 10
3	5	2

Pelos resultados acima, é possível constatar que três a cada dez (30%) das UFs analisadas possuem frequência nula, e cinco a cada dez (50%) apresentaram frequência menor que 10. Esses números demonstram que 80% das UFs analisadas não deveriam estar lematizadas, segundo o corte de frequência adotado, no dicionário Santillana (2009). Desse modo, somente 20% das UFs analisadas apresentaram

frequência maior que 10, além de ocorrerem em diferentes zonas linguísticas, o que indicaria que são unidades que poderiam ser lematizadas.

Tabela 6: Frequência absoluta das UFs a partir do corte de frequência maior ou igual a 10.

Resultados Michaelis (2006)		
UFs com frequência = 0	UFs com frequência < 10	UFs com frequência ≥ 10
3	1	6

Os resultados da Tabela 6 demonstram, proporcionalmente, que, no dicionário Michaelis (2006), três a cada dez UFs presentes (30%) podem ter frequência nula, e uma a cada dez UFs (10%) podem ter frequência menor que 10. Dessa forma, na amostra analisada, 40% das UF não deveriam estar presentes no dicionário, conforme os requisitos de Biderman (1984). Já 60% das UFs restantes, além de apresentarem um índice de frequência maior que 10, também apresentam uma variedade na distribuição geográfica, sendo representativas do espanhol atual e candidatas a constarem em dicionários bilíngues pelos padrões aqui estabelecidos.

Por fim, é interessante observar que nos casos de frequência menor que 10, a distribuição geográfica também é díspar, sugerindo que além de serem expressões pouco usadas, possuem uma baixa representatividade na língua como um todo. Por exemplo, a expressão *echar un cable* (Tabela 4) possui frequência 8, isto é, próxima a 10. Contudo, os oito registros são oriundos somente de uma região, a Espanha. Por outro lado, a expressão *hacer el caldo gordo* (Tabela 3) tem frequência 10. Entretanto, esses dez registros são oriundos de quatro zonas diferentes, apontando uma maior representatividade diatópica da UF. Por isso, consideramos que o critério diatópico possui também relevância para este tipo de estudo estatístico, uma vez que esse fato, aliado ao critério de frequência, pode ser um indicativo de uma UF representativa da língua atual.

CONSIDERAÇÕES FINAIS

A breve análise feita neste trabalho permite concluir que, apesar de haver uma carência de materiais lexicográficos específicos de fraseologia voltados para aprendizes brasileiros de língua espanhola, os dicionários bilíngues oferecem unidades fraseológicas, o que, em certa medida, representa uma oferta de informações fraseológicas aos aprendizes de espanhol. Contudo, o modo como as diferentes UFs são abordadas nos dicionários bilíngues aqui analisados parece ser

problemática. Conforme identificamos, no *front matter* das obras analisadas, embora esteja claro que as UFs constam na microestrutura com marcas tipográficas que as destacam, não estão explicitados os critérios de seleção e inclusão de tais unidades. Portanto, a informação dada aos seus usuários não é suficiente, principalmente as relativas à forma de busca das UFs. Nesse sentido, vimos que não há uniformidade em relação à entrada em que aparecem lematizadas; em um dos dicionários pode parecer em uma determinada palavra da UF e, no outro, em outra, como vimos no exemplo da UF *caja de ahorros*.

No que se refere à análise quantitativa das UFs, identificamos 73 UFs de 380 entradas analisadas no dicionário Santillana (2009) e 58 UFs em 280 entradas analisadas no dicionário Michaelis (2006); ambos apresentando índices percentuais semelhantes da presença de UFs (em relação ao número de entradas analisadas): 19% e 20% respectivamente.

Por sua vez, a análise qualitativa da microestrutura, permitiu identificar três padrões, considerando o tipo de informações oferecidas a respeito da UF lematizada. No padrão A, os dicionários oferecem o equivalente da expressão na língua portuguesa, seguida de sua definição. No padrão B, os dicionários incluem apenas a expressão equivalente no português, sem seu significado. Finalmente, no padrão C indica somente o significado da UF, sem oferecer seu equivalente em português, sugerindo, portanto, que o português não apresenta uma expressão equivalente. Verificamos que os padrões mais frequentes foram os padrões C e B e, como o menos frequente, o padrão A. Contudo, levando em conta os usuários e a função dos dicionários bilíngues, acreditamos que os padrões A e C são os mais efetivos para o consulente, dado que ambos oferecem o significado da UF na língua espanhola.

Quanto à representatividade numérica e geográfica das UFs na língua atual, os dados obtidos por meio do CORPES XXI para a amostra de UFs analisadas revelam que há divergências em relação ao uso das UFs no *corpus* e as registradas nos dicionários. Foi possível constatar que muitas das UFs inseridas nos dois dicionários – mais especificamente 80% das UFs do Santillana (2009) e 40% do dicionário Michaelis (2006) – possuem índice de frequência baixo ou nulo, além de terem uma pequena distribuição diatópica. Contudo, vale ressaltar que para sustentar melhor esta conclusão, seria ideal fazer uma análise com uma amostragem maior de verbetes de cada dicionário.

Embora a análise tenha sido feita para uma amostra pequena de UFs, os resultados sugerem a necessidade de uma maior atenção e um maior cuidado no tratamento da fraseologia em dicionários bilíngues, considerando a importância das combinações fixas no processo de aquisição de uma LE e a função pedagógica desse tipo de obra. Nesse sentido, pensamos que os dicionários bilíngues para aprendizes brasileiros de espanhol deveriam indicar de forma mais explícita e sistemática tanto os critérios de seleção como os de inclusão e apresentação das UFs. No que se refere aos critérios de inclusão, reforçamos a ideia de Biderman (1989) ao defender que os dicionários sejam elaborados com base em *corpus* e suas ferramentas, como, por exemplo, os índices de frequência e de distribuição diatópica. Além disso, os resultados também apontam para uma necessidade de estudos mais aprofundados que proponham soluções para os problemas aqui identificados.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BIDERMAN, Maria Tereza Camargo. O dicionário padrão da língua. *Alfa*, 28, p.27-43, 1984.
2. BRASIL. Ministério da Educação, Secretaria da Educação Básica. Edital de Convocação para o processo de inscrição e avaliação de dicionários brasileiros de língua portuguesa para o Programa Nacional do Livro Didático – PNLD. Brasília, MEC, 2012.
3. BOSQUE, Ignacio. Diccionario REDES: *Diccionario Combinatorio del Español Contemporáneo*. Ediciones SM. 2004.
4. BUGUEÑO, Félix. Balanço e perspectivas da Lexicografia. *Cadernos de tradução*, Florianópolis, 32/2: 16-37, 2013.
5. CORPAS PASTOR, Gloria. *Manual de Fraseología Española*. Editora Gredos. 1996.
6. CORPAS PASTOR, Gloria. *Diez años de investigaciones en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. *Linguística Iberoamericana*, nº 20, 2004.
7. CORPES XXI. *Corpus del Español del Siglo XXI*. Real Academia Española. Disponível em: <http://web.frl.es/CORPES/view/inicioExterno.view>

8. PEREIRA, Helena Bonito Couto. *Dicionário Escolar Espanhol & Português Michaelis*. São Paulo: Editora Melhoramentos. 2006.
9. DIAZ Y GARCÍA-TALAVERA, Miguel; *Dicionário Santillana para estudantes*. São Paulo: Editora Moderna. 2009.
10. DURAN, Magali Sanches. *Dicionários bilíngues pedagógicos: análise, reflexões e propostas*. São José do Rio Preto: Universidade Estadual Paulista. 2004.
11. DURAN, Magali Sanches; XATARA, Claudia Maria. Lexicografia pedagógica: atores e interfaces. *D.E.L.T.A.*, 23:2, 2007.
12. FIORAVANTI, Sueli. *As unidades heterogênicas em dicionários bilíngues de espanhol para aprendizes brasileiros: análise do tratamento lexicográfico*. Araraquara: Universidade Estadual Paulista. 2015
13. HAUSMANN, Franz Josef. Le dictionnaire de collocations. HAUSMANN, F.J. [et al.] Na *International encyclopedia of lexicography*. Vol. 1. Berlin, New York: Walter de Gruyter, 1990, p. 1010-1019.
14. KUBARTH, Hugo; VARELA, Fernando. *Diccionario Fraseológico del Español Moderno*. Editorial Gredos. 2004.
15. MATTOS, Monissa. *Proposta de macro e microestrutura para um dicionário especial de locuções verbais – português/espanhol*. Porto Alegre: Universidade Federal do Rio Grande do Sul. 2010.
16. NOIMANN, Aline. *Um olhar sobre os fraseologismos (locuções) em um dicionário bilíngue escolar espanhol-português/português-espanhol*. Porto Alegre: Universidade Federal do Rio Grande do Sul. 2007
17. RAMOS, Margarita Alonso. *DiCE: Diccionario de Colocaciones del Español*. Disponível em: <http://dicesp.com/paginas>
18. TARP, Sven. Desafios teóricos y prácticos de la Lexicografía de aprendizaje. IN: HUMBLÉ, Philippe; XATARA, Claudia. BEVILACQUA, CleciRegina (org.). *Lexicografia pedagógica: pesquisas e perspectivas*. Florianópolis: UFSC/NUT, 2008.
19. TAGNIN, Stella. *Glossário de Linguística de Corpus*. Editora HUB. 2013. Disponível em: http://www.hubeditorial.com.br/site/recursos/5_glossario/glossario_423.pdf
20. WELKER, Herbert. Lexicografia Pedagógica: definição, história, peculiaridades. IN: HUMBLÉ, Philippe; XATARA, Claudia. BEVILACQUA,

Cleci Regina (org.). *Lexicografia pedagógica: pesquisas e perspectivas*. Florianópolis: UFSC/NUT, 2008.

Recebido em: 04/06/2017

Aceito em: 21/07/2017

Apêndice

MICHAELIS (2006)			
Entrada	UF	UF em português	Significado em português da UF
1. Cabal	<i>No estar en sus cabales</i>	X	Estar louco, perturbado
2. Caballero	<i>Caballero andante</i>	X	Aventureiro
3. Caballete	<i>Nariz de caballete</i>	X	Nariz adunco
4. Caballo	<i>A mata caballo</i>	X	Atropeladamente
5. Caballo	<i>Caballo de batalla</i>	Cavalo de batalha	Motivo de uma disputa ou briga
6. Caballo	<i>Caballo de vapor</i>	Cavalo-vapor	X
7. Cabecera	<i>Médico de cabecera</i>	Médico de família	X
8. Cabello	<i>Estar pendiente de un cabello</i>	Estar com a corda no pescoço	X
9. Cabello	<i>Ponérsele a uno los cabellos en punta</i>	Deixar alguém de cabelo em pé	X
10. Cabello	<i>Por un cabello</i>	Por um triz	X
11. Caber	<i>No caber uno en si</i>	Não caber em si	Estar orgulhoso
12. Cabeza	<i>Bajar uno la cabeza</i>	X	Obedecer
13. Cabeza	<i>Levantar o Alzar Cabeza</i>	Dar a volta por cima	X
14. Cabeza	<i>No tener (algo) ni pies ni cabeza</i>	Não ter nem pé nem cabeça	X
15. Cabeza	<i>Perder la cabeza</i>	Perder a cabeça	X
16. Cable	<i>Cable submarino</i>	Cabo submarino	X
17. Cable	<i>Echar un cable</i>	X	Ajudar quem está em apuros
18. Cabo	<i>Al cabo de</i>	Ao fim de/ depois de	X
19. Cabo	<i>Atar cabos</i>	X	Reunir dados
20. Cabo	<i>De cabo a rabo</i>	Do princípio ao fim	X
21. Cabra	<i>Cargar las cabras a uno</i>	X	Culpar um inocente
22. Cacha	<i>Hasta las cachas</i>	Até as tampas	Farto
23. Cachimbo	<i>Chupar cachimbo</i>	X	Fumar cachimbo
24. Cachondo	<i>Estar cachondo</i>	Meio tarado	Estar dominado pela libido
25. Cadena	<i>Romper las cadenas</i>	X	Libertar
26. Cadena	<i>Trabajo en cadena</i>	Trabalho em série	Mecanização
27. Caer	<i>Caer enfermo</i>	X	Ficar doente
28. Caer	<i>Caer en gracia</i>	X	Conquistar simpatias (de alguém)

29. Caer	<i>Caer en la cuenta</i>	Dar-se conta	Compreender
30. Caer	<i>Caiga quien caiga</i>	A pesar de tudo	X
31. Caja	<i>En caja</i>	Em forma	X
32. Cal	<i>Dar una de cal y otra de arena</i>	X	Alternar coisas extremas
33. Calabazas	<i>Dar calabazas</i>	X	Ser reprovado (na escola)
34. Calentar	<i>Calentar las orejas</i>	X	Repreender severamente
35. Callar	<i>Quien calla otorga</i>	Quem cala consente	X
36. Calle	<i>Estar a la calle</i>	X	Despedir
37. Calle	<i>Irse a la calle</i>	X	Sair
38. Calle	<i>Quedarse en la calle</i>	Ficar a zero	Ficar sem dinheiro
39. Calma	<i>Tomarlo con calma</i>	X	Ter paciência
40. Cantar	<i>Cantar victoria</i>	X	Exaltar, louvar
41. Cantar	<i>Esse es otro cantar</i>	Isso é outra conversa	X
42. Canto	<i>De canto</i>	De lado	X
43. Capa	<i>Capa aguadeira</i>	X	Capa impermeável
44. Capote	<i>Decir algo para su capote</i>	Falarcomseusbotões	X
45. Capote	<i>Para mi capote</i>	Na minha opinião	X
46. Cara	<i>Cara de pascua</i>	X	Cara muito alegre
47. Cara	<i>Cara dura</i>	Cara de pau	X
48. Cara	<i>Cara larga</i>	X	Cara triste
49. Cara	<i>Cara y cruz</i>	Cara e coroa	X
50. Cara	<i>Dar la cara</i>	X	Assumir uma situação
51. Carambola	<i>Por carambola</i>	X	Indiretamente
52. Careta	<i>Quitar la careta</i>	Tirar a máscara	X
53. Carga	<i>Volver a la carga</i>	X	Insistir
54. Cargar	<i>Cargar demasiado</i>	X	Beber demais
55. Cargar	<i>Cargar en cuenta</i>	X	Debitar
56. Cargo	<i>Hacerse cargo</i>	Tomar conta	X
57. Carne	<i>Echar carnes</i>	X	Engordar
58. Carrerrilla	<i>De carrerrilla</i>	X	De memória e sem vacilação

SANTILLANA (2009)			
Entrada	UF	UF em português	Significado em português da UF
1. Cabal	<i>No estar en sus cabales</i>	Estar fora de si	Estar sem juízo
2. Caber	<i>No cabe duda</i>	Sem dúvida alguma	X
3. Cabeza	<i>Cabeza abajo</i>	De ponta cabeça	X
4. Cabida	<i>Tener cabida</i>	X	Ter influência ou prestígio
5. Cable	<i>Cruzársele (a alguien) los cables</i>	X	Perder a paciência

6. Cabo	<i>Al cabo de</i>	Ao fim/cabo de	X
7. Cabra	<i>Estar como una cabra</i>	X	Estar louco
8. Cabra	<i>La cabra siempre tira al monte</i>	X	Cada um se manifesta e se comporta de acordo a sua origem e educação
9. Cachondo	<i>Estar cachondo</i>	X	Sentir desejo sexual
10. Cada	<i>Cada loco con su tema y cada lobo con su senda</i>	Cada macaco no seu galho	X
11. Cada	<i>Cada oveja con su pareja</i>	Cada macaco no seu galho	X
12. Cadena	<i>Cadena de sonido</i>	Aparelho de som	Eletrônico com funções de rádio, toca-fitas, toca-cd, entre outras.
13. Cadena	<i>Cadena perpetua</i>	Prisão perpétua	X
14. Caer	<i>Caerle bien/mal (a alguien) una persona</i>	X	Desagradar/agradar alguém a uma pessoa.
15. Caer	<i>Caer enfermo</i>	X	Ficar doente
16. Caer	<i>Caerse la cara de vergüenza</i>	Estar a ponto de acontecer	X
17. Café	<i>Café solo</i>	Café preto	X
18. Café	<i>Café cortado</i>	X	Com um pingo de leite
19. Café	<i>Café ahogado</i>	X	Café americano, com bastante água.
20. Caja	<i>Cajá de ahorros</i>	Caixa econômica	X
21. Cajero	<i>Cajero automático</i>	Caixa eletrônico	X
22. Caída	<i>Caída de cabellos</i>	Queda de cabelos	X
23. Caída	<i>Caída de tensión</i>	X	Queda na corrente elétrica;
24. Cajón	<i>Ser de cajón</i>	X	Ser evidente
25. Calabaza	<i>Dar calabaza</i>	X	Reprovar em um exame
26. Calabaza	<i>Nadar sin calabaza</i>	X	Saber se virar sozinho na vida
27. Calabaza	<i>Salir calabaza</i>	X	Não corresponder às expectativas
28. Caldo	<i>Hacer el caldo gordo</i>	X	Ganhar bem em um negócio.
29. Caldo	<i>Haz de esse caldo tajadas</i>	X	Expressa pedido de algo impossível de ser feito.
30. Caliente	<i>En caliente</i>	No calor dos fatos	No momento, sem deixar para depois.
31. Calar	<i>El que calla otorga</i>	Quem cala consente	X
32. Calle	<i>Doblar la calle</i>	X	Virar a esquina
33. Calle	<i>Calle peatonal</i>	X	Calçada
34. Calle	<i>Echar a la calle</i>	X	Expulsar alguém do trabalho, de casa, etc. Demitir.
35. Calma	<i>Después de la tormenta siempre viene la calma</i>	Depois da tormenta sempre vem a bonança	X
36. Calza	<i>En calzas prietas</i>	No aperto	X
37. Cambiar	<i>Cambiar de chaqueta</i>	X	Mudar de opinião
38. Camisa	<i>Camisa de fuerza</i>	Camisa de força	X
39. Camisa	<i>Camisa de once varas</i>	X	Situação complicada

40. Camisa	<i>Cambiar de camisa</i>	X	Mudar de opinião
41. Camisa	<i>Dejarle sin camisa</i>	X	Arruinar alguém totalmente.
42. Camisa	<i>Jugarse hasta la camisa</i>	X	Apostar tudo
43. Campamento	<i>Levantar el campamento</i>	Levantar acampamento	Desmontar as instalações do acampamento a fim de partir para outro lugar.
44. Campo	<i>Campo de concentración</i>	Campo de concentração	X
45. Campo	<i>Salir a campo</i>	Ir à luta	X
46. Canelo	<i>Hacer el canelo</i>	X	Deixar-se enganar facilmente
47. Canje	<i>Canje de prisioneros</i>	Troca de prisioneiros	X
48. Cántaros	<i>Llover cántaros</i>	Chover cântaros	X
49. Caña	<i>Caña de azúcar</i>	Cana de açúcar	X
50. Caña	<i>Caña de pescar</i>	Vara de pescar	X
51. Caña	<i>Dar caña</i>	Dar uma surra	Bater
52. Capote	<i>Decir para su capote</i>	Falar com seus botões	X
53. Cara	<i>Cara de acelga</i>	X	Pálido
54. Cara	<i>Cara de pascua</i>	X	Risonha e tranquila
55. Cara	<i>Cara de pocos amigos</i>	Cara amarrada	X
56. Cara	<i>Echar a cara o cruz</i>	Tirar no cara ou coroa	X
57. Cara	<i>Caérsele la cara de vergüenza</i>	Morrer de vergonha	X
58. Cara	<i>Echar en cara</i>	Jogar na cara	X
59. Cara	<i>Partirle/romperle la cara</i>	Quebrar a cara	X
60. Cara	<i>Poner buena/mala cara</i>	Fazer cara boa/feia	X
61. Carne	<i>Cobrar carne</i>	X	Engordar
62. Carne	<i>Ser de carne y hueso</i>	Ser de carne e osso	Ser sensível, humano. Ter sentimentos.
63. Carrera	<i>Hacer carga</i>	Fazer carreira	Prosperar profissionalmente
64. Carrera	<i>Dar carrera</i>	X	Pagar os estudos
65. Carrera	<i>Partir carrera</i>	Colocar o carro na frente dos bois	X
66. Carrete	<i>Dar carrete (a alguien)</i>	X	Enrolar alguém
67. Cartilla	<i>No estar en su cartilla</i>	Fora do comum	X
68. Cartilla	<i>No saber la cartilla</i>	X	Ignorar os rudimentos de alguma coisa
69. Casa	<i>Caérsele (a alguien) la casa a cuestras/ encima</i>	Desabar o céu sobre a cabeça	X
70. Casa	<i>Echar/tirar la casa por la ventana</i>	Jogar dinheiro fora	Gastar mais do que se pode.
71. Casado	<i>El casado casa quiere</i>	Quem casa quer casa	X
72. Casco	<i>Calentarle (a alguien) los cascos</i>	X	Inquietar por preocupação
73. Casco	<i>Metérsele en los cascos</i>	X	Encasquetar