As bases de dados verbais ADESSE e ViPEr: uma análise constrastiva das construções locativas em espanhol e em português

Roana Rodrigues¹, Oto Vale¹, Laura Alonso Alemany²

¹ Programa de Pós-Graduação em Linguística – Univ. Federal de São Carlos (UFSCar) Caixa Postal 676 – 13.565.905 - São Carlos - SP – Brasil

³ Facultad de Matemática, Astronomía y Física – Universidad Nacional de Córdoba, Córdoba, Argentina

{rroanarodriques,otovale,lauraalonsoalemany}@gmail.com

Abstract. In this work, we contrasted Spanish locative verbs with Portuguese locative verbs from ADESSE and ViPEr databases. Having the meanings in Spanish from the ADESSE database, we searched for their equivalents in Portuguese in the ViPEr database, using our introspective knowledge as researchers and other resources, such as monolingual and bilingual dictionaries. From the 693 ADESSE's space verbs, 352 are locative constructions in the ViPEr, presenting different pairs of correspondence. We connected this resource to EuroWordNet. As a result, we obtained a bilingual resource with a relatively big granularity and a wealth of syntactic-semantic description.

Resumo. Neste trabalho, contrastamos os verbos locativos do espanhol e do português das bases de dados ADESSE e ViPEr. Partindo dos sentidos em espanhol da base de dados ADESSE, procurou-se seus equivalentes em língua portuguesa na base de dados ViPEr, a partir dos conhecimentos introspectivos dos pesquisadores e de outros recursos, como dicionários monolíngues e bilíngues. Dos 693 verbos de espaço presentes no ADESSE, 352 são construções locativas no ViPEr, apresentando diferentes pares de correspondência. Conectamos esses recursos à EuroWordNet. Como resultado, obtivemos um recurso bilíngue com uma granularidade relativamente grande e uma riqueza de descrição sintático-semântica.

1. Introdução

Este trabalho é um estudo inicial de análise contrastiva das construções verbais locativas do espanhol e do português, com o intuito de estabelecer os aspectos comuns e divergentes entre as duas línguas. Para tanto, partimos dos verbos *espaciais* da base de dados verbais do espanhol ADESSE (García-Miguel, 2003, 2006) e os contrastamos aos verbos *locativos* do português da base de dados verbais ViPEr (Baptista, 2012). Desse modo, enriquecemos as duas bases de dados em questão e contribuimos na construção de recursos linguísticos que podem ser utilizados em diferentes aplicações didáticas e no Processamento de Língua Natural (PLN).

Para o desenvolvimento de uma concepção de construções locativas, nos pautamos, sobretudo, nos trabalhos de Guillet e Leclère (1992) para o francês, Baptista (2012) para o português e nas descrições sintático-semânticas do espanhol de García-Miguel et al. (2003). As construções verbais locativas aqui analisadas são as que estabelecem uma relação de localização entre os elementos constituintes da frase e, de um ponto de vista sintático, respondem adequadamente às perguntas com o advérbio interrogativo (*Prep*) onde, cujos complementos são selecionados pelo próprio verbo – e não meros complementos circunstanciais selecionados pela frase. Os exemplos de (1) a (5) ilustram algumas dessas construções que apresentam estruturas similares em língua portuguesa e em língua espanhola.

- (1) O Pedro mora em Buenos Aires. (Pedro vive en Buenos Aires)
- (2) O Pedro entrou na sala. (Pedro entró en la sala)
- (3) O Pedro descolou o papel da parede. (Pedro despegó el papel de la pared)
- (4) O Pedro atravessou a rua. (Pedro cruzó la calle)
- (5) O Pedro engavetou os documentos. (Pedro encajonó los documentos)

Em (1), o verbo *morar* seleciona um complemento locativo *estativo* de lugar (*Buenos Aires*). Já nas frases de (2) a (5), têm-se uma construção verbal locativa dinâmica. Em (2), o nome que ocupa a posição de sujeito da frase é o mesmo que se desloca para o complemento locativo de destino (na sala). Em (3), o verbo descolar seleciona o complemento direto (papel) e um complemento locativo de origem (da parede). Em (4) e (5) têm-se uma construção transitiva direta. Na frase (4), o complemento direto (rua) é interpretado como um lugar de destino. Em (5), por sua vez, o verbo engavetar/encajonar se constitui pelo nome cognato de lugar gaveta/cajón e o complemento direto (documentos) é o objeto que ocupa esse lugar.

Sabe-se que um grande número de construções verbais em língua portuguesa e em língua espanhola exprimem o conceito de localização. Trata-se de um fenômeno já descrito ou mencionado em trabalhos anteriores, como as propostas de classificações sintático-semânticas de Macedo (1987), Cançado et al. (2013) e Baptista, (2012), para o português; e Rojas Nieto (1988), Crego García (1995) e García-Miguel (2003, 2006), para o espanhol. No entanto, não se tem notícia de estudos contrastivos dessas construções que considerem o português e o espanhol. Desse modo, comparamos as construções verbais espaciais do espanhol, disponível na base de dados sintáticosemânticos ADESSE (García-Miguel, 2006) às construções verbais locativas do português, descritas na base de dados de construções léxico-sintáticas de verbos do português ViPEr (Baptista, 2012). Embora as duas bases de dados em questão partam de aspectos teórico-metodológicos diferenciados, as selecionamos pelo fácil acesso aos dados e, principalmente, pela quantidade relevante de verbos e pela qualidade de sua descrição e classificação. A classe espaço do ADESSE apresenta a classificação de 674 construções verbais. O ViPEr, uma base de dados mais granular, contém a classificação de 1.142 empregos locativos.

Ressalta-se que, em um primeiro momento, os verbos do ADESSE foram relacionados à WordNET, afim de identificar com maior profundidade os sentidos veiculados pelas construções espaciais ali descritas, além de estabelecer uma conexão entre as bases de dados, enriquecendo-as.

2. Classificações sintático-semânticas dos verbos locativos

O ADESSE (Base de datos de verbos, alternancias de diátesis y esquemas sintácticosemánticos del español) é uma base de dados de construções verbais do espanhol, em que é possível encontrar a classificação dessas construções de acordo com as suas propriedades sintático-semânticas. O projeto ADESSE, da Universidad de Vigo (Galícia, Espanha), se construiu sob a base de dados sintáticos do espanhol atual (BDS), que contém a análise sintática de um corpus do espanhol contemporâneo com cerca de 1,5 milhão de palavras. Na BDS está anotada a função sintática (sujeito, objeto direto, objeto indireto, etc.), categoria sintática (tipo da frase: verbal, nominal), preposição, animacidade e número. Partindo desses dados da BDS, acrescentou-se, ao ADESSE, informações semânticas pertinentes: sentidos verbais, classes semânticas de processos e papéis semânticos. A construção das classes verbais tem inspiração nas classificações sintático-semânticas de Levin e dialoga, em alguma medida, com a WordNet e a FrameNet, embora apresente níveis mais generalizados. De acordo com García-Miguel et al. (2003, p.14), são agrupados nas mesmas classes verbos que estabelecem uma relação de parassinonímia (lexemas com mesmo significado, mas diferente distribuição) e co-hiponímia (lexemas que possuem um hiperônimo comum). A classificação hierárquica de base conceitual dessa base de dados apresenta 6 macro-classes, 20 classes e 38 subclasses. É na macro-classe material que se encontra nosso objeto de estudo, os verbos de espaço, subdividido em: espaço, deslocamento, localização, postura-posição, orientação, maneira movimento e união, como se verifica na Tabela 1.

Tabela 1. Classificação dos verbos de espaço do ADESSE

Subclasses	Definição ¹	Verbo	Exemplo	#	
Espaço	Uma entidade possui uma determinada localização, configuração ou orientação espacial. Ou ainda realiza algum tipo de mudança em sua localização, configuração ou orientação espacial. Engloba construções não classificadas nas demais classes de espaço.	mover	Pedro se movió para la izquierda.	5	
Deslocamento	Uma entidade se desloca de uma localização inicial a uma localização final, percorrendo um trajeto. Pedro fue a escuela.		9	230	
Localização	Uma entidade apresenta una determinada localização no espaço (ou, por extensão, no tempo).	entrar	Pedro entró en la iglesia.	219	
Postura-Posição	Uma entidade apresenta uma determinada configuração espacial ou sofre um processo de modificação da mesma.	sentar	Pedro sentó a su hijo en la silla	42	
Orientação	Uma entidade se situa de tal modo que indica uma determinada direção.	volver	Pedro volvió a su casa.	10	
Maneira- Movimento	Uma entidade realiza um movimento que não supõe uma mudança de localização nem de configuração espacial.	agitar	Pedro agitó la botella.	43	
União	Uma entidade se encontra em um estado ou realiza um processo de união, reunião, integração ou separação com outra entidade.		Juan añadió sal a la comida.	125	
Total					

¹ As definições foram retiradas do ADESSE: http://adesse.uvigo.es/index.php>, acesso em agosto de 2017 (tradução nossa).

O ViPEr (Baptista, 2012) é uma base de dados dos verbos do português europeu², que tem como arcabouço teórico-metodológico o Léxico-Gramática (Gross, 1975, 1981) e se assenta na proposta de classificação sintático-semântica dos verbos locativos do francês de Guillet e Leclère (1992). Até o momento, o ViPEr contém a análise de 130 propriedades sintáticas, semânticas, estruturais, transformacionais e distribucionais de aproximadamente 7.000 construções verbais, categorizadas em 71 classes. As construções verbais locativas compreendem 1.142 verbos e estão distribuídas por 12 classes distintas, como se observa na Tabela 2.

Tabela 2. Classificação dos verbos locativos do ViPEr (Baptista, 2012)

Classe	Estrutura ³	Verbo	Exemplo	#
35LD	N _o V-din Loc ₁ Nloc ₁	entrar	O Pedro entrou na sala	193
35LS	N _o V-stat Loc ₁ Nloc ₁	viver	O Pedro vive em Lisboa	32
37LD	No Vdin Loc-s, Nloc, Loc-d, Nloc,	viajar	O Pedro viajou daqui para ali	117
38L1	$N_o \ V \ Nloc_i$	invadir	O Pedro invadiu a sala	205
38L2	N _o Nloc-v Nobj, [V=pôr em Nloc]	enjaular	O Pedro enjaulou o leão	39
38L3	$Nloc_{\scriptscriptstyle 0} \ V \ Nobj_{\scriptscriptstyle I}$	encerrar	A jaula encerrava a fera	12
38L4	N _o Nobj-v Nloc-d _i [V=pôr Nobj]	apimentar	O Pedro apimentou a comida	122
38L5	N _o Nobj-v Nloc-s _i [V=tirar Nobj]	desengordurar	O Pedro desengordurou o prato	11
38LD	No Vdin N1 Loc-d2 Nloc2	pousar	O Pedro pousou o livro na mesa	281
38LS	No Vdin N1 Loc-s2 Nloc2	retirar	O Pedro retirou o livro da mesa	77
38LT	No Vdin N. Loc-s: Nloc: Loc-d: Nloc:	transferir	O Pedro transferiu o livro daqui para ali	50
38R	N _o Vstat N ₁ Loc ₂ N ₂	situar	O Pedro situou o Butão no mapa.	3
			Total	1.142

Respeitando as abordagens teórico-metodológicas das bases de dados ADESSE e ViPEr, assim como suas motivações e objetivos, realizamos a análise comparativa dos verbos de espaço do espanhol com os verbos locativos do português.

3. Análise dos dados

Inicialmente, foram conectadas as 674 construções verbais de espaço do ADESSE a seus sentidos correspondentes na WordNet, uma base de dados lexicais, que, a partir de uma abordagem relacional, é organizada de acordo com o significado e conceito, e não com a forma. Trata-se de uma ontologia linguística, amplamente utilizada na área de PLN, na qual estão armazenados conceitos lexicalizados (nomes, verbos, adjetivos e advérbios) que, como apresentado por Di Felippo (2008, p. 47), "se organizam sob a

² Embora o trabalho de Baptista (2012) se refira às construções verbais do português europeu, segundo Rodrigues (2016), o comportamento sintático-semântico dos verbos locativos do português europeu assemelha-se muito às mesmas construções do português brasileiro.

³ Notações: N0, N1, N2, N3: sujeito e complementos; Prep: preposição; N: nome ou grupo nominal; Nloc: nome locativo (papel semântico); Nobj: "objeto" (papel semântico); Loc: preposição locativa, -d de destino, -s de origem; V: verbo, Vdin: verbo locativo dinâmico; Vstat: verbo locativo estativo.

forma de *synsets* (abreviação do termo em inglês *synonym set*, isto é, conjunto de unidades sinônimas)".

O estabelecimento da relação ADESSE/WordNET ocorreu através do Multilingual Central Repository⁴, ferramenta multilíngue (inglês, basco, espanhol, galego, catalão e português), baseada na WordNet de Princeton, associada ao projeto EuroWordNet. Ressalta-se que usamos a WordNet apenas como referência e não como fonte de informação, ou seja, não consideramos todas as entradas léxicas e distinções de sentidos da WordNet, já que são muito granulares. Dessa intersecção, 99 sentidos verbais não constam na WordNet. Essas lacunas decorrem principalmente por serem lexemas: (i) pouco usuais em língua espanhola (callejear, eslabonar); (ii) com prefixação (desorbitar, reintroducir); (iii) relacionados a partes do corpo ou postura corporal (gallear, manotear, acodarse); e (iv) relacionados a verbos que apresentam uma rigorosa restrição do nome que ocupa a posição de complemento (levar, adosar). Além disso, é importante ressaltar que as redes wordnets - como o Projeto EuroWordNet - se baseiam na WordNet de Princeton, que considera o inglês americano. Desse modo, acredita-se que muitos desses casos não descritos na WordNet para o espanhol se referem a verbos não lexicalizados ou não usuais em língua inglesa, daí a sua omissão.

Após familiarizar-se com os sentidos estabelecidos nas classes de *espaço* do ADESSE, iniciou-se a segunda etapa do trabalho: as associações entre as 674 construções verbais do ADESSE às construções do ViPEr, de maneira manual, com o auxílio de pesquisadores lusofalantes e hispanofalantes, além do uso da WordNet, de dicionários monolíngues e bilíngues e de corpora. Neste trabalho, a definição de *locativo* relaciona-se diretamente às noções de *locativo* da base de dados ViPEr (BAPTISTA, 2012). Por este motivo, consideramos construções locativas do ADESSE apenas as que apresentam correspondência com as construções locativas do ViPEr. Como mencionado, o ViPEr é uma base de dados mais granular, por isso um mesmo verbo, como por exemplo *subir*, é apresentado em diferentes entradas, devido às suas construções sintático-semânticas particulares:

(6) O Pedro subiu as malas do quinto ao oitavo andar.

(Pedro subió las valijas del quinto piso al octavo.)

(6a) O Pedro subiu do quinto ao oitavo andar.

(Pedro subió del quinto piso al octavo.)

(6b) O Pedro subiu as escadas.

(Pedro subió las escaleras.)

No ADESSE, o verbo *subir* apresenta a definição *mover de abajo arriba* e está classificado como uma construção espacial da classe Deslocamento. No ViPEr, no entanto, as construções são multiplicadas devido às diferentes construções sintáticas: em (6), tem-se uma construção com o verbo *subir* em que é selecionado um objeto na posição de complemento direto (*malas*) que se desloca de um ponto de origem (*do quinto*) a um ponto de destino (*ao oitavo andar*), classe 38LT do ViPEr; em (6a) é o nome que ocupa a posição de sujeito (*Pedro*) que se desloca, classe 37LD do ViPEr; já

⁴ Ferramenta de léxico multilíngue baseada na WordNet de Princeton – Multilingual Central Repository disponível em: http://adimen.si.ehu.es/cgi-bin/wei/public/wei.consult.perl, acesso em agosto de 2017.

em (6b), verifica-se uma construção transitiva direta, na qual o objeto na posição de complemento direto (*escadas*) é interpretado como um lugar, classe 38L1 do ViPEr.

Sendo assim, da comparação ADESSE/ViPEr, foram acrescentados 19 lexemas verbais locativos por terem suas construções sintáticas descritas no ViPEr, a saber: mover, desplazar, llevar, navegar, saltar, subir¹, subir², venir, desperdigar, albergar, apoyar, cobijar, evacuar, guardar, habitar, ingresar, acomodar, tender e orientar. Desse modo, ao invés da análise de 674 construções verbais do espanhol da base de dados ADESSE, passou-se a analisar 693 lexemas verbais. Deste valor, 352 lexemas se relacionam a construções verbais locativas do ViPEr e o restante (341 lexemas verbais) refere-se a: (i) 251 construções descritas em classes não locativas do ViPEr; (ii) 75 construções não descritas no ViPEr por serem, de maneira geral, pouco usuais em língua portuguesa; e (iii) 15 lexemas verbais que se referem a verbos não lexicalizados em língua portuguesa (gallear, corretear, arracimar, brujulear, regazar, desaparroquiar, etc.). Sobre as relações locativas, 352 ao todo, observam-se diferentes pares de correspondência – do espanhol para o português - como se verifica na Tabela 3.

Tabela 3. Correspondências entre as construções locativas ADESSE/ ViPEr

Correspondências	Verbos	Classes (ADESSE – ViPEr)	Exemplos	Total	
Uma única entrada locativa no ADESSE	emigrar	Deslocamento – 37LD	Pedro emigró de Portugal hacia Francia.	326	
- Uma única entrada locativa no ViPEr	cruzar Deslocame	Deslocamento – 38L1	Pedro cruzó la calle.	320	
Duplicação no ADESSE	localizar	Localização/Percepção – 38R	Pedro localizó el origen del problema.	1	
 Uma única entrada no ViPEr 	locanzar	Localização – 38R	Buenos Aires se localiza en Argentina.	1	
Uma única entrada no ADESSE - Duplicação no ViPEr	albergar	Localização – 38LD	Pedro albergó a Ana en su casa.		
		Localização - 38L3	Esta casa alberga el museo de un gran poeta.	19	
	habitar	Localização – 38L1	Pedro habita una aldea.		
		Localização – 35LS	Pedro habita en una aldea.		
Duplicação no ADESSE - Duplicação no ViPEr	dejar/deixar	Localização – 38LD	Pedro dejó el libro en la mesa.		
		Deslocamento - 38L1	Pedro dejó Buenos Aires.		
	despegar/descolar, descolar	União – 38LS	Pedro despegó el papel de la pared.	6	
		Deslocamento – 37LD	El avión despegó de la pista del aeropuerto.		

Dos 352 lexemas verbais locativos da correspondência ADESSE/ViPEr: (i) 326 possuem apenas uma construção verbal locativa em cada base de dados; (ii) apenas o verbo *localizar* é duplicado no ADESSE e apresenta uma única entrada no ViPEr; (iii) 19 verbos são duplicados no ViPEr e apresentam apenas uma entrada no ADESSE, cuja grande maioria dos casos refere-se a mudanças sintáticas dessas construções; e (iv) 6 verbos apresentam mais de uma construção verbal nas duas bases de dados, a saber: *apuntar/apontar*; *dejar/deixar*; *despegar/descolar*, *decolar*; *enmarcar/emoldurar*, *enquadrar*; *tirar/lançar*, *andar*; e *volver/voltar*.

Ainda sobre as 352 correspondências locativas ADESSE/ViPEr, pode-se afirmar que 258 (73.3%) são consideradas cognatas, ou seja, compartilham o mesmo sentido e a mesma forma (navegar/navegar) – ou formas muito similares (dejar/deixar,

apoyar/apoiar) — nas duas línguas; e 103 verbos (29%) não são cognatos nas duas línguas, a saber: 82 verbos (23.3%) não possuem semelhanças de forma com seus correspondentes em língua portuguesa (*Pedro buceó en el mar / Pedro mergulhou no mar*); e 12 verbos (3.4%) são considerados falsos cognatos, pois possuem forma idêntica ou semelhante nas duas línguas, mas sentidos diferentes (*Pedro brincó hasta el balcón/O Pedro pulou até a sacada*). Apesar de se tratar de uma lista delimitada de verbos analisados, foi possível observar o alto grau de similaridade dos lexemas, totalizando 73,3% de cognatos das duas línguas, o que corrobora, de certa maneira, com a citação de Almeida Filho (2001, p. 14), que afirma que mais de 85% dos vocábulos em português e em espanhol têm uma origem comum.

4. Considerações Finais

No presente trabalho, descrevemos e associamos as construções verbais locativas das bases de dados verbais em língua espanhola (ADESSE) e em língua portuguesa (ViPEr), além de as conectarmos à WordNet. Como se verifica nos dados apresentados, nem todas as construções espaciais do ADESSE são consideradas locativas no ViPEr, o que reitera as diferenças teórico-metodológicas entre as duas bases de dados contrastadas.

Dos 693 verbos da análise contrastiva ADESSE/ViPEr, 258 são cognatos, apresentando forma e sentido idênticos ou muito similares nas duas línguas. Ressalte-se no entanto que nem sempre esses verbos apresentam os mesmos padrões sintáticos, nem a mesma seleção de preposição e de argumentos para a constituição da frase de base. Deste modo, espera-se dar continuidade à esta pesquisa, analisando, manualmente, o comportamento sintático-semântico de cada um dos 352 lexemas verbais locativos ADESSE/ViPEr. Pretende-se incorporar informações de corpus para validar algumas propriedades dos verbos, como a co-ocorrência do lema com preposições, com padrões sintáticos e com seleções restritivas dos argumentos. Além disso, serão avaliados os pontos comuns e divergentes entre as duas línguas. Assim, poderemos utilizar os dados gerados tanto em aplicações na área da tradução automática, quanto no ensino de português e espanhol como segundas línguas.

Da relação ADESSE/ViPEr, obtivemos um recurso bilíngue com uma granularidade e riqueza de descrição sintático-semântica muito maior que a de um dicionário bilíngue, com a descrição dos diferentes pares de correspondências dos verbos locativos em espanhol e em português e sua conexão às informações semânticas da WordNet.

Agradecimentos. Este trabalho foi financiado pelo Programa de Doutorado Sanduíche da Capes, processo nº 99999.000238/2016-03 e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo nº 2016/20545-0.

5. Referências

Almeida Filho, J. C. P. (2001). Uma metodologia específica para ensino de línguas próximas? In: *Português para estrangeiros - interface com o espanhol*. Campinas: Pontes

Baptista, J. (2012). ViPEr: A Lexicon-Grammar of European Portuguese Verbs. In: *31e Colloque International sur le Lexique et la Grammaire*. České Budějovice: Université de Bohéme du Sud, pp. 10 – 16.

- Cançado, M.; Godoy, L.; Amaral, L. (2013). Catálogo de verbos do português brasileiro. Classificação verbal segundo a decomposição de predicados: Verbos de Mudança. Belo Horizonte: Editora UFMG.
- Crego García, M. V. (1995). El complemento locativo en español. In: *Moenia, Revista lucense de lingüística e literatura*. Lugo: Universidad de Santiago de Compostela. pp. 331-344.
- Di Felippo, A. (2008). Ontologias linguísticas aplicadas ao processamento automático das línguas naturais: o caso das redes wordnets. In: *Múltiplas perspectivas em Linguística*. Uberlândia: Edufu. pp. 469-479.
- García-Miguel, J. M.; Costas, L.; Martínez, S. (2003). Diátesis verbales y esquemas construccionales: Verbos, clases semánticas y esquemas sintáctico-semánticos en el proyecto ADESSE. In: VI Congreso Internacional de Lingüística Hispánica. Leipzig.
- García-Miguel, J. M. (2006) Los complementos locativos. In: Sintaxis histórica de la lengua española. Disponível em http://weba575.webs.uvigo.es/jmgm/public/locativos.pdf, acesso em agosto de 2017.
- Gross, M. (1975). Méthodes en syntaxe. Paris: Hermann.
- Gross, M. (1981). Les bases empiriques de la notion de prédicat sémantique. *Langages*, v. 63, p. 7-52.
- Guillet, A.; Leclère, C. (1992). La structure des phrases simples en français: constructions transitives locatives. Genebra: Librairie Droz S.A.
- Macedo, M. E. (1987). Construções Transitivas Locativas. Centro de Linguística da Universidade de Lisboa, Lisboa.
- Rodrigues, R. (2016). *Análise contrastiva dos verbos locativos do português do Brasil e do português europeu*. Dissertação de Mestrado. Universidade Federal de São Carlos. São Carlos: UFSCar.
- Rojas Nieto, C. (1988). Verbos Locativos en español: aproximación sintácticosemántica. Instituto de investigaciones filológicas. México: Universidad Nacional de México.