

ENTRE CRUZAR, LLEGAR Y ENTRAR: CONSTRUÇÕES LOCATIVAS EM ESPANHOL E PORTUGUÊS COM BASE EM UM CORPUS PARALELO

BETWEEN CRUZAR, LLEGAR AND ENTRAR: LOCATIVE CONSTRUCTIONS IN SPANISH AND PORTUGUESE BASED ON A PARALLEL CORPUS

Lavinia Karolayne dos SANTOS¹

Ana Caroline dos Santos VIEIRA²

Roana RODRIGUES³

RESUMO: As construções locativas são aquelas que se constituem por um verbo que seleciona um argumento interpretado como lugar nas posições de sujeito ou complemento da frase. A fim de analisar, descrever e comparar o comportamento sintático-semântico desse tipo de construção em português brasileiro (PB) e em variedades do espanhol hispano-americano (ESP), construímos e exploramos um corpus/amostra paralelo/a do par linguístico espanhol-português. Desse modo, nesta pesquisa, descrevemos a etapa de compilação do corpus, que se constitui de três obras literárias e suas traduções, além de apresentarmos a análise sintático-semântica de construções locativas transitivas diretas (com os verbos *llenar*, *cruzar* e *salvar*) e preposicionadas (*llegar*, *entrar* e *ingresar*). Ao todo, foram analisadas e comparadas 154 construções verbais locativas, que, de maneira geral, apresentam um comportamento lexical e sintático-semântico semelhante. Destacamos como pontos de distanciamento: (i) a seleção distinta de preposição para a introdução do argumento locativo (*entrar*^{ESP}–*entrar em*^{PB}, *ingresar*^{ESP}–*ingressar em*^{PB}); e (ii) a variação lexical entre os equivalentes, para além dos pares cognatos (*cruzar*^{ESP}–*atravessar*^{PB}, *ingresar*^{ESP}–*entrar*^{PB}, *llegar*^{ESP}–*vir*). A partir dos dados descritos, acreditamos que esta pesquisa, que é fruto de uma iniciação científica, contribui com os estudos na área de Linguística de Corpus e com a pesquisa comparada entre as línguas portuguesa e espanhola.

PALAVRAS-CHAVE: Construções Locativas. Corpus Paralelo. Linguística de Corpus. Espanhol-Português. Análise Comparada.

ABSTRACT: Locative constructions are those that are formed from a verb that selects an argument interpreted as a place in the positions of subject or complement of the sentence. In order to analyze, describe, and compare the syntactic-semantic behavior of this type of construction in Brazilian Portuguese (BP) and varieties of Latin American Spanish (ESP), a parallel corpus of the Spanish-Portuguese language pair was built and explored. In this way, this research describes the stage of compiling the corpus, which consists of three literary works and their respective translations, and also presents the syntactic-semantic analysis of direct transitive locative constructions

1. Mestranda no Programa de Pós-Graduação em Letras da Universidade Federal de Sergipe (UFS). Graduada em Letras (Português/Espanhol) também pela UFS. E-mail: santoslavinia20019@gmail.com. [ORCID](#)

2. Graduada em Letras-Espanhol pela Universidade Federal de Sergipe (UFS). E-mail: anac.santosvieira2023@gmail.com. [ORCID](#)

3. Doutora em Linguística pela Universidade Federal de São Carlos (UFSCar). Professora do Departamento de Letras Estrangeiras e do Programa de Pós-Graduação em Letras da Universidade Federal de Sergipe (UFS). E-mail: roana@academico.ufs.br. [ORCID](#)

(with the verbs *llenar*, *cruzar*, and *salvar*) and prepositional locative constructions (with the verbs *llegar*, *entrar*, and *ingresar*). In total, 154 locative verbal constructions were analyzed and compared, which, in general terms, exhibit similar lexical and syntactic-semantic behavior. Points of divergence include: (i) the different selection of prepositions for the introduction of the locative argument (*entrar a^{ESP}–entrar em^{PB}*; *ingresar a^{ESP}–ingresar em^{PB}*); and (ii) the lexical variation between the equivalents, beyond the cognate pairs (*cruzar^{ESP}–atravesar^{PB}*; *ingresar^{ESP}–entrar^{PB}*; *llegar^{ESP}–vir^{PB}*). Based on the data described, it is considered that this research, resulting from a scientific initiation, contributes to studies in the field of Corpus Linguistics and to comparative research between the Portuguese and Spanish languages.

KEYWORDS: Locative constructions. Parallel corpus. Corpus linguistics. Spanish-Portuguese. Comparative analysis.

Palavras iniciais

Nesta pesquisa, definimos as construções verbais locativas como aquelas em que o argumento de lugar é selecionado, em caráter obrigatório, pelo verbo, fazendo parte de sua valência.

⁴ Nos exemplos em (1) e (2), retirados de Rodrigues (2019), temos casos de construções locativas do espanhol peninsular, com traduções livres ao português brasileiro propostas pela autora:

- (1) a. *Las jóvenes se asentaron en el estado de Virginia.*
b. As jovens assentaram-se no estado de Virgínia.
- (2) a. *El inmigrante escaló un edificio.*
b. O imigrante escalou um edifício.

Na construção locativa reflexiva exemplificada em (1), tem-se o verbo *asentar/assentar* que seleciona o argumento locativo (*el estado de Virgínia*) introduzido pela preposição *en*, configurando uma construção de mudança de localização com leitura resultativa. Em (2), o verbo locativo dinâmico *escalar* seleciona o locativo (*un edificio*) constituindo uma construção transitiva direta. Em (1) e (2), os argumentos de lugar são fundamentais para a constituição e compreensão do que é dito, sendo, portanto, um elemento necessário e obrigatório para a constituição da frase.

Conforme já estudado anteriormente, para o português (Cançado, Godoy, Amaral, 2013; Rodrigues, 2016; Baptista, 2020) e para o espanhol (Lamiroy, 1991; García-Miguel, 2013), os verbos locativos possibilitam diferentes tipos de construções sintático-semânticas, com o locativo ocupando a posição de sujeito (3a e 3b), de complemento direto (4a e 4b) ou, ainda, de complemento preposicionado (5a e 5b), conforme replicamos abaixo, com exemplos retirados de Rodrigues (2019):

4. É importante salientar – e em consonância com o que pontua Rodrigues (2019) – que algumas construções podem selecionar o locativo como argumento acessório, como *na biblioteca* e *em BH* nas frases: (a) *João leu o livro na biblioteca* e (b) *Maria cantou em BH*. Nesses casos, os verbos *ler* e *cantar* não selecionam o lugar como elemento constituinte de sua valência e, portanto, não são entendidos como verbos locativos, nem estudados nesta pesquisa.

- (3) a. *La caja contiene las joyas.*
b. A caixa contém as joias.
- (4) a. *El ejército alcanzó el monte Sinyar.*
b. O exército alcançou o monte Sinyar.
- (5) a. *La hija internó a su padre en un asilo de ancianos.*
b. A filha internou o seu pai em um asilo.

Em Rodrigues (2019), foram analisadas 318 construções verbais locativas do espanhol peninsular e classificadas em 12 classes, com adaptações da tipologia proposta por Baptista (2013) para o português europeu. A autora teve como base teórico-metodológica o modelo do Léxico-Gramática (Gross, 1981), cuja unidade mínima de análise é a frase de base (constituída pelo núcleo e os argumentos por ele selecionados) e com foco na descrição das propriedades estruturais (como o número de argumentos e tipo de preposição), distribucionais (tipo de argumento: humano, não humano, nome parte do corpo, nome locativo, entre outros) e transformacionais (como apassivação, fusão e nominalização) das construções analisadas. A Tabela 1 resume a proposta de tipologia realizada, a qual Rodrigues (2019) nomeou de LGLE (Léxico-Gramática dos verbos Locativos da Língua Espanhola):

Tabela 1 – Tipologia da base de dados LGLE

Classe ⁵	Estrutura ⁶	Verbo	Exemplo	#
35LD	N _o V _{din} Loc ₁ Nloc ₁	acceder	Los migrantes accedieron a costas españolas.	66
35LS	N _o V _{stat} Loc ₁ Nloc ₁	reposar	Una obra reposó en los almacenes de la pinacoteca.	15
37LD	N _o V _{din} Loc-s ₁ Nloc ₁ Loc-d ₂ Nloc ₂	saltar	El hombre saltó desde la plataforma hasta la pista.	46
38L1	N _o V Nloc ₁	abandonar	El jugador abandonó el recinto.	43
38L2	N _o Nloc-v [prep-a] N ₁ [V=poner en Nloc]	empaquetar	El actor empaquetó sus cosas.	7
38L3	Nloc _o V [prep-a] N ₁	albergar	El pequeño hotel albergaba a un grupo.	4
38LD	N _o V _{din} [prep-a] N ₁ Loc-d ₂ Nloc ₂	instalar	Las Vegas instaló cámaras en el pasillo.	101
38LS	N _o V _{din} [prep-a] N ₁ Loc-s ₂ Nloc ₂	evacuar	La policía evacuó a los turistas del lugar.	19
38LT	N _o V _{din} [prep-a] N ₁ Loc-s ₂ Nloc ₂ Loc-d ₃ Nloc ₃	transportar	Thomas transportó millones de pasajeros de la ciudad hasta el archipiélago.	14
38R	N _o V _{stat} [prep-a] N ₁ Loc ₂ N ₂	localizar	Eduardo localizó el poblado minero en el mapa.	3
Total				318

Fonte: Rodrigues (2019, p. 123).

5. Os nomes das classes se baseiam na tipologia de Baptista (2013) que também segue os pressupostos do modelo do Léxico-Gramática. No caso das classes de verbos locativos, LS refere-se às construções Locativas eStativas e LD, às construções Locativas Dinâmicas.

6. Notações: N_o, N₁, N₂, N₃: sujeito e complementos; Prep: preposição; [prep-a]: possibilidade da preposição *a*, quando N₁ é um nome humano; Nloc: nome locativo; Loc: preposição locativa, -d de destino, -s de origem; V: verbo, V_{din}: verbo locativo dinâmico; V_{stat}: verbo locativo estativo.

De acordo com os dados da Tabela 1, as construções com verbos locativos possuem diferentes comportamentos sintático-semânticos em língua espanhola. Em português europeu (Baptista, 2013, 2020) e em português brasileiro (Cançado, Godoy, Amaral, 2013; Rodrigues, 2016), o cenário não parece ser muito diferente; o que nos traz até a presente pesquisa: avaliar quais são efetivamente os pontos de encontro e distanciamento entre algumas construções verbais locativas em espanhol (de variedades hispano-americanas) e em português (do Brasil), com base em dados reais de uso das línguas, ou seja, a partir da análise de dados extraídos de um corpus paralelo. Como se trata de um estudo inicial, que contempla diversas etapas, conforme será melhor descrito adiante, optamos por selecionar verbos classificados na classe 38L1, ou seja, que constituem construções transitivas diretas (*llenar, cruzar e salvar*) e verbos da classe de construções preposicionadas 35LD (*llegar, entrar, ingresar*), a fim de verificar o seu comportamento em língua espanhola (ESP) e os equivalentes no português do Brasil (PB).

Defendemos que este trabalho possui, desde a sua idealização, três objetivos principais: o primeiro, se refere aos conhecimentos básicos relacionados aos estudos linguísticos baseados em corpus: da compilação ou seleção do corpus para o empreendimento descritivo; o segundo, se relaciona à descrição linguística em nível sintático-semântico; e o terceiro, tem como foco a formação acadêmico-científica das estudantes de graduação envolvidas no projeto. Neste artigo, ampliamos e apresentamos os resultados dos trabalhos desenvolvidos no Programa Institucional de Bolsas de Iniciação Científica, que ocorreram na Universidade Federal de Sergipe, no período de 2022 e 2023. Parte dos dados analisados está disponível no relatório desenvolvido e publicado por Santos (2023).

Sendo assim, neste artigo, apresentaremos as etapas de compilação do corpus para posterior análise dos dados, com base, sobretudo, nos preceitos metodológicos da Linguística de Corpus, os quais serão descritos na próxima seção. Em seguida, apontaremos o passo a passo da pesquisa, de cunho descritivo-qualitativo e, então, discutiremos os resultados levantados e a análise sintático-semântica das construções verbais locativas selecionadas, comparando as duas línguas. Por fim, apresentaremos as considerações finais e perspectivas de trabalhos futuros.

Por onde começar? Princípios gerais da Linguística de Corpus

Fillmore (1992) descreve a famosa dicotomia entre os *linguistas de poltrona* (que elaboram suas pesquisas apenas com base na introspecção) e os *linguistas de corpus* (que focam na descrição com base em uma grande quantidade de dados). Segundo o autor, ambos os perfis possuem limitações: os primeiros por não possuírem dados para comprovar suas teorias; e os segundos por acumularem dados que podem não ser representativos e/ou apontar para direcionamentos interessantes na descrição linguística. Como conclusão, o autor aponta que os dois linguistas deveriam coexistir em um único corpo.

O modelo do Léxico-Gramática, adotado nesta investigação, é de base estruturalista e, em seus primórdios, tinha como uma de suas principais diretrizes a descrição de fenômenos relacionados à língua materna do/da pesquisador/a, já que este/a deveria recorrer aos seus conhecimentos introspectivos para a descrição linguística. Mais recentemente, no entanto, ainda com base no Léxico-Gramática, pesquisadores passaram a considerar o uso de corpus na pesquisa descritiva, pois, segundo Laporte (2015), “corpora são importantes para a identificação de formas que de outra maneira podem passar despercebidas, enquanto a introspecção é necessária para distinguir formas raras daquelas que de fato não se usam.”⁷

Desta maneira, em nossa investigação, vislumbramos como único meio de descrição do fenômeno em questão a utilização de corpus linguístico para a obtenção dos dados a serem descritos em nível sintático-semântico. Isso se deu porque concordamos com os autores no sentido de compreender a importância de analisar a língua em efetivo uso. Portanto, a primeira etapa da pesquisa foi definir qual seria o nosso corpus para a realização da análise comparada do comportamento sintático-semântico de construções verbais locativas nas duas línguas. Sendo assim, começamos pelo início: entendendo o que é corpus para a Linguística de Corpus (LC).

De acordo com García-Miguel (2022, p. 11), “um corpus pode se definir [...] como uma coleção de textos orais ou escritos produzidos em um contexto comunicativo natural, representativos de uma língua ou variedade de língua, armazenados em suporte informático e destinados à análise linguística”⁸. Sendo assim, e como bem alertam Aluísio e Almeida (2006), para a LC, este conjunto de textos, que deve ser criteriosamente coletado, precisa estar em formato eletrônico, de maneira que possa ser legível por computadores.

Entendida como uma metodologia, a LC nos apresenta ferramentas para o trabalho com textos autênticos representativos de uma língua ou variedade linguística, desde a sua seleção, passando pela sua organização até chegar à exploração dos dados com o auxílio de ferramentas computacionais. Além disso, trata-se de uma área que estabelece pontes e interfaces com várias outras, sendo basilar, por exemplo, para os estudos desenvolvidos na Linguística Computacional (Freitas, 2015; Souza, Cardoso, Rodrigues, 2025), em propostas pedagógicas para o ensino de línguas (Rocha, 2017; Bocorny, Welp, 2021) e em pesquisas descritivas em diferentes níveis, com base em dados de usos reais da língua (Batista, 2022; Cabré, Da Cunha, 2022).

Em seu texto precursor, Berber Sardinha (2004) aponta a existência de uma grande variedade de corpora, classificados segundo diferentes critérios, os quais sucintamente reproduzimos: (i) *finalidade* (corpus de referência: representa a língua de maneira ampla; corpus especializado: com textos de um domínio específico; corpus de aprendizes: com produções de estudantes de uma segunda língua/língua estrangeira; corpus monitor: em constante atualização); (ii) *modo*

7. Tradução livre de: “Corpora are important for forms that might otherwise go unnoticed, while introspection is needed to distinguish rare forms from those that are not in use.” (Laporte, 2015).

8. Tradução livre de: Un corpus puede definirse [...] como una colección de textos orales o escritos producidos en un contexto comunicativo natural, representativos de una lengua o variedad de lengua, almacenados en soporte informático y destinados al análisis lingüístico. (García-Miguel, 2022, p. 11).

(formado por textos escritos, orais ou multimodais); (iii) *composição* (paralelos: com textos originais e suas traduções; comparados: com textos em duas ou mais línguas comparáveis, mas não traduzidos entre si); (iv) *seleção e extensão* (equilibrado: com amostras selecionadas para representar a diversidade de uma língua; amostra: constituída por amostras de textos maiores, entre outros); e (iv) *tempo* (sincrônico: representa o uso da língua em um período específico; diacrônico: abrange diferentes períodos históricos; entre outros).⁹

O trabalho com corpus, na perspectiva da LC, conforme destaca García-Miguel (2022), possibilita documentar e quantificar as (frequências de) ocorrências de uso e identificar colocações (elementos que co-aparecem/co-ocorrem na língua) e padrões sintagmáticos de maneira sistemática, automatizada e com precisão.

Como nosso objetivo foi a análise do comportamento sintático-semântico de algumas construções verbais locativas em espanhol e seus equivalentes em língua portuguesa, optamos pela utilização de um corpus *paralelo*, ou seja, com textos originais em uma língua (espanhol) e suas traduções na outra língua (português). Na próxima seção, apresentaremos as etapas de compilação e exploração do referido corpus.

Passo a passo da pesquisa: da obtenção dos textos à sua exploração

Podemos organizar o procedimento metodológico desta pesquisa em três etapas: (i) a seleção das construções verbais locativas a serem estudadas; (ii) a compilação do corpus paralelo do par espanhol-português; e (iii) a exploração dos dados e a descrição do comportamento sintático-semântico das construções selecionadas a partir da análise em corpus.

No que se refere à primeira etapa, caracterizamos esta pesquisa como sendo um estudo preliminar e exploratório, já que se restringe à proposta de descrição sintático-semântica de apenas 6 construções verbais locativas, sendo 3 da classe transitiva direta 38L1 (*llenar, cruzar, salvar*) e 3 da classe preposicionada 35LD (*llegar, entrar, ingresar*). Em Rodrigues (2019), havia indícios de que as construções constituídas com esses verbos poderiam apresentar particularidades interessantes se comparadas à língua portuguesa, seja por seu uso aparentemente pouco frequente em espanhol (*salvar* como *ultrapassar*); por seu uso abundante (*llenar, cruzar, entrar*); pelo seu traço fórico, cabendo ao complemento preposicionado o papel de aclarar se o locativo se refere a um lugar de origem ou destino, como ocorre com *llegar*; ou ainda por supostamente haver preferências distintas de uso em cada língua e/ou se tratar de um caso de falso cognato (*ingresar*). A seleção de poucas construções se deu, principalmente, devido ao trabalho já esperado para a compilação do corpus, ação relacionada à segunda etapa da pesquisa.

9. O corpus pode estar cru, ou seja, disponível com a compilação dos textos em seu formato original; ou pode estar *anotado*. Segundo Hovy e Lavid (2010), “[a] anotação de corpus, também chamada de *etiquetagem*, pode ser amplamente concebida como o processo de enriquecer um corpus acrescentando informações linguísticas e outras, inseridas por humanos, por máquinas (ou por ambos em conjunto), com o objetivo de atender a uma finalidade teórica ou prática”. Neste trabalho, utilizamos o corpus cru.

Embora existam diversos corpora compilados e disponíveis online, com interface amigável e gratuitos¹⁰, optamos pela compilação de um corpus de textos literários porque acreditamos que, devido a sua amplitude, esse gênero é propício para a utilização dos verbos selecionados e, sobretudo, porque queríamos verificar o comportamento das construções em um corpus paralelo, ou seja, um corpus constituído com os textos originais (TO) em língua espanhola e os textos traduzidos (TT) em língua portuguesa. O trabalho com o corpus paralelo nos permite avaliar também possíveis seleções de tradução que têm a ver, obviamente, com as preferências/motivações do/a tradutor/a, o que, em nossa visão, parece indicar caminhos produtivos para a compreensão (e variações) do fenômeno locativo nos dois idiomas.

Ressaltamos que a tarefa de compilação do corpus teve uma importância particular para a formação das jovens pesquisadoras envolvidas no projeto, visto que tiveram a oportunidade de lidar com referências fundamentais da área de LC e conhecer e aplicar conhecimentos metodológicos para a seleção, compilação e organização dos textos, com rigorosos critérios e severa organização. A mesma afirmação se pode fazer sobre o uso da ferramenta de exploração, o AntConc: apesar de existirem outras opções de softwares, consideramos que a seleção realizada foi a mais acessível e simplificada para o primeiro contato com a LC.

Selecionamos, inicialmente, 8 romances de autores hispano-americanos e suas traduções para o português. O critério inicial era que os textos já estivessem disponibilizados em .pdf editável e tivessem suas respectivas traduções ao PB, a fim de evitar a árdua tarefa de digitalização e tratamento de textos para reconhecimento óptico de caracteres por meio de algum *software* OCR. Diante do tempo demandado para o processo de alinhamento dos textos, houve a necessidade de restringir o nosso corpus a 3 obras, que foram selecionadas segundo o critério de origem, a fim de incluir amostras de diferentes países, conforme apontamos no Quadro 1.

Quadro 1 – Textos selecionados para a composição do corpus

Título	Autor(a)	Ano/TO	Tradutor/a	Ano/TT	Código¹¹
¿Quién mató a Palomino Molero? Quem matou Palomino Molero?	Mario Vargas Llosa (Peru)	1986	Remy Gorga Filho	1987	[P1-QMPM]
Cien años de soledad Cem anos de solidão	Gabriel García Márquez (Colômbia)	1982	Eliane Zagury	2008	[C2-CAS]
Las venas abiertas de América Latina As veias abertas da América Latina	Eduardo Galeano (Uruguai)	1978	Sergio Faraco	2012	[U3- LVAAL]

Fonte: Elaborado pelas autoras.

10. Como exemplos, podemos citar, sobre a língua portuguesa, os corpora disponíveis na Linguatca (<https://linguateca.pt>), o Corpus do Português (<https://www.corpusdoportugues.org>); e sobre a língua espanhola o Corpes del Siglo XXI (<https://apps2.rae.es/corpes>) e o Corpus del Español (<https://www.corpusdelespanol.org>). Páginas acessadas em out. 2025.

11. Ao longo deste artigo, os exemplos retirados das obras serão identificados com o código de referência, o qual se constitui pela inicial do país de origem, seguido da numeração (P1 / C2 / U3) e da sigla formada pelo título da obra.

Salientamos que, por questões de direitos autorais, as obras foram usadas unicamente para fins acadêmicos, de descrição do fenômeno aqui proposto, sem intenções comerciais e sem o seu acesso livre e compartilhado. Os dados tampouco foram reproduzidos na íntegra, pois, como será visto na próxima seção, foram utilizados apenas fragmentos das obras para a análise, em nível oracional, das construções locativas.

Tendo como base a metodologia adotada por Novodvoski (2013), fomos muito cuidadosas na organização dos dados em cada etapa de compilação do corpus, de modo que adotamos os seguintes procedimentos: (i) nomeação dos arquivos .pdf para facilitar a identificação; (ii) conversão dos arquivos do formato .pdf para .txt; (iii) revisão manual dos textos para limpeza dos arquivos, corrigindo eventuais desvios comuns no processo de conversão; (iv) alinhamento automático de TO e TT através da ferramenta Auto Aligner, disponível na plataforma COPA-TRAD¹²; (v) revisão e limpeza dos arquivos alinhados, disponibilizados em formato de planilha; (vi) conversão do arquivo de planilha a extensão *csv utf-8* e, por fim, (vii) exploração do corpus no concordanciador AntConc¹³.

O processo de limpeza e separação parágrafo por parágrafo realizado previamente, na etapa de conversão do .pdf para o .txt, garantiu um alinhamento textual satisfatório gerado automaticamente no COPA-TRAD, o que deu origem ao nosso corpus/amostra paralelo/a. Em cada procedimento da compilação do corpus, os arquivos foram salvos em diferentes pastas, com o objetivo de manter a organização da pesquisa e facilitar o acesso aos arquivos em suas diferentes extensões.

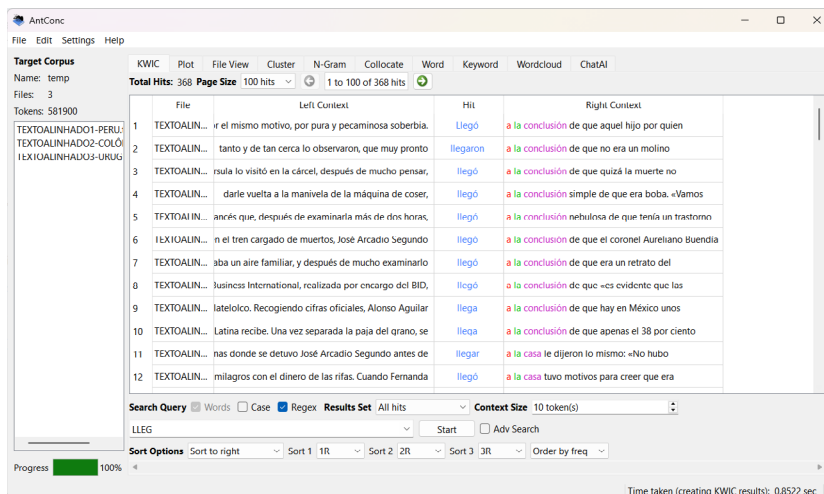
Passamos, assim, para a terceira etapa metodológica da pesquisa: a da exploração dos dados compilados em nossa amostra paralela. Com a utilização do AntConc, subimos todos os textos alinhados de uma só vez, o que totalizou 3 arquivos e 581.900 *tokens* (ocorrências de palavras no corpus).

Utilizamos o sistema de busca *regex*, que nos permite encontrar padrões - e não necessariamente palavras exatas. Sendo assim, colocamos as sequências de caracteres dos verbos que nos levariam a encontrar sua forma em diferentes conjugações (*entr, lleg, ingres, llen, cruz e salv*). Ressaltamos que este tipo de busca também gerou resultados em língua portuguesa, já que o corpus paralelo gerado estava em um só arquivo e é lido como um “bloco único” pelo AntConc. No entanto, manualmente, selecionamos os casos em espanhol para, a partir dele, comparar com o seu uso em português. A Figura 1 ilustra esse processo de exploração dos dados no AntConc.

12. A COPA-TRAD (Corpus Paralelo de Tradução) é uma plataforma online e gratuita, desenvolvida na Universidade Federal de Santa Catarina, que dispõe de várias ferramentas e funcionalidades que auxiliam no processo de construção e exploração de corpus. Ela pode ser acessada em: <https://copa-trad.ufsc.br>. Último acesso em dez. 2025.

13. O AntConc é um concordanciador de análise textual, de acesso gratuito, desenvolvido por Laurence Anthony. Disponível em: <https://www.laurenceanthony.net/software/antconc>. Último acesso em: dez. 2025.

Figura 1 – Interface do AntConc utilizada para exploração do corpus paralelo



Fonte: Elaborado pelas autoras, com o processamento dos dados na ferramenta AntConc.

Os casos identificados como locativos foram transpostos a uma nova planilha para análise e comparação detalhada. O processo metodológico adotado possibilitou resultados assertivos e o uso de ferramentas tecnológicas otimizou e facilitou a construção e exploração dos dados. Na próxima seção, apresentaremos a análise sintático-semântica realizada.

Llenar é Encher? Ingresar é Ingressar? Descrição sintático-semântica e comparação das construções verbais locativas do par espanhol-português

Com o uso do AntConc, realizamos as buscas específicas pelos verbos locativos do espanhol, a fim de observar seu comportamento e a tradução realizada para a língua portuguesa. Na Tabela 2, apresentamos os dados quantitativos e um exemplo dos resultados encontrados.

Tabela 2 – Dados quantitativos dos verbos locativos do espanhol no corpus

Verbo - Espanhol	Casos encontrados	Casos locativos	Exemplo TO - TT
Llenar	128	21	La casa se <u>llenó</u> de amor. A casa se <u>encheu</u> de amor.
Cruzar	180	15	Ya los barcos negreros no <u>cruzan</u> el océano. Os navios negreiros já não <u>cruzam</u> o oceano.
Salvar	178	0	--
Llegar	368	70	Cuando Fernanda <u>llegó</u> a la casa Quando Fernanda <u>chegou</u> em casa
Entrar	2.437	44	Una vez que <u>entra</u> en la casa, nadie escapa a la peste. Uma vez que a peste <u>entra</u> em casa, ninguém escapa.
Ingresar	59	4	Al <u>ingresar</u> a la Universidad, los estudiantes juran... Ao <u>ingressar</u> na universidade, os estudantes juram por escrito...
Total	3.350	154	

Fonte: Elaborado pelas autoras.

Conforme observado na Tabela 2, foram realizados recortes para a análise exclusiva de construções entendidas prototipicamente como locativas. Sendo assim, desconsideramos, por exemplo, resultados que fugiam completamente da análise em questão por coincidir com a sequência de caracteres procurada (como *gallego* que compartilha a sequência de caracteres do verbo *llegar* ou ainda *centro*, *concentración* e vários outros que compartilham a sequência do verbo *entrar*) ou que, embora apresentasse características locativas, o argumento que deveria ser entendido como um lugar físico/prototípico, era um nome não-locativo, como exemplificamos em (6) e (7) com o verbo *llegar*:

- (6) a. Había llegado a la vejez con todas sus nostalgias vivas.
b. Tinha chegado à velhice com todas as suas saudades vivas. [C2-CAS]
- (7) a. Aureliano Segundo llegó a ser un virtuoso del acordeón.
b. Aureliano Segundo chegou a ser um virtuose do acordeão. [C2-CAS]

Em (6) tem-se uma construção muito parecida às construções locativas prototípicas, no entanto, o argumento de lugar é na verdade um conceito temporal de idade (*vejez/velhice*) e não um lugar físico. Por isso, não foi levada em consideração para a análise. O mesmo ocorreu em casos como (7), em que o verbo faz parte de uma construção gramatical que, nesse caso específico, nos remete a mudança de estado (*llegar a ser / chegar a ser - tornar-se*).

A coluna *casos locativos* da Tabela 2 aponta as construções que realmente foram estudadas (154 ao todo) por se constituírem pelos verbos selecionando argumentos locativos-físicos.

Os verbos da classe 38L1 (*llenar*, *cruzar* e *salvar*) selecionam um nome humano na posição de sujeito e um argumento interpretado como lugar na posição de complemento direto. No caso de *llenar*, todas as construções locativas encontradas do ESP mantiveram uma regularidade e foram traduzidas por *encher* em PB, como em (9) e (10). Os casos que destoam, possuem usos de *llenar* metafóricos e não locativos, como apresentado em (8), com o equivalente *embarrigar*:

- (8) a. ¿Qué le hiciste? ¿La llenaste?
b. Que foi que fez? Você a embarrigou? [P1-QMPM]
- (9) a. Aquellas visitas fuleron llenando la casa de juguetes prodigiosos.
b. Aquelas visitas foram enchendo a casa de brinquedos prodigiosos. [C2-CAS]
- (10) a. El estrecho taller de orfebrería se llenó de emisarios.
b. A estreita oficina de ourivesaria se encheu de emissários. [C2-CAS]

Em (9), tem-se o verbo (*llenar/encher*) selecionando um nome humano na posição de sujeito (*visitas*) e um nome locativo na posição de complemento (*la casa*), além de explicitar o argumento que *enche* este espaço (*juguetes/brinquedos*), introduzido pela preposição *de*. Essa construção nos leva a questionar se não seria mais adequada a sua inserção na classe 38LD – e não 38L1 –, visto que se trata de uma seleção triargumental e não biargumental: o complemento preposicionado (*de juguetes*) é um argumento importante para a constituição da frase.

Consideramos que a construção poderia se assemelhar à estrutura da classe 38LD, no sentido de selecionar três argumentos: um argumento na posição de sujeito (*visitas*), um argumento na posição de complemento direto (*los juguetes*) e outro na posição de complemento preposicionado de lugar de destino (*en la casa*): *Aquellas visitas fueron llenando los juguetes prodigiosos en la casa*. No entanto, esse tipo de construção não é natural nem espanhol, nem em português e, inclusive, não foi identificada em nosso corpus, o que nos leva a acreditar que *llenar/encher* possui particularidades interessantes e que merecem atenção em um trabalho específico sobre a sua multifuncionalidade (ou seja, sobre a maneira como eles, em cada idioma, apresentam diferentes construções sintáticas, mantendo-se o seu valor locativo). A construção em (10), por sua parte, representa o processo transformacional de passiva pronominal, por isso o locativo (*taller/oficina*) ocupa a posição de sujeito da frase e o argumento agentivo (que realiza a ação) é omitido. Na voz ativa, a construção apresentaria que *alguien llenó el estrecho taller de orfebrería de emisarios* – assemelhando-se à construção em (9).

A respeito do verbo *cruzar*, as construções apresentaram traduções para três verbos do PB: o equivalente *cruzar* (que obteve a maior quantidade de dados), o verbo *atravessar* e o verbo *passar*, como se verifica nos exemplos de (11) a (13):

- (11) a. *El guardia lo siguió. Cruzaron la pista de baile.*
b. O guarda o seguiu. Atravessaram a pista de dança. [P1-QMPM]
- (12) a. *Ya los barcos negreros no cruzan el océano.*
b. Os navios negreiros já não cruzam o oceano. [U3-LVAAL]
- (13) a. *Cuando cruzaban las últimas viviendas del caserío...*
b. Quando passavam pelas últimas moradias do casario... [P1-QMPM]

As construções analisadas são locativas e pertencentes à classe 38L1, com a seleção de um nome humano na posição de sujeito e um argumento locativo na posição de complemento direto. À princípio, com base em nossa introspecção acreditávamos que o verbo *atravessar* seria mais produtivo que *cruzar* em língua portuguesa, no entanto, no corpus de análise, *cruzar* ainda apresentou a maior quantidade de usos, o que pode ser justificado pelo gênero analisado (literário) e o seu caráter de tradução.

Por fim, no caso de *salvar* não houve nenhum dado em que o verbo apresentou sentido locativo, conforme esperado pelas pesquisadoras por desconhecerem, antes do estudo de Rodrigues (2019) o uso do verbo *salvar* com o sentido locativo de *ultrapassar* e o considerarem pouco produtivo, daí, inclusive, o interesse em sua investigação.

Na primeira versão deste trabalho o corpus analisado se constituiu também da obra *Inés del alma mía*, de Isabel Allende (Chile), com tradução de Ana Mendes Lopes. Essa obra foi posteriormente excluída devido ao fato de a tradutora ser portuguesa e não brasileira. Ressaltamos, no entanto, que nesse livro encontramos uma ocorrência de *salvar* traduzido por *ultrapassar*, embora o lugar tampouco tenha sido físico/prototípico¹⁴.

14. A frase encontrada na obra de Allende é a seguinte: *Se convirtió en un obstáculo imposible de salvar*. | revelou-se um obstáculo impossível de ultrapassar.

Por sua vez, as construções preposicionadas na classe 35LD (*llegar, entrar e ingresar*) se caracterizam por selecionar um argumento na posição de sujeito e um argumento locativo preposicionado na posição de complemento. No caso do verbo *llegar* foram identificadas traduções para quatro verbos diferentes em PB, conforme é possível observar nos exemplos de (14) a (18), sendo eles: *alcançar* (3 casos), *desembarcar* (1 caso), *vir* (2 casos) e seu equivalente, *chegar* (64 casos).

- (14) a. *El carro llegó a Amotape.*
b. O carro chegou a Amotape. [P1-QMPM]
- (15) a. *Entre 1503 y 1660 llegaron al puerto de Sevilla.*
b. Entre 1503 e 1660, desembarcaram no porto de Sevilha. [U3-LVAAL]
- (16) a. *No alcanzó a llegar al camino de la ciénega.*
b. Não chegou a alcançar a estrada do pantanal. [C2-CAS]
- (17) a. *Los obreros... también llegaron desde Europa.*
b. Os operários... também vieram da Europa. [U3-LVAAL]
- (18) a. *Llegó hasta la puerta con un paso de baile.*
b. Chegou até a porta com um passo de dança. [C2-CAS]

Observamos que *llegar* selecionou quatro preposições diferentes entre os dados e o mesmo ocorreu com as traduções ao PB. As proposições identificadas em ESP foram *a, de, desde e hasta*, já em PB foram *a, em, de e até*, conforme é possível observar nos exemplos apresentados acima.

No exemplo (15), a preposição *en* não foi empregada em espanhol, embora apareça na tradução para o português, porque, apesar de possuir valor locativo, não expressa a mesma noção espacial atribuída à construção original. Em espanhol, *en* tende a indicar localização, posicionamento ou referência espaço-temporal, e não propriamente deslocamento ou direção, função normalmente desempenhada pela preposição *a* com o verbo *llegar*. Esse fenômeno reforça as reflexões feitas sobre o comportamento das preposições em construções locativas em PB e ESP realizada por Rodrigues (2022).

Um processo semelhante ocorreu com o uso de *desde* no exemplo (17), que foi traduzido para *de* em PB, isso porque a preposição *desde*, no PB, possui sobretudo sentido de localização temporal e não de deslocamento espacial. Além disso, ainda sobre o exemplo (17), enquanto em ESP a estrutura *llegar desde* é utilizada para indicar procedência, em PB a origem é indicada pelo uso de *vir de*.

Destacamos ainda os casos em que *llegar* é traduzido pelo verbo *vir* como está representado nos exemplos (17) e (19):

- (19) a. *Era uno de los primeros que habían llegado a la casa para el bautismo.*
b. Era um dos primeiros que tinham vindo à casa para o batismo. [C2-CAS]

Este tipo de seleção nos leva a refletir sobre o caráter *fórico* do verbo *llegar*. Segundo Rodrigues (2019), alguns verbos possuem caráter *fórico* por se referir a um local que considera o *aquí* e o *agora* da situação comunicativa. Para a autora, o verbo *venir* (*vir* em português) é *fórico*, já que a noção de local está intrínseca nele e se refere à localização daquele que está falando. O verbo *llegar* não possui um caráter *fórico* como *venir*, por ser possível dissociar o referente locativo daquele que fala (é possível *llegar a* ou *de* algum lugar que não é aquele ocupado pelo falante). No entanto, no exemplo (17), o verbo *llegar* indica a presença do sujeito em um lugar ocupado pelo falante (o *aquí*), tendo partido de um destino anterior (*desde Europa*/da Europa).

Sobre o verbo *entrar*, as traduções ao português foram todas para o verbo *entrar*, conforme acontece em (20):

- (20) a. *Entró a la casa con mucha familiaridad...*
b. *Entrou na casa com muita familiaridade...* [C2-CAS]

No caso de *entrar*, os dados variaram entre as preposições *a* e *en*, em ESP, enquanto no PB todos os casos analisados tiveram o argumento locativo introduzido pela preposição *em*, como demonstramos em (21) e (22):

- (21) a. *Un grupo de hombres entró en el Riobar y Moisés fue a atenderlos.*
b. Um grupo de homens entrou no Riobar e Moisés foi atendê-los. [P1-QMPM]
(22) *Entró a la casa arrastrando la batea llena de granos.*
Entrou na casa arrastando a bacia cheia de grãos de milho. [P1-QMPM]

Por fim, a respeito do verbo *ingresar*, foram obtidos apenas 4 casos retirados do corpus. Dentre esses dados, três construções apresentaram traduções em PB para o verbo equivalente, *ingressar* e uma para o verbo *entrar*. Verificamos também variações na seleção das preposições, sendo em ESP a preferência pela preposição *a*, enquanto em PB, opta-se pela preposição *em*, como se pode ver nos exemplos (23) e (24).

- (23) a. *En el primer año se sienten los primeros síntomas, y en diez años se ingresa al cementerio.*
b. No primeiro ano, sentem-se os primeiros sintomas, e em dez anos entra-se no cemitério. [U3-LVAAL]
(24) *Al ingresar a la Universidad, los estudiantes juran por escrito.*
Ao ingressar na universidade, os estudantes juram por escrito. [U3-LVAAL]

Considerando os dados analisados, sintetizamos nossa análise no Quadro 3.

Quadro 3 – Comparação do comportamento sintático-semântico de construções locativas em ESP e PB

ESP	PB	Prep ESP	Prep PB	Comentários gerais
<i>Llenar</i>	<i>Encher</i>	-	-	Em todos os dados locativos encontrados, <i>llenar</i> apresentou tradução para o verbo <i>encher</i> em PB, os poucos dados cuja tradução destoava disso, produziam sentidos metafóricos, não locativos.
<i>Cruzar</i>	<i>Atravessar</i> <i>Cruzar</i> <i>Passar</i>	-	-	Os dados do verbo <i>cruzar</i> apresentaram 3 possibilidades de tradução ao PB: <i>cruzar</i> , <i>atravessar</i> e <i>passar</i> .
<i>Salvar</i>	-	-	-	Não foram encontrados dados locativos na análise.
<i>Llegar</i>	<i>Alcançar</i> <i>Desembarcar</i> <i>Chegar</i> <i>Vir</i>	<i>a</i> <i>de</i> <i>desde</i> <i>hasta</i>	<i>a</i> <i>em</i> <i>de</i> <i>até</i>	Em ESP, <i>llegar</i> acompanha a preposição <i>a</i> , <i>de</i> , <i>desde</i> e <i>hasta</i> . Em PB, foi traduzido por <i>chegar</i> (majoritariamente), <i>desembarcar</i> , <i>vir</i> e <i>alcançar</i> , com as preposições: <i>a</i> , <i>em</i> , <i>de</i> e <i>até</i> .
<i>Entrar</i>	<i>Entrar</i>	<i>a</i> <i>en</i>	<i>em</i>	Em ESP, <i>entrar</i> pode ser sucedido pela preposição <i>a</i> ou <i>en</i> . Em PB, <i>entrar</i> é acompanhado apenas pela preposição <i>em</i> .
<i>Ingresar</i>	<i>Ingressar</i> <i>Entrar</i>	<i>a</i>	<i>em</i>	Em ESP, <i>ingresar</i> apresentou a seleção da preposição <i>a</i> e os casos analisados nos levaram à interpretação de <i>ingresso</i> com o sentido de <i>matrícula</i> ou da ação de <i>entrar</i> . Em PB, as traduções foram <i>ingressar</i> e <i>entrar</i> , com o argumento locativo sempre introduzido pela preposição <i>em</i> .

Fonte: Elaborado pelas autoras.

Considerações finais

Como apresentado em nossas palavras iniciais, objetivamos com este trabalho avaliar quais são efetivamente os pontos de encontro e distanciamento entre algumas construções verbais locativas em espanhol (de variedades hispano-americanas) e em português (do Brasil), a partir da exploração de dados retirados de corpus paralelo.

Ao longo do nosso artigo, apresentamos os processos metodológicos adotados para as três fases da pesquisa: o recorte da análise sintático-semântica a ser realizada; a seleção, compilação e exploração dos textos que constituem nosso corpus paralelo; e a análise sintático-semântica efetiva.

A pesquisa contou com a ampliação de conhecimentos relacionados à Linguística de Corpus, área da Linguística caracterizada pelo seu alto teor metodológico e pelo uso de tecnologias para a obtenção e investigação dos dados; além da reflexão, análise e descrição dos fenômenos em nível sintático-semântico estudado.

Os dados demonstraram que as construções verbais analisadas apresentam muitos pontos comuns nas duas línguas, com a seleção dos mesmos números e tipos de argumentos para a construção da frase de base. Como pontos de distanciamento, verificamos pequenas mudanças na seleção dos verbos para tradução (*llegar* também admitiu *vir* em português e *ingresar* como

entrar); e uma variação mais expressiva na seleção da preposição que antecede o argumento locativo, sendo em espanhol majoritariamente introduzido pela preposição *a*, em predicados dinâmicos, e em português, a preposição *em*, o que reitera a investigação de Rodrigues (2022) a respeito das características semânticas das preposições locativas nas duas línguas.

É importante relembrar que se trata de uma pesquisa que analisou um número limitado de dados, de um gênero específico (literário) e em textos originais e traduzidos (paralelos) o que, por conseguinte, não representa a totalidade de uso desses verbos nas duas línguas. De qualquer maneira, a pesquisa aponta indícios do comportamento dessas construções, que foram analisadas com rigor científico.

Em decorrência desta pesquisa, já foi realizado o trabalho de conclusão de curso sobre o comportamento sintático-semântico do verbo *llenar* em língua espanhola (Santos, 2024). Além disso, acreditamos ser muito produtivo a continuação deste estudo, com a ampliação do corpus e/ou a análise dessas construções em outros gêneros textuais, como *notícias*.

Com este trabalho, conseguimos destacar as principais contribuições desta pesquisa: a construção e análise de um corpus, assim como a descrição do fenômeno das construções verbais locativas em uma perspectiva comparada, atendendo ao objetivo principal de nosso projeto; e o compromisso com a formação crítica e científica de pesquisadoras da área de estudos linguísticos, cujo trabalho com a Linguística de Corpus proporcionou um olhar atento e apurado para as questões metodológicas da pesquisa científica.

Referências

- ALUÍSIO, S. M.; ALMEIDA, G. M. B. O que é e como se constrói um corpus? Lições apreendidas na compilação de vários corpora para pesquisa linguística. *Caleidoscópio*, v. 4, n. 3, 2006.
- BAPTISTA, J. ViPER: uma base de dados de construções léxico-sintáticas de verbos do Português Europeu. In: *Textos Seleccionados*. XXVIII Encontro Nacional da Associação Portuguesa de Linguística. Coimbra: APL, 2013.
- BAPTISTA, J.; MAMEDE, N. *Dicionário Gramatical de Verbos do Português*. Faro: Universidade do Algarve Editora, 2020.
- BATISTA, H. R. Uso de artigos definidos em redações escolares. *Revista Linguística*, v. 17, 2022.
- BERBER SARDINHA, T. *Linguística de Corpus*. São Paulo: Manole, 2004.
- BOCORNÝ, A. E. P.; WELP, A. K. S. O desenho de tarefas pedagógicas para o ensino de inglês para fins acadêmicos: conquistas e desafios da Linguística de Corpus. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 29, n. 2, 2021.
- CABRÉ, M. T.; DA CUNHA, I. El papel de los corpus en la terminología: Una mirada específica a la terminología del español. In: *Lingüística de corpus en español*. Routledge, 2022.
- CANÇADO, M., GODOY, L.; AMARAL, L. *Catálogo de verbos do português brasileiro: classificação segundo a decomposição de predicados: verbos de mudança*. Belo Horizonte: Editora UFMG, 2013.

- FILLMORE, C. J. “Corpus linguistics” or “computer-aided armchair linguistics”. In: SVARTVIK, J. (org.). *Directions in corpus linguistics*. Berlin: Mouton de Gruyter, 1992.
- FREITAS, C. Corpus, linguística computacional e as humanidades digitais. In: LEITE, M. S.; GABRIEL, C. *Linguística, Pesquisa e Educação*. Rio de Janeiro: DP et ali, 2015. Disponível em: <https://www.researchgate.net/publication/280078754>. Acesso em: out. 2025.
- GARCÍA-MIGUEL, J. M. Los complementos locativos. In: *Sintaxis histórica de la lengua española*. México: Fondo de Cultura Económica, 2006.
- GARCÍA-MIGUEL, J. M. Lingüística de corpus: de los datos textuales a la teoría lingüística. *Estudios de lingüística del español*, v. 45, 2022.
- GROSS, M. Les bases empiriques de la notion de prédicat sémantique. *Langages*, 1981. p. 7-52.
- HOVY, E.; LAVID, J. Towards a ‘science’ of corpus annotation: a new methodological challenge for corpus linguistics. *International journal of translation*, v. 22, n. 1, 2010.
- LAPORTE, E. The Science of Linguistics. *Inference: International Review of Science*, 2015. Disponível em: <https://inference-review.com/article/the-science-of-linguistics>. Acesso em: out. 2025.
- LAMIROY, B. *Léxico y gramática del español*: Estructuras verbales de espacio y de tiempo. Barcelona: Anthropos, 1991.
- NOVODVORSKI, A. *Estilo das traduções de Sergio Molina de obras de Ernesto Sabato*: um estudo de corpora paralelos Espanhol/Português. Tese (Doutorado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2013.
- ROCHA, J. M. P. Abordagem orientada por dados: desafios e aplicações no ensino de LE. *Revista do GEL*, v. 14, n. 3, 2017.
- RODRIGUES, R. *Análise contrastiva dos verbos locativos do português do Brasil e do português europeu*. Dissertação (Mestrado em Linguística) – Universidade Federal de São Carlos (UFSCar), São Carlos, 2016.
- RODRIGUES, R. *Contribuições para um léxico-gramática das construções locativas do espanhol*. Tese (Doutorado em Linguística) – Universidade Federal de São Carlos, São Carlos, 2019.
- RODRIGUES, R. Estudos comparados Português e Espanhol: o lugar das preposições locativas. In: MATOS, D.; COLAÇA, J. P. (Org.). *Diálogos interculturais e linguísticos em Espanhol*: entre teorias e práticas docentes. 1ed. São Cristóvão: Editora UFS, 2022.
- SANTOS, L. K. *Construção e análise de corpora paralelos: Análise de construções locativas transitivas diretas*. Relatório final de Iniciação Científica – Universidade Federal de Sergipe, São Cristóvão, 2023.
- SANTOS, L. K. *A sintaxe do sujeito locativo nas construções locativas transitivas diretas do espanhol*. Trabalho de Conclusão de Curso – Universidade Federal de Sergipe, São Cristóvão, 2024.
- SOUZA, J. W. C.; CARDOSO, P. C. F.; RODRIGUES, R. Taxonomy of Discourse Signals for RST Relations: A Study in a News Corpus. *Revista Brasileira de Linguística Aplicada*, v. 25, p. 1-31, 2025.