



Caracol

ISSN: 2178-1702

ISSN: 2317-9651

Universidade de São Paulo

Novodvorski, Ariel; Lima, Fernanda Ravazzi  
Fraseologia com léxico tabu: uma análise contrastiva em corpus  
paralelo espanhol/ português de legendas de filmes argentinos  
Caracol, núm. 19, 2020, Janeiro-Junho, pp. 172-198  
Universidade de São Paulo

Disponível em: <https://www.redalyc.org/articulo.oa?id=583766868006>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em redalyc.org

UNEM  redalyc.org

Sistema de Informação Científica Redalyc  
Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal  
Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa  
acesso aberto

# Fraseologia com léxico tabu: uma análise contrastiva em corpus paralelo espanhol/ português de legendas de filmes argentinos

Ariel Novodvorski  
Fernanda Ravazzi Lima

Recebido em: 30 de julho de 2019  
Aceito em: 29 de agosto de 2019

Doutor em Estudos Linguísticos  
pela UFMG. Professor do Curso  
de Letras/Espanhol e do Programa  
de Pós-Graduação em Estudos  
Linguísticos do ILEEL/UFU.  
Atualmente realiza Pós-doutorado  
na UFRGS.  
Contato: arivorski@ufu.br

Mestre em Estudos Linguísticos  
pela Universidade Federal de  
Uberlândia (2019).  
Contato: fernandaravazzi@  
hotmail.com  
Brasil

## PALAVRAS-CHAVE:

léxico tabu; fraseologia contrastiva; corpus paralelo espanhol/português; linguística de corpus; legendas de filmes argentinos.

## KEYWORDS:

taboo lexicon; contrastive phraseology; Spanish/Portuguese parallel corpus; corpus linguistics; Argentinean movies' subtitle corpus.

Resumo: O objetivo deste artigo é identificar e investigar os diferentes tipos de fraseologismos em torno do vocábulo tabu *cagar*, presente em filmes argentinos, e analisar como foram realizadas suas respectivas traduções ao português brasileiro, num corpus paralelo de legendas alternativas. Nosso corpus de estudo é composto por cinco filmes argentinos, em que prevalece o espanhol rio-platense. A metodologia adotada segue os pressupostos da Linguística de Corpus para a compilação e análise do corpus, com suporte da fundamentação teórica existente na área de Fraseologia e dos Estudos da Tradução. Os resultados apontam para a produtividade de significações do vocábulo *cagar* e para a importância da consideração integrada de aspectos sintático-semânticos nas análises, essencial para a adequada percepção dos sentidos nos fraseologismos e sua consequente tradução.

Abstract: The objective of this paper is to identify and investigate the different types of phraseologisms around the taboo word *cagar*, present in Argentinean movies, and to analyze how they were translated into Brazilian Portuguese, in a parallel corpus of alternative subtitles. Our corpus of study is composed by five Argentinean movies, in which the River Plate Spanish prevails. The methodology adopted follows the assumptions of Corpus Linguistics, both for the compilation and for the analysis of the corpus, supported by the existing theoretical foundation in the area of Phraseology and Translation Studies. The results point to the productivity of meanings of the word *cagar* and to the importance of integrated consideration of syntactic-semantic aspects in analyses, which are essential for the proper perception of the senses in phraseologisms and their consequent translation.

## 1. INTRODUÇÃO

Este trabalho está inserido nas áreas de Fraseologia Contrastiva (Corpas Pastor, 2010; 1996) e dos Estudos da Tradução (Hurtado Albir, 2008), a partir de uma investigação com base na Linguística de Corpus (Berber Sardinha, 2004; Parodi, 2010; Novodvorski, 2015; 2017) dos diferentes tipos de fraseologismos em torno de léxico tabu (Preti, 1984), presentes num corpus paralelo espanhol/português de legendas de filmes argentinos com o ator Ricardo Darín. Nosso objetivo é analisar as respectivas traduções dos fraseologismos em torno do léxico tabu, mais especificamente do vocábulo *cagar*, num corpus paralelo de “legendas alternativas” traduzidas para o português brasileiro, em vistas de aguçar a subcompetência fraseológica, na compreensão mais aprimorada da riqueza semântica desse vocábulo e da percepção necessária atrelada ao ato tradutório.

Entendemos pelo rótulo “legendas alternativas” aquelas que são disponibilizadas de forma gratuita em páginas da internet, geralmente feitas por tradutores amadores, intitulados como *legenders* (tradutores independentes). Neste trabalho, denominamos léxico tabu ao uso de palavrões e expressões de baixo calão, consideradas de cunho grosseiro e vulgar. Em particular, analisaremos as traduções dos fraseologismos formados em torno do léxico tabu *cagar*. A expressão “corpus paralelo” corresponde à disposição alinhada (em paralelo) dos segmentos originais e traduzidos das legendas, o que possibilita sua análise por meio de linhas de concordância paralelas.

Uma das principais motivações para a realização deste trabalho são dificuldades implicadas no ato tradutório, principalmente no tocante à

manutenção ou adaptação das fraseologias com referências culturais locais próprias no corpus original. Segundo Novodvorski (2015, 16) “os tradutores convivem em uma realidade em que os espaços para informações sobre fraseologia ainda são restritos e com informações confusas, no auxílio às investigações sobre essas estruturas”.

Os filmes escolhidos para a compilação do corpus são aqueles em que atua o ator Ricardo Darín, considerado um grande ator do cinema latino-americano<sup>1</sup>. Ricardo Darín nasceu em Buenos Aires, Argentina, em 1957. Iniciou a carreira muito jovem e começou a atuar no cinema em 1969. Por se tratar de um dos mais populares atores de seu país e pelo fato de os filmes que contam com sua atuação terem uma ampla circulação e reconhecimento no Brasil, entendemos que analisar um corpus de filmes argentinos com o ator Ricardo Darín seja um modo de alcançar os usos autênticos e mais recorrentes de fraseologismos com léxico tabu, que poderiam oferecer dificuldades para a tradução ao português.

Como suporte para a análise do corpus, utilizaremos dois programas computacionais, a saber: *Word Smith Tools 6.0* (Scott, 2012), para levantamento do léxico do corpus (itens e formas, isto é, palavras totais e diferentes), por meio de listas de palavras para ocorrências individuais e de linhas de concordância para identificação de seus fraseologismos; e *ParaConc* (Barlow, 2004), para alinhamento do corpus e análise das concordâncias paralelas. A seguir, apresentamos uma breve contextualização do marco teórico deste trabalho.

---

1 Fonte: <<http://www.revistastatus.com.br/2012/10/22/ricardo-darin/>>. Acesso em: 4 jun. 2019.

## 2. FUNDAMENTAÇÃO TEÓRICA

Tendo em vista seu efeito comunicativo, a compreensão dos usos do léxico deve levar em consideração o contexto de ocorrência dos vocábulos. Tal percepção é determinante quando as unidades lexicais em questão correspondem a expressões caracterizadas por marcas tabuísticas próprias de determinada região, ainda mais quando se trata de expressões em línguas estrangeiras e em relação tradutória com o português. Segundo Zavaglia e Welker (2013),

O léxico é entendido como o conjunto de todas as palavras de uma língua, também chamadas de *lexias*. [...] Embora possa parecer um conjunto finito, o léxico de cada uma das línguas é tão rico e dinâmico que mesmo o melhor dos linguistas não seria capaz de enumerá-lo. Isto ocorre porque dele faz parte a totalidade das palavras, desde as preposições, conjunções ou interjeições, até os neologismos, regionalismos ou terminologias, passando pelas gírias, expressões idiomáticas, provérbios ou *palavrões* (Zavaglia; Welker, 2013).

Como se observa, o estudo dos *palavrões* também faz parte da Lexicologia. Guérios (1956, 11) define o tabu linguístico como “a proibição de dizer certo nome ou certa palavra, aos quais se atribui poder sobrenatural, para evitar infelicidade ou desgraça. Impropriamente, o tabu linguístico é a proibição de dizer qualquer expressão imoral ou grosseira”. Este aspecto também é definido por Orsi (2011), que atesta que o tabu linguístico

é decorrente das sanções, restrições e escrúpulos sociais; atua na não permissão ou na interdição de se pronunciar ou dizer certos itens lexicais aos quais se atribui algum poder e que, se violados, poderão trazer perseguições e castigos para quem os emprega. E, por estar em si também o impulso por ultrapassá-los, o homem reverte as imposições e usa os *palavrões* e outras construções lexicais como forma de expressão de seus sentimentos e meio de subversão das proibições (Orsi, 2011, 336).

Orsi (2011, 335 e 340) concebe *palavrão* “por aquele item que ultrapassa o limite da considerada boa decência e da moralidade”. Além de também serem definidos “como injúrias, que são, por definição, um atentado a outrem, uma ofensa”. Monteiro (1986) aponta que o contexto determina os usos da linguagem, já que existem situações específicas para se dizer algo, devendo estar de acordo com a cultura, ambiente e classe social. Em relação a isso, Preti (1984) também comenta que o contexto deve ser levado em conta, já que, “é a situação (condições extraverbais que cercam o ato de fala) que nos permitirá caracterizar o que vulgarmente costuma chamar-se de ‘palavrão’, empregado como blasfêmia ou injúria” (1984, 41). Isto vem ao encontro de Guedelha (2011), que concluiu que os palavrões são ditos por pessoas de quase todas as idades e classes sociais.

O palavrão tem uma estranha força de choque e atração que nem sempre é fácil de se explicar. A atração é responsável pelo uso crescente desse tipo de expressão; já o choque, que responde pela inconveniência do uso, a qual atenta contra a “boa moral” e os “bons costumes”, ocasiona a tabuização do próprio palavrão, para o qual se criam expressões substitutivas (Guedelha, 2011, 60).

Na presente pesquisa, pressupomos que os eufemismos possam ser empregados como um recurso, nas traduções de palavrões nas legendas de filmes argentinos, com o intuito de amenizar o impacto dos intensificadores que seriam percebidos como chulos ou grosseiros e que poderiam provocar impressões desagradáveis no espectador. Cabe, aqui, reforçar o caráter cultural do que é tido por tabu. Determinadas expressões que seriam recebidas naturalmente pela cultura de origem, pelo fato de serem marcas

de identidade regional, próprias de um grupo de falantes, numa eventual tradução literal, poderiam causar efeitos muito diferentes na cultura de chegada, se comparado ao contexto original. Isto é, o vocábulo ou expressão correspondente, na língua de chegada, não é garantia de resultado do mesmo ato de fala. As situações comunicativas são marcadas culturalmente; assim, dependendo do contexto de situação, o uso de uma expressão tabu ou outra, ou inclusive de sua ausência, estará determinada pelo grau de aceitabilidade da comunidade de recepção.

A esse respeito, no âmbito dos Estudos da Tradução, Hurtado Albir (2008, 25) considera a tradução um “*saber fazer* que consiste em saber percorrer o processo tradutor, sabendo resolver os problemas de tradução que aparecem em cada caso<sup>2</sup>”. A autora destaca:

As traduções são feitas porque as línguas e as culturas são diferentes; a razão da tradução é, portanto, a diferença linguística e cultural. As traduções são feitas para comunicar, para ultrapassar a barreira da incomunicação devido a essa diferença linguística e cultural; a tradução tem, portanto, uma finalidade comunicativa. As traduções são feitas para alguém que não conhece a língua e, geralmente, nem mesmo a cultura em que está formulado um texto (escrito, oral ou audiovisual). O tradutor não traduz para si próprio (exceto em raras ocasiões), traduz para um destinatário que precisa dele, como mediador linguístico e cultural, para ter acesso a um texto<sup>3</sup> (Hurtado Albir, 2008, 28).

- 2 No original: “un *saber hacer* que consiste en saber recorrer el proceso traductor, sabiendo resolver los problemas de traducción que se plantean en cada caso”. Observação: todas as traduções são de nossa autoria.
- 3 No original: “Se traduce porque las lenguas y las culturas son diferentes; la razón de ser de la traducción es, pues, la diferencia lingüística y cultural. Se traduce para comunicar, para traspasar la barrera de incomunicación debida a esa diferencia lingüística y cultural; la traducción tiene, pues, una finalidad comunicativa. Se traduce para alguien que no conoce la lengua, y generalmente tampoco la cultura, en que está formulado un texto (escrito, oral



Sendo assim, a tradução tem lugar enquanto mediação de conhecimento linguístico e histórico-cultural, para uma comunidade que desconhece a língua, a história e a cultura de partida. As escolhas linguísticas feitas durante o processo tradutório interferem diretamente na recepção dos textos traduzidos no contexto de chegada. Newmark (2006) aponta que o objetivo primordial de toda tradução seria conseguir um “efeito equivalente”, ou seja, a produção no leitor da tradução (ou no espectador, no caso do presente trabalho) do mesmo efeito que seria produzido no leitor do original. Para o autor, esse efeito seria o resultado desejado de toda tradução.

Ainda a respeito da linguagem obscena e sua relação social, Preti (1984) assegura que

[...] pertence ao campo dos tabus linguísticos e, por isso, são claras as suas ligações com os estudos sociolinguísticos. Opõe-se à linguagem corrente (e disso o falante guarda consciência), servindo à descarga afetiva, à injúria, quer como índice de coloquialismo, quer como expressão carinhosa, perdida sua conotação injuriosa, em determinadas situações, onde se pretende forçar uma intimidade maior com o ouvinte. Esse último enfoque do vocabulário obsceno vem-se acentuando no contexto histórico moderno e constitui um verdadeiro processo desmistificador do chamado “palavrão” (Preti, 1984, 27).

Já com relação à Fraseologia, entendida como uma disciplina ou subdisciplina da Lexicologia, que se ocupa das combinações de palavras e unidades fraseológicas (Corpas Pastor, 2010), a autora afirma

---

o audiovisual). El traductor no traduce para sí mismo (excepto en raras ocasiones), traduce para un destinatario que necesita de él, como mediador lingüístico y cultural, para acceder a un texto”.

Partimos de uma concepção ampla de Fraseologia, que engloba todas aquelas combinações formadas por pelo menos duas palavras, cujo limite superior se situa na oração composta, caracterizadas por uma alta frequência de ocorrência na língua e de coocorrência de seus elementos integrantes, assim como a institucionalização, a estabilidade, a idiomaticidade e a variação que tais unidades apresentam em diferente grau<sup>4</sup> (Corpas Pastor, 2010, 126).

Corpas Pastor (1996) propõe uma classificação para as unidades fraseológicas (UFs), num primeiro nível de estruturação, em três esferas distintas: colocações, locuções e enunciados fraseológicos. As **colocações** se caracterizam pela frequência de determinadas combinações de palavras, são fixadas pelo uso e possuem algum grau de restrição combinatória. À segunda esfera correspondem as **locuções**, que são fixadas no sistema. Por sua vez, **os enunciados fraseológicos** (parêmsias e fórmulas) constituem enunciados e atos de fala em si mesmos e, além disso, estão presentes na fala formando parte dos aspectos sociais e culturais do falante. As colocações e locuções se diferenciam dos enunciados fraseológicos, pelo fato de não formarem enunciados completos em si mesmos e de não realizarem, portanto, atos de fala, já que precisam da combinação com outros elementos no discurso (Corpas Pastor, 2010). No presente trabalho, nossa atenção estará voltada para os fraseologismos formados a partir do léxico tabu, identificado no corpus de estudo de legendas de filmes argentinos.

---

4 No original: “Partimos de una concepción amplia de la fraseología, que engloba todas aquellas combinaciones formadas por al menos dos palabras y cuyo límite superior se sitúa en la oración compuesta, caracterizadas por una alta frecuencia de aparición en la lengua y de coaparición de sus elementos integrantes, así como la institucionalización, la estabilidad, la idiomaticidad y la variación que dichas unidades presentan en diverso grado.”

Para identificação e posterior análise e descrição dos fraseologismos, tanto pelo viés da abordagem quanto da metodologia, adotamos os princípios norteadores da Linguística de Corpus, numa perspectiva aplicada a pesquisas empírico-descritivas (Novodvorski, 2015; 2017). Berber Sardinha (2004), um dos principais autores responsáveis pela disseminação da Linguística de Corpus no Brasil, apresenta a seguinte definição:

A Linguística de Corpus ocupa-se da coleta e da exploração de *corpora*, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística. Como tal, dedica-se à exploração da linguagem por meio de evidências empíricas, extraídas por computador (Berber Sardinha, 2004, 3).

Partindo desse mesmo estudo, com relação ao suporte para análise do corpus, Parodi (2010, 15) informa que “[...] a LC contribui para o estudo de *corpora* textuais digitais preferentemente de tamanho amplo e com suporte em tecnologias computacionais de variados tipos com ênfase em uma aproximação empírica, baseada em conjuntos de dados reais [...]”<sup>5</sup>. Sendo assim, os programas computacionais para análise lexical são uma ferramenta indispensável para o armazenamento de dados, a descrição e a análise linguística.

De acordo com Novodvorski (2015; 2017), entre alguns benefícios na utilização de computadores para investigação de aspectos da linguagem em uso, para além da contagem de palavras, da identificação das palavras

---

5 No original: “[...] la LC aporta al estudio de corpus textuales digitales preferentemente de tamaño amplio y con soporte en tecnologías computacionales de variada índole, con énfasis en una aproximación empírica, basada en amplios conjuntos de datos reales [...]”.

mais frequentes e do reconhecimento das ocorrências de um termo, podemos apontar a consulta de concordâncias individuais ou em corpus paralelo e a extração de palavras chave para análises específicas e muitos outros recursos. Todas essas funções podem ser realizadas por programas de computador, que conferem rigor e confiabilidade à pesquisa, em que os *corpora* são preparados e sistematizados para análise. Entre outros programas, mencionamos o *ParaConc* (Barlow, 2004), fundamentalmente para análise de concordâncias paralelas entre texto original e sua tradução, e o *Word Smith Tools* (Scott, 2012), pois serão utilizados neste trabalho, como será descrito na próxima seção.

### 3. CORPUS E METODOLOGIA

Como já apontado, o corpus de estudo deste trabalho está formado pelas legendas de cinco filmes argentinos com o ator Ricardo Darín, em espanhol rio-platense, e suas respectivas traduções alternativas ao português brasileiro. O corpus é paralelo, pois está composto pelas legendas originais e traduzidas, e unidirecional, na direção espanhol/português. A escolha dos filmes foi definida pelo caráter dramático das histórias da recorrência da linguagem tabuística. As legendas alternativas foram coletadas na internet, num *site* de livre acesso *on-line*, que está disponível há mais de dez anos, podendo ser acessado através do endereço <<http://legendas.tv/>>. O Quadro abaixo informa o nome dos filmes que compõem o corpus de estudo, os gêneros, os diretores e as datas de lançamento.

**Quadro 1:** Corpus de estudo

<b>Filmes</b>	<b>Gênero</b>	<b>Diretor</b>	<b>Lançamento</b>
<i>Nueve reinas</i>	Drama/Suspense	Fabián Bielinsky	2000
<i>El hijo de la novia</i>	Comédia/Drama	Juan José Campanella	2001
<i>El secreto de sus ojos</i>	Drama/Suspense	Juan José Campanella	2009
<i>Tesis sobre un homicidio</i>	Suspense/Policial	Hernán Goldfrid	2013
<i>Relatos salvajes</i>	Comédia dramática/ Suspense	Damián Szifron	2014

Fonte: elaboração dos autores

A seguir, a Figura 1 mostra a extensão do corpus de estudo com os dados estatísticos gerais e de cada arquivo de legendas em espanhol e português, a saber: o número de itens (*tokens*), palavras totais, e o número de formas (*types*), palavras diferentes. Para a obtenção desses dados, foi utilizada a função *Statistics* da ferramenta *WordList* do programa *WordSmith Tools*® (Scott, 2012) na versão 6.0.

**Top Screenshot (Spanish Corpus):**

N	text file	file size	tokens (running words) in text	tokens used for word list	sum of entries	types (distinct words)	type/token ratio (TTR)	standardised TTR
1	Overall	322.953	54.720	47.477		6.225	13,11	36,92
2	El Hijo de la Novia [ESP].txt	81.434	13.957	12.569		2.422	19,27	37,47
3	El secreto de sus ojos [ESP].txt	70.748	12.255	10.826		2.145	19,81	36,73
4	Nueve Reinas [ESP].txt	71.291	12.031	10.121		1.939	19,16	34,81
5	Relatos Salvajes [ESP].txt	54.847	8.956	7.484		1.894	25,31	37,64
6	Tesis Sobre un Homicidio [ESP].txt	44.633	7.521	6.477		1.657	25,58	39,03

**Bottom Screenshot (Portuguese Corpus):**

N	text file	file size	tokens (running words) in text	tokens used for word list	sum of entries	types (distinct words)	type/token ratio (TTR)	standardised TTR
1	Overall	285.218	46.719	39.387		5.420	13,76	38,38
2	El secreto de sus ojos [PTB].txt	64.635	10.793	9.349		1.919	20,53	38,52
3	El Hijo de la Novia [PTB].txt	63.357	10.417	8.985		2.000	22,26	39,53
4	Nueve Reinas [PTB].txt	64.223	10.432	8.503		1.768	20,79	35,86
5	Relatos Salvajes [PTB].txt	50.790	8.115	6.641		1.751	26,37	38,46
6	Tesis Sobre un Homicidio [PTB].txt	42.213	6.962	5.909		1.599	27,06	40,32

At the bottom of the WordList window, the status bar shows: 6 entries, Row 1, T, S, 39.387.

**Figura 1:** *Extensão do corpus de estudo*

Fonte: elaboração dos autores

A seguir, listamos os procedimentos metodológicos empregados desde a compilação e preparação do corpus até o levantamento, descrição e análise dos dados.

- Compilação, seleção e armazenamento do corpus;
- Conversão dos arquivos das legendas ao formato *txt*;
- Levantamento dos dados gerais do corpus;
- Utilização da ferramenta *WordList* do programa WST para extração de lista de palavras, em ordem de frequência e alfabética;

- e) Identificação e extração do léxico tabu mais frequente com o programa WST;
- f) Lematização das palavras tabu no WST, por meio da lista de palavras em ordem alfabética;
- g) Exame preliminar da produtividade do corpus, através das linhas de concordâncias geradas no WST, para análise dos fraseologismos com o vocábulo *cagar*;
- h) Preparação, limpeza e alinhamento do corpus paralelo espanhol/português no programa *ParaConc*(2004);
- i) Estudo contrastivo do corpus paralelo no *ParaConc* (2004), a partir da busca pelo vocábulo *cagar*;
- j) Identificação dos fraseologismos em torno do vocábulo tabu *cagar*;
- k) Descrição e classificação das ocorrências dos fraseologismos identificados nas concordâncias, por meio de percepção e análise sintático-semântica;
- l) Análises contrastivas com as traduções, a partir dos resultados obtidos nos *corpora* paralelos e da consulta a dicionários de uso;
- m) Análise interpretativa final dos dados.

A partir das etapas metodológicas descritas, identificamos o léxico tabu no corpus de estudo e empreendemos a análise das 56 ocorrências da palavra *cagar*, especificamente para este artigo, como será descrito na próxima seção.

## 4. ANÁLISE DOS DADOS

Por meio da ferramenta *Word List* do programa WST, fizemos a extração da lista de palavras do corpus de legendas originais, primeiro, com o objetivo de identificar e lematizar o léxico tabu, isto é, agrupar as flexões e derivações do léxico identificado. Organizada por ordem de frequência, fizemos a limpeza da lista, no intuito de deixarmos apenas os vocábulos de interesse para a pesquisa. A Figura 2 exibe uma vista parcial dos resultados.

N	Word	Freq	%	Tests	%	Lemmas	Set
1	PUTA	76	0,14	5	100,00		
2	BOLUDO	56	0,10	0		boludo[47] bol[1] boluda[1] boludeces[1] boludito[1] boludos[5]	
3	<b>CAGAR</b>	56	0,10	0		cagar[10] cagara[1] cagarme[1] cagaron[2] cagáste[3] cagazo[1] cago[8] cago[3] cagón[3]	
4	PELOTUDO	55	0,10	0		pelotudo[43] pelotuda[2] pelotudeces[5] pelotudito[1] pelotudos[4]	
5	MIERDA	49	0,09	4	80,00		
6	JODER	35	0,06	0		joder[11] joda[4] jodan[2] jodas[3] jode[3] joderme[2] jodete[1] jodido[1] jodidos[2] jodiendo	
7	CARAJA	23	0,04	4	80,00		
8	CULO	14	0,03	4	80,00		
9	PELOTA	14	0,03	1	20,00	pelota[1] pelotas[13]	
10	PEDO	10	0,02	3	60,00		
11	QUILOMBO	9	0,02	3	60,00		
12	CONCHA	8	0,01	0		concha[8] conchuda[1] conchudo[1]	
13	PENDEJO	7	0,01	0		pendejo[3] pendejas[1] pendeja[3]	
14	COGI	6	0,01	1	20,00	cogi[3] cogiste[1] cogieras[2]	
15	MEAR	4		1	20,00	meiar[1] meando[3]	
16	ORTO	4		3	60,00		

frequency | alphabetical | statistics | filenames | notes  
5,983 entries Row 3 CAGAR T S CAGAR

**Figura 2:** Léxico tabu no corpus de estudo

Fonte: elaboração dos autores

Uma rápida consulta ao *Diccionario coloquial de los argentinos* (Governatori; Larocca, 2014) já permite apreciar a criatividade no uso do vocábulo *cagar* nesse país. Apenas em expressões iniciadas por esse vocábulo, o dicionário oferece 29 entradas, com acepções variadas, muitas delas com o traço semântico de intensificação, como em *cagar a + goles/palos/pedos/trompadas* (respectivamente, *ganhar um time por muitos gols* no futebol, *dar uma paulada*,



*xingar muito e encher de porrada*). O *Diccionario de Uso del Español* (Moliner, 2008) oferece um quantitativo bastante menor de acepções, se comparado ao outro dicionário. Resumidamente, aponta que se trata de uma expressão de uso vulgar, própria de exclamações grosseiras de irritação, contrariedade, por algo ou contra alguém, ou em blasfêmias. A próxima figura ilustra uma vista parcial da ferramenta *Concord* do WST, em que se apreciam alguns dos fraseologismos formados com as flexões e derivações de *cagar* no corpus.

The screenshot shows the Concord software interface with a concordance table. The table has columns for line number, text, and various statistical fields. The text column shows concordance lines for the word 'cagar' and its derivatives. The statistical fields include Set, Tag, Word #, Sent. #, Para. Pos., Para. Hear, and Sect. Sect.

N	Concordance	Set	Tag	Word #	Sent. #	Para. Pos.	Para. Hear	Sect. Sect
1	935 ¡Estoy bien, sí, estoy bien! <b>Qué cagada</b> . 936 Pobre viejo... ahora	CAGADA		9.756	1.577	100%	0	70' El F
2	qué 805 ¿Qué te pasa? Ay, nada, <b>esta cagada</b> , la puse en vibrador, 806 casi	CAGADA		8.357	1.370	33%	0	60' El F
3	a 819 armar algo. Bah, me terminó <b>de cagar</b> el matrimonio, pero me fue bien	CAGAR		6.493	1.032	100%	0	10' El F
4	, 1194 eso es parte de la solución. <b>La cagada</b> es que no le dicen qué parte	CAGADA		12.32	1.918	27%	0	88' El F
5	unos años, pero... nada. 611 No, la <b>cagada</b> fue cuando largué abogacía,	CAGADA		6.388	1.024	33%	0	48' El F
6	yanqui. 1017 Así parece que <b>me cagó</b> mi abuelo. ¿Y no te dijeron nada	CAGÓ		10.571	1.673	78%	0	76' El F
7	a que te acrediten la culpa porque <b>te cago</b> a trompadas!!! 95 ¡Pero pará un	CAGO		990	153	86%	0	7% El F
8	la gaseosa. ¿Te mandaste <b>alguna cagada</b> ? 609 No. Las normales. 610	CAGADA		6.372	1.020	100%	0	46' El F
9	. Con queso crema es una <b>reverenda cagada</b> . 32 El Mascarpone es muy	CAGADA		277	49	100%	0	2% El F
10	, un juicio 1052 y vos me <b>sacaste cagando</b> . 1053 Pero eso fue hace	CAGANDO		9.746	1.197	100%	0	80' El S
11	afuera! 418 A ver, si viene la <b>vieja, cagamos</b> . No va venir la vieja, relájate	CAGAMOS		4.031	496	100%	0	10' El S
12	con Irene. 613 capaz que me <b>saca cagando</b> , perdón, capaz que me	CAGANDO		5.820	680	61%	0	47' El S
13	quiere 666 y como decís vos, me <b>sigo cagando</b> la vida viniendo a tugorios	CAGANDO		6.319	734	77%	0	52' El S
14	no me oyó... 465 ¡sino que usted <b>se caga</b> en la orden que yo le di! 466	CAGA		4.466	546	100%	0	10' El S
15	que el Mascal Espósito. 112 lo va <b>a cagar</b> a trompadas. 113 Bien. ¿A	CAGAR		990	128	100%	0	10' El S
16	que es un genio. 140 se manda <b>mil cagadas</b> y uno tiene que andar detrás	CAGADAS		1.308	159	100%	0	10' El S
17	nunca me lo arreglaron. 1056 ¡ <b>Que cagada!</b> No te puedes quedar a dormir	CAGADA		9.774	1.200	100%	0	80' El S

**Figura 3:** Linhas de concordância no *Concord* com *cagar* lematizado

Fonte: elaboração dos autores

Na figura anterior, é possível observar algumas expressões com sentido de contrariedade (linhas 1, 2, 4, 5 e 17), ou de prejudicar (linhas 3 e 6), e outras como *cagar a trompadas* (linhas 7 e 15) e *sacar cagando* (linhas 10 e 12), com sentido de bater muito em alguém no rosto e de mandar alguém embora de modo ostensivo, respectivamente. Na próxima figura, uma vista

parcial do *ParaConc* (Barlow, 2004) ilustra as concordâncias paralelas, para análise contrastiva com as traduções.

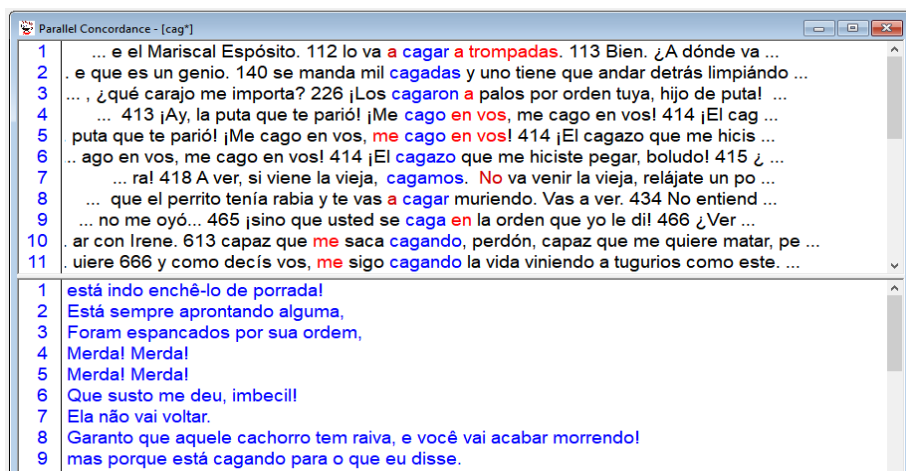


Figura 4: Linhas de concordâncias paralelas no *ParaConc* com *cagar* lematizado

Fonte: elaboração dos autores

Na figura anterior, tomando apenas as três primeiras linhas de concordâncias paralelas, podemos observar tanto as ocorrências no corpus em espanhol como as respectivas traduções para o português: *lo va a cagar a trompadas* (está indo enchê-lo de porrada), *se manda mil cagadas* (está sempre aprontando alguma) e *Los cagaron a palos* (foram espancados). Isto é, com o corpus alinhado no *ParaConc*, foi possível obter as concordâncias paralelas pelo item de busca, para análise contrastiva das ocorrências. A seguir, e por questão de espaço, analisamos uma seleção representativa das 56 ocorrências identificadas no corpus de estudo.

**Quadro 2:** *cagar(se) a trompadas/patadas/palos(a/con alguien)*

1	A menos que quieras que <b>te caguen a trompadas</b> . ( <i>Nueve reinas</i> )	A menos que queira que te caguem na cabeça.
2	<b>...cagarme a trompadas</b> si alguien me hincha las pelotas. Me gusta. ( <i>El secreto de sus ojos</i> )	...brigar com alguém que me sacaneie, eu gosto disso!
3	Espero que te acuerdes de la cantidad de veces que <b>me dejé cagar a trompadas</b> para hacerte feliz. ( <i>Tesis sobre un homicidio</i> )	Só espero que se lembre quantas vezes dei um jeito para te deixar feliz.
4	¡Andá a que te acrediten la guita porque <b>te cago a trompadas</b> ! ( <i>El hijo de la novia</i> )	Vá cuidar do depósito, agora mesmo!
5	Decile que el Mariscal Espósito <b>lo va a cagar a trompadas</b> . ( <i>El secreto de sus ojos</i> )	Diga-lhe que o “Marechal” Espósito está indo enchê-lo de porrada!
6	Si fuera otro momento salgo y <b>te cago a patadas</b> . ( <i>Nueve reinas</i> )	Se fosse outro momento, te encheria de porrada.
7	<b>¡Los cagaron a palos</b> por orden tuya, hijo de puta! ( <i>El secreto de sus ojos</i> )	Foram espancados por sua ordem, filho da puta!

Fonte: elaboração dos autores

As ocorrências registradas no quadro anterior reúnem os usos de *cagar* como intensificador, em estruturas que denotam agressão física, formadas por i) *cagar a + parte do corpo humano alvo da agressão* (exemplos 1 a 5); por ii) *cagar a + parte do corpo humano utilizada para bater* (6); ou por iii) *cagar a + instrumento da agressão* (7). Na linguagem coloquial, em espanhol rio-platense

e pelo contexto de uso, *trompa* corresponde a *nariz*; assim, *trompada* seria a ação de ‘socar o nariz, a cara’. Também *pata* é o nome dado em linguagem informal aos ‘pés’ e *patada* ao ‘chute’. Desse modo, pelas características de intensificação, *cagar a trompadas* corresponde a *encher de socos no nariz, no rosto*; já *cagar a patadas*, a *encher de chutes*. Por sua vez, *palos* corresponde a ‘paus’; consequentemente, *cagar a palos* equivaleria a *dar uma paulada*.

Considerando as traduções, em (1) temos “A menos que queira que **te caguem na cabeça**”. Apesar da opção pela manutenção do tabu, a compreensão do fraseologismo *cagar a trompadas* não foi adequada, pois a tradução literal de *cagar* e de *trompadas* por *cabeça* originou um sentido muito discrepante do original. Em *cagarme a trompadas* (2), o uso do dativo de interesse *me* denota a ênfase posta na capacidade de *cagarse a trompadas con alguien*, bem interpretada na tradução *brigar com alguém*. Também em (5) a tradução é acertada, pelo uso do tabu em português, *está indo enchê-lo de porrada*, para *lo va a cagar a trompadas*.

Por outro lado, em (3) e (4), observa-se, além da omissão do tabuísmo, a não tradução dos fraseologismos. Só em (3), *dei um jeito* é utilizada como tradução para *me dejé cagar a trompadas*; a opção poderia ser simplesmente *apanhei* ou *leveí uma surraluma porrada*. Por sua vez, *te encheria de porrada* (6) é uma adaptação de *te cago a patadas*, que mantém o intensificador e a informalidade do original. Em (7), *foram espancados* é a tradução para *los cagaron a palos*; outra opção poderia ser *levaram uma paulada*. O próximo quadro registra ocorrências de *cagar* com sentido de ‘prejudicar alguém ou de prejudicar-se’.

**Quadro 4.** *estar/venir/seguir cagando (a alguien ou a sí mismo)*

8	Oíme, yo <b>no estoy cagando a nadie</b> . ( <i>Nueve reinas</i> )	Não estou sacaneando ninguém.
9	<b>Me estás cagando</b> en algo, no sé. ( <i>Nueve reinas</i> )	Está me enganando. Em algo, não sei.
10	¡Hace un año que <b>me venís cagando</b> con la misma historia! ( <i>Nueve reinas</i> )	Um ano me enganando com a mesma história!
11	...y como decís vos, <b>me sigo cagando</b> la vida viniendo a tugurios como este. ( <i>El secreto de sus ojos</i> )	...e como diz você, continuo estragando minha vida vindo em lugares como este.

Fonte: elaboração dos autores

Nos quatro fragmentos anteriores, a estrutura fraseológica está formada pelo léxico tabu como base, em perífrase de gerúndio com os verbos auxiliares *estar*, *venir* e *seguir*. O sentido é de prejudicar alguém ou a si próprio, denotando uma ação continuada no tempo presente (8 a 11) e iniciada em algum ponto do passado (10 e 11). Nas traduções, foram acertadas as escolhas lexicais por *sacanear*, *enganar* e *estragar* (11), em perífrases também de gerúndio, com o auxiliar *estar* ou sem (10) e com o auxiliar *continuar* (11), para manutenção da equivalência semântica de prejudicar alguém ou prejudicar-se.

Em (8), “Não estou sacaneando ninguém”, além da manutenção da estrutura perifrástica presente no original, também o registro informal é utilizado na tradução. *Sacanear* é um tabuísmo, conforme o dicionário

Houaiss (2009): “agir como sacana (‘devasso’, ‘espertalhão’, ‘trocista’)”. Nos outros fragmentos (9 a 11), os usos de *enganar* e *estragar*, apesar de não denotarem a informalidade presente nas legendas originais, possuem sentidos equivalentes aos do espanhol. O próximo quadro ilustra outras ocorrências com *cagar*.

**Quadro 5.** *sacar cagando*

12	...capaz que <b>me saca cagando</b> , perdón, capaz que me quiere matar... ( <i>El secreto de sus ojos</i> )	Talvez ela vá me mandar à merda, desculpe, talvez ela vá querer me matar...
13	Yo quería hacerles quilombo, un juicio y vos <b>me sacaste cagando</b> . ( <i>El secreto de sus ojos</i> )	Eu lhe disse mil vezes, Benjamín, que tinha que fazer uma queixa, e você me mandava aos diabos!

Fonte: elaboração dos autores

Em ambos os fragmentos, o fraseologismo *sacar cagando* remete a expulsar, mandar alguém embora, de maneira bruta, violenta (Governatori; Larocca, 2014). A expressão tem o sentido de mandar alguém às favas, mandar cagar no mato. Na tradução temos, respectivamente, *mandar à merda* e *mandar aos diabos*, também com características de informalidade e sentido chulo no ato de desprezar alguém. Entendemos que essas escolhas tradutórias caracterizam adaptações à cultura de chegada, de acordo com os usos locais, e que têm a aceção correspondente às dos fragmentos em espanhol. Para finalizar, o próximo quadro recopila as últimas ocorrências com o léxico tabu que analisaremos neste artigo.

**Quadro 6.** Outras ocorrências com *cagar*

14	¡Mira! ¡Yo <b>me cago en los catálogos!</b> ( <i>Nueve reinas</i> )	Estou cagando pros catálogos!
15	¡Ay, la puta que te parió! ¡ <b>Me cago en vos</b> , me cago en vos! ( <i>El secreto de sus ojos</i> )	Ah, puta que pariu! Merda! Merda!
16	Pero que <b>me cago golpeando</b> siempre. ( <i>El secreto de sus ojos</i> )	Estou sempre esbarrando.
17	¡Aaay! ¡Qué hacés, boludo, ¿ <b>me querés cagar de otro infarto?</b> ! ( <i>El hijo de la novia</i> )	O que está fazendo? Quer que eu tenha outro enfarte?
18	<b>...se manda mil cagadas</b> y uno tiene que andar detrás limpiándoles el culo. ( <i>El secreto de sus ojos</i> )	Está sempre aprontando alguma, e alguém tem que limpar sua merda.

Fonte: elaboração dos autores

Os dois primeiros fragmentos do quadro anterior caracterizam ocorrências dos fraseologismos por colocação *cagarse en algo* (14) ou *cagarse en alguien* (15), que correspondem em português a *não dar a mínima* (importância) *para algo* ou *para alguém*. Como se observa em ambas as traduções desses fragmentos, o tabuísmo está presente. Em (14), a escolha lexical pelo fraseologismo *estou cagando pros catálogos* denota tanto a manutenção do sentido quanto do registro informal, em comparação com o original. Já em (15) observamos alteração no sentido, uma vez que seria viável a opção *não dou a mínima pra você* ou *estou cagando pra você*.

O fraseologismo de (16) está formado pela colocação *cagarse + gerúndio* e expressa uma ação ocorrida de modo acidental, *me cago golpeando siempre*,

cujo sentido foi bem captado pela tradução em *estou sempre esbarrando*, embora não tenha sido denotado o registro vulgar do original. Em (17), a estrutura fraseológica por colocação (*querer*) *cagar a alguien de algo* equivale a *querer perjudicar* ou *destruir alguém com/por alguma coisa*. Assim, a tradução *quer que eu tenha outro enfarte?* traduz a ideia que é exatamente essa em *¿me querés cagar de otro infarto?* A diferença reside no uso do registro vulgar em espanhol, em que *querer cagar* levaria à compreensão de *você quer que eu morra* ou *que eu acabe tendo outro enfarte?*

A última ocorrência (18) analisada neste artigo corresponde a um fraseologismo presente no filme *El secreto de sus ojos*. A colocação *mandarse una cagada* ou, como no fragmento selecionado, *se manda mil cagadas*, são de uso recorrente no espanhol rio-platense. A expressão tabu tem o sentido de fazer várias (mil) coisas erradas, ou seja, cometer vários erros. Podemos comprovar esse significado no *Diccionario de Uso del Español* (Moliner, 2008), que define *cagada* como “Error grave, acción muy desafortunada” ou no *Diccionario coloquial de los argentinos* (Governatori; Larocca, 2014), que define exatamente a colocação *mandarse una cagada* como “cometer un error grave”. A tradução *está sempre aprontando alguma*, embora isoladamente não expresse o léxico tabu por meio de uma expressão chula equivalente, na continuação da frase observamos que o registro é mantido. O fragmento *e alguém tem que limpar sua merda* complementa o sentido, como tradução de *y uno tiene que andar detrás limpiándoles el culo* (literalmente, *e a gente tem que ir atrás limpando a bunda*) dá a entender que *aprontar alguma* corresponderia a *fazer uma cagada*, e que depois



seria necessário alguém para limpá-la. A seguir, algumas ponderações a respeito deste trabalho.

## 5. CONSIDERAÇÕES FINAIS

Ao finalizar esta pesquisa, é possível perceber a importância da Linguística de Corpus na análise das legendas originais e traduzidas dos filmes argentinos, a partir do objetivo de identificar o léxico tabu e seus fraseologismos, alinhado ao desenvolvimento da subcompetência fraseológica, que exige uma percepção fina no jogo das significações. Mediante as análises sintático-semânticas realizadas com o auxílio das ferramentas computacionais, percebemos a complexidade e a importância da interpretação adequada dos vocábulos analisados separadamente e nos agrupamentos lexicais, para o sucesso da tradução. A produtividade do uso do vocábulo *cagar* no espanhol rio-platense ficou em evidência, por meio da análise feita neste artigo, ainda mais se considerado que se trata de um corpus de pequena extensão. A palavra *cagar* denota significados variados, com conotações diversas, e pode ser utilizada para expressar diferentes situações e sentimentos, ofensivos ou não. A consideração das questões culturais e sociais, por outro lado, também fica em evidência, uma vez que o ato de traduzir se configura no estabelecimento de uma mediação com a comunidade de chegada, pela utilização de expressões equivalentes, que poderão amenizar o tom do texto original ou omitir determinado vocábulo ou fraseologismo tabu, conforme a cultura na

língua de chegada e os valores dessa sociedade. Isto é, na tradução se faz presente e necessário esse sentido de percepção aguda para a compreensão e reexpressão dos sentidos, em que joga um papel preponderante a subcompetência fraseológica.

Esperamos, por meio deste trabalho, ter trazido alguma contribuição para os estudantes na área de tradução, pela ampliação dos conhecimentos em fraseologia contrastiva em espanhol/português, e em torno do léxico tabu, uma área ainda pouco explorada, principalmente neste par linguístico.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Barlow, Michael. *ParaConc*, 1.0 (Build 269). Parallel Concordance Software. Houston, USA: Programming, ELF, Ltd., 2004.
- Berber Sardinha, Tony. *Linguística de Corpus*. Barueri, SP: Manole, 2004.
- Corpas Pastor, Gloria. *Diez años de investigación en fraseología: Análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Iberoamericana, 2010.
- Corpas Pastor, Gloria. *Manual de fraseología española*. Madrid: Gredos, 1996.
- El hijo de la novia*. Direção: Juan José Campanella. Produção: Adrián Suar. Argentina, 2001.
- El secreto de sus ojos*. Direção: Juan José Campanella. Produção: Mariela Besuievski, Juan José Campanella, Carolina Urbieto. Argentina, 2009.
- Governatori, Gerardo; Larocca, Roberto. *¿Qué lo parió, che! Diccionario coloquial de los argentinos*. Buenos Aires: Continente, 2014.
- Guedelha, Carlos Antônio Magalhães. “Tabus linguísticos como motivação na formação de palavras do PB”. In: *Working Papers em Linguística*, v. 12, n. 2,

- p.49-68. 2011. Disponível em: <<https://periodicos.ufsc.br/index.php/workingpapers/article/viewFile/19848420.2011v12n2p49/21337>>. Acesso em: 16 dez. 2017.
- Guérios, Mansur. *Tabus Linguísticos*. Curitiba: Editora “Organização Simões”, 1956.
- Houaiss, A. *Dicionário eletrônico Houaiss da língua portuguesa*. Versão 3.0, 2009.
- Hurtado Albir, Amparo. *Traducción y Traductología: Introducción a la Traductología*. 4. ed. Madrid: Cátedra, 2008 (2001).
- Legendas.TV*. Disponível em: <<http://legendas.tv/>>. Acesso em: 17 set. de 2018.
- Moliner, María. *Diccionario de uso del español*. Edición electrónica, versión, 3.0. Madrid: Editorial Gredos, S.A.U., 2008.
- Monteiro, José Lemos. “As palavras proibidas”. In: *Revista de Letras*, Fortaleza, n. 2, v. 11, p. 11-22, 1986. Disponível em: <[http://www.revistadeletras.ufc.br/revista%20vol.11,%20n.2\\_artigos.pdf](http://www.revistadeletras.ufc.br/revista%20vol.11,%20n.2_artigos.pdf)>. Acesso em: 16 dez. 2017.
- Newmark, Peter. “Métodos de traducción”. In: Newmark, Peter. *Manual de traducción*. Trad. de Virgilio Moya. 5. ed. Madrid: Cátedra, 2006. p. 69-80.
- Novodvorski, Ariel. “Estudo de fraseologia contrastiva em corpus paralelo de filmes argentinos”. In: Zavaglia, Cláudia; Simão, Angélica Karim Garcia. *Reflexões, tendências e novos rumos dos estudos fraseoparemiológicos*. São José do Rio Preto: UNESP/IBILCE, 2017. p. 72-87.
- Novodvorski, Ariel. *Linguística de Corpus aplicada a pesquisas empírico-descritivas: tradução, interlíngua, fraseologia, discurso referido e transitividade em contraste*. 2015. 30 f. Projeto de pesquisa – Instituto de Letras e Linguística, Universidade Federal de Uberlândia, Uberlândia, 2015.
- Nueve reinas*. Direção: Fabián Bielinsky. Produção: Cecilia Bossi, Pablo Bossi. Argentina, 2000.
- Orsi, Vivian. “Tabu e preconceito linguístico”. In: *ReVEL*, v.9, n. 17, 2011, p. 334-348. Disponível em: <[http://www.revel.inf.br/files/artigos/revel\\_17\\_tabu\\_e\\_preconceito\\_linguistico.pdf](http://www.revel.inf.br/files/artigos/revel_17_tabu_e_preconceito_linguistico.pdf)>. Acesso em: 19 dez. 2017.

Parodi, Giovanni. *Lingüística de Corpus*: de la teoría a la empiria. Madrid: Iberoamericana, 2010.

Preti, Dino. *A gíria e outros temas*. São Paulo: T. A. Queiroz, 1984.

Preti, Dino. *A linguagem proibida*: um estudo sobre a linguagem erótica. São Paulo: Queiróz, 1984.

*Relatos Salvajes*. Direção: Damián Szifron. Produção: Hugo Sigman, Agustín Almodóvar, Pedro Almodóvar. Argentina, 2014.

Scott, Mike. *Word Smith Tools*, version 6.0. Liverpool: Lexical Analysis Software, 2012.

*Tesis sobre un homicidio*. Direção: Hernán Goldfrid. Produção: Gerardo Herrero, Diego Dubcovsky. Argentina, 2013.

Zavaglia, Adriana.; Welker, Herbert. “O que é Léxico”. [2008] 2013. In: *GTLEX. ANPOLL*. Disponível em: <[www.letras.ufmg.br/gtlexNovo](http://www.letras.ufmg.br/gtlexNovo)>. Acesso em: 14 dez. 2017.