

## EXPRESSÕES IDIOMÁTICAS EM TEXTOS JORNALÍSTICOS: UM ESTUDO PRELIMINAR COM BASE EM DADOS CULTURAIS E ESPORTIVOS

Marina Luz<sup>1</sup>; Elizabete Aparecida Marques<sup>2</sup>

**Resumo:** As unidades fraseológicas (CORPAS PASTOR, 1996), combinações indissociáveis de unidades léxicas de uma língua, podem se classificar em várias categorias, dentre as quais estão as expressões idiomáticas – fraseologismos presentes na comunicação rotineira que, dentre outras características, possuem sentido figurado. Nesse viés, partindo dos pressupostos da Fraseologia, este trabalho visa a apresentar e discutir os resultados do levantamento de expressões idiomáticas extraídas de um corpus de textos jornalísticos brasileiros da versão on-line do jornal Folha de São Paulo. Para atingir tal objetivo, a metodologia da pesquisa seguiu procedimentos manuais de armazenamento diário de textos das seções de esporte e cultura e, posteriormente, seleção e tabelamento dos candidatos à expressão idiomática. Considerando o recorte do corpus estudado neste trabalho, os resultados da pesquisa mostram que a incidência de expressões foi baixa, reforçando a hipótese de que o uso dessas unidades lexicais está associado às características dos textos pesquisados. As expressões idiomáticas inventariadas ocorreram com maior frequência nos textos de temática descontraída, como, por exemplo, as colunas culturais humorísticas e as reportagens esportivas que trouxeram discursos espontâneos dos participantes envolvidos na ação verbal. Espera-se que a pesquisa contribua com os estudos fraseológicos e aponte novas perspectivas para estudos futuros.

**Palavras-chave:** Fraseologia. Expressões idiomáticas. Linguística de Corpus.

### **IDIOMATIC EXPRESSIONS IN JOURNALISTIC TEXTS: A PRELIMINARY STUDY BASED IN CULTURAL AND SPORTS DATA**

**Abstract:** *The phraseology units (CORPAS PASTOR, 1996), inseparable combinations of language lexical units, can be classified in many categories, including the idiomatic expressions – phraseologies present in the routine communication that, among other characteristics, present figurative language. In this sense, based on phraseology assumptions, this paper aims to present and discuss the research results of idiomatic expressions extracted from Brazilian journalistic texts corpus contained in the digital version of the newspaper Folha de São Paulo. In order to reach this goal, the research methodology attended to manual procedures of daily text storage from the sport and culture sections, and then the selection and tabulation*

---

<sup>1</sup> Mestranda do Programa de Pós-graduação em Estudos de Linguagens – FAALC/UFMS. ORCID ID: 0000-0002-0599-9058.

<sup>2</sup> Doutora em Linguística Aplicada pela Universidad de Alcalá de Henares (Espanha, 2007), docente do Programa de Pós-graduação em Estudos de Linguagens – FAALC/UFMS, Chefe da Divisão Editora UFMS e vice coordenadora do Grupo de Trabalho em Lexicologia, Lexicografia, Terminologia e Terminografia (GTLEX) da Associação Nacional de Pós-Graduação e Pesquisa em Linguística e Letras (ANPOLL). ORCID ID: 0000-0002-6308-9597.

*of the candidates to idiomatic expressions. Considering the corpus sample studied in this paper, the results show that expressions incidence was low. This result reinforces that the occurrence of this group of phraseology units is directly associated to the searched texts characteristics since the idiomatic expressions found appeared more often in informal texts as humorous cultural columns and sportive interviews that brought spontaneous speeches of the involved participants in the verbal action. This research is expected to contribute to phraseological studies and provide new perspectives for further studies.*

**Keywords:** *Phraseology. Idiomatic Expressions. Corpus Linguistics.*

## **Introdução**

Como em outras línguas existentes, os falantes da Língua Portuguesa só conseguem obter comunicação linguística graças ao acoplamento das palavras. É desse processo que nasce a possibilidade de expressar ideias e o processo de interação social.

Do ponto de vista lexical, essa união pode ser formada por unidades léxicas simples ou combinações fixas, nos termos de Biderman (1999), lexias simples e complexas, sendo as simples ou livres aquelas que aceitam qualquer tipo de variação dos elementos das combinações, e as fixas ou complexas as que tem algum tipo de restrição na seleção dos constituintes. Segundo Corpas Pastor (1996), essas combinações restritas são chamadas de unidades fraseológicas e são estudadas pela Fraseologia, disciplina que faz parte da Linguística.

Segundo sua natureza, as unidades fraseológicas podem ser distribuídas, conforme Corpas Pastor (1996), em três subdivisões: colocações; enunciados fraseológicos (citações, parênteses e fórmulas de rotina); e locuções. É nesse último grupo que se encontra a expressão idiomática, foco deste trabalho, que Xatara (1998, p.149) define como “uma lexia complexa indecomponível, conotativa e cristalizada em um idioma pela tradição cultural”.

Aprofundando o que já foi previamente comentado, a citação acima se refere aquele grupo de palavras que, geralmente, não podem ser separadas nem terem posição alterada em sua ordem sequencial e, em bloco indissociável, têm sentido figurado, isto é, sentido diferente daquele analisado literalmente. Esses agrupamentos são consolidados pela comunidade que os usa. Um bom exemplo de expressão idiomática é *bater as botas*, porque é totalmente cristalizada na sociedade brasileira, é composta por mais de um elemento lexical, não pode sofrer alteração na ordem das lexias nem substituição (*\*botas bater, \*bater o chinelo, \*tocar as botas*), e não significa literalmente o movimento de quando botas são batidas, mas sim “morrer”.

Tendo como base esses preceitos, pretende-se, neste artigo, apresentar e discutir os resultados de uma pesquisa de iniciação científica realizada nos anos de 2015 e 2016. O trabalho consistiu em elaborar um inventário de unidades fraseológicas, mais especificamente expressões idiomáticas, extraídas de textos jornalísticos, de diferentes gêneros, do acervo de notícias da versão on-line do jornal *Folha de São Paulo*. Considerando que a vastidão do campo da Fraseologia quanto às possibilidades de pesquisa, espera-se que a abordagem adotada neste trabalho traga contribuições para os estudos fraseológicos, somando-se às investigações realizadas a partir de *corpora* de textos jornalísticos.

## 1 Pressupostos teóricos

A Fraseologia é concebida, neste trabalho, como o ramo da Linguística que se ocupa do estudo das unidades fraseológicas. Essa concepção, por sua vez, está envolta por distintas definições, de modo que se mostra necessário delimitar que os fundamentos desta pesquisa passam pela conceituação de Corpas Pastor (1996) ao entender que “as unidades fraseológicas, objeto de estudo da fraseologia, são unidades léxicas formadas por mais de duas palavras gráficas em seu limite inferior, cujo limite superior se situa no nível da oração composta”.

Como já mencionado na Introdução, os fraseologismos se agrupam em diferentes categorias, segundo sua natureza. Corpas Pastor (1996), utilizando como premissas os critérios do enunciado e o da fixidez na norma, no sistema ou na fala, agrupa os fraseologismos em três esferas:

- a. as colocações: engloba unidades fixas na norma e correspondem a sintagmas livres que, devido ao uso, adquiriram certo grau de fixidez. São exemplos de colocações construções como *colocar em funcionamento*, *dar um passeio*, *compor uma música*, *lançar um míssil*, *discutir calorosamente* etc.;
- b. as locuções: correspondem as unidades fixas no sistema e, geralmente, apresentam um alto grau de idiomaticidade: *abotoar o paletó*, *bater as botas*, *cabeça de vento*, *matar cachorro a grito*, *na maciota*, *rodar a baiana*, *virar/estar uma onça*; etc.;
- c. enunciados fraseológicos: englobam unidades que constituem enunciados por si mesmos, são fixos na fala e fazem parte do acervo

cultural do falante de uma determinada cultura. São exemplos de enunciados fraseológicos os provérbios, os ditos populares, as citações, as fórmulas conversacionais e de cortesia, os slogans.

A autora salienta que as duas primeiras classes não constituem enunciados nem atos de fala completos, pois equivalem a sintagmas que necessitam combinar-se com outras unidades da língua para se constituírem como ato de comunicação, enquanto a terceira categoria corresponde a enunciados que não dependem de outras unidades linguísticas para se atualizarem na fala para formar um ato de comunicação. Os elementos dessa última categoria dependem fortemente de fatores pragmáticos e contextuais. O objeto deste estudo são as locuções idiomáticas que, no Brasil, são conhecidas como expressões idiomáticas.

## 2 Metodologia

Fundado em 1921, o jornal *Folha de São Paulo* é, atualmente, o jornal de maior circulação<sup>3</sup> no território brasileiro editado na cidade de São Paulo – SP. Sua versão on-line foi fornecida pela primeira vez em 1995 e, no ano da pesquisa aqui relatada, contava com 36 seções atualizadas diariamente com conteúdos sobre assuntos que variam desde política até tecnologia.

Como efeito da grande extensão do acervo de notícias on-line do jornal, foi necessário delimitar as seções para a composição do *corpus* de estudo. Para tal escolha, o primeiro critério foi pensar no conteúdo e na linguagem que seriam encontrados nas diferentes seções do jornal. Pensando, então, nas características já comentadas das expressões idiomáticas, e entendendo que textos descontraídos e informais facilitariam o aparecimento delas, preconizou-se as reportagens de esporte e as colunas culturais.

Objetivando a maior variedade possível de colunas culturais, realizou-se um levantamento de informações pertinentes sobre a frequência de publicação de cada coluna disponível. Assim os seguintes autores foram selecionados:

- a. José Simão: Jornalista que na época utilizava sua coluna, publicada de terça à domingo, para fazer comentários satíricos sobre assuntos cotidianos que surgiam nos meios de comunicação. Dentre eles estava a situação política, econômica e o comportamento de personalidades.

---

<sup>3</sup> De acordo com pesquisa realizada pelo Instituto Verificador de Comunicação (IVC) em junho de 2020.

- b. Mônica Bergamo: Jornalista que relatava diariamente, na maioria das vezes de forma imparcial, os acontecimentos políticos e de entretenimento mais relevantes para o dia.
- c. Contardo Calligaris: Psicanalista e escritor italiano que em sua coluna, publicada toda quinta-feira, refletia sobre o comportamento humano nas questões culturais e modernas.
- d. Vladimir Safatle: Professor de filosofia que toda sexta-feira escrevia em sua coluna textos sobre política e o comportamento da sociedade.
- e. Samy Dana: Ph.D em business, o economista publicava todo sábado reflexões sobre negócios. Em meados de outubro do ano de 2015 sua coluna teve que ser substituída em nossa pesquisa porque ele se tornou ex-colunista. O estudo prosseguiu com Maurício Meireles.
- f. Maurício Meireles: Jornalista especializado em literatura que publicava todo sábado acontecimentos literários, em especial questões sobre mercado editorial.
- g. Luciana Coelho: Jornalista que trazia em sua coluna de domingo opiniões sobre séries de televisão.
- h. Gregorio Duvivier: Ator e escritor que toda segunda publicava reflexões sobre a sociedade e os assuntos em alta.
- i. Luiz Felipe Pondé: Filósofo e escritor que escrevia toda segunda sobre questões de comportamento, religião e ciência.

Já no que se refere ao esporte, existem períodos de campeonatos das mais distintas modalidades, o que implica em um alto índice de reportagens de determinado esporte em um período do ano e sua posterior ausência por muitos meses. Dessa maneira, ficou determinado que as reportagens esportivas seriam coletadas por ordem aleatória, visto que, nesse caso específico, tentar provocar uma variedade de assuntos seria inaplicável.

A partir do estabelecimento desses critérios, a metodologia da pesquisa consistiu em: formação do *corpus* textual a partir das seções de esporte e as colunas culturais da versão on-line do jornal *Folha de São Paulo*; análise dos textos com o intuito de realizar a identificação, seleção e extração dos candidatos a expressões idiomáticas; constituição do inventário de dados e tabelamento das informações recolhidas; e análise dos dados à luz da teoria fraseológica.

Todos os procedimentos metodológicos foram fundamentados no arcabouço teórico da Fraseologia, em estudos de autores como Biderman (1999), Corpas Pastor (1996), Montoro Del Arco (2006), Tagnin (2011), Tristá (1988) e Xatara (1998). Já o procedimento de constituição do *corpus* foi realizado por sistema de coleta manual que compreende a ação diária de entrar no site do jornal em questão e recolher uma reportagem esportiva e dois textos de colunas culturais. É válido ressaltar que *softwares* específicos atualmente utilizados em pesquisa lexicais foram dispensados por, no momento em que se realizou a pesquisa, não reconhecerem satisfatoriamente as lexias complexas. Nesse sentido, a pesquisadora precisou analisar cada contexto para confirmar ou refutar os dados e, em caso de confirmação, realizar a seleção manual da expressão idiomática.

Os textos coletados no site do jornal foram etiquetados com suas informações pertinentes como data, seção e autor e, então, armazenados de maneira sistemática e cuidadosa em pastas com identificação prévia. Esse processo evitou que informações fossem perdidas e, por consequência, garantiu que a pesquisa não fosse comprometida em sua conclusão.

Por fim, os textos coletados em formato PDF foram analisados por, no mínimo, três vezes e as candidatas a expressões idiomáticas foram transpostas para tabelas específicas de cada mês. As informações tabeladas foram: data da publicação do texto; seção escolhida (Cultura/Esporte); subseção (Coluna/Reportagem e modalidade); autor do texto; localização do candidato no texto (Título/Corpo); e candidato a expressão idiomática (conforme aparece no texto).

A figura 1 apresenta um excerto da base de dados formada com as expressões idiomáticas encontradas nos textos de esporte e cultura do jornal on-line. Mostra-se relevante ressaltar que o material original é formado por quatro planilhas, uma para cada um dos últimos meses do ano de 2015.

**Figura 1:** Excerto da base de dados formada.

DATA	SEÇÃO	SUBSEÇÃO	AUTOR DO TEXTO	LOCALIZAÇÃO NO TEXTO	CANDIDATO A EI
13/10/2015	Esporte	Reportagem - Tênis	Folha de <u>sp</u>	Corpo	se dar bem
14/10/2015	Cultura	Coluna	José Simão		
14/10/2015	Cultura	Coluna	Mônica Bergamo	Corpo	fazer o meio de campo
14/10/2015	Cultura	Coluna	Mônica Bergamo	Corpo	botar o pé no freio
14/10/2015	Esporte	Reportagem - Tênis	Folha de <u>sp</u>		
15/10/2015	Cultura	Coluna	José Simão	Corpo	puxar o tapete
15/10/2015	Cultura	Coluna	José Simão	Corpo	puxando o saco
15/10/2015	Cultura	Coluna	José Simão	Corpo	moeda de troca
15/10/2015	Cultura	Coluna	Contardo Calligaris		
15/10/2015	Esporte	Reportagem - Tênis	Folha de <u>sp</u>		
16/10/2015	Cultura	Coluna	José Simão		
16/10/2015	Cultura	Coluna	Vladimir Safatle	Corpo	sangue nos olhos

Fonte: Elaboração da autora.

Na sequência, são apresentados os resultados da pesquisa, bem como a discussão dos dados fraseológicos.

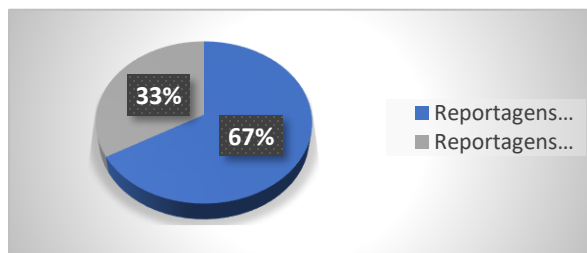
### 3 Resultados e discussão dos dados

No período de um ano de pesquisa, foram armazenados 300 textos em formato PDF, dos quais 200 são produtos de colunas culturais e 100 de reportagens esportivas. Desse acervo, foram colhidos apenas 25 candidatos a expressão idiomática, 16 nas colunas culturais - *a chapa estar/ser quente; acabar-se em lágrimas; botar/colocar o pé no freio; fazer o meio de campo; matar-se para ficar (magérrima); mudar de figura (situação); passar à linha de frente; puxar o saco; puxar o tapete; ressurgir das cinzas; rota de fuga; ser uma luz no fim do túnel; ser/ter cara de pau; ser/ter moeda de troca; tempos de vacas magras; ter sangue nos olhos* - e 9 nas reportagens esportivas - *cair de pau; cereja do bolo; dar-se bem; entregar o jogo; ganhar o mundo; levar um pneu; matar o jogo; ponto de equilíbrio; tirar um peso das costas*.



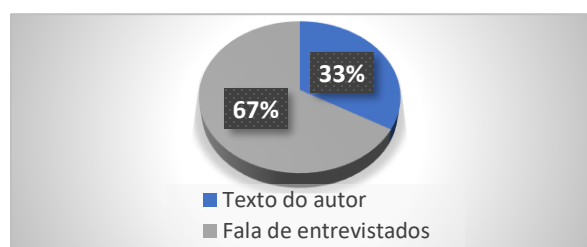
A baixa ocorrência foi analisada como uma consequência de aspectos inesperados dos textos arquivados. As reportagens esportivas disponíveis on-line, ao contrário da versão impressa, sempre eram curtas e com intuito bastante informativo, mostrando alguma personalidade apenas em citações ou nas reportagens assinadas por jornalistas mais experientes (gráficos 1 e 2), visto que a grande parte das reportagens são feitas por estagiários e levam, portanto, o nome do jornal inteiro. Esse aspecto pode ser justificado pelas características da sociedade atual que, pela pressa do dia-a-dia, demanda textos que informem de maneira rápida. Já as colunas culturais eram relativamente mais longas, porém mesmo as que possuíam tom humorístico apresentavam linguagem formal e, quando utilizavam alguma expressão cotidiana informal, essa vinha marcada por aspas ou artifícios similares.

**Gráfico 1:** Expressões idiomáticas em reportagens esportivas



Fonte: Elaboração da autora.

**Gráfico 2:** Ocorrência de EI em falas de entrevistados nas reportagens

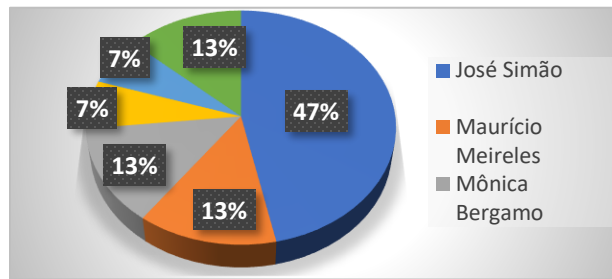


Fonte: Elaboração da autora.

Com o gráfico 3 que segue é possível perceber que embora a linguagem de todas as colunas seja formal e distante, a coluna humorística de José Simão foi a responsável pela maior quantidade de ocorrências de expressões idiomáticas. Esse caso pode ser explicado pelo fato de que muitas vezes essas unidades fraseológicas são usadas para tentar uma aproximação com o leitor, item importantíssimo quando se trata de humor.



Gráfico 3: Ocorrência das EI nas colunas culturais



Fonte: Elaboração da autora.

É válido ressaltar, ainda, que, inesperadamente, foram encontradas mais ocorrências de outra subdivisão das unidades fraseológicas, as citações e parêmiat. Esse fato pode indicar uma grande oportunidade de pesquisa nesse outro ramo da Fraseologia, colocando os textos jornalísticos, principalmente na parte cultural, como um campo potencialmente rico de enunciados fraseológicos.

Como exemplo dessas aparições, é apresentada uma amostra – figura 2 - da coluna do filósofo e escritor Luiz Felipe Pondé do dia 5 de outubro de 2015. O colunista, para dar início às reflexões do dia sobre o comportamento humano, utiliza a citação “Sem Deus, tudo é permitido” do personagem Ivan Karamázov, que fará com que o leitor de Dostoiévski rapidamente se recorde do livro *Irmãos Karamázov* e construa a ponte entre a frase e o tema do texto.

Figura 2: Exemplo de citação encontrada no material armazenado.



Fonte: Elaboração da autora

## Considerações finais

Ao longo do período da pesquisa foi possível realizar o objetivo proposto previamente, que consistia em coletar textos jornalísticos da *Folha de São Paulo* e procurar neles expressões idiomáticas para, mais tarde, formar uma base de dados com todas as informações pertinentes que orientaram a pesquisa. O planejamento estabelecido na metodologia foi fundamental para que a pesquisa tivesse êxito e o *corpus* constituído, além de ter sido utilizado na investigação apresentada, pode ainda ser compartilhado com outras pesquisas com base em textos jornalísticos, o que engrandece a área da Fraseologia.

Outrossim, foi possível, ainda, comprovar que a ocorrência das expressões idiomáticas depende das características dos textos analisados, como linguagem, tema e gênero. Condição que é confirmada, por exemplo, pelo alto índice de expressões na coluna de José Simão, que tinha caráter humorístico.

Assim, além de ter contribuído para o melhor entendimento das aparições das expressões idiomáticas e garantido um material que pode ser usado em estudos futuros, a pesquisa relatada abre as portas para uma investigação sobre parêmiás e citações em textos jornalísticos, outro tipo de unidade fraseológica que ainda pode se configurar como objeto de estudo relevante em pesquisas de caráter fraseológico no Brasil.

## Referências

BIDERMAN, M. T. C. O Conceito Linguístico de Palavra. In: BASÍLIO, M. (org.). **A Delimitação de Unidades Lexicais**. Rio de Janeiro: Grypho, v. 1, 1999.

CORPAS PASTOR, G. **Manual de fraseología española**. Madrid: Gredos, 1996.

MONTORO DEL ARCO, E. T. Clasificaciones de las Ufs: el lugar de las locuciones. In: \_\_\_\_\_ **Teoría fraseológica de las locuciones particulares. Las locuciones prepositivas, conjuntivas y marcadoras del español**. Frankfurt am Main: Peter Lang, 2006.

SILVA, M. E. O. O. Dicionários: armas de dois gumes no estudo da Fraseologia: o caso das locuções. In: ALVAREZ, M. L. O.; UNTERNBÄUMEN, E. H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas**. Campinas: Pontes, 2011, pp. 161-182.

TAGNIN, S. E. O. Linguística de *corpus* e Fraseologia: uma feita para a outra. In: ALVAREZ, M. L. O.; UNTERNBÄUMEN, E. H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas**. Campinas: Pontes, 2011, pp. 277-302.

TRISTÁ PÉREZ, A. M<sup>a</sup>. “Teoría fraseológica: visión general del problema. In: \_\_\_\_\_ **Fraseología y contexto**. La Habana: Ciencias Sociales, 1988, pp. 2-43.

XATARA, C. M. O campo minado das expressões idiomáticas. **Alfa**. São Paulo, 42 (n. esp.), 1998, pp. 147-159.