

Advérbios compostos do Português do Brasil

Izabela Müller^{1,3}, Nuno Mamede^{2,3}, Jorge Baptista^{1,3}

¹ Universidade do Algarve, Faculdade de Ciências Humanas e Sociais, Faro, Portugal

² Universidade de Lisboa, Instituto Superior Técnico, Lisboa, Portugal

³ INESC-ID Lisboa - Human Language Technology Lab, Lisboa, Portugal

Resumo

Este artigo apresenta um estudo em curso que tem como objetivo identificar, classificar e descrever advérbios compostos do português brasileiro, com foco em suas propriedades sintáticas e semânticas. A meta é criar um léxico abrangente de advérbios compostos do português, que possa ser aplicado em vários domínios, como processamento de linguagem natural (PLN), tradução e ensino de idiomas. Essa classificação ajudará a estabelecer os padrões de distribuição específicos da variedade brasileira, permitindo uma distinção clara entre expressões adverbiais brasileiras e europeias. Além de suas aplicações práticas, o estudo visa investigar vários aspectos dessas expressões adverbiais. Isso inclui examinar as suas frequências, possíveis ambiguidades, *faux-amis* e diferenças gramaticais quando comparadas ao português europeu. Além disso, esta pesquisa visa contribuir para o desenvolvimento de um léxico abrangente de expressões adverbiais, com vista a uma melhor compreensão do funcionamento e uso desses elementos linguísticos.

Palavras-chave: advérbios compostos, sintaxe, português brasileiro, expressão idiomática, MWE.

Abstract

This ongoing study aims to identify, classify, and describe multiword/compound adverbs in Brazilian Portuguese, focusing on their syntactic-semantic properties. The goal is to create a comprehensive lexicon of compound adverbs in Portuguese, which can be applied in various domains such as Natural Language Processing (NLP), translation, and language teaching. This classification will help establish the distribution patterns specific to the Brazilian variety, allowing for a clear distinction between Brazilian and European Portuguese adverbial expressions. In addition to their practical applications, the study aims to investigate various aspects of these adverbial expressions. This includes examining their frequency, potential ambiguities, *faux-amis*, and grammatical differences when compared to European Portuguese. Furthermore, this research seeks to contribute to the development of a comprehensive lexicon of adverbial expressions, in view of enhancing the understanding of the functioning and usage of these linguistic elements.

Keywords: compound adverbs, syntax, Brazilian Portuguese, idioms, MWE.

1. Introdução

Os *advérbios* são considerados uma classe gramatical complexa e heterogênea¹ de palavras, sem variação flexional e classificações semânticas diversas (Bechara, 2009; Costa, 2008; Cunha & Cintra, 2017; Raposo, 2013). Podem funcionar como modificadores de verbos, adjetivos e outros advérbios e podem, ainda, modificar frases. Semanticamente, fornecem sobretudo informações adicionais sobre circunstâncias de tempo, modo, lugar, quantidade, intensidade, entre outros aspectos.

¹ Este texto foi escrito na variedade do português do Brasil.



Os *advérbios compostos* (ou *expressões multipalavra* adverbiais, do inglês *multiword expressions* (MWE); ou, ainda, *locuções adverbiais*, de acordo com a terminologia gramatical do português)² são expressões formadas por duas ou mais palavras que funcionam nessa combinação como uma só unidade lexical, cuja constituição é fixa, ou seja, seus elementos ocorrem em uma determinada ordem, fortemente restrita, geralmente não sendo permitida a permutação, a inserção, a transposição nem a redução (elipse) de seus elementos (Gross, 1982, 1986; Guimier, 1996; Mel'čuk, 2023). Têm como característica a não-composicionalidade semântica, ou seja, a interpretação global não pode ser derivada do significado individual de seus elementos, ou seja, em sua maioria, são idiomáticos. Os exemplos abaixo mostram primeiramente, (1) expressões adverbiais idiomáticas em português brasileiro (PB); (2) uma expressão similar, em português europeu (PE); e (3) um advérbio derivado terminado em *-mente*, equivalente às anteriores (3)³:

- (1) [O Leo fez a prova] *na maciota; com o pé nas costas* (PB)
- (2) [O Leo fez a prova] *com uma perna às costas* (PT)
- (3) [O Leo fez a prova] *facilmente*

Este artigo tem como objetivo apresentar um estudo em andamento sobre as construções adverbiais no português brasileiro, mais especificamente os advérbios compostos/multipalavra. Nos últimos anos, houve um considerável aumento nas pesquisas dedicadas ao estudo de expressões multipalavra (MWE), expressões idiomáticas, fraseologia, e outras construções de natureza não composicional (Ramisch, 2015) que constituem um problema para as aplicações de processamento de língua natural, principalmente no que tange a sua identificação e reconhecimento automático em textos. Os advérbios compostos são formados por elementos lexicais (preposições, nomes, adjetivos, etc.) que em sua maioria são detectados pelos sistemas como unidades individuais. Experiências realizadas (Gonçalves et al., 2020) demonstram os limites dos léxicos computacionais disponíveis para o processamento desta classe.

A pesquisa tem como principal propósito identificar, classificar e descrever os advérbios compostos, levando em consideração suas propriedades sintáticas e semânticas, com o intuito de enriquecer um léxico de advérbios compostos do português. Além disso, busca-se fornecer uma descrição formal e sintática destes advérbios, bem como uma descrição linguística voltada para o Processamento de Linguagem Natural (PLN) e outras aplicações. Por fim, pretende-se também determinar as fronteiras lexicais entre o português brasileiro e europeu, no que se refere às diferenças idiomáticas, gramaticais e respectivas frequências destas expressões.⁴

No âmbito deste estudo, realizou-se, até ao momento, um levantamento de aproximadamente 3.300 expressões adverbiais em português. Dentro deste conjunto, muitas destas expressões são comuns a ambas as variedades linguísticas, enquanto outras são exclusivas de cada uma das variedades.

Este artigo está organizado do seguinte modo: Começamos esta breve introdução por apresentar sucintamente os conceitos aqui empregues de advérbio e advérbio composto, bem como apresentar as motivações subjacentes a este estudo. Em seguida, na seção 2, são apresentados alguns estudos similares sobre construções de advérbios compostos em diferentes línguas. Na seção 3, descrevemos genericamente a metodologia empregada para a construção do léxico de expressões adverbiais, juntamente com os critérios

² Neste artigo, usamos o termo *advérbio composto* no sentido em que este é definido no texto. Tal corresponde, em grande medida, ao conceito do termo mais tradicional *locução adverbial*, bem como ao do termo mais recente (e genérico) de *expressão multipalavra* (adverbial). Não atribuímos importância a esta diferença terminológica.

³ Nos exemplos, a frase sobre a qual o advérbio opera é representada entre parênteses retos. No caso das expressões (mais) idiomáticas, acrescenta-se uma glosa entre aspas simples ‘...’, sobretudo quando uma dada interpretação é exclusiva de uma das variedades do português. Quando relevante, essas variedades são assinaladas, em expoente, pelas siglas PT (=português europeu) e PB (= português brasileiro). Os exemplos retirados do corpus *TenTen2018* do Sketch Engine (Wagner et al., 2018), <https://app.sketchengine.eu>, vêm assinalados com a indicação [SE].

⁴ O estudo das construções adverbiais noutras variedades do português, nomeadamente de Angola e de Moçambique, deverá ser prosseguido noutros trabalhos, para os quais este recurso linguístico certamente será útil.



estabelecidos para inclusão e exclusão de determinadas expressões. A seção 4 apresenta o enquadramento teórico-metodológico adotado neste estudo e prossegue com a classificação formal das expressões adverbiais. Por fim, na seção 5, são apresentados os resultados provisórios deste estudo e indicados os próximos passos.

2. Trabalhos relacionados

Embora os advérbios simples mais frequentes, precisamente aqueles terminados em *-mente*, no contexto do português brasileiro, tenham já sido previamente estudados e classificados de acordo com suas características sintáticas e semânticas (Fernandes, 2011), os advérbios compostos ainda não receberam a mesma atenção quando se trata de análise detalhada e classificação descritiva de suas propriedades sintáticas e semânticas.

No que diz respeito ao português europeu, Palma (2009) desenvolveu um léxico para as construções adverbiais, tendo descrito suas propriedades sintáticas com o objetivo de compará-las a estruturas semelhantes em espanhol. A autora coletou e classificou cerca de 1.800 advérbios compostos, que foram posteriormente contrastados com construções equivalentes em espanhol. Mais tarde, estes advérbios foram adicionados ao léxico da cadeia de PLN para português, STRING (Mamede et al., 2012). Este léxico computacional de advérbios compostos nos serve de base para este projeto. Um trabalho mais recente (Català et al., 2020) aborda os desafios das traduções dos advérbios compostos entre o português e o espanhol.

Encontramos estudos semelhantes sobre a construção de léxicos de advérbios compostos em diferentes línguas, além do português. Por exemplo, Gross (1996) desenvolveu um léxico de advérbios compostos em francês, contendo cerca de 6.800 entradas, com suas respectivas descrições sintáticas e semânticas. Este mesmo dicionário foi mais tarde incorporado em sistemas de Processamento de Linguagem Natural (Laporte et al., 2008; Laporte & Voyatzí, 2008). Para a língua japonesa, o *Dictionary of Multi-Word Expressions* (JDMWE) (Shudo et al., 2011) possui aproximadamente 104.000 entradas de expressões multipalavra, das quais 6.000 são expressões adverbiais. Este léxico fornece informações sobre as funções sintáticas, estrutura, mobilidade nas frases e frequências destas expressões. Em relação à língua Tcheca, um estudo realizado por Žižková (2018) analisou cerca de 470 construções adverbiais e, dentre estas, 103 expressões adverbiais foram acrescentadas a um dicionário morfológico de MWE já existente naquela língua. Quanto à língua espanhola, Català (2003) construiu um léxico de expressões adverbiais fixas, contendo cerca de 6.000 entradas. Essas expressões foram formalizadas e classificadas em 11 classes, seguindo os critérios de M. Gross (1986). Adicionalmente, para o espanhol, o léxico *SentiText*, (Moreno-Ortiz et al., 2013), criado especificamente para a análise de sentimentos, inclui 2.255 construções adverbiais. Para a língua inglesa, (Shigetó et al., 2013) marcaram aproximadamente 1.500 MWE adverbiais do Wiktionary e 468 do Penn Treebank.

3. Metodologia

Esta seção apresenta a metodologia para a construção de um léxico de advérbios compostos, contendo tanto as expressões do português europeu como da variedade do Brasil.

3.1. Recursos linguísticos para a construção do léxico

O foco principal do estudo consistiu na expansão de um recurso lexical computacional já existente, originalmente desenvolvido para o português europeu, o qual compreende aproximadamente 2.750 advérbios, incluindo advérbios simples, derivados em *-mente* e advérbios compostos. Os advérbios compostos deste léxico são baseados no trabalho de Català et al. (2020), que foram inicialmente estudados e descritos por Palma (2009). O léxico inclui as seguintes informações:



1. cerca de 1.000 advérbios derivados e terminados em *-mente* acompanhados de sua base adjetival correspondente e sua classificação sintático-semântica (Fernandes, 2011), feita de acordo com propriedades apresentadas por Molinier e Levrier (2000);
2. aproximadamente 1.750 advérbios compostos (Català et al., 2020; Palma, 2009), classificados com base em suas estruturas formais (Gross, 1986) e também de acordo com propriedades sintático-semânticas de Molinier e Levrier (2000);
3. aproximadamente 6.000 advérbios derivados terminados em *-mente* com informação sobre o adjetivo de base, porém sem informação complementar sobre a classe sintática-semântica.

Muitos advérbios dos indicados em 2., acima, existem igualmente no português do Brasil, o que foi verificado manualmente por um dos autores, que é falante nativa, e validado em *corpora* ou em textos da internet (ver adiante).

Para completar e aprimorar os recursos lexicais existentes com advérbios compostos do português do Brasil, examinamos algumas fontes para a coleta de novas expressões, mais precisamente o *Dicionário de Locuções e Expressões da Língua Portuguesa* (Rocha & Rocha, 2011), que contém por volta de 18.000 expressões, das quais, aproximadamente, 3.000 adverbiais. Neste dicionário, as expressões adverbiais estão em meio a outras construções multpalavra, tanto adjetivais quanto verbais, pelo que é necessária uma consulta manual e meticulosa para extrair somente as que fazem parte deste estudo. Embora o dicionário não especifique a variedade do português à qual essas expressões pertencem, uma razoável porcentagem se alinha com as encontradas no português europeu, portanto algumas já estavam integradas ao léxico computacional inicial. Houve de se ter o cuidado em verificar que, mesmo quando as expressões são aparentemente as mesmas, poderia haver diferenças distribucionais, morfológicas ou morfossintáticas (por exemplo, em que uma variedade use a expressão mas com diferentes preposições; ou só admita algumas flexões) ou até mesmo a ocorrência de *faux-amis*, os chamados falsos amigos. Dentre os demais recursos para a coleta das expressões, utilizamos ainda o livro *Locuções Adverbiais* (Schwab, 1985), e o *Dicionário Brasileiro de Fraseologia* (Silva, 2013), cuja análise ainda está em curso.

3.2. Critérios para o recenseamento dos advérbios compostos

Esta pesquisa envolve, como se disse, o recenseamento, a classificação formal e a descrição das propriedades sintáticas e semânticas de advérbios compostos no português brasileiro. Com o intuito de desenvolver um léxico abrangente de advérbios compostos para o português, partimos de um trabalho previamente elaborado para o português europeu, (Palma, 2009), o qual foi examinado minuciosamente para identificar aquelas expressões que fossem, a princípio, comuns às duas variedades da língua, assinalando-as como tal. Logo nessa etapa, foi possível detectar alguns casos em que cada variedade apresenta pequenas diferenças para uma expressão de significado idêntico (e.g., *com uma perna às costas^{pt}/com o pé nas costas^{pb}*). Em seguida fizemos o recenseamento das expressões nos recursos citados acima. Estabelecemos alguns critérios a fim de determinar a inclusão de expressões que fossem relevantes ao nosso estudo ou a sua exclusão, devidamente fundamentada.

Decidimos, assim, incluir os seguintes tipos de construções adverbiais:

1 - Expressões adverbiais idiomáticas, e.g., *com a cara e a coragem*

(4) [O Leo explorou as ruas] *com a cara e a coragem*

Estas expressões adverbiais são semanticamente não-composicionais, ou seja, exibem um certo grau de fixidez formal interna, apresentando restrições quanto a:



- (i) *permutação* dos elementos coordenados: *[O Leo explorou as ruas] *com a coragem e a cara*;⁵
- (ii) a *variação* de gênero e/ou número de seus elementos: *[O Leo explorou as ruas] *com as caras e as coragens*;
- (iii) a *substituição* de seus elementos por sinônimos/antônimos: *[O Leo explorou as ruas] *com a cara e o coração*;
- (iv) *inserções* de determinantes e/ou modificadores livres ou outros elementos lexicais, mesmo que tais determinantes e/ou modificadores possam modificar os elementos do composto em outro lugar: *[O Leo explorou as ruas] *com a cara limpa e a destemida coragem* e, finalmente
- (v) a *exclusão* de alguns de seus elementos: *[O Leo explorou as ruas] *com a cara*

2 - Construções adverbiais multipalavra equivalentes a uma única palavra, nomeadamente os advérbios derivados com o sufixo *-mente* e.g., *em geral* = *geralmente*, *de súbito* = *subitamente*.

3 - Construções adverbiais que permitem algum grau de variação em seus componentes: e.g., *a certa altura*, a qual permite variação dos pronomes demonstrativos usados, como, por exemplo: *a esta / essa / aquela altura*; ou a variação de outro elemento da construção, neste caso, o adjetivo *em um (num) determinado / certo / dado / momento*

- (5) *A essa altura*, [o Leo já desistiu de ir ao show]
- (6) *Em um determinado momento*, [tudo mudou na vida do Leo]

Seria pouco interessante representar num léxico (uma listagem) todas estas famílias de expressões adverbiais, com variação de elementos gramaticais relativamente previsíveis. Estas expressões deverão ser representadas por meio de gramáticas locais (GL), como sugerem, para as expressões temporais, Baptista (1999), Hagège et al. (2010) e Maurício (2011). Nestes casos, somente uma entrada lexical é registrada.

4 - Expressões temporais idiomáticas que denotam referências temporais como *no tempo das vacas gordas/magras*, *à hora das galinhas*^{pb}, *ao toque das ave-marias*.

- (7) [Se a obra não foi realizada] *no tempo das vacas gordas* [, também não será feita agora] *no tempo das vacas magras* [SE]
- (8) [O Leo se deita] *à hora das galinhas* ‘muito cedo’
- (9) [Todos se recolhiam] *ao toque das ave-marias* ‘ao entardecer’

5 - Construções comparativas fixas idiomáticas, semelhantes às já estudadas por Ranchhod (1991) para o português, a partir dos trabalhos homólogos para o francês de Gross (1984, 1986), mas que sejam exclusivas do português brasileiro, como por exemplo:

- (10) [... os candidatos falam como santos, mas se comportam] *como o diabo gosta* [SE]
- (11) [O menino correu] *como uma flecha* [ao sentir que estava em perigo]
- (12) [O Leo e a irmã vivem] *como cão e gato* (cf. *dão-se como o cão e o gato*)

⁵ Os símbolos ‘*’ e ‘?’ indicam a inaceitabilidade da expressão ou a sua aceitabilidade duvidosa, respectivamente. O símbolo ‘°’ indica que a expressão é aceitável mas que o significado é diferente do que está em discussão.



- (13) [O Leo vestiu-se impecavelmente], *como dita o figurino* (cf. *como manda o figurino*)
- (14) [O Leo fuma] *como uma caipora*
- (15) [A saudade amarga] *que nem jiló*

Resolvemos, porém, excluir os seguintes tipos de construções:

1 - Construções preposicionais e conjuncionais: algumas destas expressões têm valor adverbial, porém selecionam elementos livres e variáveis, por exemplo, *aos cuidados de*, *ao som de*, *em harmonia com*.

- (16) [O Leo deixou os filhos] *aos cuidados de* a enfermeira/a mãe/a avó/a tia
- (17) [O Leo acordou] *ao som de* as ondas/a viola/o violino/o vento
- (18) [O Leo fez algo] *em harmonia com* a lei/a natureza/o universo

2 - Construções adverbiais associadas a frases com nomes predicativos e o verbo-suporte *estar*. Excluímos, assim, expressões como [estar] *com a corda no pescoço*, dentre outras, muitas delas já previamente estudadas por Ranchhod (1990) para o português europeu, na medida em que são analisadas como construções com verbo-suporte. Efetivamente, estas expressões dão frequentemente origem a modificadores adverbiais por redução do verbo-suporte (Ranchhod 1983; Ranchhod, 1990, pp. 90 ss.):

- (19) [O Zé estava] *com a corda no pescoço*. [classe EPC; Ranchhod, 1990]
- (20) [Depois de ter jogado] *com a corda no pescoço* [até abril, o Arsenal simplesmente perdeu o fôlego]⁶

Do mesmo modo, foram excluídas as construções adverbiais com nomes predicativos que selecionam o verbo-suporte *ter* (e que por vezes apresentam construções equivalentes com *estar com*):

- (21) [Sou português] *com muita honra!*
- (22) Tenho muita honra de/em ser português. /É uma honra para mim ser português.

3.3. Metodologia da descrição linguística

Para fazer a descrição das expressões e aferição das propriedades linguísticas das construções adverbiais nas variedades europeia e brasileira do português, recorreremos a dois tipos de fontes: (i) a intuição linguística de falantes nativos de cada variedade; juntamente com (ii) consultas em *corpora*. Efetivamente, a elicitação de juízos de aceitabilidade quanto à boa formação e interpretabilidade de exemplos construídos desempenha um papel essencial na descrição linguística (Laporte, 2015), enquanto o recurso de forma crítica a dados textuais obtidos de *corpora* de dimensões apreciáveis pode servir não só para validar essas intuições, como também suscitar novas direções de pesquisa.

Inicialmente, consultamos o CETEMFolha/NILC (Pinheiro & Aluísio, 2003),⁷ um *corpus* com aproximadamente 24 milhões de palavras em textos jornalísticos do periódico Folha de São Paulo (1999), coletados em 1994 para o português brasileiro, e disponível na plataforma da Linguateca. Consultamos também

⁶ <https://www.flashscore.pt/noticias/futebol-premier-league-arsenal-perdeu-o-titulo-devido-aos-jogos-mentais-de-guardiola-e-a-falta-de-profundidade/YVh9iWjJ/> [consultado em 31/05/2023].

⁷ https://www.linguateca.pt/cetenfolha/index_info.html



o CETEMPúblico (Rocha & Santos, 2000),⁸ um *corpus* de aproximadamente 180 milhões de palavras em português europeu, também disponível na plataforma Linguateca. Ao analisarmos os dados disponíveis, principalmente no NILC, notamos que somente algumas das expressões daquelas que coletamos apareciam nas buscas, o que nos leva a crer que o volume deste *corpus* não era suficiente para o escopo da nossa pesquisa.

Tivemos depois acesso ao *corpus ptTenTen18* (Kilgarrif et al., 2014; Wagner et al., 2018), um *corpus* de grandes dimensões (precisamente 5,542,074,775 tokens para o português brasileiro, e 206,869,477 tokens para o português europeu), disponível na plataforma do Sketch Engine. Aí verificamos que praticamente todas as buscas de expressões apresentavam resultados positivos, o que nos levou a considerar que este recurso já respondia à necessidade de validar em textos as expressões aqui estudadas, tornando-se, assim, uma ferramenta indispensável para o desenvolvimento deste trabalho. Outro ponto positivo do *corpus ptTenTen18*, além da abrangência e cobertura das variedades europeia e brasileira do português, é o fato de ser um *corpus* relativamente recente, que nos permite explorar expressões contemporâneas e observar alguns fenômenos interessantes, como, por exemplo, o uso de expressões tipicamente brasileiras em textos das mídias portuguesas e vice-versa.

Na seção seguinte, apresentamos os critérios de classificação formal, de classificação sintático-semântica e de representação das propriedades linguísticas das construções adverbiais recenseadas neste estudo.

4. Descrição linguística

Adotamos a perspectiva teórico-metodológica do Léxico-Gramática proposta por Maurice Gross (1975, 1981) e fundada na gramática transformacional de operadores de Zellig S. Harris (1991). De acordo com essa perspectiva, a *frase elementar*, constituída pelo elemento predicativo e seus argumentos – sujeito e eventuais complementos essenciais, é considerada a unidade mínima de análise linguística. Gross (1984, p. 275) afirma que somente numa frase elementar se pode determinar claramente as propriedades sintáticas e o significado preciso das expressões linguísticas. Consequentemente, a classificação dos advérbios, tanto simples como compostos, é estabelecida com base nas propriedades sintáticas dessas unidades lexicais, dentro do contexto da frase elementar.

4.1. Classificação formal

Gross (1986, p. 12) introduz o conceito de *advérbio generalizado* (em francês *adverbe généralisé*), que engloba tanto expressões composicionais e sintaticamente analisáveis quanto expressões fixas, cristalizadas, não-composicionais e idiomáticas. O conceito de advérbio generalizado inclui as seguintes estruturas:

(i) advérbios simples, incluindo os advérbios derivados (em *-mente*):

(23) [O Leo trabalha] *diariamente*

(ii) complementos preposicionais circunstanciais (adjuntos), geralmente de estrutura morfossintática livre:

(24) [O Leo trabalha] *com afinco/nos feriados/em casa/à beça*⁸

(iii) orações subordinadas circunstanciais:

⁸ <https://www.linguateca.pt/CETEMPUBLICO/>



(25) [O Leo trabalha] enquanto eu descanso/enquanto o diabo esfrega um olho

Gross (1986) também propõe uma organização taxonômica das expressões adverbiais compostas que é determinada com base em sua sequência interna de categorias gramaticais. Com base nessa estrutura, Gross (1986) classifica as expressões adverbiais do francês em 16 classes, dentre as quais utilizamos 10 para a formalização das expressões em português, dada a semelhança entre as estruturas das expressões do francês e do português. Palma (2009) também formalizou os advérbios compostos do português europeu fazendo uso destes critérios. A Tabela 1 apresenta de forma compacta esta classificação.

As classes são identificadas por siglas que indicam a estrutura interna de cada tipo de advérbio composto. Por exemplo, **PAC** é uma expressão composta tipicamente por uma *preposição*, um *adjetivo* e um *nome*, como se indica na coluna da estrutura interna do composto; e.g., *de má vontade*: [O Pedro fez isso] *de má vontade*. Cada classe é ilustrada por um exemplo. Os valores da coluna **PT**, indicam advérbios que são exclusivos da variedade europeia, seguidos da porcentagem em cada uma das classes. Igualmente, os valores da coluna **BR** indicam expressões exclusivas da variedade brasileira, e sua respectiva porcentagem nessa classe. Finalmente, a coluna **PTBR** indica o número de expressões comuns às duas variedades e a porcentagem dessa classe. Na coluna **Total** apresenta-se a soma dos valores PT+BR+PTBR e a porcentagem do total das expressões desta classe relativamente ao total de entradas do léxico. A última linha apresenta o número de expressões e porcentagem do léxico de cada variedade (ou comum a ambas as variedades).



Tabela 1. Classificação dos advérbios compostos em português europeu e brasileiro. As classes são indicadas por códigos convencionais

| Classe | Estrutura Interna | Exemplos | PT | % | BR | % | PTBR | % | Total | % total |
|--------------|-------------------|-----------------------------|------------|------------|--------------|------------|--------------|------------|--------------|---------|
| PC | Prep C | <i>em vão</i> | 79 | 8% | 451 | 46% | 441 | 46% | 971 | 29% |
| PDETC | Prep Det C | <i>pelo menos</i> | 112 | 16% | 236 | 32% | 381 | 52% | 729 | 22% |
| PAC | Prep Adj C | <i>de má vontade</i> | 27 | 11% | 109 | 44% | 113 | 45% | 249 | 7% |
| PCA | Prep C Adj | <i>por maioria absoluta</i> | 52 | 15% | 104 | 33% | 158 | 52% | 314 | 9% |
| PCDC | Prep C1 de C2 | <i>por conta da casa</i> | 54 | 21% | 105 | 41% | 98 | 38% | 257 | 8% |
| PCPC | Prep C1 Prep C2 | <i>da cabeça aos pés</i> | 60 | 16% | 140 | 39% | 166 | 45% | 366 | 11% |
| PCONJ | Prep C1 Conj C2 | <i>em verso e prosa</i> | 14 | 6% | 87 | 36% | 138 | 58% | 239 | 7% |
| PF | Frase fixa | <i>dito isso</i> | 5 | 5% | 63 | 67% | 26 | 28% | 94 | 3% |
| PV | Prep V W | <i>até dizer chega</i> | 1 | 4% | 10 | 40% | 14 | 56% | 25 | 1% |
| PJC | Conj C | <i>e por aí vai</i> | 3 | 3% | 61 | 69% | 24 | 28% | 88 | 3% |
| Total | | | 407 | 12% | 1.366 | 41% | 1.559 | 47% | 3.332 | |

Nota. Na representação da estrutura interna, usa-se a notação seguinte: Adj – adjetivo, Conj – conjunção, Det – determinante, Prep – preposição; C – indica um elemento nominal fixo da combinação; W – indica qualquer sequência (não especificada, eventualmente nula) de complementos.

Para fins desta pesquisa, determinamos que excluiríamos certas classes propostas por Gross (1986), nomeadamente, as classes **PCDN**, que inclui construções como *à margem de, por ocasião de*, e a classe **PCPN**, que inclui construções como *de acordo com, em relação a*, por entendermos que se trata de expressões preposicionais e conjuncionais, na medida em que introduzem um elemento distribucionalmente livre, as primeiras exclusivamente um grupo nominal (e.g., *à margem de*_as negociações) e as segundas podendo introduzir uma oração subordinada ou um grupo nominal de conteúdo proposicional/nome predicativo (e.g., *Em relação a essa tarefa/a comprar os medicamentos, o Leo sabe o que fazer*), apesar de se reconhecer o estatuto adverbial de todo o constituinte que estas locuções introduzem.

Além destas, também deixamos de fora as classes de construções comparativas (Gross, 1984): **PECO**, que incluem construções comparativas modificadoras de um adjetivo, como, por exemplo, [*ser magro*] *como um palito*; **PVCO**, que são construções comparativas que modificam verbos, como, por exemplo, [*fumar*] *como uma chaminé*; e, finalmente, a classe **PPCO**, formada por expressões comparativas introduzidas por preposição, e.g., [*surgir*] *como (que) por encanto*. Estas expressões já foram previamente descritas para o português europeu (Ranchhod, 1991) e também revisadas e formalizadas no desenvolvimento do sistema STRING, como mencionado anteriormente. Consideramos e acrescentamos ao léxico somente expressões que são exclusivas da variedade brasileira, como, por exemplo, [*fumar*] *como uma caipora* ‘muito’ (a *caipora* é uma entidade mitológica da cultura brasileira).



Finalmente, tratando-se de um processo de recenseamento em curso, os números apresentados na Tabela 1 devem considerar-se em permanente atualização.

4.2. Propriedades sintáticas e semânticas

Para apresentar um exame detalhado do funcionamento dos advérbios compostos quando integrados em frases, adotamos, genericamente, a classificação sintático-semântica proposta por Molinier e Levrier (2000). O estudo destes autores concentrou-se na classificação de advérbios derivados terminados em *-ment* em francês. Acreditamos que esta estrutura de classificação não só se aplica genericamente aos advérbios correspondentes, derivados e terminados em *-mente* em português (já previamente estudados por Fernandes (2011), embora apenas para um conjunto de advérbios do português do Brasil); mas também pode ser estendida e adaptada para a descrição das construções das expressões adverbiais compostas.

Em essência, Molinier e Levrier (2000) estabeleceram um conjunto de critérios para distinguir diferentes classes de construções adverbiais com base na relação que estes elementos estabelecem com a frase de base (ou frase elementar) em que se encontram e/ou com constituintes dessa frase. Define-se, assim, os dois tipos principais de construções adverbiais, (P) e (M):

- (P) advérbios modificadores *externos* das frases; e
- (M) advérbios modificadores dos elementos *internos* das frases.

Seguindo os princípios metodológicos do Léxico-Gramática (Gross 1975, 1981), a determinação das propriedades linguísticas (sintático-semânticas) dos elementos lexicais (neste caso, dos advérbios) só se pode fazer adequadamente no quadro sintático das frases de base (ou frases elementares) da gramática. As frases elementares são concebidas, nesta perspetiva, como a expressão linguística de um predicado semântico reduzido à sua expressão mais simples (com os seus elementos essenciais expressos). Como os advérbios constituem (geralmente) *operadores de segunda ordem* -- no sentido de Harris (1991), ou seja, operam sobre o resultado de outro operador -- o contexto sintático adequado para determinar as suas propriedades é uma frase elementar. Vejamos os exemplos a seguir, para ilustrar estes conceitos:

- (26) *No entanto*, [o Leo acredita que possa ganhar o jogo]
- (27) [O Leo chegou à festa] *de mãos abanando/de mãos a abanar*

Para determinar se os advérbios são modificadores *internos* ou *externos* à proposição, Molinier e Levrier (2000) propõem as seguintes duas propriedades, (A) e (B):

A. Topicalização do advérbio e negação do predicado principal:

- (26a) *No entanto*, [o Leo **não** acredita que possa ganhar o jogo]
- (27a) **De mãos abanando / de mãos a abanar*, [o Leo **não** chegou à festa]

B. Clivagem (ou extração *ser...que*):

- (26b) **É no entanto **que*** [o Leo acredita que possa ganhar o jogo]
- (27b) *Foi de mãos abanando / de mãos a abanar **que*** [o Leo chegou à festa]



De modo geral, os modificadores externos das frases podem ser *topicalizados* quando a frase é colocada na forma negativa, pois eles operam sobre a frase inteira e não são por isso afetados pela negação do predicado dessa frase. No entanto, eles não podem sofrer a *clivagem*, que é uma operação reservada aos constituintes internos de uma frase. Por outro lado, os advérbios internos à frase geralmente não podem ser topicalizados quando esse predicado se encontra sob uma negação; mas podem ser extraídos. Estas duas propriedades são, em grande medida, independentes. A classificação das construções adverbiais isola, em primeiro lugar, os advérbios modificadores externos à frase (P), por apresentarem *coordenação* da propriedade (A) com a *inaceitabilidade* da propriedade (B): $P = A+ \rightarrow B$. Todas as outras possibilidades são classificadas como advérbios *internos* da frase (M).

Assim, nos exemplos acima, com o advérbio *no entanto* (26a), a topicalização não afeta a relação entre este advérbio e o predicado, pelo que a modificação que o advérbio exerce é independente da negação. Por outro lado, em (26b), a clivagem não pode ser aplicada a esse advérbio, o que demonstra não se tratar de um modificador/complemento do verbo *acreditar* ou de outro elemento dessa frase de base. O advérbio funciona, portanto, como um modificador externo da frase.

Em contrapartida, a expressão adverbial *de mãos abanando / de mãos a abanar* não admite a *topicalização* com a frase na forma negativa, como visto em (27a) mas permite a *clivagem* (ou extração *ser...que*), como se vê em (27b). Trata-se, portanto, de um advérbio modificador interno à frase.

Na classificação proposta pelos autores, a esta divisão de ordem mais geral (P ou M), vem juntar-se uma subclassificação mais fina, que distingue diferentes classes de construções adverbiais. Quanto aos advérbios que operam como modificadores externos da frase, consideram-se três classes principais:

(i) Advérbios conjuntivos (PC)

Estes advérbios têm como característica principal a função de conectar a frase em que se encontram à frase anterior e, em consequência disso, nunca ocorrem em início absoluto de discurso, por exemplo, *consequentemente, portanto, a propósito, em suma, não obstante, pelo contrário*, entre outros:

(28) *Consequentemente/portanto/a propósito/em suma/não obstante/pelo contrário*, [isso não é assim]

Esta classe reúne alguns subtipos de advérbios, agrupados por seus valores semânticos, e tendo como critério principal o fato de serem conectores, portanto não existe uma subclassificação ou organização especial. Na proposta de Molinier e Levrier (2000), alguns dos subtipos conjuntivos são: (a) enumerativos (*primeiramente, em seguida*); (b) justificadores (*pois, com efeito, aliás*); (c) transicionais (*a propósito, de resto, por esta ordem de ideias*) (d) reformulativos (*em suma, em conclusão*); (e) apositivos (*a saber, quer dizer, isto é, por exemplo*), entre outros.

Por oposição aos advérbios conjuntivos, Molinier e Levrier (2000) definem outro conjunto de advérbios de frase a que chamam *disjuntivos* (termo sem qualquer relação com a operação lógica). Neles distinguem em primeiro lugar a classe dos:

(ii) Advérbios disjuntivos de estilo (PS)

Esta classe de advérbios define-se por operar sobre o operador metalinguístico harrissiano *eu digo* (Harris 1991, p. 141), e que corresponde àquilo a que Molinier e Levrier (2000, p. 65) designam por “un verbe du type « dire » placé dans une phrase supérieure” (um verbo do tipo “dizer” colocado numa frase superior). Estes advérbios são, pois, modificadores de modo particularmente apropriados a verbos *de dizer* «*verba dicendi*», (Baptista, 2010) ou de expressões nominais como *termos, palavras*, podendo ser identificados pela paráfrase com uma construção desse tipo:

(29) *Sinceramente/numa palavra/ por outras palavras*, [é um erro fazer isso]



- (30) Eu digo sinceramente/numa palavra/ por outras palavras que [é um erro fazer isso]

Trata-se, portanto, de advérbios que estabelecem algum tipo de relação entre o locutor e o destinatário ou o próprio enunciado. Molinier e Levrier (2000, p. 65 ss.) identificam ainda, neste conjunto, vários conjuntos, de que se destacam: (a) os de atitude do locutor relativamente ao destinatário (*sinceramente, com toda a sinceridade*); (b) os que caracterizam a forma do enunciado (*concretamente, em concreto*); (c) a origem do enunciado (*oficialmente, segundo N, de acordo com N*); (d) expressão do ponto de vista do locutor e implicação no próprio enunciado (*pessoalmente, no que me diz respeito, da/pela minha parte*); entre outros.

Para os restantes advérbios disjuntivos, a que chamam *de atitude (PA)*, os autores procedem a uma subclassificação, sem que, contudo, indiquem um critério que os distinga, no seu conjunto, destes advérbios disjuntivos de estilo (**PS**). Haveria, talvez, que considerar uma classificação que opusesse os advérbios *conjuntivos (PC)* dos restantes (“disjuntivos”, não seria um termo muito motivado) e, entre estes, considerar uma única classe, com critérios particulares para a identificação de cada subclasse. No entanto, por uma questão de clareza, mantemos aqui a estrutura taxonômica e a nomenclatura proposta por Molinier e Levrier (2000). Assim, na grande classe dos:

- (iii) Advérbios disjuntivos de atitude (**PA**), que se dividem em quatro subclasses:

— Advérbios *de hábito (PAh)*

Podem ser definidos pela seguinte propriedade: eles só são compatíveis com o presente e o imperfeito do indicativo, em sua interpretação aspectual **habitual**. Como se sabe, o presente e o imperfeito se prestam a duas grandes interpretações aspectuais: a interpretação referencial (ou de evento) e a interpretação habitual. Os advérbios *de hábito*, de que é modelo o advérbio *habitualmente*, não devem ser confundidos com os advérbios *de frequência* (uma das subclasses dos advérbios de tempo), de que é modelo o advérbio *frequentemente*, nomeadamente por serem modificadores de toda a frase e não poderem, por isso ocorrer destacados no início de frase com o predicado desta na negativa:

- (31) *Habitualmente*, [o Leo (não) bebe/bebia álcool]

- (32) **Frequentemente*, [o Leo (*não) bebe/bebia álcool]

Estes advérbios *de hábito* são, pois, incompatíveis com os tempos-modos verbais que denotam um aspecto pontual/perfectivo, nomeadamente os pretéritos perfeito, mais-que-perfeito e perfeito composto:

- (33) **Habitualmente*, [o Leo bebeu/tinha bebido/bebera álcool]

Alguns exemplos de advérbios compostos deste tipo:

- (34) *Por norma / por hábito / em regra (geral)* [o Leo (não) bebe/bebia álcool]

— Advérbios avaliativos (**PAa**)

Estes advérbios, como o nome indica, exprimem uma avaliação subjetiva por parte do enunciador relativamente ao conteúdo do enunciado e em particular a percepção do caráter favorável ou desfavorável desse conteúdo. Alguns exemplos incluem *felizmente, curiosamente e surpreendentemente*; entre os compostos, *por sorte, por estranho que pareça, por um acaso*. Dado exprimirem um juízo do falante sobre uma proposição, estes advérbios não podem ser usados em frases interrogativas nem nas imperativas, como nos exemplos:

- (35) *Curiosamente/por estranho que pareça*, [o Leo (não) bebe álcool]



- (35a) *É curiosamente/por estranho que pareça, que [o Leo não bebe álcool]
 (35b) Eu digo que curiosamente/por estranho que pareça [o Leo não bebe álcool]
 (35c) *Curiosamente/por estranho que pareça, [o Leo não bebe álcool?]
 (35d) *Curiosamente/por estranho que pareça, [não beba álcool, Leo!]

Alguns destes advérbios aceitam a paráfrase por um verbo dito de *percepção mental* (Baptista, 2013; Baptista & Mamede, 2020; classe 06) como *achar, considerar, etc*

- (35e) *Eu acho curioso/estranho que* [o Leo (não) beba álcool]

— Advérbios modais (PAm)

Os advérbios desse grupo, também chamados de advérbios *assertivos* ou de *modalização da asserção* (Borillo, 1976) constituem uma classe relativamente heterogênea mas, como o nome indica, modificam globalmente uma frase conferindo-lhe uma determinada modalidade, concretamente, a atitude/perspectiva do locutor relativamente ao conteúdo da proposição, como, por exemplo, a verossimilhança, probabilidade, dúvida, etc. São exemplos típicos desta classe:

- (36) *Certamente / fatalmente / indubitavelmente / provavelmente / presumivelmente*, [o Leo chegará hoje].
 (37) *Salvo melhor opinião^{pt}/juízo^{pt} / até prova em contrário^{pt} / sem dúvida (alguma) / em hipótese alguma* [o Leo fará isso].

Estes advérbios podem ser caracterizados por não aceitarem que a frase seja colocada nos modos imperativo (a) ou exclamativo (b) ou numa interrogativa (c), dado já se encontrar modalizada pelo advérbio:

- (36a) **Certamente / fatalmente / indubitavelmente / provavelmente / presumivelmente*, [faz/faça isso, Leo]!
 (37a) **Salvo melhor opinião^{pt}/juízo^{pt} / até prova em contrário^{pt} / sem dúvida (alguma) / em hipótese alguma*, [faz/faça isso, Leo]!
 (36b) **Certamente / fatalmente / indubitavelmente / provavelmente / presumivelmente*, [Leu fez isso]!
 (37b) **Salvo melhor opinião^{pt}/juízo^{pt} / até prova em contrário^{pt} / sem dúvida (alguma) / em hipótese alguma*, [Leu fez isso]!
 (36c) **Certamente / fatalmente / indubitavelmente / provavelmente / presumivelmente*, [Leo chegará hoje]?
 (37c) **Salvo melhor opinião^{pt}/juízo^{pt} / até prova em contrário^{pt} / sem dúvida (alguma) / em hipótese alguma*, [Leo fará isso]?

— Advérbios de frase orientados para o sujeito (PAs)

Trata-se de um conjunto de expressões, definido para os advérbios franceses terminados em *-ment*, que, além de se comportarem como a generalidade dos advérbios de frase (possibilidade de topicalização do advérbio com a frase na negativa e inaceitabilidade da extração *ser...que*), apresentam ainda um escopo sobre o sujeito, geralmente parafraseadas pela construção adjetival associada ao advérbio (38b). Em português, podemos referir, por exemplo, advérbios como *inteligentemente*:



- (38) *Inteligentemente*, [o Leo (não) instalou o computador]
 (38a) °Foi *inteligentemente* que [o Leo (não) instalou o computador]
 (38b) O Leo foi *inteligente* em [(não) instalar o computador]

A frase do exemplo (38a) é, obviamente, aceitável, mas o seu significado é subtilmente diferente (o que se assinala com o símbolo “°”, tratando-se, então, de um modificador do verbo (advérbio de modo orientado para o sujeito; ver adiante). Nesta construção, a frase de base não pode ser negada, já que o advérbio indica o modo como se fez a ação (a instalação do computador). No significado que nos interessa aqui, o locutor exprime uma opinião sobre a ação descrita na frase e, simultaneamente, faz uma apreciação sobre o respectivo sujeito. Nesse sentido, as frases adjetivais associadas às construções destes advérbios admitem ser encaixadas sob um verbo de **opinião/percepção mental**, tais como *achar*, *considerar*, etc.

- (38c) Eu acho que Leo foi *inteligente* [em (não) instalar o computador]

Por essa razão, estas construções não admitem que a frase seja colocada nos modos imperativo (38d) ou exclamativo (38e) ou numa interrogativa (38f):

- (38d) **Inteligentemente*, [instala/instale o computador, Leo]!
 (38e) **Inteligentemente*, [o Leo (não) instalou o computador]!
 (38f) **Inteligentemente*, [o Leo (não) instalou o computador]?

dado que as modalidades imperativa, exclamativa e a formação da interrogativa são incompatíveis com a expressão de uma opinião sobre todo o conteúdo da frase. Até ao momento, não se encontrou exemplos de advérbios compostos desta classe.

Quanto aos advérbios modificadores internos da proposição, estes podem apresentar qualquer outra configuração das propriedades definitórias -- exceto a que determina os advérbios de frase, ou seja, apresentam comportamentos sintáticos variados relativamente à topicalização com a frase na negativa ou a extração *ser...que*.

Estes advérbios são classificados segundo Molinier e Levrier (2000) em seis grandes classes:

- (i) os advérbios de modo (**MV**), e.g., *à moda antiga*, *lentamente*, *de mala e cuia*^{pb}

Essa classe de advérbios é, sem dúvidas, a mais numerosa, e estes respondem adequadamente à interrogativa *como?* São definidos por meio de quatro propriedades, nomeadamente: (1) a impossibilidade de ocorrer destacado na posição inicial de uma frase na negativa (39); (2) a possibilidade de aplicar a extração *ser...que* (39a); (3) a possibilidade de associar o adverbial ao advérbio interrogativo *como?* (39b); (4) no caso dos advérbios simples terminados em *-mente*, a impossibilidade de associar à frase em que estes aparecem uma frase predicativa na qual o adjetivo de que é derivado o advérbio aparece a qualificar o sujeito, quando este é humano (*O Pedro procurou diligentemente o livro# O Pedro foi diligente (a procurar o livro)*). Esta propriedade permite distinguir os advérbios de modo modificadores do verbo (classe **MV**) dos advérbios de modo orientados para o sujeito (classe **MS**, ver adiante); esta propriedade não se aplica portanto diretamente à classificação dos advérbios compostos; (5) impossibilidade de associação com um adverbial de quantidade, quer de completude (e.g., *completamente*) quer de extensão qualitativa (e.g., *predominantemente*); esta propriedade permite diferenciar os advérbios de modo dos de quantidade (39c).

- (39) **De mala e cuia*, [o Leo não se mudou para o Japão]
 (39a) **Foi de mala e cuia que** [o Leo se mudou para o Japão] 'definitivamente, com todos os seus pertences'



- (39b) P: **Como** foi que [o Leo se mudou para o Japão]? R: de mala e cuia
 (39c) [O Leo se mudou para o Japão] ?*^ocompletamente/*predominantemente

(ii) os advérbios orientados para o sujeito (**MS**), e.g., *de bom grado, aos berros, do fundo do coração*

Estes advérbios são definidos através de três propriedades, nomeadamente: (1) a inaceitabilidade na posição inicial e destacada de uma frase negativa (40); (2) a possibilidade de extração com *ser...que* (40a); (3) como os restantes advérbios de modo, estes podem também ser associados ao advérbio interrogativo *como?* (40b);

- (40) **Aos berros*, [o Leo não entrou no carro]
 (40a) **Foi aos berros que** [o Leo entrou no carro]
 (40b) P: **Como** foi que [o Leo entrou no carro]? R: *aos berros*

Tal como se disse atrás, no caso dos advérbios simples terminados em *-mente*, distingue-se no seio dos advérbios de modo, um subconjunto, que autoriza uma paráfrase adjetival com sujeito humano (e.g., [O Pedro entrou no carro] *cuidadosamente* # *O Pedro foi cuidadoso* [a entrar no carro]): esta propriedade não se aplica diretamente, portanto, à classificação dos advérbios compostos. Contudo, é possível conceber uma adaptação e, assim, generalizar este critério a outras formas linguísticas, incluindo as expressões fixas idiomáticas aqui tratadas. Assim, no caso de o advérbio envolver um nome predicativo (e.g., *berro*), é possível construir a frase de base desse nome em que se reencontra o sujeito (humano) da construção analisada, e.g., [O Pedro entrou no carro] *aos berros* # *O Pedro deu (uns) berros* (sobre as construções com verbo-suporte *dar*, *ver*, entre outros, Baptista, 1997; Rassi, 2015; Vaza, 1988). Noutros casos, a possibilidade de inserção de um pronome possessivo demonstra a correferência restrita entre este e o sujeito da construção, e.g., [Nós perdoamos o João] *do fundo do nosso/*seu coração*.

(iii) os advérbios de tempo (**MT**), e.g., *na calada da noite, todo santo dia*

Os advérbios de tempo podem ser verificados a partir das seguintes propriedades: (1) a aceitabilidade em início destacado de uma frase negativa (41); (2) a possibilidade de extração *ser...que* (41a); (3) a possibilidade de responderem adequadamente à interrogativa com o advérbio interrogativo *quando?* (41b):

- (41) *No tempo das vacas magras*, [o Leo não investiu nessa empresa]
 (41a) **Foi no tempo das vacas magras que** [o Leo investiu nessa empresa]
 (41b) P: **Quando** (é que) [o Leo investiu nessa empresa]? R: *No tempo das vacas magras*

Os advérbios de tempo subclassificam-se, por sua vez, em expressões de **data** (MTd, como o exemplo acima, *no tempo das vacas magras*), de **duração** (MTu, *o dia todo*) ou de **frequência** (MTf, *todo santo dia^{pb}*).

(iv) os advérbios de ponto de vista (**MP**), e.g., *em teoria, de direito, em tese*

Estes advérbios podem ser verificados através das seguintes propriedades: (1) a possibilidade de ser aceito em início destacado de uma frase negativa (42); (2) a possibilidade de se aplicar, em geral, a extração *ser...que* (42a); (3) no caso dos advérbios derivados e terminados em *-mente*, a possibilidade de paráfrase do advérbio por um adjetivo a modificar o nome *ponto de vista* (42b); *mutatis mutandis*, esta propriedade pode ser estendida a várias locuções, incluindo algumas expressões idiomáticas aqui tratadas:

- (42) *Em teoria*, [o Leo (não) fez o trabalho]
 (42a) ?**Foi em teoria que** [o Leo fez o trabalho]
 (42b) *Em teoria/do ponto de vista teórico/teoricamente*, [o Leo fez o trabalho]



Como definem Molinier e Levrier (200, pp. 222) estes advérbios “restringem o domínio sobre o qual uma afirmação é válida ou verdadeira”. Por esse motivo, estas construções admitem uma continuação de significado contraditório, desde que essa seja restringida por um advérbio de ponto de vista diferente:

(42c) *Em teoria*, [o Leo (não) cometeu o crime] mas, *na prática*, [foi ele quem tudo fez].

Neste sentido, dado só se combinarem com proposições, as construções com MP não aceitam o imperativo:

(42d) *Teoricamente/em teoria/de um ponto de vista teórico, [Faz/Faça o trabalho, Leo]!

Esta adaptação nos permite incluir nesta classe advérbios quase-sinônimos como *em tese*, mesmo que estes não possam ser associados a *ponto de vista*.

(vi) os advérbios quantificadores (MQ), e.g., *à beça*^{pb}, *para dar e vender*

Estes advérbios funcionam como quantificadores sobre um predicado e essa construção pode ser verificada a partir de três propriedades: (1) inaceitabilidade de ocorrerem destacados em início de frase na negativa (43a); (2) a possibilidade de, na grande maioria, admitirem a extração *ser...que* (43b); (3) a possibilidade de associar o adverbial a uma frase interrogativa com o advérbio *muito* (43c):

(43) [O Leo chorou] *à beça*

(43a) **À beça*, [o Leo não chorou]

(43b) *Foi à beça que* [o Leo chorou]

(43c) P: [O Leo chorou **muito**]? R: Sim, [ele chorou] *à beça*

Os advérbios quantificadores foram organizados por Molinier e Levrier (2000) em três subclasses: (a) os advérbios de *intensidade* (MQi), como *à beça*, acima; (b) os advérbios de *completude* (MQc), *por completo*, *de todo*, *em parte*; e (c) os advérbios de *extensão qualitativa* (*tipicamente*, *verdadeiramente*). Dada a sua diversidade e complexidade de comportamentos sintáticos, uma descrição pormenorizada desta classe terá de ficar para outro momento.

(vii) advérbios focalizadores (MF), e.g., *em especial*, *especialmente*

Estes advérbios formam uma classe caracterizada pelas seguintes propriedades: (1) inaceitabilidade de ocorrerem destacados em início de uma frase na negativa (44a); (2) a impossibilidade de se aplicar a extração *ser...que* (44b); e (3) a possibilidade de extração do advérbio na companhia de um grupo nominal ou outro constituinte maior de uma frase (sujeito, complemento) (44c):

(44) [O Leo (não) gostou] *em especial/especialmente* [da sopa]

(44a) **Em especial/especialmente*, [o Leo (não) gostou da sopa]

(44b) **Foi em especial/especialmente que* [o Leo (não) gostou da sopa]

(44c) *Foi ?em especial/especialmente da sopa que* [o Leo (não) gostou]

A estas classes, acrescentamos a proposta de uma classe para:

(viii) os advérbios locativos (ML), e.g., *à beira-mar*, *onde o vento faz a curva*^{pb}

Apesar da natureza sistemática da classificação dos advérbios derivados terminados em *-ment* para o francês de Molinier e Levrier (2000), verificamos que estes autores não constituem uma classe específica para acolher os advérbios locativos. Ora, tendo os advérbios locativos sido mencionados em vários trabalhos



anteriores (Bechara, 2009; Costa, 2008, p. 44; Raposo, 2013), propomos, como parte desta pesquisa, a constituição de uma classe de *advérbios locativos* compostos, com o intuito de preencher esta lacuna na classificação das construções adverbiais que nos serviu de base. De um modo geral, ainda que marcadamente idiomáticos, estes advérbios podem ser parafraseados por *onde?* numa interrogativa, envolvendo construções verbais locativas, isto é, com um complemento essencial locativo:

- (45) [Leo vive] em Lisboa / *onde o vento faz a curva*
 (45a) P: **Onde** vive o Leo? R: em Lisboa / *onde o vento faz a curva*

A Tabela 2 apresenta as classes sintático-semânticas brevemente identificadas acima e o estado atual do seu recenseamento (e respectiva porcentagem):

Tabela 2. Classes semântico-sintáticas dos advérbios compostos em português

| Classe | Exemplos | Total | % |
|-----------------------------|-----------------------------|--------------|-------|
| PC (conjuntivos) | <i>afinal de contas</i> | 213 | 0,065 |
| PS (disjuntivos de estilo) | <i>com toda a franqueza</i> | 52 | 0,015 |
| PA (disjuntivos de atitude) | <i>em geral</i> | 55 | 0,016 |
| MV (modo) | <i>por amor à camisa</i> | 2.071 | 0,633 |
| MS (modo orient. sujeito) | <i>de boa vontade</i> | 108 | 0,033 |
| MT (tempo) | <i>ao romper do dia</i> | 429 | 0,131 |
| MP (ponto de vista) | <i>na prática</i> | 6 | 0,001 |
| MQ (quantitativos) | <i>aos montes</i> | 162 | 0,049 |
| MF (focalizadores) | <i>em especial</i> | 19 | 0,005 |
| ML (locativos) | <i>nos confins do mundo</i> | 155 | 0,047 |
| Total | | 3.270 | |

5. Considerações finais e trabalhos futuros

Este estudo é parte de uma pesquisa mais ampla que tem como objetivo descrever os advérbios compostos do português brasileiro quanto às suas propriedades sintáticas e semânticas. Após o recenseamento de expressões adverbiais, utilizando vários recursos (sobretudo dicionários), com o intuito de construirmos um léxico abrangente destes advérbios, começamos por descrever suas propriedades de acordo com os princípios teóricos e metodológicos do Léxico-Gramática. Tal consistiu em (a) atribuir-lhes uma classificação formal, com base na estrutura interna das expressões, definida a partir da sequência das categorias morfossintáticas por que são formadas; (b) classificá-las com base nas propriedades que estes advérbios apresentam relativamente às frases (elementares) que modificam; e (c) identificar a variedade do português (europeu ou brasileiro) em que ocorrem, recorrendo sobretudo a *corpora* de dimensões apreciáveis.



Até o momento, examinamos aproximadamente 3.300 advérbios compostos. Os exemplos abaixo (Figura 1) exemplificam a representação de algumas destas expressões na construção do léxico.

Figura 1. Algumas entradas do léxico dos advérbios compostos do português, no formato DELA (Paumier et al. 2021).

de ora em diante, .ADV+PCPC+MTd+PT+BR+EN= "from now on"
de orelha a orelha, .ADV+PCPC+MV+PT+BR+EN= "from ear to ear"
de pai para filho, .ADV+PCPC+MV+PT+BR+EN= "from father to son"
de papel passado, .ADV+PCPC+MV+PT+BR+EN= "on paper"
de passagem, .ADV+PC+MV+PT+BR+EN= "in passing"
de par em par, .ADV+PCPC+MV+PT+BR+EN= "(open) wide"
de pé em pé, .ADV+PCPC+MV+BR+EN= "slowly, carefully"
de ponta a ponta, .ADV+PCPC+MV+PT+BR+EN= "end to end"
de porta em porta, .ADV+PCPC+MV+PT+BR+EN= "door to door"
de preferência, .ADV+PC+PC+PT+BR+EN= "preferably"
de propósito, .ADV+PC+MV+PT+BR+EN= "on purpose"
de quando em quando, .ADV+PCPC+MTf+PT+BR+EN= "from time to time"
de quando em vez, .ADV+PCPC+MTf+PT+BR+EN= "from time to time"

Os resultados deste recenseamento permitem desde já construir um mapa que dá uma dimensão aproximada das diferentes estruturas formais por que são constituídos os advérbios compostos (Tabela 1), a dimensão das diferentes classes sintático-semânticas (Tabela 2), bem como a distribuição destas expressões pelas duas variedades (portuguesa e brasileira) do português.

Algumas das limitações deste trabalho deverão, porém, ser referidas. Muitos advérbios compostos são idiomáticos e relevam de situações comunicativas próprias da oralidade. O recurso a *corpora* de textos escritos, nomeadamente para (a) determinar as propriedades sintático-semânticas dos advérbios e (b) quantificar a frequência da ocorrência destas expressões e daí tirar conclusões quanto à sua distribuição nas variedades do português, poderá, de algum modo, estar enviesado. Contudo, consideramos que o uso do *corpus TenTen18* como fonte poderá mitigar esta limitação, dado a sua extensão e abrangência de conteúdos.

Naturalmente, as restrições distribucionais que se observa entre os advérbios e os predicados que eles modificam, embora tenham sido parcialmente levadas em conta na descrição das entradas do léxico-gramática, nomeadamente pela sua ilustração com um exemplo característico e claro (isto é, não ambíguo), seja ele retirado do *corpus* ou de exemplos da internet, seja ele construído ou adaptado a partir de exemplos reais, é um trabalho que deverá ser desenvolvido de forma sistemática em futuros trabalhos.

Os próximos passos deste trabalho incluem, assim, a descrição mais precisa e detalhada das propriedades de cada uma das classes sintático-semânticas propostas por Molinier e Levrier (2000). Este trabalho já foi iniciado a partir da proposta de uma classe, que não tinha sido considerada no trabalho de Molinier e Levrier (2000), que descreve os advérbios locativos (ML). Seguiremos com a descrição minuciosa das demais classes, incluindo o desdobramento lexical de algumas entradas, uma vez que, ao decorrer do estudo, identificamos que algumas expressões podem, em determinados casos, ser classificadas em mais de uma classe. Por exemplo, a expressão adverbial *no tempo* possui valor temporal (classe MT), e.g., [O estudante terminou a prova] *no tempo* 'dentro do tempo previsto/destinado'; mas também um valor locativo (classe ML), como no exemplo [As crianças brincam] *no tempo*³⁸ 'ao ar livre', exclusivo da variedade brasileira. Esse desdobramento de sentidos é essencial para diversas aplicações.

A recolha das expressões poderá ser prosseguida sistematicamente em outros recursos, já identificados, como por exemplo o dicionário de da Silva (2013), e que ainda não foram integralmente examinados.

Para algumas construções relativamente produtivas e que, por isso, se prestam mal a uma representação por listagem, e.g., em *DET circunstância: nesta/nessa/naquela circunstância, nestas/nessas/naquelas circunstâncias*, será necessário construir um conjunto bastante apreciável de gramáticas locais, por forma a representá-las adequadamente nos léxicos computacionais existentes, especificamente nos léxicos do sistema



de processamento computacional do português STRING (Hagège et al., 2010; Mamede et al. 2012; Maurício, 2011). Neste momento, estas expressões produtivas são representadas por uma única entrada no léxico-gramática. Pelo contrário, casos já identificados de forte fixidez morfossintática (e.g., *em circunstância alguma/nenhuma* vs. **em circunstâncias algumas/nenhumas*) foram representados no léxico-gramática como entradas independentes, e consideradas construções autônomas daquelas expressões produtivas. A análise dessas “famílias” de expressões e das condições que presidem às variações observadas será objeto de outro trabalho.

Pretendemos ainda determinar a frequência e a distribuição das expressões pelas duas variedades, brasileira e europeia, através de consulta em *corpora* e, se necessário, validação dessas observações com falantes nativos de cada variedade.

Esperamos, com o avanço deste estudo, contribuir para um maior conhecimento dos advérbios compostos em português, de forma a construir uma base sólida para futuros estudos, eventualmente para outras variedades do português.

Agradecimentos

A pesquisa para este trabalho foi parcialmente apoiada pelo programa de doutoramento em Ciências da Linguagem da Faculdade de Ciências Humanas e Sociais da Universidade do Algarve e através da Fundação para a Ciência e a Tecnologia, pelo INESC-ID Lisboa, Human Language Technology Laboratory (Ref. 50021/2021).

Referências

- Baptista, Jorge (1997) *Sermão, tarefa e facada*. Uma classificação das construções conversas *dar - levar*. *Seminários de Linguística* 1, pp. 5–37.
- Baptista, Jorge (1999) *Manhã, tarde, noite*. Analysis of temporal adverbs using local grammars. *Seminários de Linguística* 3, pp. 1–27.
- Baptista, Jorge (2010) *Verba dicendi: A structure looking for verbs*. In Takuya Nakamura, Éric Laporte, Anne Dister & Cédric Fairon (orgs.), *Les tables. La grammaire du français par le menu. Mélanges en hommage à Christian Leclère*. Université Catholique de Louvain & Presses Universitaires de Louvain, pp. 11–20.
- Baptista, Jorge (2013) ViPER: Uma base de dados de construções léxico-sintáticas de verbos do Português Europeu. In *Actas do XXVIII Encontro da APL - Textos Seleccionados*. APL & Colibri, pp. 111–129.
- Baptista, Jorge & Nuno Mamede (2020) *Dicionário gramatical de verbos do português*. Universidade do Algarve Editora.
- Bechara, Evanildo (2009) *Moderna gramática da língua portuguesa*. (37.^a ed.). Nova Fronteira.
- Borillo, Andrée (1976) Adverbs and the modalization of assertion. *French Language* 30, pp. 74–89.
- Català, Dolors (2003) *Les adverbes composés. Approches contrastives en linguistique appliquée*. Tese de doutoramento, Universitat Autònoma de Barcelona.
- Català, Dolors, Jorge Baptista & Cristina Palma (2020) Problèmes formels concernant la traduction des adverbes composés (espagnol/portugais). *Langues & Parole* 5, pp. 67–82. <https://doi.org/10.5565/rev/languesparole.64>
- Costa, João (2008) *O advérbio em português europeu*. Colibri.
- Cunha, Celso & Luís Filipe Lindley Cintra (2017) *Nova gramática do português contemporâneo. de acordo com a nova ortografia* (7.^a ed.). Lexikon Editora Digital.
- da Silva, José (2013) *Dicionário brasileiro de fraseologia* [em linha]. Disponível em http://www.josepereira.com.br/DBF_2013.pdf [consultado em 31/05/2023].
- Fernandes, Gaia (2011) *Automatic disambiguation of -mente ending adverbs in Brazilian Portuguese*. Tese de mestrado, Universidade do Algarve / Universitat Autònoma de Barcelona.



- Folha de São Paulo (1999) *CD-ROM Folha - Edição 99* [Texto integral da Folha de S. Paulo de 1994 a 1998]. Publifolha. Disponível em <http://bd.folha.uol.com.br/cdrom.html>
- Gonçalves, Matilde, Luísa Coheur, Jorge Baptista & Ana Mineiro (2020) Avaliação de recursos computacionais em Português. *Linguamática* 12 (2), pp. 51–68. <https://doi.org/10.21814/lm.12.2.331>
- Gross, Maurice (1975) *Méthodes en syntaxe*. Hermann.
- Gross, Maurice (1981) Les bases empiriques de la notion de prédicat sémantique. *Langages* 63, pp. 7–52.
- Gross, Maurice (1982) Une classification des phrases figées du français. *Revue québécoise de linguistique* 11 (2), pp. 151–185. Presses de l'Université du Québec à Montréal. <https://doi.org/10.7202/602492ar>
- Gross, Maurice (1984) A linguistic environment for comparative romance syntax. In *Papers from the XIIIth Linguistic Symposium on Romance Languages*, (4) 26, pp. 373–446.
- Gross, Maurice (1986) *Grammaire transformationnelle du français : 3 - Syntaxe de l'adverbe*. ASSTRIL.
- Gross, Maurice (1996) Lexicon-Grammar. In Keith Brown & Jim Miller (eds.), *Concise encyclopedia of syntactic theories*. Pergamon, pp. 244–259.
- Guimier, Claude (1996) *Les adverbes du français : Le cas des adverbes en -ment*. Editions Ophrys.
- Hagège, Carolina, Jorge Baptista & Nuno Mamede (2010) Caracterização e processamento de expressões temporais em português. *Linguamática* 2 (1), pp. 63–76.
- Harris, Zellig S. (1991) *Theory of language and information: A mathematical approach*. Clarendon Press.
- Kilgarriff, Adam, Miloš Jakubiček, Jan Pomikalek, Tony Berber Sardinha & Pete Whitelock (2014) PtTenTen: a corpus for Portuguese lexicography. In Tony Berber Sardinha & Telma de Lurdes São Bento Ferreira (eds.), *Working with Portuguese corpora*, pp. 111–130. <https://doi.org/10.5040/9781472593641.ch-006>
- Laporte, Éric, Takuya Nakamura & Stravoula Voyatzi (2008) A French corpus annotated for multiword expressions with adverbial function. In *Language Resources and Evaluation Conference (LREC). Linguistic Annotation Workshop*, pp. 48–51. Disponível em <https://shs.hal.science/halshs-00286541>
- Laporte, Éric, & Voyatzi, Stravoula (2008) An electronic dictionary of French multiword adverbs. In *Language Resources and Evaluation Conference (LREC). Workshop towards a shared task for multiword expressions*, pp. 31–34 Disponível em <https://shs.hal.science/halshs-00286546>
- Laporte, Éric (2015) The science of Linguistics. *Inference: International Review of Science* 1 (2). <https://inference-review.com/article/the-science-of-linguistics>
- Mamede, Nuno, Jorge Baptista, Cláudio Diniz & Vera Cabarrão (2012) STRING: A hybrid, statistical and rule-based natural language processing chain for Portuguese. In *Proceedings of the 10th International Conference on Computational Processing of Portuguese (PROPOR'12)*. Springer-Verlag, Disponível em <http://www.inesc-id.pt/ficheiros/publicacoes/8578.pdf>
- Maurício, Andreia (2011) *Identificação, classificação e normalização de expressões temporais*. Tese de mestrado, Instituto Superior Técnico, Universidade Técnica de Lisboa.
- Mel'čuk, Igor (2023) *General phraseology: Theory and practice*. John Benjamins Pub. Co.
- Molinier, Christian, & Françoise Levrier (2000) *Grammaire des adverbes : Description des formes en -ment*. Librairie Droz.
- Moreno-Ortiz, Antonio, Chantal Pérez-Hernández & Maria Del-Olmo (2013) Managing multiword expressions in a lexicon-based sentiment analysis system for Spanish. In *Proceedings of the 9th Workshop on Multiword Expressions*, pp. 1–10. Disponível em <https://aclanthology.org/W13-1001.pdf>
- Palma, Cristina (2009). *Estudo contrastivo português-espanhol de expressões fixas adverbiais*. Tese de mestrado, Universidade do Algarve.
- Paumier, Sébastien, Franz Guenther, Eric Laporte, Friederike Malchok, Clemens Marschner, Claude Martineau, Cristian Martínez, Denis Maurel, Sebastian Nagel, Alexis Neme, Maxime Petit, Johannes Stiehler & Gilles Vollant (2021) *United 3.3. User Manual*. Paris: Université de Marne-la-Vallée. Disponível em <https://unitexgramlab.org/releases/3.3/man/Unitex-GramLab-3.3-usermanual-en.pdf>



- Pinheiro, Gisele & Sandra Aluísio (2003) *Corpus NILC: Descrição e análise crítica com vistas ao projeto Lacio-Web*. São Paulo: Instituto de Ciências Matemáticas e de Computação – ICMC/USP. Disponível em <http://www.nilc.icmc.usp.br/nilc/projects/lacio-web.htm>
- Ramisch, Carlos (2015) *Multiword expressions acquisition. A generic and open framework*. Springer International Publishing.
- Ranchhod, Elisabete (1983) On the support verbs *ser* and *estar* in Portuguese. *Linguisticae Investigationes* 7 (2), pp. 317–353. <https://doi.org/10.1075/li.7.2.07ran>
- Ranchhod, Elisabete (1990) *Sintaxe dos predicados nominais com estar*. INIC – Instituto Nacional de Investigação Científica.
- Ranchhod, Elisabete (1991) Frozen adverbs—Comparative forms *Como C* in Portuguese. *Linguisticae Investigationes* 15 (1), pp. 141–170. <https://doi.org/10.1075/li.15.1.07ran>
- Raposo, Eduardo (2013) Advérbio e sintagma adverbial. In Eduardo Raposo, Maria F. B. Nascimento, Maria A. C. Mota, Luísa Segura & Amália Mendes (orgs.), *Gramática do português* (Vol. 2). Fundação Calouste Gulbenkian, pp. 1569–1675.
- Rassi, Amanda (2015) *Descrição, classificação e processamento automático das construções com o verbo DAR em Português Brasileiro*. Tese de doutoramento, Universidade Federal de São Carlos.
- Rocha, Paulo Alexandre & Diana Santos (2000) CETEMPúblico: Um corpus de grandes dimensões de linguagem jornalística portuguesa. In *V Encontro para o processamento computacional da língua portuguesa escrita e falada (PROPOR 2000)*. ICMC/USP, pp. 131–140. Disponível em <http://www.linguatca.pt/documentos/RochaSantosPROPOR2000.pdf>
- Rocha, Carlos Alberto & Carlos Eduardo Rocha (2011) *Dicionário de locuções e expressões da língua portuguesa*. LEXIKON Editora.
- Schwab, Artur (1985) *Locuções adverbiais* (2.^a ed). Fundação da Universidade Federal do Paraná.
- Shigeto, Yutaro, Ai Azuma, Sorami Hisamoto, Shuhei Kondo, Tomoya Kose, Keisuke Sakaguchi, Akifumi Yoshimoto, Frances Yung & Yuji Matsumoto (2013) Construction of English MWE dictionary and its application to POS tagging. In *Proceedings of the 9th Workshop on Multiword Expressions*, pp. 139–144. Disponível em <https://aclanthology.org/W13-1021.pdf>
- Shudo, Kosho, Akira Kurahone & Toshifumi Tanabe (2011) A Comprehensive Dictionary of Multiword Expressions. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT'11)*, pp.161–170. Disponível em <https://aclanthology.org/P11-1017.pdf>
- Vaza, Aldina (1988) *Estruturas com nomes predicativos e o verbo-suporte dar*. Dissertação de mestrado, Universidade de Lisboa.
- Wagner Filho, Jorge, Rodrigo Wilkens, Marco Idiart & Aline Villavicencio (2018) The brWaC corpus: a new open resource for Brazilian Portuguese. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, pp. 4339–4344. Disponível em <https://aclanthology.org/L18-1686.pdf>
- Žižková, Hana (2018) Improving compound adverb tagging. In *Recent Advances in Slavonic Natural Language Processing (RASLAN 2018)*, pp. 103–109. Disponível em <https://nlp.fi.muni.cz/raslan/raslan18.pdf>

