



Malationa, malation ou malatiom? A variação denominativa no processo de criação de um glossário bilíngue da área de química de pesticidas

José Victor de Souza^{1*}, Paula Tavares Pinto² e Marcela Marques de Freitas Lima³

¹Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista 'Júlio de Mesquita Filho', Rua Cristóvão Colombo, 2265, 15054-000, São José do Rio Preto, São Paulo, Brasil. ²Departamento de Letras Modernas, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista 'Júlio de Mesquita Filho', São José do Rio Preto, São Paulo, Brasil. ³Departamento de Química e Ciências Ambientais, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista 'Júlio de Mesquita Filho', São José do Rio Preto, São Paulo, Brasil. *Autor para correspondência. E-mail: jose.victor@unesp.br

RESUMO. A presente pesquisa visa a atender às necessidades de um estudo terminológico na área de química de pesticidas, tendo em vista o fato de o Brasil ser um dos países mais produtivos em atividades agrícolas no mundo. O estudo parte do problema tradutório e terminológico da área, previamente observado por Pinto e Lima (2018) e Souza (2019), e propõe o levantamento de termos da área para a elaboração de um glossário português ⇌ inglês. A fundamentação teórica deste trabalho se baseia em estudos sobre Terminologia (Krieger & Finatto, 2004; Finatto, 2004; Bevilacqua, 2013), Química (Abakerli et al., 2003) e Linguística de Corpus (Sardinha, 2000; Tagnin, 2002). Neste artigo, explicamos o processo de criação do *corpus* de estudo, bem como os critérios para seleção dos candidatos a termos em inglês. A fim de ilustrar as ocorrências de variação denominativa durante a seleção dos equivalentes em português, apresentamos os passos utilizados para a seleção do termo mais adequado dentre as variantes vernaculares em português 'malation', 'malationa' e 'malatiom' e mostramos por que 'malation' é a forma escolhida, levando-se em consideração sua frequência absoluta, *Average Reduced Frequency*, coeficiente de variação, número de resultados em pesquisa na *web*, bem como critérios linguísticos, conceituais e normatizadores. Concluímos que a parceria entre ferramentas informatizadas e conhecimento especializado, tanto em linguística quanto em química, trouxe benefícios à discussão que não seriam alcançados somente por meio de cálculos computacionais.

Palavras-chave: química de pesticidas; terminologia; linguística de corpus; agrotóxico; variação interlinguística.

Malationa, malation or malatiom? Denominative variation in the process of creating a bilingual glossary in the area of pesticide chemistry

ABSTRACT. This research aims to meet the needs of a terminological study in the field of pesticide chemistry, since Brazil is one of the most productive countries in agricultural activities in the world. The study starts from the translation and terminological problem of the area, previously observed by Pinto and Lima (2018) and Souza (2019), and proposes the search of terms in the area for the elaboration of a Portuguese ⇌ English glossary. The research is based on the areas of Terminology (Krieger & Finatto, 2004; Finatto, 2004; Bevilacqua, 2013), Chemistry (Abakerli et al., 2003) and Corpus Linguistics (Sardinha, 2000; Tagnin, 2002). In this paper, we demonstrate the process of creating the study corpus, as well as the criteria for selecting candidates for terms in English. In order to illustrate occurrences of denominative variation during the selection of equivalents in Portuguese, we present the steps used for the selection of the most appropriate term among the Portuguese vernacular variants 'malation', 'malationa', and 'malatiom', and show why 'malation' is the form chosen, taking into account its absolute frequency, Average Reduced Frequency, coefficient of variation, number of results in web research, as well as linguistic, conceptual and normative criteria. We concluded that the partnership between computerized tools and specialized knowledge, both in linguistics and in chemistry, brought benefits to the discussion that would not be achieved only through computational calculations.

Keywords: pesticide chemistry; terminology; corpus linguistics; pesticide; interlinguistic variation.

Received on September 18, 2020.

Accepted on August 6, 2021.

Introdução

O Brasil é o terceiro maior exportador agrícola no mundo (Estadão Conteúdo, 2018) e, consequentemente, um dos maiores consumidores de pesticidas mundiais. Por esse motivo, pesquisas sobre os efeitos dos agrotóxicos e propostas mais sustentáveis de seu uso tornaram-se necessárias.

Trazendo foco para a questão terminológica, sabe-se que a linguagem científica, na qual surgem ou são criados os termos, serve de base para a comunicação entre pares (Wüster, 1998; Cabré, 1999; Krieger, 2000; Barros, 2004; Krieger & Finatto, 2004). Entretanto, alguns dos termos empregados na área da química de pesticidas, campo dedicado à pesquisa dos agrotóxicos, são originários do inglês e muitas vezes não possuem termos equivalentes na língua portuguesa (Abakerli et al., 2003; Pinto & Lima, 2018; Souza, 2019). Desse modo, quando utilizados por brasileiros em seus textos sobre a área, esses termos figuram como empréstimos ou traduções que não suprem por completo o significado de seus equivalentes na língua inglesa ou que são pouco convencionalizados. Essa falta de regularidade terminológica prejudica a escrita científica sobre o assunto, pois a torna menos fluida e menos natural uma vez que ainda não há uma descrição dos contextos ideais para uso de cada variante (Bevilacqua, 2013).

Dessa maneira, estamos desenvolvendo pesquisa que pretende contribuir para a discussão e enriquecimento dos estudos terminológicos dessa área de especialidade por meio da elaboração de um glossário bilíngue que servirá tanto para alunos de Química quanto para pesquisadores da área, bem como tradutores, principalmente no sentido de oferecer termos na língua portuguesa que ainda não apresentem consenso de uso no Brasil.

Como recorte desta pesquisa, discutiremos no presente artigo o processo de criação do *corpus* de estudo, seguindo para a explanação dos critérios para seleção dos candidatos a termos em inglês. No que tange à seleção dos equivalentes em português, mostraremos os passos utilizados para a seleção do termo preferido dentre as variantes vernaculares em português ‘malation’, ‘malationa’ e ‘malatium’.

Fundamentação teórica

Neste estudo, partimos de dois pressupostos interligados que servem de base norteadora para a postura aqui assumida em relação à análise linguística do léxico especializado: o primeiro é a concepção de língua como um sistema probabilístico que pode ser estudado por meio de uma abordagem empirista. Empirismo, na linguística, como afirma Sardinha (2000, p. 350), “[...] significa dar primazia aos dados provenientes da observação da linguagem, em geral reunidos sob a forma de um *corpus*”. Por *corpus*, tomamos a seguinte definição:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise. (Sanchez, 1995, p. 8-9 apud Sardinha, 2000, p. 338).

O segundo pressuposto, diretamente ligado ao primeiro, está no tratamento da natureza do termo técnico em relação ao sistema linguístico em que este está inserido. Neste sentido, seguimos a proposta da Teoria Comunicativa da Terminologia (TCT) de Cabré (1999), que leva em consideração os aspectos pragmáticos e comunicativos da linguagem. Dentro do arcabouço da TCT, aceita-se a variação como parte inerente à linguagem especializada, agora observada a partir de níveis de especialização e, dessa maneira, com diferentes níveis de descrição (Krieger, 2000).

Dessa maneira, em consonância com as tendências mais atuais da Terminologia, Finatto (2007, p. 222, grifos da autora) define o termo técnico como uma:

[...] unidade poliédrica [que], ocupando o seu devido lugar nessa hierarquia de recortes do nosso objeto-texto, tem sido foco de maiores atenções, bem sabemos. Esse ‘termo’ cada vez mais é entendido como uma condição especial da palavra, um signo linguístico dotado de significado e significante, e atrelado a uma determinada unidade e corpo de conhecimentos historicamente estabelecidos. Desse modo, terminologias deixam de ser unidades ‘estranhas’ ou ‘artificiais’, índices de uma língua à parte da língua-sistema, e passam a ser vistas como palavras que têm ou adquirem um estatuto peculiar em uma dada situação de comunicação.

Dessa busca pelo reconhecimento da chamada ‘terminologia da realidade’, surge a necessidade de análise de coleções extensivas de textos, os *corpora*, a partir dos quais pode-se retirar amostras da língua em uso.

Desse modo, retomando o primeiro pressuposto de que a língua é um sistema probabilístico empiricamente observável, o segundo pressuposto, dado pela TCT, é o de que, assim como a língua geral, o léxico especializado também está passível de variação (Faulstich, 2001; Freixa, 2005; 2006). Desse encontro, surge a interface da Linguística de Corpus com a Terminologia.

A Linguística de Corpus colabora com o trabalho terminológico, pois, por meio de suas ferramentas, mostra os termos mais prováveis e frequentes de determinada área de especialidade. Logo, a criação de glossários e dicionários com os termos mais recorrentes, e não com os possíveis, gera maior fluência e naturalidade para os textos produzidos ou traduzidos com o seu auxílio (Tagnin, 2002; Bevilacqua, 2013; Krieger & Santiago, 2014).

Além da interface Terminologia-Linguística de Corpus, é indispensável para o nosso trabalho a interface Tradução-Terminologia, a qual é discutida por Barros (2004), Krieger e Finatto (2004) e Paiva (2006; 2009) e Finatto (2007). As autoras, que embasam o trabalho aqui proposto, ressaltam os benefícios da parceria entre as áreas, dentre eles a naturalidade e, principalmente, a aceitabilidade que um texto possui ao ser produzido ou traduzido respeitando o uso de termos e das fraseologias pertencentes à área de especialidade

Sketch Engine: ferramentas utilizadas

A plataforma de análise lexical escolhida para nosso projeto foi o Sketch Engine¹ (SE – Kilgariff et al., 2014), que oferece um serviço na *web* para criação e manutenção de *corpora*, por reunir quase todas as funcionalidades necessárias para a compilação do glossário. A plataforma SE foi utilizada desde a seleção de textos para compor o *corpus* até a seleção de candidatos a termos da área de especialidade. Esta ferramenta apresenta vários utilitários, dentre os quais utilizamos as *Keywords* para o levantamento dos termos do estudo.

Keywords

A funcionalidade *Keywords* identifica palavras-chave no *corpus* de estudo (especializado) ao compará-lo a um *corpus* de referência (língua geral), gerando uma lista de termos simples (no SE, chamada de *Keywords*) e uma lista de termos complexos, que são termos formados por mais de um item lexical (no SE, chamada de *Terms*).

Além disso, o SE utiliza um método chamado *simple maths* (Kilgariff, 2009; Kilgariff et al., 2014), que gera um *keyness score* que ranqueia as palavras do *corpus* a partir da seguinte fórmula:

$$\frac{fpm_{focus} + N}{fpm_{ref} + N}$$

Onde fpm_{focus} representa a frequência relativa (por milhão) de determinada palavra no *corpus* de estudo, e fpm_{ref} representa a frequência relativa (por milhão) de determinada palavra no *corpus* de referência. A frequência relativa (*fpm*), de maneira geral, é representada pela fórmula abaixo:

$$fpm = \frac{numberofhits \cdot 1.000.000}{corpus}$$

Ou seja, para determinada palavra de um *corpus*, seja o de estudo ou o de referência, há um dado *number of hits* (frequência da palavra no *corpus*) que é normatizado em partes por milhão e dividido pelo *corpus size* (o número total de *tokens*, ou itens, do *corpus* em questão).

Já N é o que os autores chamam de *smoothing parameter*, que define a raridade das palavras que vão compor a lista, além de evitar que $fpm_{ref} + N$ resulte em zero, o que impossibilitaria a divisão. O valor recomendado para N , como demonstrado por Kilgariff (2009), pode ser representado por 10^n , onde $\{n \in \mathbb{Z}\}$.

Na interface do usuário no SE (Figura 1), na aba *Advanced*, há um controle deslizante que apresenta as seguintes opções como representantes de 10^n : 0,001 – 0,01 – 0,1 – 1 – 10 – 100 – 1.000 – 10.000 – 100.000 – 1.000.000, onde 1 é a opção *default*, sendo os números à esquerda deste recomendados para priorizar palavras mais raras e os números à direita para palavras mais comuns.

Para nossa pesquisa, definimos que $N = 1$, a opção padrão, seria o *smoothing parameter* adotado, pois valores acima deste geravam listas com palavras pouco específicas, e, em contrapartida, valores abaixo disso resultavam em listas muito específicas, deixando de fora alguns termos já reconhecidos como centrais ao estudo.

¹ <https://www.sketchengine.eu/>

Corpus e subcorpora: compilação manual e automatizada pela web

Neste trabalho, utilizamos um *corpus* de especialidade que foi compilado com a ajuda de especialistas em linguística de *corpus* e em química de pesticidas (Pinto & Lima, 2018). Os textos que figuram neste *corpus* foram selecionados, em inglês e português, consultando-se fontes especializadas e revistas científicas da área. Procuramos montar um *corpus* que fosse representativo da área de especialidade da química de pesticidas, especificamente a subárea dos organofosforados.

Dessa maneira, para dar início à elaboração do glossário, foi utilizado um *corpus* bilíngue comparável. Sendo o *subcorpus* em português, doravante CPt, composto por 342.205 *tokens*, ou seja, palavras no total, e 247.021 *word forms*, que representam o total de palavras diferentes no *corpus*. Os textos deste *corpus* são artigos retirados de revistas científicas originalmente escritos em português, sendo 56 ao todo.

Para a criação do *corpus* em inglês, doravante CEn, experimentamos a ferramenta do SE de criação e ampliação de *corpus* que permite ao usuário selecionar textos da *web* a partir de filtros como *seeds* e URL. As *seeds* têm a função de palavras-chaves, escolhidas pelo usuário, que se unem para criar resultados de pesquisa no buscador Bing da Microsoft®. A ferramenta, então, extrai os textos desses resultados automaticamente. Já o filtro de URL permite que o usuário limite as páginas de onde os textos podem ser retirados. A partir desses parâmetros, a ferramenta coleta somente textos que atendam aos critérios estabelecidos pelo usuário para figurarem no *corpus*.

Para este estudo, como mostra a Figura 2, selecionamos as *seeds*: ‘*organophosphate*’ e ‘*organophosphorus*’, que, como aponta Souza (2019), são palavras utilizadas para referir-se a compostos organofosforados, e ‘*pesticide*’, uma palavra-chave representativa da área. No filtro de URL, delimitamos a quantidade de *links* para no máximo 100, pois queríamos uma quantidade razoável de fontes para que pudéssemos escolher as melhores. Além disso, restringimos os domínios de busca para ‘.edu’ (próprio de *sites* de instituições educacionais).

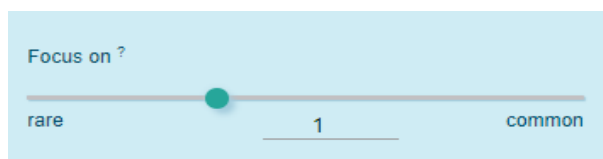


Figura 1. Controle deslizante para definir raridade das palavras ao alterar *N*. Fonte: Captura de tela feita pelos autores.

Figura 2. Compilação de *corpus* a partir de textos da *web* no Sketch Engine. Fonte: Captura de tela feita pelos autores.

Todos esses critérios foram selecionados a fim de aumentar as chances de conseguirmos textos acadêmicos do escopo desejado.

Apesar de o limite estabelecido ter sido 100 *links*, a ferramenta encontrou apenas 50. Verificamos cada um dos *links* encontrados para observar os gêneros dos textos e selecionar aqueles que mais se encaixassem no registro acadêmico. Tipificamos cada um dos textos encontrados como mostra o Tabela 1.

Mantivemos os *links* que continham os gêneros 1) artigo, 2) *dissertation*, 3) *abstract* e 4) relatório, por terem registro acadêmico e descreverem experimentos. Acreditamos, assim, ter conseguido uma proximidade de gênero e registro com os textos de Cpt. Também mantivemos em nosso *corpus* 5) entradas em dicionários especializados, por serem fontes úteis de pesquisa para definições. No total, 21 textos foram inseridos ao *corpus*.

Após a compilação do *subcorpus*, realizamos uma comparação entre as palavras-chave de Cpt e de CEn com os mesmos parâmetros anteriores para verificar o escopo, como pode ser visto nos Figuras 3 e 4.

Tabela 1. Tipologia dos gêneros encontrados nos *links*.

Gênero	Quantidade
Artigo	18
<i>Dissertation</i> (Tese)	5
Centro de informações	8
Capítulo de livro	4
Entrada em dicionário especializado	2
Página de iniciativas	1
Página de laboratório	1
Página de referências	1
<i>Abstract</i>	2
Relatório	1
Entrada em página de biblioteca	1
Notícia	2
Resultado de busca em site de universidade	1
<i>Newsletters</i> arquivados	2
Propaganda	1
Total:	50

Fonte: os autores.

1 organofosforados	11 pesticida	21 quim	31 pesq	41 suscetibilidade
2 l-1	12 µl	22 malationa	32 spe	42 metílico
3 nd	13 acetilcolinesterase	23 intoxicação	33 c18	43 cromatografia
4 µg	14 clorpirifós	24 toxicol	34 cl50	44 chim
5 crossref	15 organofosforado	25 environ	35 tabaci	45 °c
6 inseticida	16 cipermetrina	26 pesticides	36 pesticide	46 pubmed
7 chem	17 temefós	27 quinquiefasciatus	37 oximas	47 lc-ms
8 carbamatos	18 chromatogr	28 agrotóxico	38 acefato	48 biossensores
9 analitos	19 anovar	29 organoclorados	39 orfs	49 homoscedástico
10 ache	20 piretróides	30 acetonitrila	40 vet	50 diclorvós

Figura 3. Amostra das 50 primeiras palavras-chave de Cpt com possíveis equivalentes em CEn destacados. Fonte: os autores.

1 parathion	11 LC50	21 methyl	31 NL	41 kcat
2 LD50	12 OP	22 immobilize	32 CL	42 Chem
3 OPH	13 VAP	23 Toxicity	33 Ca-alginate	43 Organophosphorus
4 organophosphate	14 APCI	24 SH-SAW	34 toxicity	44 bioconcentration
5 DETP	15 adsorption	25 no.	35 biosensor	45 spectra
6 paraoxon	16 insecticide	26 Dietary	36 pesticide	46 demeton-S
7 chlorferon	17 AChE	27 BPA-HMTS	37 toxicological	47 OPAA
8 Q1MS	18 hydrolysis	28 mM	38 Dermal	48 mL
9 analyte	19 carbamate	29 hydrolyze	39 Raushel	49 Mallard
10 organophosphorus	20 ppm	30 ESI	40 Formulations	50 acaricide

Figura 4. Amostra das 50 primeiras palavras-chave de CEn (*Web*) com possíveis equivalentes em CPT destacados. Fonte: os autores.

Tendo em vista a quantidade de possíveis equivalentes tradutórios, além da presença de palavras que circulam pelo mesmo campo semântico (*toxicity* [CEn] e intoxicação [CPT] / *pesticide* [CEn] e agrotóxicos [CPT]), acreditamos que o *subcorpus* em língua inglesa compilado está adequado aos objetivos de nossa pesquisa, contendo 365.178 *tokens* e 240.285 *word types*.

Seleção de termos

Tendo em vista que o objetivo da presente pesquisa é a produção de um glossário com termos da área de química de pesticidas, iniciamos a seleção gerando uma lista de termos simples de CEn, utilizando a função *Keywords* com os parâmetros já mencionados, tendo o ‘enTenTen13’ como *corpus* de referência. Decidimos iniciar nossa busca terminológica a partir do *corpus* em inglês devido ao fato de a grande maioria dos termos surgirem primeiramente nesta língua. Esse fato se dá graças à inovação em pesquisas em pesticidas estar centralizada nos países mais desenvolvidos. Além disso, na área da química, os parâmetros terminológicos em inglês são quase sempre o ponto de partida para tradução de nomes de substâncias e de processos técnicos para o português.

Para a seleção dos termos simples, separamos os candidatos entre ‘centrais’ e ‘periféricos’. ‘Centrais’ caracterizam-se por termos diretamente ligados à pesquisa com pesticidas organofosforados, sendo assim, os nomes dessas substâncias. Já ‘periféricos’ são candidatos coadjuvantes, como termos que indicam técnicas analíticas (instrumentação) ou fenômenos relacionados à análise dos pesticidas organofosforados. Foram excluídos itens lexicais considerados pouco específicos para a química de pesticidas, por exemplo, procedimentos laboratoriais comuns a qualquer campo da pesquisa química.

Tal seleção foi feita com a ajuda da especialista, cuja participação foi essencial para um desenvolvimento mais rápido da seleção de candidatos a termos, bem como definição de critérios para tal.

Após selecionados os termos centrais, utilizamo-los como palavras-chave na procura por combinações de palavras por meio da ferramenta *Word Sketch* a fim de encontrarmos candidatos a termos complexos. A imagem mostrada na Figura 5 ilustra o procedimento adotado, onde a cor verde representa os termos centrais e a amarela os termos periféricos. Além disso, em azul-claro estão marcadas as siglas que denotam variações possíveis de termos. Na coluna ‘Termos complexos’, são destacados com a cor de fonte azul possíveis candidatos a termos complexos para o glossário, marcados durante discussão com a especialista.

Conforme explicado anteriormente, em verde temos os termos ‘*parathion*’, ‘*organophosphate*’, ‘*paraoxon*’ e ‘*organophosphorus*’, que representam os termos centrais, ou seja, os mais representativos no *corpus* e mais específicos da área de química de pesticidas. Já em azul-claro estão destacadas as siglas utilizadas pelos autores da área, como ‘OPH’ e ‘OP’, que denotam variações possíveis de termos e que deverão também compor o glossário da área. Em amarelo temos ‘LD50’ e ‘LC50’, termos marginais relacionados à análise dos níveis de letalidade de uma substância. Em relação aos termos complexos, selecionamos os possíveis candidatos em azul, que constam na Figura 5, como ‘*organophosphate pesticides*’, ‘*organophosphate insecticide*’ e ‘*organophosphate poisoning*’.

Term	Score	Freq	Ref freq	Rel freq	Rel ref freq	Termos complexos
parathion	1318.59	507	1220	1388.364	0.054	methyl parathion; parathion-methyl; parathion adsorption; parathion sorption; adsorption of parathion
LD50	1064.52	437	2843	1196.677	0.125	LD50 value
OPH	996.15	368	287	1007.728	0.013	
organophosphate	850.22	356	3359	974.867	0.148	organophosphate pesticides; organophosphate insecticide; organophosphate poisoning; organophosphate compounds; organophosphate exposure
DETP	786.33	287	17	785.918	0.001	
paraoxon	753.55	276	98	755.796	0.004	methyl paraoxon; paraoxon concentration; paraoxon toxicity; paraoxon hydrolysis;
chlorferon	732.15	267	0	731.15	0	
Q1MS	723.94	264	0	722.935	0	
analyte	700.64	316	5375	865.331	0.236	
organophosphorus	693.49	259	549	709.243	0.024	organophosphorus insecticide; organophosphorus compounds; organophosphorus pesticides; organophosphorus nerve agents; organophosphorus systemic nematocide; organophosphorus neurotoxins; organophosphorus herbicide; organophosphorus poisoning; organophosphorus plant-growth regulator
LC50	687.13	260	855	711.982	0.038	LC50 values; Dietary LC50
OP	624.42	560	33127	1533.499	1.457	OP pesticides; OP compounds; OP exposure; OP poisoning; OP neurotoxins; OP concentrations; OP metabolites; OP degrading enzymes

Figura 5. Amostra da seleção de candidatos a termos. Fonte: Captura de tela feita pelos autores.

Para alguns termos selecionados, já se iniciou a identificação dos equivalentes tradutórios em CPt para a criação do glossário e criação das fichas terminológicas, na qual constarão o equivalente tradutório e os contextos de onde foram retirados os termos e suas traduções. Dado que CPt é comparável em relação a CEn, mas não paralelo, a busca por equivalentes dependeu da intuição e repertório do tradutor-pesquisador e das sugestões da especialista. A seguir, relatamos algumas ocorrências que dizem respeito aos equivalentes encontrados, discutindo um caso de variação denominativa ilustrativo.

Variação denominativa

A tradução dos nomes usuais² para o português tende a variar no que diz respeito a dígrafos e sufixos. Dessa maneira, por meio da linguagem de *regular expressions* no SE, pudemos substituir as partes variáveis de determinado item lexical por (.*), que permite à ferramenta procurar quaisquer caracteres em qualquer quantidade na posição indicada. Na Figura 6, ilustramos o uso desse recurso com o termo ‘malathion’.

Conforme demonstrado, o uso das *regular expressions* revelou três equivalentes utilizados em português, sendo o último deles um empréstimo.

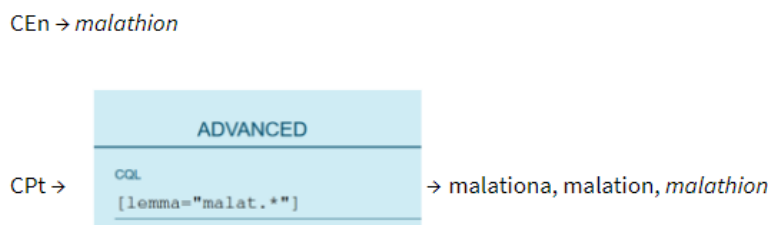


Figura 6. Corpus Query Language (CQL) para a busca de variação denominativa. Fonte: Captura de tela feita pelos autores.

Para a seleção da forma mais adequada, em vez de utilizarmos apenas a frequência absoluta (FA), que não leva em consideração a densidade das ocorrências no *corpus*, utilizamos o recurso estatístico *Average Reduced Frequency* (ARF), também disponível no SE. Este recurso gera uma nova frequência para o termo, considerando ocorrências muito concentradas como apenas uma, dando, assim, uma ideia de quão bem distribuído o termo está e diminuindo as probabilidades de um texto que use muito determinado termo adulterar os resultados.

Ainda utilizando o termo ‘malathion’ como exemplo, pudemos notar que, enquanto, em termos de FA, os equivalentes ‘malation’ e ‘malationa’ apresentam, respectivamente, 70 e 97 ocorrências (a variante estrangeira ‘malathion’ apresenta somente 25 *hits* no *corpus*, sendo descartada como preferencial já nesta etapa), o ARF de ‘malation’ é 16,7 e o de ‘malationa’ é 12,9. Dessa maneira, se uma observação que levasse em consideração apenas FA escolheria ‘malationa’ como termo favorito, uma análise que leve o ARF em consideração aponta para uma inversão nessa preferência em um *corpus* homogêneo ideal. Além disso, fora de nosso *corpus* de estudo, numa breve busca on-line, foi possível encontrar uma quarta forma equivalente: ‘malatiom’.

Quando o ARF não era suficiente para determinar o equivalente favorito, lançamos mão do coeficiente de variação (CV), cálculo que mostra o quão bem distribuído no *corpus* um item está (quanto mais próximo de 0 o resultado for, mais bem distribuído o termo está). Entretanto, este recurso não está disponível no SE. Para isso, utilizamos outro programa, o #LancsBox3. Usando ainda ‘malation’ e ‘malationa’ como exemplos, notou-se a reiteração de que este (CV: 3,18) está mais bem distribuído que aquele (4,84) no *corpus*. Por fim, quando tanto o ARF quanto o CV não eram conclusivos, consultamos a *web* utilizando a ferramenta de busca do Google. Esta decisão foi embasada em estudos que defendem o uso da *web* como um *corpus* (Xatara, 2008; Valença & Sabino, 2016).

Tendo por base Faulstich (2001), pode-se afirmar que essas quatro formas são variantes resultantes do processo de empréstimo linguístico. Itens lexicais emprestados de outra língua, como é o caso de ‘malathion’, ao serem incorporados à língua vernacular podem manter a forma estrangeira (‘malathion’) ou “[...] gerar uma forma adaptada às regras morfofonêmicas do PB” (Faulstich, 2001, p. 35). No caso do termo aqui em questão, há mais de uma forma adaptada (‘malation’, ‘malationa’ e ‘malatiom’). Dentre essas formas, é possível identificar o que a autora chama de variantes terminológicas gráficas, tipo de variação que reside na forma escrita de um termo, o que é o caso dos itens ‘malation’ e ‘malatiom’ que, por terem a mesma pronúncia em PB, caracterizam variação gráfica. Abaixo, trechos retirados do *corpus* de estudo, evidenciando o uso das variantes ‘malation’ e ‘malationa’. Além disso, um trecho retirado da *web*, ilustrando o uso da variante gráfica ‘malatiom’.

Paration, paraoxon, carbofuran, aldicarb e ‘malation’ são exemplos de pesticidas que têm sido detectados por este tipo de biossensor, em limites que alcançam 0,03 ppb. (Marques & Yamanaka, 2008, s.p., grifo nosso).

O inseticida ‘malationa’ em calda oleosa é utilizado no controle do *Aedes aegypti* e a sua aplicação é feita por meio de nebulização. Essa atividade exige o uso de equipamento de proteção individual (EPI) pelos aplicadores (Leme, Papini, Vieira, & Luchini, 2014, s.p., grifo nosso).

² Nomes utilizados para se referir a pesticidas de maneira simples e resumida, muitas vezes destacando seu principal grupo molecular, em detrimento dos nomes padronizados pela IUPAC (União Internacional de Química Pura e Aplicada), que tendem a ser precisos e invariáveis, porém, o custo disso é a complexidade aparente e extensão dos termos.

³ <http://corpora.lancs.ac.uk/lancsbox/>

Carbaril e ‘malatiom’ foram seletivos em favor de *P. versicolor versicolor* (18,6 e 6,7% de mortalidade, respectivamente), e deltametrina em favor de *P. versicolor versicolor* (3,7% d mortalidade) e *P. exigua* (11,9% de mortalidade) (Gonring, Picanço, Moura, Bacci, & Bruckner, 1999, s.p., grifo nosso).

Abakerli et al. (2003), numa interpretação da norma 257 (*International Standards Organization* [ISO], 1988), que diz respeito à criação de nomes comuns para pesticidas em inglês, recomendam que a tradução do sufixo *-thion* seja *-tiona*, o que resultaria em um dos equivalentes encontrados: ‘malationa’. Entretanto, um problema observado nessa tradução está no fato de o segmento *-ona* (contido em *-tiona*) é tradicionalmente conhecido como identificador da função cetona em uma substância. Dessa maneira, pelo fato de o pesticida em questão não ser uma cetona e para evitar que o nome leve a uma interpretação errônea da função química na substância, as formas variantes ‘malation’ e ‘malatiom’ podem ser optadas por alguns pesquisadores e tradutores. Contudo, tendo em vista que uma pesquisa no Google pelo item ‘malatiom’ gera por volta de 1.960 resultados, número baixo em comparação aos outros termos, que geram resultados mais expressivos (‘malation’, por volta de 240.000, e ‘malationa’, por volta de 23.000), não a consideramos como a mais adequada.

Como fruto dessa discussão, surgiu o seguinte modelo de ficha terminológica (Tabela 2), no qual indicamos o tipo de variante denominativa (cf. Faulstich, 2001) e suas possíveis causas (cf. Freixa, 2006). À direita da Tabela 2, o termo em inglês, e à esquerda os seus equivalentes em português, com indicação do gênero do item lexical. Além disso, são mostrados o ARF entre colchetes e o CV entre parênteses quando pertinente. Os contextos de uso, bem como sua respectiva fonte, ainda serão mais bem discutidos e selecionados para o produto final. Por fim, para a forma não normatizada, acrescentamos notas com recomendações.

Tabela 2. Ficha terminológica de *malathion*.

variante competitiva estrangeira causas interlinguísticas	
‘malathion’ m fonte: autor (ano) contexto: a definir	
variantes competitivas vernaculares causas interlinguísticas	
‘malathion’ fonte: autor (ano) contexto: a definir	‘malation’ m ‘[16,7] (3,18)’ fonte: autor (ano) contexto: a definir
variante concorrente gráfica ‘malatiom’ m fonte: autor (ano) contexto: a definir	
‘malationa’ f ‘[12,9] (4,84)’ fonte: autor (ano) contexto: a definir	
notas: forma não normatizada, recomenda-se ‘malation’	
Fonte: os autores.	

Considerações finais

Neste artigo, apresentamos o processo de seleção de termos da área de química de pesticidas para compor um glossário bilíngue inglês ⇔ português. Para tanto, apresentamos os pressupostos teórico-metodológicos que embasam a pesquisa que vem sendo desenvolvida de forma interdisciplinar ao contar com pesquisadores das áreas de linguística e química. Faz-se relevante enfatizar que, por ser a química de pesticidas uma área que ainda conta com pouca produção acadêmica no Brasil, muitos autores têm esbarrado na dificuldade de selecionar os termos adequados para a escrita dos textos em português e acabam empregando empréstimos, ao manterem o termo na língua inglesa, ou variando o emprego na língua portuguesa, sem levar em conta as normas da IUPAC ou as diretrizes para a tradução dos termos. Por esse motivo, apresentamos neste artigo a primeira etapa de seleção de termos de química de pesticidas, que contou com a compilação de um corpus em inglês, levantamento dos termos a partir de palavras-chave, elucidando um caso de variação denominativa em português com os equivalentes ao termo ‘*malathion*’.

O trabalho conjunto com a especialista, com o auxílio do arcabouço teórico-metodológico da linguística de *corpus*, mostrou-se enriquecedor para ambos os lados. Para o terminólogo, ganha-se com a testagem rápida de hipóteses para variação e até mesmo a compreensão dos conceitos na área de especialidade. Para a especialista, uma oportunidade de conscientizar-se de como a terminologia de sua área está organizada, compreendendo melhor os fenômenos linguísticos que envolvem a criação e tradução de termos. Em suma, parece-nos que o trabalho terminológico quando somente feito pelo linguista ou somente pela especialista da área careceria de pontos de vistas essenciais para os resultados aqui obtidos.

Como encaminhamento, estamos em fase de seleção de equivalentes no corpus em língua portuguesa, assim como contextos de uso empregados pela comunidade científica em foco, com o intuito de elucidar a questão e apresentar um glossário harmonizado e de fácil manuseio para profissionais e estudantes de química, assim como demais interessados, como tradutores especializados.

Agradecimentos

Agradecemos o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), processo nº 2019/14752-0.

Referências

- Abakerli, R. B., Fay, E. F., Rembischevski, P., Vekic, A. M., Godoy, K., Maximiano, A. A., & Bonifácio, A. (2003). Regras para nomenclatura dos nomes comuns dos agrotóxicos. *Pesticidas: Revista de Ecotoxicologia e Meio Ambiente*, 13, 29-36. DOI: <http://dx.doi.org/10.5380/pes.v13i0.3162>
- Barros, L. A. (2004). *Curso Básico de Terminologia*. São Paulo: Editora da USP.
- Bevilacqua, C. R. (2013). Por que e para que a linguística de *corpus* na terminologia. In S. Tagnín & C. R. Bevilacqua (Orgs.), *Corpora na terminologia* (p. 11-45). São Paulo, SP: HUB.
- Cabré, M. T. (1999). *La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona, ES: Institut Universitari de Lingüística Aplicada.
- Estadão Conteúdo. (2018, 17 de setembro). Brasil passa a ser 3º maior exportador agrícola, mas clima ameaça futuro. *Globo Rural*. Recuperado de <https://revistagloborural.globo.com/Noticias/Agricultura/noticia/2018/09/globo-rural-brasil-passa-a-ser-3o-maior-exportador-agricola-mas-clima-ameaca-futuro.html>
- Faulstich, E. (2001). Aspectos de terminologia geral e terminologia variacionista. *Tradterm*, 7, 11-40. DOI: <https://doi.org/10.11606/issn.2317-9511.tradterm.2001.49140>
- Finatto, M. J. B. (2004). Terminologia e linguística de corpus: Da perspectiva enunciativa aos novos enfoques do texto técnico-científico. *Letras de Hoje*, 39(4), 97-106.
- Finatto, M. J. B. (2007). Exploração terminológica com apoio informatizado: diálogos entre terminologia e linguística de corpus. In M. Lorente, R. Estopà, J. Freixa, J. Martí, & C. Tebé (Eds.), *Estudis de lingüística i lingüística aplicada en honor de M. Teresa Cabré: volum 2. De deixebles* (p. 221-230). Barcelona, ES: Institut Universitari de Lingüística Aplicada.
- Freixa, J. (2005). Variación terminológica: ¿Por qué y para qué? *Meta : Journal des Traducteurs*, 50(4), 1-13. DOI: <https://doi.org/10.7202/019917ar>
- Freixa, J. (2006). Causes of denominative variation in terminology: a typology proposal. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 12(1), 51-77. DOI: <https://doi.org/10.1075/term.12.1.04fre>
- Gonring, A. H. R., Picanço, M., Moura, M. F., Bacci, L., & Bruckner, C. H. (1999). Seletividade de inseticidas, utilizados no controle de *Grapholita molesta* (Busch) (Lepidoptera: Olethreutidae) em pêssego, a Vespidae predadores. *Anais da Sociedade Entomológica do Brasil*, 28(2), 301-306. DOI: <https://doi.org/10.1590/S0301-80591999000200013>
- International Standards Organization [ISO]. (1988). *ISO 257:2018. Pesticides and other agrochemicals: principles for the selection of common names* (2nd ed.). Geneva, CH: ISO.
- Kilgariff, A. (2009). Simple maths for keywords. In M. Mahlberg, V. González-Díaz, & C. Smith (Eds.), *Proceedings of Corpus Linguistics Conference CL2009*. Liverpool, GB: University of Liverpool.

- Kilgarrieff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... Suchomel, V. (2014). The sketch engine: ten years on. *Lexicography*, 1, 7-36. DOI: <https://doi.org/10.1007/s40607-014-0009-9>
- Krieger, M. G. (2000). Terminologia revisitada. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 16(2), 209-228. DOI: <https://doi.org/10.1590/S0102-44502000000200001>
- Krieger, M. G., & Finatto, M. J. B. (2004). *Introdução à terminologia: teoria e prática*. São Paulo, SP: Contexto.
- Krieger, M. G., & Santiago, M. S. (2014). Estudos de terminologia para a tradução técnica. *Revista de Letras*, 2(33), 42-52.
- Leme, T. S., Papini, S., Vieira, E., & Luchini, L. C. (2014). Avaliação da vestimenta utilizada como equipamento de proteção individual pelos aplicadores de malationa no controle da dengue em São Paulo, Brasil. *Cadernos de Saúde Pública*, 30(3), 567-576. DOI: <https://doi.org/10.1590/0102-311X00144912>
- Marques, P. R. B. O., & Yamanaka, H. (2008). Biossensores baseados no processo de inibição enzimática. *Química Nova*, 31(7), 1791-1799. DOI: <https://doi.org/10.1590/S0100-40422008000700034>
- Paiva, P. T. P. (2006). *Estudo em corpora de traduções e três glossários bilíngues nas subáreas de anestesiologia, cardiologia e ortopedia* (Dissertação de Mestrado). Universidade Estadual Paulista, São José do Rio Preto.
- Paiva, P. T. P. (2009). *Uma investigação de traduções de textos da área médica sob a luz dos estudos da tradução baseados em corpus* (Tese de Doutorado). Universidade Estadual Paulista, São José do Rio Preto.
- Pinto, P. T., & Lima, M. F. (2018). A tradução na área de química orgânica: da adaptação à tradução literal. *Estudos Linguísticos (São Paulo, 1978)*, 47(2), 573-585. DOI: <https://doi.org/10.21165/el.v47i2.2050>
- Sardinha, T. B. (2000). Linguística de Corpus: histórico e problemática. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, 16(2), 323-367. DOI: <https://doi.org/10.1590/S0102-44502000000200005>
- Souza, J. V. (2019). A questão terminológica dos organofosforados na química de pesticidas: uma abordagem baseada em corpus. *Estudos Linguísticos (São Paulo, 1978)*, 48(3), 1620-1638. DOI: <https://doi.org/10.21165/el.v48i3.2270>
- Tagnin, S. E. O. (2002). Os corpora: instrumentos de auto-ajuda para o tradutor. *Cadernos de Tradução*, 1(9), 191-219. DOI: <https://doi.org/10.5007/%25x>
- Valença, E. M., & Sabino, M. A. (2016). O uso da web como corpus em pesquisas fraseológicas: uma prática prejudicial ou um recurso valioso? *Calidoscópio*, 14(3), 480-488. DOI: <https://doi.org/10.4013/cld.2016.143.11>
- Wüster, E. (1998). *Introducción a la teoría general de la terminología y a la lexicografía terminológica*. Barcelona, ES: PPU.
- Xatara, C. M. (2008). A web para um levantamento de frequência. In J. S. Magalhães, & L. C. Travaglia (Orgs.), *Múltiplas perspectivas em linguística* (p. 770-777). Uberlândia, MG: UFU.