

# EXPRESSÕES ANUNCIADORAS DE PARÁFRASE COMO CARACTERÍSTICA DE GÊNEROS TEXTUAIS: DO MANUAL DIDÁTICO DE QUÍMICA À LEGISLAÇÃO AMBIENTAL BRASILEIRA

Maria José B. FINATTO  
Susana de AZEREDO  
Edna Regina Hornes de LIMA  
(Universidade Federal do Rio Grande do Sul)

*ABSTRACT: This paper deals with the presence and the use of EAP (Expressões Anunciadoras de Paráfrase, paraphrase introducing expressions) in Chemistry academic texts. The aim is to provide support to identify factors that contribute to comprehension difficulties of this kind of text. First, this research describes the frequency and distribution of EAP in a corpus formed by two General Chemistry academic manuals translated into Portuguese. Afterwards, the translation of EAP is discussed by comparing the original text in English to the translated text in Portuguese. This research also compares Chemistry manuals to other textual genres by analyzing the presence of EAP aiming to evaluate how much the presence of EAP characterizes a textual genre. The aim is to provide a means of gaining a better understanding of reading comprehension difficulties of a didactical material largely used by academic students of Chemistry (sponsored by CNPq).*

*KEYWORDS: Expressões Anunciadoras de Paráfrase, Educação Química, Lingüística de Corpus, Terminologia, Tradução.*

## 1. Introdução

O projeto TextQuim do Instituto de Letras e a Área de Educação Química (AEQ) do Instituto de Química da UFRGS realizam uma série de estudos que têm buscado reconhecer diferentes perfis da linguagem da Química em uso no Brasil. São explorados temas de Educação e de Lingüística Aplicada, especialmente os que dizem respeito ao texto científico. A parceria Letras-Química foi inicialmente motivada por relatos esparsos, mas constantes, sobre dificuldades de compreensão de leitura de textos acadêmicos de Química por parte estudantes brasileiros do ensino superior.

Em meio uma tal associação, nossa apresentação visa descrever uma parcela das características que temos percebido no texto didático de Química em português. Como a grande maioria do texto didático brasileiro dessa área é fruto de tradução do inglês, trataremos aqui também de identificar algumas características dos gêneros textuais *texto didático* e *texto traduzido*, partindo, especificamente, da presença de expressões anunciadoras de paráfrase (EAP) em um *corpus* de estudo. Esse *corpus* é integrado por uma amostra de manuais acadêmico-didáticos de Química Geral traduzidos do inglês para o português.

Ao propormos uma caracterização de gênero textual recorrendo ao auxílio dos preceitos e métodos da Lingüística de Corpus e da Terminologia de perspectiva textualista<sup>1</sup>, interessa-nos descrever em que medida a presença maior ou menor de EAP contribuiria para singularizar o gênero *manual acadêmico didático de Química Geral traduzido* frente a outros gêneros. Para melhor perceber o que se passa em Química, utilizamos também alguns *corpora* de contraste. Além do *Corpus Possamai*, reunido por Possamai (2004), recorreremos também ao ACERVO TERMISUL de legislação brasileira. Esse *corpus*, disponível em [www.ufrgs.br/termisul](http://www.ufrgs.br/termisul) (veja “Acervo”), é um recorte da Base LEGIS, um conjunto de

<sup>1</sup> A Terminologia é uma área de estudos que se ocupa dos diferentes fenômenos da comunicação técnico-científica. Avançando dos termos para os textos, temos a sua perspectiva textualista. Menos centrada na observação de termos técnicos e mais preocupada, visa descrever e explicar a configuração da linguagem especializada concretizada em textos.

*corpora* formado pelos textos da Constituição Brasileira, legislação ambiental brasileira, Código de Defesa do Consumidor e Estatuto da Criança e do Adolescente.

O ponto de partida para o reconhecimento da linguagem da Química, frente à linguagem da Informática, linguagem legislativa e linguagem do jornalismo de divulgação científica para leigos, será a verificação da presença das EAP OU SEJA, ISTO É e EM OUTRAS PALAVRAS. Especialmente importantes no texto didático, sua função é inter-relacionar um dito e sua reformulação, geralmente antecedendo uma explicação ou esclarecimento que visa a garantir a compreensão de um determinado termo ou tópico por parte do leitor.

Esta apresentação se dividirá em duas partes. A primeira procura fazer uma observação de EAP em manuais acadêmicos didáticos de português. Em seguida, oferece um contraponto entre o texto traduzido e texto original em inglês a fim de observar a origem das EAP em português. A segunda apresenta oferece um contraste entre manuais didáticos de Química e outros gêneros textuais, também via observação da presença de EAP. Para dar conta de nossos objetivos, reunimos, neste trabalho, perspectivas teóricas da Linguística de *Corpus* (Berber Sardinha, 2004), da Terminologia Textual, estudos de Gênero Textual (Swales, 1990), Tradução (Rónai, 1981, 1987) e Linguística da Enunciação (Flores e Teixeira, 2005). Essas referências, por questão de espaço, não serão aqui comentadas. Entretanto, à guisa de uma brevíssima situação teórica, vale dizer que reconhecer uma enunciação em ciências é um dos principais elementos que nos move.

## 2. Materiais e Métodos

Os textos do *corpus* de estudo deste trabalho, como já dissemos, representam uma amostra do *corpus* do projeto TextQuim,<sup>2</sup> o qual inclui, entre outros textos, manuais acadêmicos de Química Geral. Os manuais A e B (em português), bem como suas versões originais, os manuais A<sub>(To)</sub> e B<sub>(To)</sub> (em inglês). Estão detalhados a seguir.

### Em português:

**Manual A** – ATKINS, Peter & Jones, Loretta, *Princípios de Química, questionando a vida moderna e o meio ambiente*, 1.ed, Porto Alegre: Artmed, 2002. 914p.

**Manual B** – MAHAN, Bruce M. & MYERS, Rollie J. *Química, um curso universitário*, 4.ed, São Paulo: Edgard Blücher, 1995. 582p.

### Em inglês – texto original (To)<sup>3</sup>

**Manual A<sub>(To)</sub>** - ATKINS, Peter; JONES, Loretta. *Chemical Principles, The Quest for Insight*. First Printing. United States of America: Ed. Freeman, 1998.

**Manual B<sub>(To)</sub>** - MAHAN, Bruce M.; MYERS, Rollie J. *University Chemistry*. Forth Edition. United States of America: World Student Series Edition, 1987.

A dimensão do *corpus* de estudo de Química, nesse segmento de gênero, em português, é de 185.887 *tokens*; em inglês é de 183.350 *tokens*.

---

<sup>2</sup> O *corpus* do projeto TextQuim, parcialmente disponível em [www.ufrgs.br/textquim](http://www.ufrgs.br/textquim), não é composto única e exclusivamente de manuais acadêmicos didáticos de Química Geral. Há também textos de artigos científicos, textos de popularização, entre outros tantos, todos relacionados com Química. Neste trabalho tratamos e descrevemos apenas a parte do *corpus* TextQuim relacionada com manuais acadêmicos didáticos de Química Geral traduzidos do inglês para o português.

<sup>3</sup> To abrevia Texto Original.

Desses materiais, tanto em português como em inglês, foi necessário selecionar uma amostragem, visto que a totalidade do texto pareceu-nos um impedimento em livros com até 1.200 páginas.<sup>4</sup> Os capítulos/segmentos selecionados para exame foram apontados por professores entrevistados em pesquisa da Área de Educação Química (Silva, Eichler, Del Pino, 2003) e correspondem aos temas mais importantes e de maior convergência conceitual em Química Geral. São eles: *Equilíbrio Químico*, *Equilíbrio Iônico*, *Ligação Química* e *Termodinâmica*.

Para um contraste entre o *corpus* de estudo e *corpora* de diferentes gêneros textuais, a fim de observar se a presença de EAP poderia caracterizar uma especificidade de gênero, utilizamos cinco *corpora* de contraste (CC), identificados como CC1, CC2, CC3, CC4 e CC5. Esses corpora estão detalhados a seguir.

<b>Corpus</b>	<b>Especificação</b>	<b>Dimensão</b>
CC1 (Pilla)	Capítulo de <i>Termodinâmica</i> de um manual de Físico-Química originalmente escrito em português. (1979)	53.100 <i>tokens</i>
CC2 (Química Nova)	Artigos científicos de Química, retirados da revista Química Nova, ano 2004.	427.632 <i>tokens</i>
CC3 (SuperInteressante)	Textos de popularização, retirados da revista SuperInteressante. Ano: 1987-2002	59.585 <i>tokens</i>
CC4 (Corpus Possamai)	Artigos científicos de Informática, apresentados nos Congressos da Sociedade Brasileira de Computação em 2002 e 2003.	1.287.260 <i>tokens</i>
CC5 (Base LEGIS)	Textos legislativos: Legislação Ambiental Brasileira, Estatuto da Criança e do Adolescente, Código de Defesa do Consumidor e Constituição Federal Brasileira	2.037.844 <i>tokens</i>

A ocorrência de EAP no *corpus* de estudo e nos *corpora* de contraste foi observada com auxílio do programa *Wordsmith Tools*, que seleciona e visualiza os contextos.

Em um primeiro momento, o corpus em português precisava estar devidamente preparado para que pudesse ser lido pelo programa *Wordsmith Tools*. Isso exigiu que os textos fossem escaneados, corrigidos e transformados de arquivo .doc para arquivo .txt. Além disso, também foi necessário retirar tabelas, figuras e suas legendas, fórmulas que ficavam em caixas de texto separadas do próprio texto e a lista de exercícios que figura ao final de cada capítulo. No caso das tabelas, quando retiradas, colocamos a palavra tabela entre parênteses: (tabela); o mesmo foi feito com as figuras e com as fórmulas. Indicamos no local: (figura) e

<sup>4</sup> A totalidade parecia um impedimento pois os textos não estavam em formato digital. Considerar a totalidade envolveria escanear e revisar a digitalização de todas as obras. Assim, foi uma necessidade fazer uma escolha de alguns capítulos que valessem como uma amostra representativa da totalidade do texto de cada obra.

(fórmula). Essa marcação permite visualizar, no caso dos contextos coletados no *WordSmith Tools*, onde estavam, originalmente, esses itens. Além disso, foi necessário alinhar o *corpus* em português com o *corpus* em inglês. Isso exigiu que os textos originais em inglês também fossem escaneados, corrigidos e transformados de arquivo .doc para .txt.

Após esses estágios de preparação dos *corpora*, utilizamos a ferramenta *Concordance* do *software WordSmith Tools* para o início da nossa coleta de dados que, inicialmente, limitou-se somente ao texto em português. Com a ajuda da *Concordance*, digitávamos a EAP e a ferramenta nos retornava tanto a quantidade do elemento no texto quanto os contextos em que aparecia.<sup>5</sup> Assim, registrávamos as ocorrências em tabelas, para que pudessem ser, posteriormente, colocadas em gráficos, e armazenávamos os contextos encontrados em um arquivo devidamente identificado.

As quantidades de EAP encontradas no *corpus* foram registradas em tabelas que armazenavam o *total bruto*, o *total percentual* e o *total por dez mil*. O *total bruto* é a quantidade de EAP obtida na busca no programa *WordSmith Tools*. O *total percentual* corresponde à quantidade de EAP a cada 100 palavras do *corpus*. Esse valor é obtido através de aplicação de uma regra de três simples, que leva em conta o total de *tokens* do *corpus* e o total de EAP encontrado. Por fim, temos o *total por dez mil*. Esse valor representa o total de EAP a cada dez mil palavras do *corpus*. Para encontrar esse valor, multiplica-se o *total percentual* por 100<sup>6</sup>.

A coleta dos contextos seguiu alguns critérios. Levamos em conta que nosso enfoque é textual e que a EAP é um conector que permite um movimento de retroação e progressão no texto. Assim, nosso contexto não compreendia somente a frase em que se encontrava a EAP. Estabelecemos que todos os nossos contextos deveriam incluir a frase anterior e a frase posterior à frase em que estava presente a EAP. Isso foi necessário, uma vez que, muitas vezes, para entender a frase em que havia a presença de EAP, era preciso ter a informação da frase anterior. Assim, estabelecemos que todos os nossos contextos seriam compostos pela frase em que havia EAP mais a frase anterior e a posterior.

Após a coleta dos dados em português, cada um dos contextos foi lido. Foi possível verificar que alguns deles eram difíceis de entender; que algumas das EAP encontradas pareciam mais funcionar como conectores de conclusão ou de finalidade e não, propriamente, como elementos que estavam introduzindo uma paráfrase.

Essa inquietação quanto à função de algumas EAP com que nos deparamos levou a pensar sobre a origem dos contextos encontrados em português. Buscamos, então, os contextos correspondentes no texto original em inglês. Para isso, utilizando o *corpus* alinhado inglês-português TEXTQUIM. Identificamos os originais de todos os contextos coletados em português. Esse procedimento permitiu fazer uma apreensão qualitativa das EAP, e antever suas condições de tradução.

Em um segundo momento, organizamos os *corpora* de contraste. O objetivo do contraste, conforme já emnionamos, é avaliar se a presença de EAP, maior ou menor, pode caracterizar o gênero textual em foco.

---

<sup>5</sup> No caso da EAP *ISTO É*, foi necessário fazer uma “limpeza” manual dos contextos, pois percebemos que nem sempre *ISTO É* era uma EAP. Quando colocávamos a expressão *ISTO É*, a ferramenta *Concordance* retornava contextos como “*ISTO É* possível porque...” ou “*ISTO É* necessário devido a...”. O *ISTO É*, nesses contextos, não é considerado uma EAP, mas um pronome demonstrativo. Por isso, retiramos todos os contextos em que o *ISTO É* apareceu como um pronome demonstrativo.

<sup>6</sup> O *total por dez mil* é um cálculo estatístico linear; é uma projeção de que a cada dez mil palavras há uma determinada quantidade de EAP. No entanto, visto que é linear, assume-se que a quantidade de EAP prevista não está uniformemente distribuída, mas, sim, aleatoriamente distribuída. Isso significa que, se separarmos uma amostra de 10.000 palavras do *corpus*, há a possibilidade de não haver nenhuma EAP. No entanto, em uma outra amostra de 10.000 palavras do mesmo *corpus*, é possível encontrar mais do que o previsto para dez mil palavras.

O CC1, manual acadêmico didático de Físico-Química originalmente escrito em português, foi escaneado, corrigido e transformado de arquivo .doc para arquivo .txt para que pudesse ser lido pelo programa *Wordsmith Tools*. O arquivo do CC1 recebeu um cabeçalho com a referência bibliográfica, o ano do texto, o código do arquivo, a língua, o gênero, o autor, o assunto, o público-alvo, *tokens*, *types* e número e título do capítulo observado. Essas informações permitiam identificar o arquivo e facilitaram a consulta em qualquer momento da pesquisa.

O CC2, artigos científicos de Química, foi coletado a partir do *site* da revista Química Nova. Os artigos estão disponíveis no formato .pdf. Foi necessário transformar os artigos de .pdf para formato .doc. Depois disso, foram retiradas as figuras e as tabelas de cada texto, com a devida indicação da supressão. Após, a revisão dos arquivos, foram salvos e transformados em formato .txt.

O CC3, artigos da revista *SuperInteressante*, foi coletado do CD lançado em comemoração aos 15 anos da revista *SuperInteressante*. Inclui as revistas publicadas desde o seu lançamento, de 1987 até 2002. Manualmente, um bolsista do projeto TextQuim, procurou, através de uma ferramenta de busca disponível no próprio CD, artigos relacionados com o assunto Química. Foi necessário, obviamente, fazer uma seleção dos artigos retornados, pois o CD contém muitos artigos que incluíam segmentos, como, por exemplo, “a química do amor”. Esses artigos foram excluídos e foram selecionados apenas os que tinham relação com a Química como área de conhecimento. Após a seleção, os arquivos também foram transformados do formato .pdf para formato .doc, devidamente revisados e transformados para o formato .txt.

O CC4, Corpus Possamai, já estava devidamente coletado e pronto para uso. A autora, gentilmente, ofereceu o *corpus* para uso na nossa pesquisa.

O CC5, Base LEGIS, é constituído por leis relativas ao meio ambiente, coletadas no *site* do Senado Federal brasileiro, pelo Código Civil e Estatuto da Criança e do Adolescente. Já nos foi cedido em formato txt.

Após esses estágios de preparação dos *corpora* de contraste, tal como fizemos no *corpus* de estudo, utilizamos a ferramenta *Concordance* do software *Wordsmith Tools* para o início da coleta de contextos. Digitávamos uma das EAP e a ferramenta nos retornava a quantidade de ocorrências dessa expressão no texto e seus respectivos contextos. Em seguida, assim como fizemos com o *corpus* de estudo, registrávamos essa informação quantitativa em tabelas que armazenavam um *total bruto*, um *total percentual* e um *total por dez mil*.

A partir do próximo item, trazemos uma descrição dos dados encontrados no *corpus* de estudo e nos *corpora* de contraste. Nossa descrição de dados tem duas partes. A primeira privilegia a observação do *corpus* de estudo para averiguar se a presença de EAP contribui ou não para problemas de compreensão e se a tradução dessas expressões contribui para essa dificuldade. A segunda parte empreende a comparação entre dados encontrados no *corpus* de estudo e dados encontrados nos *corpora* de contraste.

## 2. Observação dos dados

### 2.1 Corpus de estudo

Focalizamos a presença de EAP no *corpus* de estudo em português. As EAP observadas foram<sup>7</sup>: *OU SEJA, ISTO É* e *EM OUTRAS PALAVRAS*; os temas observados compreendem: *Equilíbrio Químico, Equilíbrio Iônico, Ligação Química e Termodinâmica*.

---

<sup>7</sup> Fizemos uma busca, no *corpus* de estudo, de um conjunto de EAP sugerido nos estudos de Koch (2001, p. 52). Entre as EAP sugeridas pela autora estavam expressões tais como: *isto é, ou seja, quer dizer, ou melhor, em outras palavras, em síntese, em resumo*. De todo esse conjunto, encontramos somente: *OU SEJA, ISTO É* e *EM*

Para identificar a quantidade total de EAP no *corpus* de estudo em português, fizemos uma busca geral — sem distinção entre as três EAP observados —, com a ajuda da ferramenta *concordance* do software *Wordsmith*. Essa busca revelou a existência de 110 EAP no *corpus* de estudo. O quadro a seguir mostra esse resultado.

**Quadro VII:**  
*EAP no corpus de estudo (manual A + B)*

	<b>Total Bruto</b>	<b>Total Percentual</b>	<b>Total por dez mil</b>
<i>Ou seja</i>	55	0,029	2,958
<i>Isto é</i>	37	0,020	1,990
<i>Em outras palavras</i>	18	0,010	0,968
<b>TOTAL</b>	<b>110</b>	<b>0,059</b>	<b>5,916</b>

Considerando o *corpus* de estudo, a cada dez mil palavras, há quase 6 EAP. Além disso, a EAP que mais aparece no *corpus* de estudo é a *OU SEJA*: a cada dez mil palavras do *corpus*, 2,958; logo após, *ISTO É* com 1,990 a cada dez mil palavras; e, por último, a EAP *EM OUTRAS PALAVRAS* que é a menos utilizada, não havendo nem 1 EAP a cada dez mil palavras do *corpus* de estudo.

Após essa observação quantitativa das EAP, fizemos a leitura de cada um dos 110 contextos encontrados no *corpus* de estudo em português. Em muitos desses contextos, encontramos EAP que, superficialmente, parecem conectores que vinculam um dito e um redito, como os exemplos a seguir:

### **Contexto 1 – manual B**

*A teoria cinética dos gases e os experimentos realizados demonstram que a energia de um gás ideal depende do volume, portanto, o processo de expansão não faz diminuir a energia do sistema. O processo inverso, **ou seja**, a contração espontânea ou compressão das moléculas gasosas é permitida pela primeira lei da termodinâmica, mas nunca ocorre. Analogamente, observamos que o calor flui espontaneamente de um corpo mais quente para um mais frio.*

As EAP nos contextos anteriores desempenham sua função de introduzir um redizer, retomando algo já mencionado. No entanto, a leitura de outros contextos revelou alguns contextos de difícil compreensão. Para ilustrar essa dificuldade, trazemos abaixo cinco exemplos. Neles, percebemos que as expressões *OU SEJA* e *ISTO É* parecem ter uma outra função; não parecem estar introduzindo uma paráfrase.

### **Contexto 1 – manual B**

---

*OUTRAS PALAVRAS*. Assim, os resultados encontrados no *corpus* guiaram a nossa pesquisa e nosso estudo limita-se a observação dessas três EAP.

Assim, despenderemos a maior parte deste capítulo no estudo de equilíbrios ácido-base. No terceiro nível, trataremos de sistemas que apresentam vários equilíbrios simultâneos, ou seja, ácidos polipróticos fracos, formação de complexos metálicos iônicos e sais pouco solúveis. A habilidade de resolver problemas envolvendo equilíbrios é o resultado da compreensão dos princípios físicos e de uma intuição que somente pode ser adquirida com a experiência.

### Contexto II – manual A

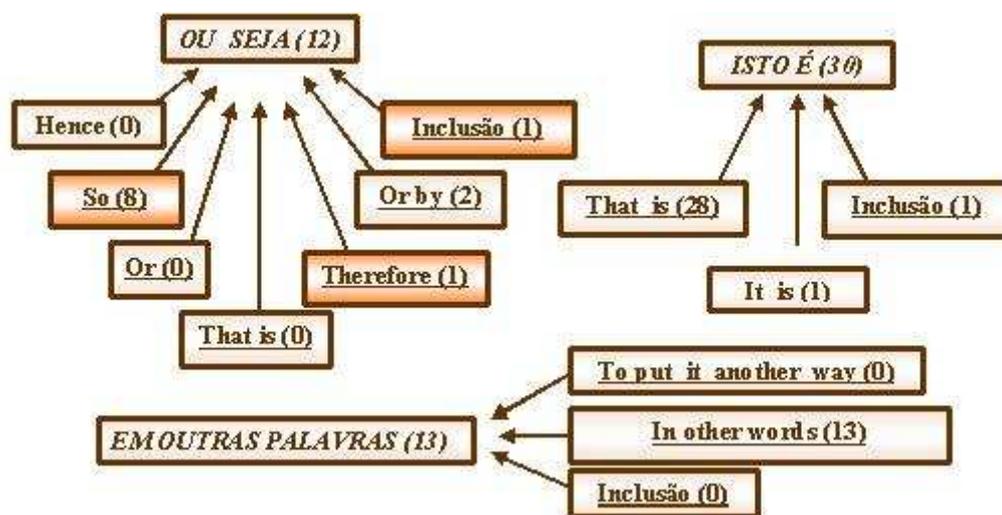
Suponha que 1 mol de H<sub>2</sub>O congele no sistema e que o sistema esteja à pressão constante. A variação de entalpia da amostra (sistema) é -6,0 kJ; ou seja, 6,0 kJ de calor fluem do sistema para as vizinhanças, e podemos escrever  $q_{viz} = +6,0 \text{ kJ}$ . Em geral, se a variação de entalpia do sistema é  $\Delta H$ , então para o calor transferido à pressão constante  $q_{viz} = -\Delta H$ .

Nos contextos 1, a EAP parece introduzir um exemplo. No contexto 2, temos a impressão de que *OU SEJA* anuncia uma conclusão ou uma explicação.

Essas impressões e dificuldades de compreensão do texto em português causaram-nos uma inquietação em torno da tradução, que nos levou a fazer um contraste das EAP entre texto traduzido e original.

A partir dos contextos com EAP alinhados no manual A, que possibilita ver a EAP em português e seus equivalentes no texto original, chegamos aos dados do esquema a seguir:

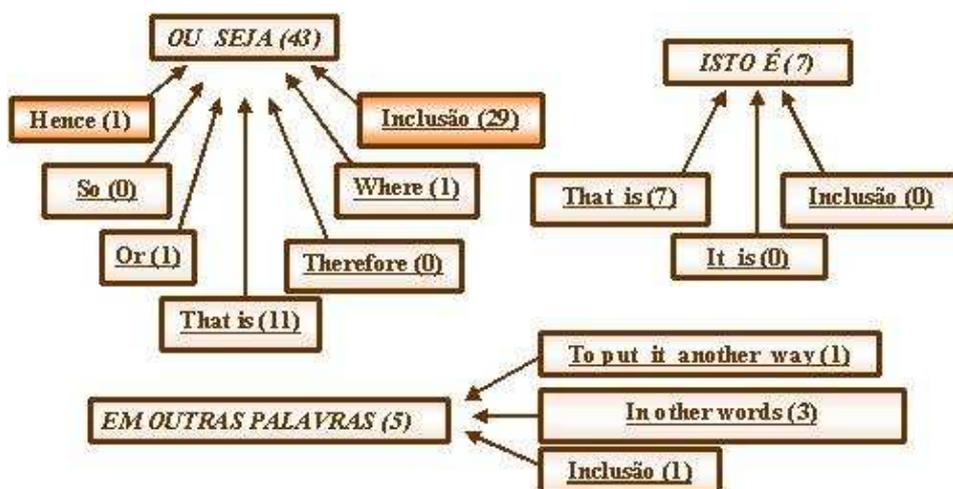
#### Correspondências das EAP no corpus de estudo alinhado: Manual A



O esquema 1 traz as EAP encontradas em português e seus correspondentes originais encontrados no texto original. Dois elementos nos chamaram atenção. Primeiro, o fato de conectores como *so* e *therefore* terem sido traduzidos por *OU SEJA*. Segundo, a existência de inclusão de EAP no texto em português quando não havia no texto original.

Com relação ao manual B, o mesmo procedimento foi realizado. Primeiro, buscamos os 55 equivalentes das EAP, observando os contextos alinhados. Os dados obtidos podem ser visualizados no esquema a seguir:

ESQUEMA 3:  
Correspondências das EAP no *corpus* de estudo alinhado: Manual B



Aqui, chama a atenção o conector *hence* traduzido por *OU SEJA* e uma grande quantidade de inclusões de EAP no texto traduzido. Das 55 EAP encontradas no texto traduzido, 30 são inclusões, o que representa mais de 50%.

Sintetizando os dados do manual A e B, são 110 contextos que apresentam EAP no *corpus* de estudo em português. Desses 110, 55 são *OU SEJA*, 37 são *ISTO É* e 18 são *EM OUTRAS PALAVRAS*. Os correspondentes dessas EAP em inglês mostram que:

- dos 55 *OU SEJA* em português, 11 são EAP (*That is*), 30 são inclusões e 14 são outro tipo de conector;
- dos 37 *ISTO É* em português, 35 são EAP, 01 é inclusão e 01 é outro tipo de conector; e
- dos 18 *EM OUTRAS PALAVRAS*, 17 são EAP e 01 é inclusão.

A EAP *OU SEJA* é a que menos apresenta correspondência com o texto original. As EAP *EM OUTRAS PALAVRAS* e *ISTO É* apresentam maior equivalência entre as EAP encontradas no texto em inglês. Isso revela que parece haver um problema concentrado nos usos de *OU SEJA*.

Passamos agora à observação de EAP nos *corpora* de contraste.

## 2.2 *Corpus* de estudo e *corpora* de contraste

Observamos, aqui, os dados encontrados nos cinco *corpora* de contraste (CC), comparando-os com os dados encontrados no nosso *corpus* de estudo. Como já dissemos, o contraponto tem como objetivo verificar se a presença de EAP caracteriza o gênero textual do manual acadêmico de Química traduzido.

O levantamento dos dados se limitou a uma observação quantitativa. Por isso, não coletamos todos os contextos nos *corpora* de contraste, nem os examinamos um a um.

A seguir, temos um quadro com os resultados encontrados.

**Quadro-síntese de quantidade de EAP I**  
*Corpus de estudo x Corpora de contraste*

	<i>Ou seja</i>	<i>Isto é</i>	<i>Em outras palavras</i>	<b>TOTAL BRUTO</b>	<b>Total percentual</b>	<b>Total por dez mil</b>
<i>Corpus de estudo</i>	55	37	18	<b>110</b>	<b>0,059</b>	<b>5,917</b>
CC1	22	58	6	<b>86</b>	<b>0,161</b>	<b>16,195</b>
CC2	81	39	4	<b>124</b>	<b>0,028</b>	<b>2,899</b>
CC3	30	14	1	<b>45</b>	<b>0,075</b>	<b>7,552</b>
CC4	470	188	24	<b>682</b>	<b>0,052</b>	<b>5,298</b>
CC5	19	8	0	<b>27</b>	<b>0,001</b>	<b>0,132</b>

Os dados acima revelam que o *corpus* de estudo **não** é o que mais apresenta EAP. O *Total por dez mil* confirma que o *corpus* de estudo é o terceiro *corpus* no *ranking*. Antes dele, há o CC1, textos de manuais didáticos de Química, originalmente escritos em português, e o CC3, textos de popularização da Revista SuperInteressante.

Um fator a considerar é a diferença entre o *corpus* de estudo e o CC4. No *corpus* de estudo, há 5,917 EAP a cada dez mil palavras, enquanto no CC4 há 5,298 — uma diferença pequena de apenas 0,619. Quando consideramos que o CC4 é composto por artigos científicos de Informática, esse resultado nos surpreendeu, pois esses textos, em tese, tendem a ser pouco didático.

Além disso, comparando os resultados do CC4 com os resultados do CC2, também composto por artigos científicos, mas da área da Química, fica claro que o CC4 apresenta quase duas vezes mais EAP do que o CC2. Esse resultado indica que o artigo científico de Informática retoma elementos de informação bem mais do que o artigo científico de Química.

Seria de se esperar que os contextos do CC4 e do CC2 fossem parecidos. No entanto, uma observação descompromissada dos contextos revela que são bem diferentes. Os contextos no CC4 parecem ser, inclusive, mais semelhantes com os contextos do CC3 (Revista SuperInteressante) do que com os contextos do CC2.

Em síntese, os dados dos *corpora* de contraste e os dados do *corpus* de estudo revelam que:

- 1) manuais didáticos de Química traduzidos não são os textos que mais apresentam EAP;
- 2) manuais didáticos de Química originalmente escritos em português apresentam mais EAP do que os manuais traduzidos do inglês;
- 3) artigos científicos de Informática apresentam uma quantidade de EAP similar à encontrada nos manuais didáticos traduzidos;
- 4) há diferentes presenças de EAP em diferentes gêneros textuais; e
- 5) o texto legislativo, representado pelo CC5, faz uso muito baixo de EAP.

Sintetizamos, no quadro a seguir, o contraste *corpus* de estudo e Base LEGIS (CC5).

### 2.3 Corpus de estudo x Base LEGIS.

<i>Base LEGIS</i>	<i>Ou seja</i>	<i>Isto é</i>	<b>Em outras palavras</b>	<b>TOTAL</b>	<b>Nº de palavras por texto</b>	<b>Total por dez mil</b>
Legislação Ambiental Brasileira	19	8	0	27	1.010.519	0,2671
Estatuto da Criança e do Adolescente	0	0	0	0	17.574	0
Código de Defesa do Consumidor	0	0	0	0	9.751	0
Constituição Federal Brasileira	0	0	0	0	46.141	0
<b>TOTAL</b>	19	8	0	27	2.037.844	*0,1324*

	<i>Ou seja</i>	<i>Isto é</i>	<i>Em outras palavras</i>	<b>TOTAL BRUTO</b>	<b>Total percentual</b>	<b>Total por dez mil</b>
<i>Corpus de estudo</i>	55	37	18	<b>110</b>	<b>0,059</b>	<b>5,917</b>
CC5	19	8	0	<b>27</b>	<b>0,001</b>	<b>0,132</b>

Verificamos maior presença de EAPs em textos da revista *Superinteressante* e em manuais de Química. Em artigos de Química, a ocorrência dessas expressões diminui, assim como nos textos legais, nos quais as EAPs ocorrem numa ínfima frequência. Isso pode sinalizar que textos legislativos caracterizam-se frente aos demais pelo reduzido emprego desses elementos coesivos. A baixa frequência de EAPs nos textos legislativos indicaria seu menor grau de didatismo, comparativamente aos demais textos pesquisados. Isso pode sinalizar também que no texto legal a incidência de EAPs seria caracteristicamente menor, como recurso estilístico próprio.

### 3. Considerações Finais

Pelo cotejo entre *corpus* de estudo e *corpora* de contraste, foi possível observar que esses elementos não são um traço de identidade exclusivo do manual acadêmico traduzido de Química Geral. Em síntese, EAP não é uma peculiaridade exclusiva do manual acadêmico didático traduzido de Química Geral. Outros gêneros textuais também fazem uso de EAP, até mesmo em quantidade muito maior, como ficou evidente ao comparar corpus de estudo e CC3 (textos de popularização de ciências). Os dados mostram que o manual didático de Química, originalmente escrito em português, quando comparado com o manual didático de Química traduzido do inglês, apresenta muito mais EAP.

Ademais, o contraste *corpus* de estudo e *corpora* de contraste permitiu perceber indícios de um “macrogênero” de textos didáticos e de um *modus dicendi* que parece modelar a maneira como o conhecimento é expresso em diferentes áreas. Ao comparar o manual

acadêmico didático de Química traduzido e um originalmente escrito em português, vemos que a EAP é, sim, uma diferença digna de nota.

O *paper* de Química e o *paper* de Informática, em tese, integram um mesmo gênero textual, o *paper* acadêmico. No entanto, não são iguais no que diz respeito às EAP, porque cada área parece exibir um *modus dicendi* próprio para expressar conhecimento. A maneira como a Química argumenta, explica, expõe suas paráfrases, parece ser bem diferente da de Informática. Nos *papers* de Informática há muito mais redizes do que nos *papers* de Química, área em que as retomadas são menos frequentes. Em *papers* de Química, talvez haja um maior nível de pressuposição de conhecimento prévio por parte do leitor, maior do que em Informática.

Gêneros nascem a partir de convenções estabelecidas de comum acordo por determinada comunidade discursiva, produtora e receptora desses textos. Essa comunidade estabelece convenções para que os textos possam operar e ser entendidos dentro de uma determinada área (Swales, 1990). Assim, alguns dos textos examinados se assemelham em características que podem permitir identificá-los como pertencentes a um determinado “macrogênero”, o didático, visto que têm em comum o uso de EAP.

Este estudo, melhor desenvolvido no trabalho de Azeredo (2007), visa contribuir para melhor apreender a linguagem química em uso. A observação das EAP nos mostrou que há, sim, fundamentação, no que tange ao uso das EAP, algumas das reclamações de dificuldade de compreensão de leitura em Química. Não se resumem, naturalmente, às EAP, restando o desafio de prosseguir o exame do *corpus*.

## Referências

- AZEREDO, Susana de. Expressões anunciadoras de paráfrase em manuais acadêmicos de Química: um estudo baseado em corpus. Diss. Mestrado. Porto Alegre: UFRGS/PPG-Letras da UFRGS, 2007.
- BERBER SARDINHA, Tony. Linguística de Corpus. Baurueri-SP: Manole, 2004.
- FLORES, Valdir do N.; TEIXEIRA, Marlene. Introdução à Linguística da Enunciação. São Paulo: Contexto, 2005.
- KRIEGER, Maria da Graça; FINATTO, Maria José B. Introdução à Terminologia: teoria & prática. São Paulo: Contexto, 2004.
- POSSAMAI, Viviane. Marcadores textuais do artigo científico em comparação português-inglês – um estudo sob a perspectiva da tradução. Porto Alegre: UFRGS/PPG-Letras da UFRGS, 2004. 121p. Diss. mestrado.
- RÓNAI, Paulo. (1987) Escola de Tradutores. 5.ed.ver.ampl. Rio de Janeiro:Nova Fronteira.
- RÓNAI, Paulo. (1981) A tradução vivida. Rio de Janeiro:Nova Fronteira, 210p.
- SILVA, S.M.;EICHLER, M.;DEL PINO, J.C. As percepções dos professores de química geral sobre a seleção e a organização conceitual em sua disciplina. Química Nova, Ago.2003, v.26, n.4, p.585-594.
- SWALES, John. Genre analysis: English in academic and research settings. Cambridge University Press, 1990.