

Terminologia e Lingüística de Corpus: da perspectiva enunciativa aos novos enfoques do texto técnico-científico

Maria José Bocorny Finatto*

Resumo – Este trabalho tem o objetivo de demonstrar a viabilidade e as potencialidades de um tríptico encontro teórico-metodológico: o encontro entre os estudos e os *fazeres* de Terminologia, de Lingüística de Corpus e os de Enunciação, colocada a perspectiva enunciativa benvenistiana como uma catalisadora de inter-relação. Nessa direção, o trabalho argumenta que essa perspectiva já tem permeado as outras duas áreas sem que tenha havido, entretanto, um reconhecimento explícito de sua presença. Após uma breve caracterização de cada um dos atores do encontro, apresenta-se um exemplo ilustrativo de um ponto de convergência possível a ser explorado. Ao final, são feitas algumas considerações sobre as perspectivas de aproveitamento mútuo para os três envolvidos, especialmente para a vertente da Terminologia que toma como objeto privilegiado o todo do texto técnico-científico e não mais terminologias em isolado.

1 Introdução

Neste trabalho pretendemos demonstrar a pertinência de um encontro entre Terminologia, Lingüística de Corpus e os estudos da Enunciação, tomando os estudos enunciativos como catalisadores dessa inter-relação. Para tanto, fazemos uma breve caracterização dos dois primeiros participantes do encontro, com destaque para a Terminologia, nossa área de atuação, e, depois, apresentamos alguns exemplos ilustrativos de pontos de convergência possível. Ao fim do trabalho, propomos algumas reflexões sobre as

* Instituto de Letras da UFRGS, PPG-Letras – UFRGS.
e-mail: mfinatto@terra.com.br

perspectivas de aproveitamento recíproco para as três áreas de estudo.

2 Terminologia, Lingüística de Corpus e o terceiro da Enunciação

2.1 Terminologia

A Terminologia, como uma disciplina organizada, sustentada por referenciais teóricos mais ou menos definidos, é algo relativamente recente no panorama dos estudos de Lingüística Aplicada. Como um estudo primordial de "termos técnicos", nasceu, nos anos 60, na Europa germânica, bastante vinculada à idéia de prescrição, de normatização das designações técnico-científicas, numa perspectiva que se convencionou chamar de Teoria Clássica ou "Escola de Viena".¹

Essa ótica normativa inicial, bastante difundida e prestigiada, durante um certo tempo, conduziu à concepção de que listagens de termos, dicionários e outros repertórios poderiam dar conta de espelhar grande parte das características de uma linguagem técnico-científica ou especializada. E, em função disso, produzir glossários e dicionários tornou-se algo muito importante. Uma tal apreensão da linguagem das tecnologias e das ciências, encerrada nos limites de um léxico marcado, de caráter predominantemente nominal, gerou também a concepção de uma língua à parte, não-natural, artificialmente construída por tecnólogos e cientistas. Diferente da língua cotidiana, tal como um tecnoleto² ou jargão, foi percebida como uma modalidade de língua exclusivamente utilizada e entendida por um segmento profissional determinado e restrito em uma situação de comunicação igualmente específica.

À medida que progrediu, dos anos 80 aos 90, a disciplina Terminologia passou a abrigar concepções lingüísticas um pouco mais amplas da comunicação técnico-científica, tendo avançado de uma fase normativa para uma fase descritiva, rumando hoje para a aceitação do objeto texto e também do todo de um discurso terminológico, ainda que não se abra mão dos enfoques lexicocêntricos, os quais permanecem dominantes.

Entretanto, na direção que vai do termo ao texto, as concepções mais ou menos textualistas ou lexicocêntricas da Terminolo-

¹ Para outros detalhes sobre a trajetória histórica da Terminologia, veja Krieger e Finatto, 2004.

² Não se aplicou, entretanto, a concepção de diassistema de E. Coseriu ao entendimento de tecnoleto.

gia da atualidade ainda se prendem a determinados formalismos de apreensão dos enunciados em que se destacam, via de regra, elementos idiossincráticos como a modalidade, a negação, as marcações de tempo. Nessa via, tendem a reproduzir tipos de cortes ou de separações que são bastante familiares aos enfoques da Lingüística: em Terminologia estudam-se termos em textos ou textos em que há termos; distingue-se entre o lingüístico e o extralingüístico, entre o sistemático e o idiossincrático, entre o que é estritamente terminológico e o que seria discursivo³, entre o que está na superfície e o que está subjacente à realização, entre outras oposições colocadas nas aproximações aos fenômenos da comunicação técnico-científica.

Em meio aos recortes e às perspectivas que ainda estão sendo delineados e firmados no cenário epistemológico em que está a Terminologia, naturalmente perpassado pelos diferentes vieses da Lingüística, mais estruturais ou mais racionalistas, pensamos que as idéias basilares de E. Benveniste de uma semântica enunciativa podem oferecer uma "luz" nova e interessante ao que se apresenta, como também conferir um matiz especial para o desenrolar de algumas das suas tendências afins. Afinal, ainda que o foco e o produto principal da pesquisa terminológica sejam a elaboração de glossários, já se percebem os encaminhamentos que visam caracterizar não só terminologias, mas sim um todo de uma linguagem que se espelha em práticas textuais e discursivas.

2.2 Lingüística de Corpus

De um outro lado do cenário que desenhamos aqui, em situação de "dialogante" que se aproxima da Terminologia ao mesmo tempo que é abordada por ela, temos a Lingüística de Corpus. Essa aproximação, em meio a tantas outras, ocorre porque a Terminologia, além de "transdisciplina", cada vez mais tem se aproximado do processamento de grandes acervos textuais com o fim de identificar especificidades lexicais e, desse processamento, obter, supostamente com menos dificuldades ou em menor tempo, dicionários baseados em um uso lingüístico "real".

Definida como uma nova perspectiva para a Lingüística (Berber Sardinha, 2004, p.35), mas não como um novo tipo de Lingüística, a Lingüística de Corpus (doravante LC) ocupa-se da explora-

³ Discursivo no sentido de pertinência ao texto, à língua e não à terminologia. Para ilustrar a dicotomia: elementos coesivos seriam do plano discursivo; já a constituição de um conjunto de termos por determinados radicais ou prefixos seria do âmbito terminológico.

ção de grandes extensões *corpora* textuais em formato digital que são "minerados" com apoio informatizado, com destaque para as explorações estatísticas de elementos lexicais e observação das frequências de combinações de palavras, o que em geral é bastante aproveitado para caracterização de gêneros textuais⁴. Todos os gêneros são tratados, inclusive o literário e o jornalístico, além de haver estudos sobre *corpora* com registros orais.

A Linguística de Corpus, dito de um modo muito simples, parte do pressuposto de que é preciso observar usos extensivos de língua para que deles se depreendam sistematicidades, especificidades, descrições e explicações sobre gramática, léxico, sobre configuração e perfis de práticas textuais. O reconhecimento de usos concretos de linguagem em amostras extensivas é condição *sine qua non*, visto que a "verdade" sobre língua e gramática estará posta num corpus: combinações e associações entre uma palavra e as demais a definirão em relação ao todo de um sistema.

Essa nova via para a Linguística se delinea também como oposição à vertente gerativo-transformacional, a qual, ao contrário dela, buscou reconhecer elementos subjacentes ao uso de um falante-ouvinte ideal, de modo que uma pesquisa de usos em extensão perderia o sentido. Afinal, esse falante sintetizará a competência que é comum a todos os falantes independentemente de eventuais heterogeneidades. Instaura-se, pela ótica da linguística de corpus, uma nova dicotomia, entre o linguista de corpus, aquele que "põe a mão na massa", e sua antítese, representada pelo "linguista de poltrona", que teoriza sobre a língua com base em sua própria competência linguística.

Basicamente, a Linguística de Corpus assenta-se numa concepção funcionalista de linguagem. Compreende a língua como um sistema probabilístico em que as diferentes combinações das unidades lexicais espelharão suas condições gramaticais em um sentido amplo, as quais serão fornecidas pelo de uso, distribuição, frequência e funcionalidade. Ecoam aqui como herança as concepções da gramática sistêmico-funcional de Halliday, as idéias de Biber, de Firth e de Swales.

Terminologia e Linguística de Corpus têm dialogado em torno de um objeto comum, o texto técnico-científico e seu léxico, os quais são objeto de exploração com apoio informatizado. A aproximação, entretanto, parece ainda um tanto tímida, apesar da vocação transdisciplinar dos parceiros envolvidos, visto que se trata

de um ponto de encontro e não de uma diluição de especificidades. A Terminologia em geral não abre mão do estatuto diferenciado das unidades denominadas termos, seu foco. A LC, por sua vez, não se afasta de observações concretas e extensivas de usos em geral, sem conceber qualquer caráter diferenciado *a priori* colocado sem uma comprovação empírica prévia.

2.3 Da Enunciação como catalisadora

Sobre o encontro antes referido, cabe aqui uma ilustração didática: em Química, no estudo das misturas e reações, mostra-se que há casos em que duas substâncias podem estar juntas por anos a fio e nada acontecer, ou só vir a acontecer algo muito lentamente. Mas, se for adicionada à mistura uma terceira substância, chamada catalisador, os resultados do encontro poderão ser precipitados, melhor aproveitados e até controlados em função de um objetivo prático que se tenha. Naturalmente, há o caso dos encontros explosivos, mas não pensamos que seja esse o caso da nossa dupla.

Aproveitando tal ilustração, podemos dizer que há, sim, uma via de aproximação entre LC e Terminologia a ser catalisada por um elemento terceiro, geralmente não explicitamente nominado nos encontros já travados, mas bastante presente e localizado num ponto de intersecção bem extenso. Trata-se do reconhecimento das sistematicidades e das especificidades de usos da linguagem por parte de um grupo social composto por tecnólogos e cientistas que têm como fim comunicação, representação e produção de conhecimentos.

A Enunciação benvenisteana, a qual se consubstancia na idéia básica de uma apropriação da linguagem por parte de um sujeito que se constitui como tal a partir dessa apropriação, é o terceiro elemento. Esse terceiro, e não um outro, firma-se pelo fato de que se reconhece, tanto em Terminologia quanto em LC, que o que se dá no texto técnico-científico é um uso diferenciado de linguagem, que esse uso se realiza em determinadas combinações depreensíveis em um corpus e que, nesse uso, os "termos técnicos" não são termos por si, mas sim *estão* termos numa determinada situação⁵.

Ao existir em comum uma visão de palavras e até de textos que adquirem um valor e/ou feição diferenciados em determinadas condições de um uso, posto que são fruto de uma determinada

⁵ Essa é a concepção da Teoria Comunicativa de Terminologia, desenvolvida por Cabré e seus colaboradores, a qual se coloca como uma alternativa à Terminologia de viés normativo. Nessa teoria abandona-se a noção de termo para adotar-se a noção de "unidade de significação especializada".

apropriação da linguagem, instaura-se um bom espaço para o reconhecimento e a observação da enunciação, a qual poderá ser também percebida como tal por cada uma das partes envolvidas.

2.4 Dificuldades

Cabe, entretanto, citar algumas das problemáticas fundamentais das observações de Terminologia para que se avaliem, um pouco, as suas atuais dificuldades de diálogo, quer com a LC, quer com a Enunciação. Essas dificuldades, vale destacar, são fruto dos recortes já estabelecidos ao longo da sua história.

Um desses recortes-problema é justamente a defesa de uma fronteira entre a linguagem comum e linguagem científica/especializada, considerando-se que o foco da Terminologia incide apenas sobre a segunda. Há, obviamente, momentos em que a linguagem da ciência e linguagem cotidiana se misturam, interpenetram. No caso da palavra *acetona*, por exemplo, vemos um termo originado da Química que se tornou, na linguagem cotidiana, apenas "produto para remover esmalte de unhas". Em Direito, temos um outro caso: a palavra *casamento* tem um valor diferenciado da linguagem do dia-a-dia e hoje, em termos legais, é sinônimo de "união estável". Um outro exemplo que mostra a interpenetração também entre as linguagens de diferentes ciências é a palavra/termo *velocidade*: em Física há um significado e em Química um outro bem diferente.

No plano do texto-corpus, instância das práticas textuais técnico-científicas, temos a percepção da LC de um gênero textual como, por exemplo, o "artigo de periódico especializado" ou o "academic writing" conformado pela presença recorrente de determinados recursos estruturais, gramaticais e lexicais e de determinadas combinatórias. O foco da LC, assim, não incidirá apenas sobre as terminologias, ainda que acabem se singularizando como uma característica, visto que são uma entre tantas outras facetas de um gênero.

A partir dessa mistura-encontro entre focos, métodos e pressupostos, nossa ideia é esboçar um quadro preliminar de como a Enunciação pode gerar aproveitamentos e mediar esse diálogo ao mesmo tempo que dele participa. É importante pensar um pouco sobre o quanto e em quais circunstâncias se pode realmente precipitar a relação entre Terminologia e Linguística de Corpus, principalmente porque há todo um jogo de exclusões em curso no cenário da Terminologia. Sem grandes ou profundos conhecimentos da obra de E. Benveniste, a tentativa de demonstração em esboço es-

pera contar com contribuições e críticas por parte de melhores e mais experientes leitores dessa obra.

3 Um exemplo ilustrativo

Quando observamos textos técnicos-científicos, tal como artigos de periódicos especializados de determinadas ciências, principalmente das chamadas ciências exatas e biológicas, salta aos olhos o padrão do efeito de impessoalização dos dizeres, o seu foco em uma terceira pessoa. Em prol de uma manutenção de cientificidade, de objetividade e de aceitação perante um grupo social determinado, são descritos ou relatados, nesses textos, experimentos e feitas considerações ou em terceira pessoa ou com o uso de um SE indeterminador ou, ainda, pelo uso massivo de voz passiva. Isso é o que se vê num trecho de um artigo de Química:

Na Figura 1 são apresentados três gráficos de índice de hidroperóxido versus tensão mecânica aplicada para EPDM 43, 57 e 65. Observa-se que a tensão mecânica acelera o processo degradativo do polímero em todos os tipos de EPDM, sendo mais pronunciado nos EPDM 43 e 57. O mesmo comportamento pode ser observado com relação ao índice de carbonila: a tensão mecânica aplicada contribui positivamente para o aumento dos processos degradativos, com as mesmas relações de intensidade observadas para a formação de hidroperóxidos (EPDM 43 > EPDM 57 > EPDM 65), embora, em ambos os casos os índices do EPDM 65 tenham atingido valores mais altos em menores tempos. Os valores mais altos de IC e IHP de EPDM 65 justificam-se por seu maior teor de dieno (V, tabela abaixo). A relação direta entre a formação de hidroperóxidos e carbonilas é conhecida há bastante tempo, sendo os primeiros os precursores da formação do segundo tipo de produtos de degradação oxidativa (Zepka e Baumhardt-Neto, 1997; grifos nossos).

Esse tipo de fenômeno, em geral, é estatisticamente bem percebido em estudos com Linguística de Corpus que se integram ao estudo de gêneros textuais, mas é um elemento ainda pouco explorado pelos trabalhos de Terminologia como uma "marca" da linguagem de uma determinada ciência em uma determinada situação textual. Em geral, o "olhar" terminológico, nesse trecho, privilegiará unidades como *EPDM*, *índice de carbonila*, *polímero*, *hidroperóxidos*, *carbonilas*, *dieno*, etc., independentemente de suas repetições maiores ou menores.

Mas, o que há neste trecho além de "termos técnicos" que causariam estranhamento a um leitor não iniciado na ciência? Da parte da LC, essa pergunta seria formulada de um outro modo: o

que há, nesse corpus, além de um conjunto de palavras que se repetem mais ou menos ou que se associam de modo mais ou menos recorrente que o particulariza em relação a outros corpora?

Sob a perspectiva da LC, buscando responder tal pergunta, o trecho exemplificado como corpus seria inicialmente observado em termos da lista das palavras que o compõe e das suas frequências e combinatórias, as quais seriam comparadas às frequências e combinatórias em outros corpora. Assim, por exemplo, seria verificado que DE ocorreu 12 vezes; EPDM 07 vezes; em terceiro lugar há OS com 06 ocorrências (sublinhados no trecho inclusive aglutinados às preposições).

Nesse corpus, entre as unidades lexicais de tipo não gramatical, o item EPDM teria destaque pela alta frequência, próxima à de DE, o qual costuma ser um elemento muito freqüente em qualquer corpus. Além de EPDM, teria estatuto diferenciado também DIENO, dada sua única ocorrência, o que o torna um *hapax*. Há, então, uma polaridade na observação e nas atenções às palavras: alta frequência e baixa frequência.

Para a Terminologia, entretanto, não seria usual a percepção, nesse mesmo corpus-trecho, de determinantes como OS, visto que estão fora de um âmbito estritamente terminológico. O foco especial apenas para EPDM e DIENO também seria uma limitação, pois há outros termos entre esses dois extremos. De outro lado, na LC, o contra-argumento é que não é usual conceber elementos de estatuto *a priori* diferenciados.

As duas percepções, cada uma a seu jeito, focam o modo de dizer da ciência e se encontram nesse modo de dizer da ciência, embora não o identifiquem expressamente como um objeto entre seus pressupostos ou princípios. Em função disso, dessa ausência, a marcação, no caso do nosso trecho-corpus, de uma proximidade com a terceira pessoa, ou do ELE do qual se fala por trás da voz passiva ou de um JUSTIFICAM-SE, pela via dos artigos definidos O/A, revelam facetas da enunciação de Química e dos químicos que, como tal, têm escapado à Terminologia e à LC. Enquanto isso, a despeito da apreensão, em artigos de periódicos de Química, em todo um extenso corpus, há uma verdadeira fuga do NÓS e do EU, mas, apesar da fuga, esse EU sublimado tende a retornar pela via dos determinantes, dos artigos definidos, os quais revelam uma proximidade ou familiaridade ou maior conhecimento do enunciador em relação aos seus objetos. E isso, essa regularidade, entre as três dialogantes, só a perspectiva da Enunciação notará.

4 De outros exemplos e da interação

Afora a observação do jogo entre efeito de subjetividade e efeito de objetividade, do ocultamento e do retorno inevitável do EU na fala da ciência, uma outra possibilidade de encontro entre Terminologia, LC e Enunciação seria a observação das modalidades verbais em textos de Química. Por exemplo, em artigos de periódicos ou em textos de comunicações em anais de congressos.

O destaque à modalidade verbal é pertinente porque a Terminologia tem voltado sua atenção também para os verbos, além dos termos, visto que muitos aparecem reconhecidos como núcleos "apontadores" de terminologias (Maciel, 2001), e também pelo fato da LC muito se ocupar dos verbos e de suas combinatórias recorrentes com determinados nomes.

No trecho do texto antes indicado, podem ser destacadas as seguintes formulações verbais: *são apresentados/ observa-se que/ sendo mais pronunciado/ pode ser observado/ a tensão mecânica aplicada contribui/ relações de intensidade observadas/ tenham atingido/ justificam-se/ é conhecida/ sendo os primeiros*. Nesse conjunto-amostra de ocorrências, as formas nominais do gerúndio e do particípio são freqüentes e o uso de passiva também se realça.

A propósito de verbos nesse corpus e em outros corpora que se quisesse examinar, Benveniste, além de propor-nos focar a diferença entre o tempo das pessoas e o tempo dos acontecimentos, tal como em "As relações de Tempo no Verbo Francês" (Benveniste, 1995, p. 260) revelará dois planos de enunciação diferentes, o da terceira pessoa e o da primeira pessoa.

De tal sorte, conforme se desprende de suas idéias, na enunciação de uma ciência, os acontecimentos parecerão narrar-se a si mesmos, sem um sujeito que os conte. Daí por que, em muitos textos de Química, muitos fenômenos também parecem ocorrer independentes de um agente que os produza, tendo-se a impressão de um animismo e de uma autonomia das coisas e substâncias em relação às pessoas que as manipulam.

Ainda tratando sobre verbos, no texto "Ativo e médio no verbo", Benveniste chama atenção para um fato que também pode ser explorado em LC ou em Terminologia, visto que "a distinção entre ativo e passivo pode fornecer um exemplo de uma categoria verbal para desencaminhar os nossos hábitos de pensamento" (Benveniste, 1995, p. 183). E, das suas colocações, poderíamos questionar: afinal, que tipo de uso de passiva tem-se em textos científicos?

Esse tipo de percepção, a percepção da enunciação multifacetada, que enseja novas perguntas tanto ao corpus quanto à lingua-

gem das ciências, acreditamos, pode acrescentar muito à LC e à Terminologia à medida que, além de ampliar o foco descritivo inerente a cada uma em suas particularidades de apreensão, seja em contato ou em separado, traz um viés explicativo completamente novo para ambas. Assim, em síntese, podemos dizer que, do encontro entre LC, Terminologia e Enunciação, lucrarmos todos com focos de estudo mais abrangentes e detalhados, além de instaurar-se uma potencialidade explicativa para os fenômenos observados altamente desejável nas duas primeiras.

Referências

- BENVENISTE, E. *Problemas de lingüística geral I*. Campinas: Pontes, 1995.
- BERBER SARDINHA, Tony. *Lingüística de corpus*. Barueri: Manole, 2004.
- KRIEGER, M. da G.; FINATTO, M. J. B. *Introdução à terminologia: teoria & prática*. São Paulo: Contexto, 2004.
- MACIEL, A. M. B. *Para o reconhecimento da especificidade do termo jurídico*. Porto Alegre: UFRGS. Tese (Doutorado em Estudos da Linguagem), 2001.
- ZEPKA, M. M.; BAUMHARDT NETO, R. Foto-degradação de EPDM sob tensão mecânica In: 4º Congresso Brasileiro de Polímeros, 1997. ABPOL - CD-ROM, 1997, v. 1, p. 1-2, 1997.