

A COMPILAÇÃO LEXICOGRÁFICA COMO ALTERNATIVA EMERGENCIAL À DESCRIÇÃO TERMINOLÓGICA SISTEMÁTICA

Francis Henrik Aubert¹

Resumo: A descrição e a sistematização de dados terminológicos com base em corpus variados de textos de especialidade, embora constitua um dos métodos mais seguros para a obtenção de um retrato fiel do uso lingüístico efetivo, exige um dispêndio de tempo para a concretização de seus resultados que pode afigurar-se longo demais tendo em vista as urgências mais imediatas – tal como se manifestam no contexto brasileiro – particular mas não exclusivamente, da tradução técnica. Propõe-se, como abordagem alternativa, a compilação de dados lexicográficos a partir de dicionários e glossários técnicos já disponíveis no mercado e pré-selecionados em termos de confiabilidade por usuários especializados (tradutores profissionais), como solução emergencial, enquanto a descrição terminológica sistemática nos principais campos do conhecimento não gerar os seus primeiros frutos. Tal compilação deverá ser

Abstract: *The description and organization of terminological data based on a varied corpus of specialty texts, although methodologically adequate for obtaining a truthful picture of actual linguistic usage, requires an investment of time which may be felt as much too long, given the most immediate and urgent needs – which are evident in the Brazilian context – specially (but not exclusively) of technical translation. This paper proposes, as an alternative approach, compiling lexicographical data from technical dictionaries and glossaries already available on the market, as pre-selected in terms of reliability by specialized users (professional translators). Such approach is to be regarded as an emergency solution, to be resorted to whilst the systematic terminological description of the main fields of knowledge does not generate its first fruits. Such compilation should be*

organizada de forma a tornar cada verbete compatível com sua futura expansão e transformação dos registros, de sua forma lexicográfica simplificada para a descrição terminológica completa.

Unitermos: Terminologia; Terminografia; Terminótica; Compilação de dados lexicográficos e terminológicos.

organized in a form compatible with further expansion and conversion of the entries from their simplified lexicographical form to a fullworthy terminological description.

Key-words: Terminology; Terminography; Terminotics; Compilation of lexicographic and terminological data.

1. CARACTERIZAÇÃO DO PROBLEMA

A descrição e a sistematização de dados terminológicos, monolíngües ou bilíngües, tal como se manifestam nos discursos orais ou escritos de seus múltiplos usuários, constitui tarefa complexa, por vezes árdua, necessariamente demorada (1). Demanda uma adequada definição de temática, de escopo, de grade conceptual. Exige a aplicação de todo um conjunto de métodos de identificação, segmentação e análise linguística, discursiva e conceptual. Requer um sistema de registro apropriado. Não prescinde de cuidadosas verificações e revisões². O próprio estabelecimento dos corpúscos vê-se, ao menos em nosso contexto brasileiro, travancado pelas dificuldades de acesso aos textos e aos consultores-especialistas pertinentes, fruto, particularmente no primeiro aspecto, da falta de bases bibliográficas e de bibliotecas adequadamente informatizadas, além dos entraves burocráticos, as preocupações – compreensíveis mas, por vezes, excessivas – com o sigilo empresarial etc.

Desta forma, é fácil prever que, diante da incipiência da pesquisa terminológica no Brasil, os trabalhos sistemáticos de descrição terminológica – que podem e devem ser encaminhados a despeito dos obstáculos apontados – somente produzirão frutos consistentes a médio e longo prazos. Sucede, no

entanto, que a tarefa de facultar o acesso a informações terminológicas bilíngües se reveste de uma evidente urgência, particularmente com a entrada em vigor da norma ISO 9000 para a padronização qualitativa dos mais variados integrantes da pauta de exportações do país, para não mencionar as crescentes necessidades da tradução técnica, canal imprescindível (e, por vezes, gargalo efetivo) nos processos de transferência tecnológica, que, pela paulatina abertura do mercado interno para o intercâmbio internacional, tendem a aguçar-se em prazo curto.

As considerações acima parecem sugerir caber uma segunda abordagem, complementar à primeira (esta mais propriamente terminológica), que gerasse, em tempo mais curto, uma base de dados utilizável no contexto bilíngüe. Tal base de dados – que denominaríamos *bases terminológicas de primeiro nível* – poderia ser montada a partir de materiais lexicográficos já existentes sob forma de dicionários especializados bilíngües e multilíngües³.

Entre tais fontes, parece conveniente nos atermos àqueles que, no consenso dos seus usuários, apresentam um nível satisfatório de confiabilidade. Com efeito, a abordagem aqui proposta resulta, de certo modo, num trabalho de compilação e, como tal, assume diversos riscos:

- a. a maioria dos materiais lexicográficos bi- e multilíngües existentes proporcionam tão-somente o termo na Língua de Partida (LP) e seu(s) possível(is) correspondente(s) na Língua de Chegada (LC),
- i. sem conter definições e/ou explicações que permitam ao usuário avaliar a pertinência do emprego do equivalente sugerido para a LC em determinado contexto,
 - ii. sem utilizar, ao menos de forma sistemática, uma convenção de indicação de domínio ou subdomínio, no caso dos dicionários técnicos "gerais",
 - iii. sem hierarquizar, de maneira consistente, o grau de generalidade/especificidade dos sinônimos e parassinônimos;
 - iv. no caso dos dicionários técnicos temáticos, sem efetivamente limitar os verbetes ao(s) tema(s) originariamente proposto(s);
- b. o grau de confiabilidade dos materiais é extremamente flutuante e, para os não-especialistas na área em questão, torna-se difícil avaliar, para determinada fonte e, mesmo, para determinado verbe, até que ponto a solução proposta pelo lexicógrafo é:
- i. correspondente ao uso atual;
 - ii. baseada em verificação efetiva do uso, a qualquer época, ou meramente proposta pelo próprio lexicógrafo, sem base outra que sua própria intuição ou inspiração de momento.

Ainda, esta mesma abordagem, caso implementada, deve ser reconhecida pelo que ela é, ou seja, um expediente que visa preencher, de maneira provisória, uma lacuna, a ser substituída, tão logo possível, por uma reorganização dos dados com base em levantamentos efetuados de acordo com uma metodologia terminológica que assegure a efetiva fiabilidade da informação.

2. REDUÇÃO DA MARGEM DE RISCO

A ausência de maiores especificações pertinentes à abordagem terminológica

descritiva – definição/explicação, falta de delimitação e explicitação temáticas – constitui uma lacuna informativa que somente poderá ser preenchida mediante uma retomada das investigações em bases efetivamente terminológicas, conforme indicadas sinteticamente acima, não cabendo, pois, a sua superação no âmbito da abordagem ora proposta.

A questão mais grave e que, caso não seja adequadamente equacionada, comprometeria o valor da compilação, tornando-a dispensável (para não dizer fonte menos de informações e mais de pandemônio terminológico), prende-se ao fator de confiabilidade das fontes. Por se tratarem de fontes bilíngües e, em muitos casos, abordarem a área técnico-tecnológica de maneira generalizante, o recurso à opinião dos especialistas de assunto parece pouco factível, salvo em prejuízo evidente do fator *tempo*, aqui considerado como justificativa para a abordagem ora sugerida.

Em contrapartida, é cabível o recurso ao conjunto de usuários mais facilmente definível de tais fontes, constituído pelos tradutores técnicos. Estes, embora tendam a concentrar a maior parte de sua produção profissional em torno de alguns campos temáticos mais limitados, são usuários necessariamente generalistas. Com efeito, o mercado da prestação de serviços de tradução no Brasil ainda não atingiu um tal nível de demanda que os tradutores possam dedicar-se a um único ramo do conhecimento técnico e/ou científico. Ainda, na realidade pragmática dos textos, são numerosos os hibridismos de especialidades⁴.

Tendo em vista esses condicionantes, efetuou-se, em meados de 1991, uma investigação junto aos tradutores técnicos de língua inglesa filiados à Sociedade de Tradutores Técnicos (STT) e à Associação Profissional dos Tradutores Públicos do Estado de São Paulo (ATPIESP)⁵, num total de 26 pessoas, solicitando-lhes que classificas-

sem os dicionários técnicos bi e multilíngües existentes, à época, no mercado editorial, para a relação *inglês-português*, numa escala de confiabilidade definida como segue:

0 – confiabilidade nula (“não vale o papel em que foi impresso”);

1 – ocasionalmente proporciona uma informação interessante, mas deve ser utilizado com extrema cautela;

2 – a informação é moderadamente confiável;

3 – a informação dada é, no geral, confiável, embora apresente grandes lacunas no rol de verbetes;

4 – a informação dada é, no geral, confiável, com poucas lacunas importantes;

5 – uma obra excelente, solucionando de maneira muito apropriada, dentro de sua especialidade, as dúvidas terminológicas que surgem na tradução;

retendo-se, como material adequado para integrar o conjunto de fontes de compilação, os títulos que obtiveram uma média mínima de 3,0 pontos e com pelo menos 6 indicações. Não constituiu objeto explícito do levantamento a mensuração do grau de atualidade das informações contidas nos dicionários em questão, bastando, para os propósitos do levantamento pretendido, o ano de publicação da primeira edição (no caso de títulos com ou sem reimpressões posteriores) ou da última edição revista, conforme o caso.

Do total de pessoas contatadas, quatorze responderam ao questionário. O resultado desta investigação apontou, como fontes primárias para uma base de dados bilíngües de e para o idioma inglês, os seguintes títulos (com a pontuação – média aritmética das respostas efetivas – indicada entre parênteses):

- Antas, L. M. – *Dicionário de Termos Técnicos Inglês/Português* (3,5)

- Cavalcante, J. C. – *Dicionário Inglês/Português de Termos Econômicos e Comerciais* (3,1)
- Fragomeni, A. H. – *Dicionário Enciclopédico de Informática* (3,6)
- Furstenau, E. – *Dicionário de Termos Técnicos Inglês/Português* (3,3)
- Krahenbuhl, H. – *Dicionário Inglês/Português para Executivos* (3,3)
- Sell, L. – *English/Portuguese Comprehensive Technical Dictionary* (3,1)
- Stedman – *Dicionário Médico* (4,5)
- SUCESU – *Dicionário de Informática Inglês/Português* (3,0)
- Taylor, J. – *Dicionário Metalúrgico Inglês/Português e Português/Inglês* (4,3).

3. O REGISTRO DOS DADOS

A intenção explícita do compilamento de dados terminológicos bilíngües já consignados em dicionários técnicos existentes é, como ficou dito, a de preencher provisoriamente uma lacuna, no aguardo da geração de informações terminológicas com base em estudos metodologicamente mais apurados (*bases terminológicas de segundo nível*).

Nesta perspectiva, propõe-se um sistema de registro que seja compatível com o sistema já desenvolvido pelo Grupo de Tradução, Lexicografia e Terminologia Informatizadas da FFLCH-USP (2) e atualmente em vias de implementação no Centro Interdepartamental de Tradução e Terminologia (CITRAT), da mesma instituição, ou seja, utilizando essencialmente o mesmo modelo de “ficha eletrônica” (TERMBASE), apenas algo mais simplificado, consistente com as limitações inerentes do presente projeto, mas permitindo sua integração com a ficha completa à medida em que for sendo efetuada a migração do primeiro para o segundo nível de elaboração.

Nestes termos, a ficha proposta – COM-PIBASE – teria a seguinte constituição:

Campo 1 – LP: Indica-se neste campo a sigla – dois caracteres – correspondente à língua de partida, isto é, a língua do termo constante do Campo 2.

Campo 2 – TERMO-LP: Este campo, correspondente ao verbete, conterá o termo (mono ou plurivocabular), tal como verificado no uso efetivo (na fonte), reduzido, quando for o caso, à sua forma morfológica básica (no caso do português, por exemplo, ao masculino singular ou ao infinitivo, salvo casos particulares: sentido diverso conforme o gênero ou número, *et simil*).

Campo 3 – FONTE: Este campo conterá, sob forma de sigla composta de 4 caracteres alfanuméricos, a indicação da fonte bibliográfica. A referência bibliográfica completa será consignada no arquivo FONSTERM. Tanto quanto possível, a forma abreviada de referência deverá ter um caráter mnemônico. Note-se que a cada fonte corresponderá uma ficha, mesmo no caso de reiteração de dados entre as diversas fontes constituintes do córcpus.

Campo 4 – ANO: O ano de publicação do dicionário fonte constitui um dos principais índices de confiabilidade da fonte, em termos de retrato atual do uso terminológico. Na ficha de arquivo TERMBASE, indicar-se-á apenas os dois últimos algarismos do ano de publicação ou entrevista, os dados completos sendo consignados à parte, em arquivo bibliográfico próprio.

Campo 5 – USO: Recupera eventuais informações constantes da fonte a respeito de restrições de uso de natureza regional, dialetal ou socioletal. Será necessário estabelecer um registro das siglas utilizadas (p. ex.: PE para “português europeu”, US para “inglês americano”, AC para “acadêmico”, de modo a evitar discrepâncias entre os diversos COMPIBASEs. Quando nada houver a assinalar a este respeito, pressupor-se-á uso não-marcado (nm).

Campo 6 – LC: Indica-se neste campo a

língua de chegada, isto é, a língua do termo constante do Campo 7.

Campo 7 – TERMO-LC: Este campo conterá o equivalente na LC ao termo assinalado no Campo 2, conforme consignado na fonte. Caso a fonte indique mais de um equivalente, para cada relação de equivalência será constituída uma ficha.

Campo 8 – ÁREA: a área nocional básica a que pertencem os termos LP e LC na relação de equivalência proposta pela fonte.

Campo 9 – DATA: Este será um campo alterável. Quando a ficha for elaborada pela primeira vez, será indicada a data de sua elaboração. Quando de suas posteriores revisões ou reelaborações, a data será atualizada a cada nova intervenção.

Campo 10 – DOC; Campo 11 – REV: Designam, respectivamente, o documentador (primeiro redator) e o revisor da ficha. Aplicam-se, *mutatis mutandis*, para esses campos, os critérios descritos para o Campo 3. As referências completas serão consignadas em arquivo próprio.

Quadro Resumo 1 – Ficha de Levantamento em Base Informatizada, com Previsão do Número de Caracteres Necessários para Cada Campo

Campo	Caracteres
1: LP	2
2: TERMO-LP	70
3: FONTE	4
4: ANO	2
5: USO	2
6: LC	3
7: TERMO-LC	70
8: ÁREA	8
9: DATA	8
10: DOC	4
11: REV	4
Total	177

Quadro Resumo 2 – Modelo de Ficha de Levantamento Manual

01	02		
03	04	05	06
07			
08	09	10	11

REFERÊNCIAS BIBLIOGRÁFICAS

- (1) AUBERT, F. H. Problemas e urgências na inter-relação terminologia/tradução. In: *Revista Alfa*, São Paulo, FUNDUNESP, nº 36, 1992.
- (2) _____ *et al.* Tradução, lexicografia e terminologia informatizadas. In: *Anais do XXXIX Seminário do GEL*. 1992.
- (3) DUBUC, R. *Manuel pratique de terminologie*. Montreal, Linguattech, 1985.

1. Professor associado, Departamento de Letras Modernas, Universidade de São Paulo.
2. Para uma apresentação detalhada do processo da pesquisa temática em terminologia descritiva, vide Dubuc (3).
3. Resguardados, evidentemente, os direitos dos detentores do copirraite dos referidos materiais.
4. Assim, por exemplo, um contrato de transferência de tecnologia conterá, enquanto texto, dados terminológicos pertinentes ao direito, ao comércio, às finanças e à tecnologia específica objeto da transferência em questão.
5. A opção por consultar os tradutores da STT e da ATPESP prende-se ao fato de que a primeira entidade é extremamente exigente no que tange às qualificações profissionais dos candidatos a filiação, incluindo, como pré-requisitos, no mínimo três anos de experiência profissional contínua, um exame escrito, análise de currículo e verificação do comportamento ético dos candidatos. Os membros da ATPESP, por sua vez, foram nomeados com base em concurso público com provas escritas e orais de tradução e versão. O concurso mais recente data de 1979-1980, assegurando, portanto, à época do questionário, que os associados a tal entidade contavam com um mínimo de onze anos de vivência profissional. Tais circunstâncias qualificam os membros destas entidades como interlocutores habilitados para o objetivo da investigação.