

**VOTEC – VOCABULÁRIO TERMINOLÓGICO
DE LINGUÍSTICA HISTÓRICA – PORTUGUÊS-INGLÊS**

Márcio Issamu Yamamoto (UFU/UFG)
issamu2009@gmail.com

RESUMO

O objetivo deste trabalho é apresentar o resultado dos passos e a metodologia usados para a construção do dicionário bilíngue, português-inglês, o *VoTec* – *Vocabulário Técnico* – disponível on-line, na área de linguística histórica. Posterior aos levantamentos de *corpora*, da construção das listas de palavras e de palavras-chave, os dados foram inseridos na plataforma do *VoTec* para construção das definições do dicionário. O “*VoTec* é uma ferramenta que se vale de *corpora* técnicos para a construção de seus verbetes e de um banco de dados para o seu funcionamento” (FROMM, 2007, p. 8). Diferente da metodologia tradicional, em relação às obras terminográficas, os dados são armazenados on-line. Trataremos das dificuldades encontradas nos passos da organização dos traços conceituais, para a construção das definições e nuances entre os *corpora* de língua portuguesa e língua inglesa. Usamos, além do padrão GPDE (gênero próximo, diferença específica), o princípio do LEAD do jornalismo nas organizações dos traços conceituais.

Palavras-chave: *VoTec*. Vocabulário. Tecnologia. Linguística histórica. Lexicografia.

1. Introdução

O objetivo deste artigo é tratar das dificuldades e nuances encontradas no desenvolvimento da dissertação de mestrado intitulada linguística histórica e linguística de *corpus* – caminhos que se cruzam para desvelar a história da linguagem: um vocabulário bilíngue português-inglês. Os passos desenvolvidos neste trabalho foram apresentados e publicados neste simpósio, SINEFIL, nas edições anteriores (YAMAMOTO, 2013, p. 691-705 e 2015, p. 1262-1273); neles foi explicado e apresentado o projeto, bem como detalhada a microestrutura deste vocabulário como parte constitutiva do fazer terminológico.

Esta pesquisa objetivou a construção de um vocabulário de especialidade bilíngue, português-inglês, na área de linguística histórica – doravante linguística histórica, ou seja, é um trabalho que se insere dentro da área de terminologia e terminografia. Os *corpora* são de caráter científico: teses, dissertações e artigos científicos, todos de acesso público e gratuito. A metodologia e abordagem usada foi a linguística de *corpus* e

nosso embasamento teórico foi a partir de Cabré (1999) e Barbosa (1990), quanto ao conceito de vocabulário e Cabré, no que tange à teoria comunicativa da terminologia e a visão poliédrica dos termos. A ferramenta usada na produção das listas de palavras, de palavras-chave e concordanciador foi o *WordSmith Tools 6.0*. (WST & SCOTT, 2012)

2. *Discussão dos resultados*

Discutiremos adiante os resultados obtidos, as dificuldades e nuances relativas aos dados proveniente dos *corpora* e usados para a construção das definições dos termos na plataforma do *VoTec*¹¹⁵.

Para a condução do primeiro objetivo, a reelaboração da árvore de domínio da linguística histórica (BARROS, 2004), fizemos uma pesquisa bibliográfica com obras de autores referência das áreas, para que pudéssemos delimitá-las com rigor metodológico e científico adequados a esta proposta de dicionário bilíngue. Este rigor se fez necessário por dois motivos: (i) porque a etimologia, a filologia e linguística histórica são subáreas da linguística; e (ii) devido ao fato de serem disciplinas que emitem um olhar diacrônico sobre a língua, o limiar que as separa às vezes parece ser tênue.

Na área da filologia, pudemos analisar, a partir daquela proposta por Fromm e Yamamoto (2013), as áreas relacionadas à linguística histórica, para que traços conceituais de cada uma delas fossem delineados; pesquisamos autores como Jordan (1982), Vidos (1996), Bassetto (2001) e Faraco (2005). A partir das pesquisas, entendemos que a Filologia é uma das ciências mais antigas e trabalha com a pesquisa de textos, no sentido de buscar o contexto autoral, geográfico, social e linguístico da produção de obras. Segundo Bassetto (2001), o filólogo era um dos sábios que tinha um conhecimento amplo em várias áreas, era, diríamos, um profissional de caráter enciclopédico. A filologia se interessa pelo estudo do texto, usa o método da crítica textual e o histórico comparativo para um texto, o contexto de produção e língua, no qual se insere. Além da filologia como área, ela se subdivide em áreas menores, na tentativa de atender línguas diversas, como filologia clássica, germânica, românica, portuguesa, dentre outras.

¹¹⁵ Disponível em: <<http://www.pos.voteconline.com.br>>. Acesso em: 08-04-2016.

A etimologia, como ciência, apresenta característica semelhante à terminologia; pode denominar uma área da ciência ou o significado das palavras. Por isso, foi comum encontrarmos linhas de concordância que traziam ocorrências como “a etimologia da palavra x é y”. Os autores Viaro (2011) e Durkin (2009) apresentam a área e mostram que o objeto central desta disciplina é o étimo. A partir dele, os gregos acreditavam que podiam resgatar o conceito que levou os primeiros profissionais da linguagem a nomear os objetos, numa tentativa de conciliar o conceito com a referente.

A linguística histórica é uma ciência que vem a se consolidar bem depois dessas duas, mencionadas anteriormente. Ela surge com a busca dos filólogos europeus para encontrar a origem das línguas, usando o método comparativo e tecendo as leis fonéticas que serviriam de base para explicar as mudanças linguísticas. Friedrich Diez será um dos filólogos renomados, responsável pela consolidação da disciplina na academia. Esses estudos se amadurecem e os linguistas propõem o indo-europeu e o sânscrito como língua que antecedem as línguas atuais como o grego, latim, português e inglês, dentre outras. No Brasil, uma grande cientista dessa área foi Mattos e Silva, professora que se dedicou aos estudos profundos do português, juntamente com Mary Kato, e contribuiu para a consolidação da disciplina no Brasil. O objeto de estudo da linguística histórica são as mudanças fonéticas, mórfica, sintáticas e semântico-lexicais que ocorrem na língua. A partir do estudo dessas mudanças, também chamadas metaplasmos, consegue-se traçar a origem da língua e explicar suas mudanças no tempo, espaço numa sociedade.

O estudo dessas áreas nos levou a concluir que, apesar de serem áreas que dialogam entre si e apresentam muitas semelhanças, elas possuem certa distinção que as impedem de serem aglutinadas sob uma única subárea da linguística. Isso se deu devido ao fato de elas terem estatutos diferentes, tais quais: são nomeadas de forma diferente, têm objeto, método e objetivo que são distintos entre si. Contudo a linguística histórica é uma área que absorve os dados provenientes dessas disciplinas para compor a subárea da linguística histórica, independentemente das outras.

O segundo objetivo deste trabalho foi a compilação de *corpora* bilíngue, para que servissem como banco de dados, de onde os traços conceituais foram extraídos, a partir dos contextos definitórios e explicativos. Este trabalho foi breve, devido à rapidez proporcionada pelos computadores e pela internet. Algumas dificuldades que enfrentamos foram:

(1) nem todos os *corpora* disponíveis puderam ser salvos em formato txt, devido ao fato de estarem em formato PDF criptografado; logo, às vezes, fez-se necessário o uso de ferramentas cibernéticas que nos ajudam a editar os textos em PDF, convertê-los para versões legíveis pelo WST, isto é, em formato txt; (2) em inglês, a possibilidade de conseguirmos artigos na área de etimologia foi menor, já que a política de popularização do conhecimento nos países anglófonos é mais dificilmente adotada. Isso nos obrigou a mudar a busca na internet, buscando obras diferentes daquelas propostas no projeto original. Optamos por aumentar o *corpus* de inglês com um dicionário de etimologia, que traz definições e comparações de diferentes propostas de etimologias dos verbetes, com um parecer final do autor. Este caráter comparativo da obra, permitiu que houvesse um material não só de cunho lexicográfico, mas também enciclopédico e também acadêmico-científico, o que contribuiu para a diminuição de interferências na busca por contextos definitórios. Tomamos o cuidado de manter os *corpora* balanceados em números de ocorrências (aproximadamente 500 mil), apesar de não ter sido possível o mesmo balanceamento em número de textos. Isso se deu devido ao fato que, dentre os arquivos em inglês, houve anais de encontros internacionais em linguística histórica (1999, 2001, 2003), mais extensos que os outros gêneros, que agregavam vários artigos nesta área. Outra razão que justifica este baixo número de artigos é a dificuldade de acesso a artigos científicos em inglês, de acesso gratuito na Internet.

A extração de candidatos a termos, provenientes das listas de palavras-chave, após processamento dos *corpora* pelo *WordSmith Tools 6.0* (SCOTT, 2012), foi um passo dado a partir da lista em português, pelo fato de o nosso público alvo ser o brasileiro, em primeiro lugar. Isto é, partimos dos termos selecionados pela lista de palavras-chave do português, para depois buscá-los na mesma lista em inglês. Nos casos em que não foi encontrado um termo correspondente no inglês, fizemos a busca na lista de palavras dessa língua. Uma ocorrência que nos chamou a atenção foi o caso do termo *ablaut* em inglês, traduzido como apofonia em português. Este foi um termo bastante recorrente no inglês, aparecendo 157 vezes, em três textos, correspondendo a uma porcentagem de 0,02% do *corpus*. Em português, o termo apofonia aparece 6 vezes na lista de palavras, não tendo a mesma correspondência de porcentagem no texto¹¹⁶.

¹¹⁶ Devido à baixa ocorrência do termo, cuja frequência é menor que 0,01%, esta porcentagem não é

A disponibilização de dados oriundos de fontes confiáveis aos usuários concretizou-se por meio da plataforma, já que as definições foram construídas a partir de traços semânticos provenientes de *corpora* acadêmicos. Esta foi uma das grandes vantagens proporcionadas pela linguística de *corpus*, adotada como metodologia e como abordagem neste trabalho. Caso não a tivéssemos adotado, com certeza, o passo de confirmação dos termos, que teria sido feito com especialistas, comprometeria a conclusão desta pesquisa, já que não haveria como certificar-nos de que o termo pertenceria ou não à linguística histórica. Outra limitação seria saber se o candidato a termo seria ou não uma palavra-chave; também não teríamos como localizar contextos definitórios e explicativos para a construção de definições adequadas para os consulentes. Os *corpora* coletados são todos provenientes da comunidade científica, conforme mencionado e especificado na seção de metodologia.

A organização de dados, na qual incluímos o termo, a área a qual pertence, a classe gramatical, a definição, o exemplo dos *corpora*, entre outros, foi feita em fichas de registro, e os dados dispostos nas macro e microestrutura do vocabulário. Primeiramente fizemos a análise componencial (ILARI, 2002), para que pudéssemos organizar os dados nas fichas terminológicas do *VoTec*. Além dos dados referentes ao termo em si, agregamos a definição de dicionários e da Wikipédia, de modo que o consulente pudesse acessar as informações de formas diversas, ou seja, customizar a plataforma com suas opções de visualização, para que o uso fosse prático e de rápido acesso. Este passo objetivou disponibilizar aos leitores uma obra de leitura acessível, organizada e padronizada. Isto se fez necessário, principalmente ao considerarmos alunos ingressantes de graduação, que ainda não dominam a terminologia da linguística e podem não acompanhar os conceitos que jazem por detrás das palavras. Esta é uma das vantagens do *VoTec*, já que as definições são redigidas de forma curta, objetiva e restritas a uma oração.

A elaboração de definições para os termos selecionados foi baseada no padrão GPDE (gênero próximo e diferença específica), adotado nas obras terminológicas em nível nacional e internacional. Esta definição partiu das noções mais amplas, da relação de hiperonímia, para aquelas mais específicas, as relações de hiponímia (ILARI, 2002, p. 39). Ao analisar os padrões adotados pela academia canadense, foi possível observar a importância do padrão GPDE, haja vista a importância do Canadá co-

disponibilizada pelo WST 6.0.

mo país bilíngue, bem como referência mundial em terminologia bilíngue e em teóricos da lexicologia, lexicografia, terminologia e terminografia.

A disponibilização dos resultados no *VoTec*, foi feita sem dificuldades, já que esta plataforma se encontra em funcionamento há mais de 7 anos. Quanto ao processo de ligar o termo principal aos termos remissivos, tivemos a seguinte limitação: os termos polissêmicos e recorrentes em várias subáreas da linguística eram disponibilizados pela plataforma para serem escolhidos como remissivos. Por exemplo, ao definir o termo *gramática*, um dos termos remissivos era *língua*. Ao solicitarmos a busca deste termo no sistema, todos os termos *língua*, pertencentes a várias subáreas da linguística eram disponibilizados pelo sistema. Não havia como identificar qual deles era o termo pertencente à linguística histórica. Logo para sanar esta limitação, tivemos que descastrar todos os termos *língua*, pertencentes às outras subáreas da linguística, e deixar somente aquele pertencente à subárea da linguística histórica. Dessa forma, conseguimos cadastrar o termo como remissivo da linguística histórica, ligando ao termo gramática.

Quanto à escrita e constituição da definição, um desafio a ser enfrentado foi o de conseguir evitar uma definição que fosse muito técnica¹¹⁷ e incompleta. Na construção da definição do termo *palavra*, deparamo-nos com a seguinte possibilidade: *unidade que expressa um conceito abstrato, composta por uma base e um sufixo*. Contudo, ao ser lida por um consulente, com certeza a definição seria muito abstrata, e o sentido do termo *palavra* ficaria confuso e obscuro. Como NOTA teríamos o seguinte: *nó formado a partir do mesmo afixo/base que identifica as estruturas morfológicas formadas pela adjunção dos sufixos de flexão e de número para os adjetivos e nomes*. Para o público-alvo do *VoTec*, que seria alunos de tradução, alunos de letras e tradutores (FROMM, 2007), o objetivo final seria somente o de checar a correspondência dos termos nas duas línguas: inglês e português. O recurso da inserção da nota foi motivado pelas leituras do padrão terminográfico proposto por Bevilacqua & Finatto (2006) para que a definição não ficasse muito longa. Além de longa, uma definição redigida por orações separadas por ponto e vírgula seria confusa e não desejável para a compreensão do aspecto concei-

¹¹⁷ A expressão "muito técnico" se explica pelo fato de o nosso público alvo ser de alunos de graduação e tradutores que não dominam muito bem o sistema linguístico inglês, nem a área científica de linguística histórica.

tual do termo.

O termo *etimologia* foi difícil de ser definido, apesar da recorrência de quase 50 vezes do termo, no *corpus* de estudo. O termo, assim como acontece com a combinação terminologia *verso* terminologia, além de ser usado como subárea da linguística, também é usado como a etimologia do termo X ou Y. Neste trabalho, a segunda opção foi mais recorrente no *corpus* de inglês. No *corpus* de português ela foi mencionada como o nome da ciência mais vezes que no *corpus* de inglês. Observamos então que, como ciência, ela se coloca mais vezes, no caso do português, que no caso do inglês. Isso mostra certa consolidação da disciplina em inglês, diferentemente do português. Chegamos a esta conclusão uma vez que o termo recorria no *corpus* de inglês sem que estivesse inserido em contextos definitórios ou explicativos. As ocorrências, em sua maioria, referem-se a etimologia de dado termo e não à explicação ou definição do termo etimologia. Essa característica induz à compreensão de que, para o leitor de fala inglesa, o termo já está difundido sem que sua explicação ou definição se faça necessária.

Nesse caso, observamos que a bibliografia disponível em português brasileiro, para o estudo da etimologia como ramo da ciência linguística, é mais limitada. A ocorrência foi em quinze textos de português e cinco de inglês. Em inglês, o termo recorreu 783 vezes, enquanto que no português ele foi recorrente 294 vezes. Em contrapartida, o que observamos no *corpus* em inglês é que os estudos de Etimologia estão mais consolidados, provavelmente pela tradição filológica, etimológica e linguística mais antiga e tradicional na Europa que no Brasil.

Uma das abas das fichas terminológicas disponibiliza um espaço onde cadastramos as relações de sinonímia, hiponímia, hiperonímia, cohiponímia, antonímia, traçadas dentro do texto, das linhas de concordância dos *corpora*. Seguem abaixo alguns exemplos encontrados nos *corpora*.

1. Sinonímia

Etimologia = origem dos vocábulos

Seu autor foi Santo Isidoro de Sevilha (c.560-636AD) que, por ter sido o compilador da primeira base de dados do Ocidente, é hoje considerado o padroeiro da Internet. “A etimologia é a origem dos vocábulos, cuja força (de significação), quer em nomes, quer em palavra é desvendada pela interpretação”, afirma o bispo da igreja católica.

2. Hiponímia

Gramática > Língua¹¹⁸

A definição de gramática histórica de Francisco J. Martins Sequeira (1938b) é mais completa no segundo trabalho publicado, salientando-se neste que no estudo da evolução de uma língua devem ser estudados todos os fenômenos e que se deve ter em consideração tanto o eixo temporal como o eixo espacial.

3. Hiperonímia

Gramática < sons e letras¹¹⁹

Na análise que faz desta corrente gramatical filosófica, Duarte (1995: 15) considera que a "ideia de que as categorias gramaticais refletem elementos e processos mentais universais, mas de que é necessário analisar cada língua particular porque tais processos se podem exprimir diferentemente de língua para língua, o afastamento do objetivo normativo (a gramática é a arte de falar e não a arte de bem falar), a distinção clara entre sons e letras e a prioridade atribuída à análise dos primeiros são características importantes e singularmente contemporâneas desta escola de pensamento que Chomsky, na década de 60, invoca centralmente ao procurar os fundamentos da gramática generativa na tradição racionalista de teorização e análise das línguas naturais".

4. Co-hiponímia

Verbo = substantivo; advérbios; conjunção de coordenação; adjetivo.¹²⁰

Outros tipos de predicação envolvem, diga-se de passagem, verbo e o que vem a seguir, substantivo ou adjetivo, que, juntos com o verbo remetem ao sujeito, por exemplo: eu sinto dor, em que sinto dor refere-se ao sujeito, não apenas sinto.

5. Antonímia

Dentro da lista de candidatos a termos, extraída de uma lista de palavras-chave de 500 palavras, não encontramos contextos que trouxessem esta ocorrência.

Um aspecto interessante encontrado durante o processamento de dados foi o termo *futuridade*, termo com contexto definitório e explicati-

¹¹⁸ Usamos o símbolo > para indicar que o termo gramática é hipônimo de língua.

¹¹⁹ Usamos o símbolo < para indicar que o termo gramática é hiperônimo de sons e letras.

¹²⁰ Usamos o símbolo = para indicar a co-hiponímia entre termos.

vo, presente em um único texto no *corpus*, somente em português, sem correspondência no inglês.

Quanto à disponibilização de dados para o consulente, um detalhe observado foi quanto à sinonímia. Na presença de mais de um sinônimo, a plataforma só permite a visualização de alguns termos. Para que haja a visualização de todos os termos, é necessário solicitar o modo descritivo de visualização.

As relações semânticas entre termos, às vezes, são difíceis de serem traçadas, mesmo analisando-se o contexto, como foi o caso da tentativa de extração dos traços conceituais para elaborar a definição da expressão *linguistic composition*, como podemos ver no excerto seguinte: "*Once one child converges on some new grammar, then the linguistic composition of the population will change, because the child with the new grammar produces different structures; this may have domino effects*".

Todas essas experiências e tentativas contribuíram para uma melhor compreensão do fazer terminológico, da compreensão do processo de produção de sentidos para a identificação de traços semânticos e construção de definições.

Abaixo podemos ver os termos que compuseram o *Vocabulário de Linguística Histórica*, produto final desta pesquisa. Podemos visualizar as duas colunas, a da esquerda com os termos em português, e a da direita, com os termos em inglês.

Quadro 1. VoTec - Termos do Vocabulário bilíngue de linguística histórica.

<i>VoTec – Vocabulário de linguística histórica – lista final.</i>	
Português	Inglês
1. Língua	Language
2. Etimologia	Etymology
3. Gramática	Grammar
4. Palavra	Word
5. Verbo	Verb
6. Linguística	Linguistics
7. Linguística histórica	Historical Linguistics
8. Mudança	Change
9. Fala	Speech
10. Linguagem	Language
11. Texto	Text
12. Variação	Variation
13. Derivação	Derivation

14. Substantivo	Noun
15. Adjetivo	Adjective
16. Prefixo	Prefix
17. Léxico	Lexicon
18. Analogia	Analogy
19. Clítico	Clitic
20. Advérbio	Adverb
21. Particípio	Participle
22. Sufixação	Suffixation
23. Prefixação	Prefixation
24. Afixo	Affix
25. Mudança linguística	Linguistic change
26. Língua românica	Romance Language

Os termos apresentados no Quadro 1 foram selecionados por terem equivalência conceitual e por serem termos recorrentes na área de linguística histórica. Dentre os termos, salientamos aqueles que trazem conceitos mais usados nessa área, bem como nas áreas que a alimentam, a etimologia e a filologia; termos tais quais: língua, etimologia, linguística histórica, mudança, mudança linguística, variação, língua românica, analogia, sufixação e prefixação.

3. Conclusão

A partir da pesquisa desenvolvida e da prática terminológica proporcionada pela plataforma do *VoTec*, podemos observar alguns resultados quanto ao que vem a ser a produção de uma obra terminográfica. Primeiramente, foi necessário conhecer o que é o campo ou quais são os limites da terminologia e terminografia. Apesar da prática já estar instituída, a área em si ainda é estranha ou nova para linguistas de outras subáreas da linguística, o que pode levar a uma confusão entre a terminografia e a lexicografia como áreas que trabalham com o registro do léxico. Daí a importância da diferenciação conceitual dos objetos dessas disciplinas, quais sejam a unidade lexical versus a unidade terminológica.

Estar inscrito em uma escola que norteia os conceitos e princípios a serem seguidos e defendidos também é crucial no fazer científico da academia. A teoria comunicativa da terminologia permite que entendamos a unidade terminográfica que pode se identificar como unidade lexical ou terminológica, dependendo do seu contexto comunicativo de uso. A linguística de *corpus*, enquanto metodologia, que parte da visão probabilística da língua, possibilita ao seu usuário o manejo de dados concretos, e significativos quanto à representatividade e quanto ao uso em

corpora autênticos.

Os resultados finais obtidos para a elaboração da árvore de domínio foram que a área da linguística histórica é difícil de ser definida como uma área estanque. Isso se deve ao fato de que ela se alimenta de subáreas da linguística, principalmente como *corpus*, e de outras para fazer sua análise de dados. Isto é, falar em linguística histórica significa envolver a etimologia, a filologia e a linguística diacrônica. Quanto às unidades terminológicas, observamos que nos *corpora* de português e inglês há termos correspondentes que servem de candidatos a termos. Há também aqueles que trazem consigo traços semânticos em uma língua e não em outras, o que nos impossibilitou de manter os candidatos a termos nas listas de termos a serem inseridos na plataforma. Os *corpora* nos permitiram ver que a linguística histórica vive momentos diferentes no Brasil e Europa, o que já era esperado, já que grande parte da tradição filológica e linguística ocidental iniciou-se na Europa, mais especificamente na Alemanha, berço da língua inglesa, e na Itália e França origem dos estudos de linguística e filologia românica.

A aplicabilidade que podemos dar a esta pesquisa é a condução de projetos em áreas de letras e outras áreas, já que a terminologia é uma disciplina de caráter interdisciplinar. Dominados e aprimorados os procedimentos do fazer terminológico e terminográfico, desde a delimitação das áreas por meio da construção da árvore de domínio até a disponibilização dos dados na plataforma do *VoTec*, a construção de novos vocabulários técnicos se tornará uma realidade.

A plataforma do *VoTec* vem comprovar a evolução tecnológica no fazer terminográfico. Ela traz vantagens e desvantagens quanto ao fazer terminográfico, mas o ganho é maior que as perdas e, uma vez que uma nova plataforma está sendo concebida, as limitações desta versão serão certamente aprimoradas. É uma plataforma que possibilita a organização eficaz de dados, por meio das fichas virtuais, de forma que a rapidez com que se conclui o fazer definitório seja incomparavelmente maior em relação ao modo como era feito antigamente. Outro aspecto relevante é a disponibilidade dos dados, bem como seu alcance em nível mundial, gratuitamente, via rede mundial de computadores. O vocabulário técnico de linguística histórica possibilita a compreensão de termos técnicos por parte dos tradutores, pesquisadores e docente que estão envolvidos com a linguística histórica e com a pesquisa científica da história das mudanças linguísticas.

REFERÊNCIAS BIBLIOGRÁFICAS

BARBOSA, M. A. Lexicologia, lexicografia, terminologia, terminografia, identidade científica, objeto, métodos, campos de atuação. In: Simpósio Latino-Americano de Terminologia e I Encontro Brasileiro de Terminologia Técnico-Científica. 2., 1990, Brasília, *Anais...* p. 152-158.

BARROS, L. A. *Curso básico de terminologia*. São Paulo: EDUSP, 2004.

BASSETTO, B. F. *Elementos de filologia românica: história externa das línguas*. São Paulo: EDUSP, 2001, p. 17-42.

BEVILACQUA, C. R.; FINATTO, M. J. B. *Lexicografia e terminografia: alguns contrapontos fundamentais*. São Paulo: Alfa, 2006. Disponível em: <<http://www.ufrgs.br/texteccc/textquim/arquivos/03-Bevilacqua-Finato.pdf>>. Acesso em: 07-04-2016.

DURKIN, P. *The Oxford Guide to Etymology*. New York: Oxford University Press, 2009.

FARACO, C. A. *Linguística histórica: uma introdução ao estudo da história das línguas*. São Paulo: Parábola, 2005.

FROMM, G. *VoTec: a construção de vocabulários eletrônicos para aprendizes de tradução*. São Paulo, 2007. Tese (Doutorado em Estudos Linguísticos e Literários em Língua Inglesa). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

FROMM, G.; YAMAMOTO, M. I. Terminologia, terminografia, tradução e linguística de *corpus*: a criação de um vocabulário bilíngue sobre linguística. In: TAGNIN, S.; BEVILACQUA, C. *Corpora na terminologia*. São Paulo: Hub, 2013.

ILARI, R. *Introdução ao estudo do léxico: brincando com as palavras*. São Paulo: Contexto, 2002.

IORDAN, I. *Introdução à linguística românica*. Trad.: Júlia D. Ferreira. Lisboa: Fundação Calouste Gulbenkian, 1982.

SCOTT, M. *WordSmith Tools*. Versão 6, 2012. Disponível em: <<http://lexically.net/wordsmith/version6>>. Acesso em: 07-04-2016.

VIARO, M. E. *Etimologia*. São Paulo: Contexto, 2011.

VIDOS, B. E. *Manual de linguística românica*. Trad.: José Pereira da Silva. Revisão técnica de Evanildo Bechara e Marlit Cavalcante Bechara.

Rio de Janeiro: Eduerj, 1996.

YAMAMOTO, M. I. Linguística histórica e linguística de *corpus*: uma proposta para dicionário bilíngue português-inglês. *Revista Philologus*. Rio de Janeiro, Ano 19, n. 55 Suplemento, p. 691-705, jan./abr. 2013. Disponível em <<http://www.filologia.org.br/revista/55supl/55supl.html>>. Acesso em: 02-04-2016.

_____. Votec – fazer terminológico: a microestrutura em dicionários virtuais. *Revista Philologus*, Rio de Janeiro, Ano 21, n. 61 Suplemento, p. 1262-1273, jan./abr.2015. Disponível em <<http://www.filologia.org.br/rph/ANO21/61supl/093.pdf>>. Acesso em: 03-04-2016.