

# Extração de conhecimento: métodos de identificação de relações terminológicas aplicados ao português do Brasil

## Extraction of knowledge: methods of identification of terminological relations applied to Brazilian Portuguese

Flávia Lamberti\*

Resumo: Este estudo trata dos métodos de identificação de relações terminológicas, usados na metodologia da pesquisa terminológica. São apresentados três tipos de métodos baseados em *corpus* que investigam contextos ricos em conhecimento, a saber: i) padrões de conhecimentos, ii) propriedades formais e iii) coocorrência, todos passíveis de indicação da ocorrência de relações terminológicas entre termos (MEYER 2001; L'HOMME; MARSHMAN 2006). Os métodos são aplicados em *corpus* de textos didáticos da área de Economia no português do Brasil. O presente trabalho busca contribuir com a etapa de construção e representação da estrutura de conhecimento (SAGER 1990: 13) e de seleção de candidatos a termo (L'HOMME 2004: 48) e, como consequência, com a aquisição de conhecimento em uma determinada área, em especial pelo estudante de tradução.

---

\* Docente Adjunta da Universidade de Brasília (UnB).

Palavras-chave: relações terminológicas; métodos baseados em corpus; contextos ricos em conhecimento; aquisição de conhecimento.

Abstract: This paper deals with the methods for discovering terminological relationships applied in the terminology work methodology. We will describe three corpus-based methods for identifying and extracting knowledge-rich contexts: i) knowledge patterns, ii) formal properties and iii) co-occurrence, each of them potential indicators of terminological relationships between terms (MEYER 2001; L'HOMME; MARSHMAN 2006). We will apply those methods to didactic texts in Brazilian Portuguese from the subject field of Economics. This study intends to contribute to knowledge construction and representation (SAGER 1990: 13) and to term candidate selection (L'HOMME 2004: 48) and, as a consequence, to knowledge domain acquisition especially by the student translator.

Keywords: terminological relationship, *corpus*-based methods, knowledge-rich contexts, domain knowledge acquisition.

## 1. Introdução

Este estudo se situa no âmbito do projeto de pesquisa “Ensino de tradução e estudos terminológicos orientados à formação do tradutor” desenvolvido no curso de graduação em Letras-Tradução e no Programa de Pós-Graduação em Estudos da Tradução (POSTRAD) da Universidade de Brasília. Consideramos que o desenvolvimento de pesquisas terminológicas pelo estudante de tradução e a compreensão do funcionamento da terminologia e de suas relações representadas no texto, é uma atividade que pode contribuir para o desenvolvimento de sua competência, mais especificamente competência cognitiva (aquisição de conhecimento na área de especialidade), competência linguística (conhecimento da língua ou línguas de trabalho), competência sociofuncional (conhecimento do uso adequado de determinada construção haja vista a intenção, o tipo de texto, a situação comunicativa, os interlocutores) e, também, competência metodológica (conhecimento do trabalho sistemático da pesquisa terminológica), tais como definidas por Cabré (1999: 195-196). Consideramos, assim, a partir de Cabré (1999: 193-195), que o desenvolvimento dessas competências depende do aumento do nível de comprometimento (em espanhol *niveles de implicación*) do tradutor com a pesquisa terminológica.

No presente trabalho, apresentaremos métodos de investigação em corpora para a identificação e extração de contextos ricos em conhecimento (KRC, na sigla em inglês - *knowledge rich contexts*), com base em Meyer (2001) e L’Homme; Marshman (2006). Apresentaremos também a aplicação desses métodos para a identificação de contextos no português do Brasil. A proposta é de que esses contextos apresentam relações terminológicas e, conseqüentemente, são indicativos da representação do conhecimento e, conseqüentemente, da ocorrência de termos.

O presente trabalho busca contribuir para a condução de pesquisas terminológicas, mais especificamente para a etapa de construção e representação da estrutura de conhecimento (SAGER 1990: 13) e seleção de candidatos a termo (L'HOMME 2004: 48) e, como consequência, para a aquisição de conhecimento em uma determinada área, com ênfase na aquisição de conhecimento pelo estudante de tradução.

## 2. Contextos ricos em conhecimento

A área de conhecimento selecionada para a pesquisa é a Economia, haja vista ser uma área temática de disciplinas de Prática de Tradução do curso de graduação em Letras-Tradução da Universidade de Brasília. A pesquisa foi realizada em textos da área pertencentes ao gênero didático e foram selecionados dois livros didáticos, ambos em português do Brasil: i) livro didático traduzido para o português do Brasil (MCCONNELL; BRUE, 2001); e ii) livro didático escrito originalmente em português do Brasil (LOPES; VASCONCELOS, 2011). Em ambos os livros, foi escolhido o mesmo tema, *moeda e a oferta de moeda*.

O texto didático é um tipo de gênero textual usado na aquisição sistematizada de conhecimento ao longo do processo de aprendizagem pelo estudante. É possível observar, em especial, nesse tipo de texto i) organização e apresentação do conteúdo temático a ser desenvolvido de forma ordenada e seriada, no sentido do conteúdo mais básico ao mais complexo, ii) uso de terminologia específica para representar a organização desse conhecimento, e iii) uso de padrões metalinguísticos para a introdução de conhecimento (quer dizer, sequências linguísticas que introduzem conceitos).

Na metodologia da pesquisa terminológica adotada, a investigação de uma área requer a representação de estruturas do conhecimento, as quais são construídas por meio de relações terminológicas (SAGER, 1990, p. 13; L'HOMME, 2004, p.25). Essas relações podem apresentar-se por meio de relações hierárquicas (e.g. hiperonímia - *publicação* é hiperônimo de *revista*), relações não hierárquicas (e.g. causa-efeito - *chuva* é causa da *inundação*, o seu efeito), e outras relações tais como, antonímia, sinonímia, predicado e argumentos, categorias gramaticais diferentes com o mesmo significado, coocorrentes e colocações, elencadas em L'Homme; Marshman (2006, p.68-69).

A identificação dessas relações pode ser conduzida por meio de métodos de investigação em corpora que identificam “contextos ricos em conhecimento” (ou KRC), um conceito introduzido em 1994 por Ingrid Meyer<sup>1</sup>. Meyer (2001, p.281) afirma que “os KRCs são considerados um contexto que indica ao menos um item do conhecimento temático passível de ser útil na análise conceitual”<sup>2</sup>.

### 3. Métodos de investigação de relações terminológicas

Os métodos de investigação apresentados a seguir auxiliam a identificação de relações terminológicas (ou relações léxico-semânticas em L'HOMME 2004), pois objetivam identificar contextos ricos em conhecimento, os quais podem potencialmente apresentar um tipo de relação terminológica.

---

<sup>1</sup> Professora e pesquisadora (1957-2004) dos estudos de lexicografia, terminologia e tradução na School of Translation and Interpretation da University of Ottawa, Canadá.

<sup>2</sup> Texto original: “By knowledge-rich context, we designate a context indicating at least one item of domain knowledge that could be useful for conceptual analysis”. (MEYER 2001: 281 apud L'HOMME; MARSHMAN 2006: 70).

São apresentados a seguir três métodos com base em L’Homme; Marshman, 2006.

O primeiro método trata da identificação de *padrões de conhecimento* (*knowledge patterns*, em inglês), definidos como “sequências linguísticas que indicam características conceituais”<sup>3</sup>. Esses padrões “devem ser “previsíveis” e “recorrentes em textos” e “manifestar as relações terminológicas de interesse pelos terminólogos”<sup>4</sup> (MEYER et al. 1999: 257 apud L’HOMME; MARSHMAN 2006: 70). Apresentam-se sob três tipos:

- 1) *Padrões lexicais* incluem itens lexicais específicos, que podem expressar todos os tipos de relações. Por exemplo: padrões de HIPERONÍMIA incluem *is the, is a, such as, and other e known as*);
- 2) *Padrões gramaticais* referem-se a um número menor de relações. O padrão SUBSTANTIVO + VERBO (com a exclusão de alguns verbos), por exemplo, é bem produtivo na indicação da relação FUNÇÃO.
- 3) *Padrões paralinguísticos* incluem pontuação e vários elementos da estrutura geral de um texto. A oração “*placenta previa* (uma placenta localizada de forma anormal na parte inferior do útero)”, por exemplo, ilustra como os parênteses podem indicar HIPERONÍMIA. (MEYER et al. 1999: 257 apud L’HOMME; MARSHMAN 2006: 70, tradução nossa)<sup>5</sup>

A partir de um tipo de padrão, é possível identificar no *corpus* demais ocorrências das referidas sequências linguísticas por meio de ferramentas de busca (e.g. concordanciador de busca por KWIC - *key word in context* -

<sup>3</sup> Texto original: “KRCs often contain knowledge patterns, linguistic sequences that indicate conceptual characteristics” (L’HOMME; MARSHMAN 2006: 70).

<sup>4</sup> Texto original: “These patterns should be ‘predictable’ and ‘recurring in text’ and manifest the relations in which the terminologist is interested”. (MEYER et al. 1999: 257 apud L’HOMME, MARSHMAN 2006: 70).

<sup>5</sup> Texto original: “1) *Lexical patterns* involve specific lexical items and can convey all types of relations. For example... HYPERONYMY patterns include *is the, is a, such as, and other and known as*. 2) *Grammatical patterns* apply to a small number of relations. The pattern NOUN + VERB (with some verbs excluded), for example, is highly productive for indicating the FUNCTION relation. 3) *Paralinguistic patterns* include punctuation, as well as various elements of the general structure of a text. The phrase ‘*placenta previa* (a placenta abnormally located in the lower part of the uterus)’, for example, illustrates how parentheses may indicate HYPERONYMY.” (MEYER et al. 1999, 257-58)

palavras-chaves em contexto<sup>6</sup>). O concordanciador identifica de forma automática todos os contextos de ocorrência de um tipo de padrão, os quais podem potencialmente apresentar um tipo de relação terminológica e, conseqüentemente, auxiliar a identificar termos reais usados nesse tipo de relação.

Outras relações terminológicas expressas por meio de *padrões de conhecimento* são apresentadas em L’Homme; Marshman (2006: 72):

meronímia	“is part of” e “contains”
função (padrão grammatical)	“needed for”, “serve as” e “designed for”
causa-efeito	“X causes Y”, “X results in Y” e “X results from Y”

Tabela 1 - Padrões de conhecimento de outras relações terminológicas

Existem também outros tipos de *padrões de conhecimento* que não expressam diretamente uma relação, mas apresentam indicações de possíveis relações terminológicas. Eles são:

- i) Os atributos (em inglês *attributes*, MEYER 2001), os quais, por apresentarem características dos conceitos, são relevantes para a análise do conceito em si. Servem de exemplo os atributos presentes em padrões lexicais, tais como *X is characterized by Y* (X caracteriza-se por Y) e *the feature of an X include Y and Z* (a característica de X inclui Y e Z) (MEYER 1994: 8 apud L’HOMME; MARSHMAN 2006: 72);
- ii) Padrões metalinguísticos, tais como em *X is defined as Y* (X é definido como Y) *X is called Y* (X é chamado(a) de Y), em que se tem a indicação de uma unidade lexical como termo.

O segundo método de identificação de KRCs baseia-se nas propriedades formais compartilhadas entre os termos. Esse método pressupõe que “os

<sup>6</sup> Esta pesquisa utilizou o programa AntConc, desenvolvido por Lawrence Anthony. Disponível em: <[www.laurenceanthony.net/software.html](http://www.laurenceanthony.net/software.html)>. Acesso em: 03 set. 2016.

termos de uma área do conhecimento que compartilham propriedades formais também compartilham componentes semânticos”<sup>7</sup>. (L’HOMME; MARSHMAN 2006: 73). Essas propriedades são expressas por meio de<sup>8</sup>:

- i) Relação sintática entre termos, na qual dois termos podem estar semanticamente relacionados se compartilham o mesmo núcleo (e.g. ácido nucleico, ácido desoxirribonucléico (DNA) e ácido ribonucléico (RNA)) ou o mesmo modificador (e.g. biblioteca de expressão, vetor de expressão, cassete de expressão<sup>9</sup>);
- ii) Relação morfológica entre termos, na qual dois termos podem apresentar relação terminológica a partir de suas características morfológicas. Servem de exemplo os termos em i): a ocorrência do mesmo elemento “*ribonucleico*” em ambos os determinantes (*desoxirribonucleico*; *ribonucleico*).

O terceiro método refere-se às unidades lexicais compartilhadas no contexto de ocorrência dos termos. A proposta considera que “se duas unidades lexicais (e, mais especificamente, dois termos) compartilham um número significativo de coocorrentes, elas devem estar semanticamente relacionadas”<sup>10</sup>. (L’HOMME; MARSHMAN 2006: 74). O exemplo apresentado pelas autoras é o seguinte: “... *program, application, editor e word processor* podem combinar-se com *launch, install, quit, and powerful*”; por essa razão,

<sup>7</sup> Texto original: “Some methods assume that terms in a specific domain that share formal properties also share semantic components” (L’HOMME, MARSHMAN 2006: 73).

<sup>8</sup> Um tipo específico de relação terminológica não está formalmente definido, mas esses termos relacionados têm grande probabilidade de conter diversos tipos de relações léxico-semânticas (e.g. sinonímia, hiperonímia, e co-hiponímia) (L’HOMME, MARSHMAN 2006: 74).

<sup>9</sup> Termos retirados de: TORRES, ANTÔNIO CARLOS et al. *Glossário de biotecnologia vegetal*. Brasília: Embrapa Hortaliças. 2000. Disponível em <<https://www.embrapa.br/hortaliças/busca-de-publicacoes/-/publicacao/769141/glossario-de-biotecnologia-vegetal>>. Acesso em: 03 set. 2016.

<sup>10</sup> Texto original: “... if two lexical units (and, more specifically, two terms) share a significant number of co-occurents, then they must be semantically related” (L’HOMME; MARSHMAN 2006: 74)

eles têm a probabilidade de apresentar um tipo de relação terminológica e são potenciais candidatos a termo<sup>11</sup>.

## 4. Contextos ricos em conhecimento no português do Brasil

Os textos escolhidos para a aplicação dos métodos de investigação em corpus são apresentados a seguir, ambos referentes a capítulos sobre o sistema monetário:

- 1) McConnell, Brue (2001), *Macroeconomia*, capítulos 13, 14 e 15, *Moeda e Bancos, Como os bancos criam moeda e Política monetária*, respectivamente<sup>12</sup>, e;
- 2) Lopes e Vasconcelos (2011), *Manual de Macroeconomia*, capítulo 2, *Sistema monetário: oferta e demanda de moeda*.

Inicialmente a pesquisa foi realizada de modo manual, quer dizer, por meio da leitura do texto. Ressalte-se que nos baseamos em aspectos gráficos e da organização da macroestrutura do texto, tais como os títulos, os subtítulos, a diferença de tamanho das letras (indicando que um título inclui subtítulos), as palavras em negrito e em itálico, as enumerações, pontuação (dois-pontos, ponto de interrogação) como tentativa inicial de identificar um conjunto de termos relacionados à temática principal<sup>13</sup>.

---

<sup>11</sup> Na literatura corrente, os termos coocorrentes podem ser analisados como unidades fraseológicas especializadas (BEVILACQUA 1998).

<sup>12</sup> Em agradecimento pela colaboração, o referente corpus foi organizado por Juliana Rolim Nobre Maia, aluna que fez parte do programa de iniciação científica (PROIC) 2015/2916 da Universidade de Brasília (UnB).

<sup>13</sup> Como fonte de verificação/validação das unidades candidatas a termos, foram consultados o glossário em McConnell, Brue (2001), livro didático usado como *corpus*, e o *Novíssimo Dicionário de Economia*, de Paulo Sandroni.

Com base nos dados obtidos pela pesquisa manual, a pesquisa foi ampliada por meio do uso do software AntConc, que possibilitou identificar todas ocorrências de possíveis contextos ricos em conhecimento ao longo do corpus selecionado. Foram identificados padrões de conhecimento, assim como propriedades formais entre termos e coocorrência de unidades lexicais, a seguir apresentados:

## 4.1 Padrões de conhecimento

Foram identificados os três tipos de padrões de conhecimentos expressos por meio de sequências linguísticas apresentados na coluna à esquerda.

### a) Padrões lexicais

Foram identificadas, a seguir, sequências linguísticas que expressam a relação de hiperonímia e de hiponímia:

Padrão lexical de hiperonímia/hiponímia	
é um	a) [...] desejam manter moeda porque a moeda <i>é um</i> meio de troca e é conveniente [...] b) [...] imediatamente utilizável em compras. A moeda <i>é um</i> ativo interessante de ser retido quando [...] c) [...] humana. 2. Unidade de conta. A moeda também <i>é uma</i> unidade de conta. A sociedade utiliza [...]. Fonte: MCCONNEL; BRUE (2001)
Como	[...] <i>Como</i> meio de troca, a moeda permite que [...] [...] No Capítulo 7, utilizamos a moeda <i>como</i> unidade de conta no cálculo do tamanho [...] Fonte: MCCONNEL; BRUE (2001)
tais como	[...] Os bancos e as instituições de poupança desenvolveram novos “produtos” de empréstimo, <i>tais como</i> empréstimos garantidos por hipotecas de residências e hipotecas com pagamentos baixos ou tendendo a zero. [...] Fonte: MCCONNEL; BRUE (2001)
tipo de	[...] Paralelamente a essa evolução dos tipos de moeda, assistimos ao desenvolvimento do Sistema Financeiro, em especial dos bancos comerciais, que trouxe consigo um novo <i>tipo de</i> moeda: a moeda

LAMBERTI, F. — Extração de conhecimento: métodos de identificação de relações terminológicas aplicadas ao português do Brasil

	escritural. [...] Fonte: MCCONNEL; BRUE (2001)
Ou	[...] é mais fácil do que a sua atrelação ao ouro <i>ou</i> a qualquer outra <i>commodity</i> cuja oferta [...]. Fonte: MCCONNEL; BRUE (2001)

Tabela 2 - Padrão lexical de hiperonímia/hiponímia

Na sequência linguística “é um”, observa-se a estrutura *X é um(a) Y*, em que *X* é *moeda*, específico (hipônimo), e *Y* é genérico (hiperônimo). Nesse caso, a *moeda* é um tipo de *meio de troca em* (a), de *ativo em* (b) e de *unidade de conta em* (c), respectivamente, sendo cada ocorrência um possível termo.

Em contextos com a sequência “como”, as ocorrências reforçam a atribuição de *moeda* como um hipônimo (um tipo de) de *meio de troca* e de *unidade de conta*.

A sequência “tais como” introduz tipos de “produtos” *de empréstimo: empréstimos garantidos por hipotecas de residências; (empréstimos garantidos por) hipotecas com pagamentos baixos*.

Já a sequência “tipo de” também indica a relação entre um termo genérico (no exemplo *moeda*) e outro específico (no caso *moeda escritural*).

A conjunção “ou” indica aqui relação entre um termo específico (no caso *ouro*) e um termo genérico (*commodity*).

A seguir, observa-se o padrão lexical de meronímia:

Padrão lexical de meronímia	
constituir	[...] Já M3 é <i>constituído de</i> M2 mais os grandes depósitos a prazo (mais de US\$100.000). Fonte: MCCONNEL; BRUE (2001)
dividir-se	[...] O Sistema Financeiro <i>divide-se</i> em dois grandes blocos: sistema bancário ou monetário, [...]. Fonte: LOPES E VASCONCELOS (2011)
estar incluso	[...] No primeiro segmento, <i>estão inclusos</i> o Banco Central (Bacen), que possui o monopólio da emissão, e as instituições autorizadas pelo Bacen a receberem depósitos a vista do público [...]. Fonte: LOPES E VASCONCELOS (2011)
compor	[...] O Conselho Consultivo Federal é constituído de 12 banqueiros

LAMBERTI, F. — Extração de conhecimento: métodos de identificação de relações terminológicas aplicadas ao português do Brasil

	comerciais, e cada 1 deles é selecionado anualmente em cada um dos 12 bancos que <i>compõem</i> o Federal Reserve. Fonte: MCCONNEL; BRUE (2001)
--	---

Tabela 3 - Padrão lexical de meronímia

Os verbos indicam uma relação partitiva (parte-todo): M2 é um merônimo (parte) de M3; *sistema bancário* ou *sistema monetário* são merônimos de *sistema financeiro*; *12 bancos* são merônimos de *Federal Reserve*.

b) Padrão gramatical

Neste padrão, ocorre a indicação da relação de função.

Padrão gramatical de Função	
ser utilizada	[...] Em primeiro lugar, e acima de tudo, a moeda é um meio de troca, pois <i>é utilizada</i> para a compra e a venda de bens e serviços. Fonte: MCCONNEL; BRUE (2001)

Tabela 4 - Padrão gramatical de Função

A relação terminológica ocorre entre a entidade e a sua função ou uso. No exemplo na Tabela 4, *compra e venda de bens e serviços* é a função da *moeda*.

c) Padrão paralinguístico

Este padrão é indicado pela pontuação ou por vários elementos da estrutura geral do texto. A seguir, tem-se o uso de parênteses:

Padrão paralinguístico de hiperonímia/hiponímia	
uso de parênteses	[...] A maior parte dos meios de reter dinheiro não gera retornos monetários tais como a reserva de valor em forma de <i>ativos reais (propriedade)</i> ou <i>ativos financeiros (ações, títulos etc.)</i> [...] Fonte: MCCONNEL; BRUE (2001)

Tabela 5 - Padrão paralinguístico de hiperonímia/hiponímia

Os parênteses parecem, neste caso, indicar hiperonímia (e.g. *propriedade* seria um tipo de *ativo real*; *ações*, *títulos* seriam tipos de *ativo financeiro*).

d) Padrão metalinguístico

Os padrões metalinguísticos abaixo apresentam a indicação de unidades lexicais consideradas como termo.

Padrão metalinguístico	
X é (a) chamado(a) de Y	[...] Esse gráfico é <i>chamado de</i> “gráfico semilog”, porque distâncias verticais iguais medem variações percentuais iguais em vez de variações absolutas [...] Fonte: LOPES E VASCONCELOS (2011)
X é definido(a) como	[...] Segundo a definição mais restrita, a oferta monetária é <i>definida como</i> M1 e compõe-se de dois itens [...] Fonte: MCCONNEL; BRUE (2001)

Tabela 6 - Padrão metalinguístico

As sequências acima indicam informação de que unidades lexicais (X) são consideradas termo, tais como os exemplos, *gráfico semilog* e M1, na Tabela 6 acima.

## 4.2 Propriedades formais

Características semelhantes no âmbito da forma das estruturas linguísticas podem indicar uma relação semântica entre termos. O fato, por exemplo, de termos complexos compartilharem o mesmo núcleo, ou o mesmo modificador, pode ser um indicativo de que compartilham informação semântica.

No exemplo da tabela 7, a seguir, os termos complexos (*conta de poupança* e *conta à vista*) compartilham o mesmo núcleo (*conta*). O núcleo, por sua vez, apresenta determinantes diferentes que especificam um tipo de conta. Esses termos podem ser hipônimos do termo mais genérico (no caso, *conta*).

construção com mesmo núcleo	[...] Você pode sacar moeda imediatamente de uma <u>conta de poupança</u> em um banco comercial ou em uma instituição de poupança. Ou pode transferir seus fundos de uma <u>conta de poupança</u> para uma <u>conta à vista</u> . Fonte: MCCONNEL; BRUE (2001)
-----------------------------	--

Tabela 7 - Construção com mesmo núcleo

No caso a seguir, na Tabela 8, observa-se que a recorrência de estruturas com morfemas formalmente semelhantes pode ser um indicativo de que são unidades lexicais semanticamente relacionadas.

Relação morfológica	<p>i) Moeda <u>lastreada</u>. Fonte: MCCONNEL; BRUE (2001).</p> <p>ii) Embora o ouro não seja mais utilizado para “<u>lastrear</u>” a oferta de moeda nos Estados Unidos, atualmente os empréstimos bancários (geração de moeda) são restritos pelo montante de reservas bancárias [...] Fonte: MCCONNEL; BRUE (2001).</p>
---------------------	--

Tabela 8 - Relação morfológica

Em i) tem-se a base verbal *lastrear*, presente no adjetivo *lastreada*, e em ii), o verbo *lastrear*. O verbo é recorrente em ambas as estruturas, o que pode indicar que compartilham o mesmo significado. Uma pesquisa no Dicionário Houaiss nos permite identificar um significado para o verbo *lastrear* (verbo t.d. m.q. *lastrar* [...] 3 t.d. *servir de lastro*, de base a<sup>14</sup>) presente em

<sup>14</sup> **Lastrar** [...] 3 t.d. *servir de lastro*, de base a <anos de experiência lastravam a sua mestria como operador> <direitos creditórios, títulos ou cessão fiduciária são o que lastra os recebíveis imobiliários> (HOUAISS 2001).

i) e ii). Os termos, por sua vez, apresentam o mesmo significado, porém são expressos linguisticamente sob a forma de termos que pertencem a categorias gramaticais diferentes, adjetivo e verbo, respectivamente.

### 4.3 Coocorrência

Neste caso, unidades lexicais identificadas no contexto de ocorrência de um termo específico podem indicar que essas unidades estão semanticamente relacionadas. Nos exemplos abaixo, identificamos que uma série de unidades lexicais compartilha o mesmo verbo, como a seguir:

[...] 2. Todos os <u>depósitos em conta corrente</u> , representando os depósitos nos bancos comerciais e nas instituições de poupança em que os mesmos podem ser <u>sacados</u> [...] Fonte: MCCONNEL; BRUE (2001)
---

[...] 1. Você pode <u>sacar moeda</u> imediatamente de uma conta de poupança em um banco comercial [...] Fonte: MCCONNEL; BRUE (2001)
---

[...] Você pode <u>sacar fundos</u> prontamente de uma conta de depósito do mercado monetário (CDMM). [...] Fonte: MCCONNEL; BRUE (2001)
--

[...] Os <u>depósitos a prazo</u> tornam-se disponíveis para o depositante apenas no vencimento. Por exemplo, um depósito de 6 meses só pode <u>ser sacado</u> sem ônus pelo menos 6 meses após a data do depósito. [...] Fonte: MCCONNEL; BRUE (2001)
--

Tabela 9 - Coocorrência

Os exemplos acima apresentam a relação entre o verbo *sacar* e seus actantes (argumentos), todos passíveis de serem considerados um tipo de instrumento financeiro.

## 5. Considerações finais

---

A aplicação dos métodos de investigação de contextos ricos em conhecimento no português do Brasil permitiu a identificação de relações terminológicas. Tais contextos são marcados por *padrões de conhecimento*, *propriedades formais* e *coocorrência*. Consideramos que o reconhecimento desse tipo de contexto pelo estudante de tradução é uma estratégia que o orienta na identificação de termos e na aquisição de conhecimento de uma determinada área ou temática em razão da possibilidade de identificar, por meios linguísticos, uma estrutura de conhecimento de uma área específica.

Uma perspectiva de ampliação deste estudo é o desenvolvimento de pesquisa terminológica bilíngue com o uso de *corpus* paralelo (com o texto original e sua respectiva tradução). Por meio de um alinhador (um *software* que realiza o alinhamento), os períodos são colocados em paralelo para facilitar a busca dos termos e equivalentes. Nessa análise, é possível a aplicação dos métodos de identificação de contextos ricos em conhecimento para auxiliar a identificação das relações terminológicas e dos termos em ambas as línguas da pesquisa.

É importante ressaltar que a descrição, o registro e o gerenciamento desse tipo de estudo em bases de dados multilíngues (e.g. Termium e os projetos terminológicos do OLST<sup>15</sup>) são um importante recurso de consulta a tradutores e demais pesquisadores da linguagem.

## Referências bibliográficas

---

<sup>15</sup> TERMIUM (Government of Canada's terminology and linguistic data bank). Disponível em: <<http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng>>. Acesso em: 8 de maio de 2017. OLST (Observatoire de Linguistique Sens Texte). Disponível em: <<http://olst.ling.umontreal.ca/>>. Acesso em: 8 de maio de 2017.

LAMBERTI, F. — Extração de conhecimento: métodos de identificação de relações terminológicas aplicadas ao português do Brasil

BEVILACQUA, C. R. “Unidades fraseológicas especializadas: novas perspectivas para sua identificação e tratamento”. In: *Organon*, v.12, n. 26. Porto Alegre, 1998. pp.119-132.

CABRÉ, M. T. Traducción y terminología: un espacio de encuentro ineludible. In: *La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Institut Universitari de Lingüística Aplicada, Univesitat Pompeu Fabra, 1999. pp. 177-202.

Corpus textual/base de dados (organização: Juliana Rolim Nobre Maia).

HOUAISS, A. et al. *Dicionário Houaiss da língua portuguesa*. Instituto Antonio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa. Rio de Janeiro: Objetiva, 2001.

L’HOMME, Marie Claude. *La terminologie: principes et techniques*. Montréal: Les Presses de l’Université de Montréal, 2004.

L’HOMME, Marie-Claude; MARSHMAN, Elisabeth. Terminological relationships and corpus-based methods for discovering them: an assessment for terminographers. In: BOWKER, Lynne. (Orgs.) *Lexicography, Terminology and Translation: text-based studies in honour of Ingrid Meyer*. Ottawa, Ontario: University of Ottawa Press, 2006. pp. 67-80.

MCCONNELL, C.R.; BRUE, S. L. *Macroeconomia: princípios, problemas e políticas*. Tradução da 14ª edição. Tradução de Flávia Dias Rangel. Revisão Técnica de José Luís da Costa Oreiro. Rio de Janeiro: LTC Editora, 2001.

LOPES, Luiz Martins; VASCONCELOS, Marco Antonio S. (Orgs). *Manual de Macroeconomia: nível básico e nível intermediário*. São Paulo: Editora Atlas S.A., 2011.

MEYER, I. Extracting Knowledge-Rich Contexts for Terminography: A Conceptual and Methodological Framework. In: BOURIGAUULT, D;

LAMBERTI, F. — Extração de conhecimento: métodos de identificação de relações terminológicas aplicadas ao português do Brasil

JACQUEMIN, C.; L'HOMME, M.C. *Recent Advances in Computational Terminology*. Amsterdam: John Benjamins, 2001. pp. 279-302.

MEYER, Ingrid; MACKINTOSH, Kristen; BARRIÈRE, Caroline; MORGAN, Tricia. Conceptual Sampling for Terminographical Corpus Analysis. In: SANDRINI, Peter (Ed.). *Proceedings of Terminology and Knowledge Engineering (TKE'99)*. Innsbruck: Ergon-Verlag, 1999. pp. 256-67.

Programa ANTConc (autoria de Lawrence Anthony). Disponível em: <[www.laurenceanthony.net/software/antconc](http://www.laurenceanthony.net/software/antconc)>. Acesso em: 1 de maio de 2017.

SAGER, Juan. *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1990.

SANDRONI, PAULO. *Novíssimo Dicionário de Economia*. São Paulo: Editora Best Seller, 2000.

*Data de submissão: 08/11/2016*

*Data de aprovação: 24/04/2017*