

# APORTE DA LINGUÍSTICA NO MAPEAMENTO DOS SENTIDOS DOS VERBOS EM TESAUROS

## AN LINGUISTICS APPROACH FOR MAPPING VERBAL MEANING IN THESAURI

Flávia Rodrigues Elias Nunes   
Universidade Federal de Minas Gerais

Benildes Coura Moreira dos Santos Maculan   
Universidade Federal de Minas Gerais

Maurício Barcellos Almeida   
Universidade Federal de Minas Gerais

---

### RESUMO

O sistema conceitual nos tesouros é formado das relações semânticas. O refinamento delas exige a sua explicitação, quando são evocadas por expressões verbais, conforme indicado por Soergel *et al.* (2004), Lauser *et al.* (2006) e Maculan (2015). Nesse contexto, é preciso considerar as valências verbais, que se referem aos valores semânticos que eles podem emanar, sendo necessário impor restrições para o controle de significação dentro de um domínio. Vista a dificuldade de compreensão sobre a valência semântica, este estudo busca contribuir no entendimento desse problema, a partir da análise de dois aportes linguísticos: Teoria da Valência Verbal e a Estrutura *Qualia*, questão inicialmente levantada por Maculan (2015). A metodologia adota a Análise de Conteúdo e critérios para construção do corpus, tecnologia utilizada para compilação, análise para atribuição da valência verbal e dos papéis *Qualia* e verificação do real sentido em um tesouro específico. Os resultados confirmaram a contribuição dos argumentos na identificação das relações semânticas entre objetos.

**Palavras-Chave:** Teoria da Valência Verbal, Estrutura *Qualia*, Refinamento das Relações Semânticas, Sistemas de Organização do Conhecimento, Tesouros.

---

### ABSTRACT

The conceptual system on thesauri is built from semantic relations. The refinement require their explanation, when they are evoked by verbal expressions, according to studies by Soergel *et al.* (2004), Lauser *et al.* (2006) e Maculan (2015). In this context, it's necessary to consider verbal valences, which refer to semantic values that verbs can emanate, being necessary to impose restrictions in order to control the meaning with a given domain. In the face of comprehension difficulty about verbal semantic valence, this study seeks to contribute to problem understanding, beginning from two Linguistic inputs: the Verbal Valence Theory and the *Qualia* Structure, a question initially raised by Maculan (2015). The methodology uses Content Analysis, corpus development criteria, compilation technology, the analysis for the attribution of verbal valency and the *Qualia* roles and the verification of the real meaning in a given thesaurus. The results confirmed the contribution of the arguments in the identification of the semantic relations between objects.

**Keywords:** Verbal Valence Theory, *Qualia* Structure, Semantic Relations, Refinement, Knowledge Organization Systems, Thesaurus.

## 1. INTRODUÇÃO

A relação entre a semântica e os estudos no âmbito da organização do conhecimento é uma questão antiga, que ficou mais evidente após a Internet e os princípios que norteiam a Web Semântica, conduzidos por Berners-Lee, Hendler e Lassila (2001). Ainda assim, as contribuições da Linguística têm se mostrado cada vez mais relevantes nesse tipo de pesquisa (ALMEIDA, 2011). Afinal, para a máquina, a complexidade da comunicação humana pode tornar o uso de linguagem natural algo abstrato, impreciso e pouco confiável.

Dentre os aspectos linguísticos identificados, que podem acarretar erros no processo de transcrição da linguagem natural para a linguagem lógica, estão a tradução, a sinonímia, a homonímia, a polissemia e a ambiguidade, para citar alguns. Além disso, o entendimento do significado pode ser interpretado de maneiras diferentes no escopo de uma comunidade discursiva (público). Tudo isso faz com que o uso de restrições na estrutura semântica - termos isolados carregam sentidos alternativos - de qualquer sistema de organização do conhecimento (SOC) seja peça fundamental no processo de tratamento da informação.

O tesouro é um tipo de SOC composto por um conjunto de termos que comunicam o significado de conceitos relevantes de um domínio, proporcionando o controle do vocabulário e contribuindo, assim, para uma recuperação da informação mais conveniente. Esse artefato opera como um mediador, uma vez que tem por natureza traduzir a linguagem natural do usuário em uma linguagem normalizada, servindo de insumo para atribuir pontos de acesso para comunicar o conteúdo de documentos.

Gomes (1990, p. 16) diz que o tesouro possui conceitos, que são representados por termos, relacionados semântica e logicamente. Esse sistema conceitual nos tesouros é formado a partir de três tipos de relações semânticas básicas: de equivalência (controle de sinonímia e de variações linguísticas), hierárquica (níveis de generalidade e de especificidade) e associativas (ligações não-hierárquicas) (MACULAN *et al.*, 2015). Sobre as relações entre conceitos, Green (2001) afirma que é necessário especificar o tipo de relação entre conceitos, para então entender a sua natureza.

A explicitação das relações semânticas no sistema conceitual dos tesouros já vem sendo tratado na literatura como “refinamento de relacionamentos” (SOERGEL *et al.*, 2004; Lauser *et al.*, 2006; MACULAN, 2015), tesouro Agrovoc, da área da Agricultura, exemplifica bem esse modelo. Nessa perspectiva, considera-se o tesouro como uma ferramenta com maior nível de semântica que requer entendimento do significado das ligações entre os conceitos e da compreensão do sentido do próprio conceito (HJORLAND, 2008). Maculan e Lima (2014,

p. 189) corroboram com essa ideia ao destacarem que os verbos (ou expressões verbais) evocam o sentido da relação entre conceitos. E esse sentido, por sua vez, se refere à valência semântica dos verbos. É necessário compreender esse valor a partir de aportes da Linguística.

Os estudos no campo da Linguística, em especial da semântica, apontam que cada verbo carrega uma variedade de acepções possíveis, ou seja, de valências (BORBA, 1996), que são uma espécie de exigência, um complemento para o entendimento de um significado. A estrutura sintática sozinha não consegue prover informações suficientes para o entendimento semântico e, por isso, é fundamental a compreensão do conjunto de diáteses, ou seja, as construções verbais que combinam a representação sintática a um papel semântico.

Visto a dificuldade de compreensão sobre a valência dos verbos utilizados na explicitação das relações semânticas em tesouros, o presente estudo contribui no entendimento do problema a partir da análise de dois aportes da Linguística: a Teoria da Valência Verbal e a da Estrutura *Qualia*. As poucas iniciativas na literatura de estudos sobre a utilização dessas teorias na argumentação das relações entre os conceitos, de modo que eles possam, de fato, estar interligados e vinculados aos seus significados, reforçam a relevância de estudos que visam o refinamento das relações semânticas entre conceitos em um tesouro (SOERGEL *et al.*, 2004; e LAUSER *et al.*, 2006; MACULAN, 2015).

A fim de contribuir teoricamente para o entendimento sobre os verbos no tratamento das relações entre os conceitos na estrutura de tesouros, evitando possíveis ambiguidades e problemas de interpretação do teor significativo dessas relações semânticas, este estudo teve como finalidade atingir o seguinte objetivo: propor um modelo de metodologia, com filtros cognitivos a partir da Teoria da Valência Verbal e dos Papéis *Qualia*, para mapear os sentidos dos verbos dentro de um domínio específico, visando restringir o significado das relações semânticas em um tesouro.

Dahlberg (1978, p. 105), discursando sobre a construção de tesouros, apresentou justificativas, formando que a “valência semântica do verbo é a soma dos lugares a serem preenchidos de acordo com a ligação deste [ou de um] conceito com outros”, sobretudo no que diz respeito à representação de relações funcionais, “que expressam processos”, e a autora indica que deve-se “conhecer o caráter semântico de tais relações tendo por base as chamadas valências semânticas dos verbos, dando atenção aos verbos e respectivos complementos [argumentos]”. Ademais, destaca-se que as “restrições semânticas impostas por um verbo a seus complementos formam um desses aspectos linguísticos que apresentam considerável complexidade, quando tentamos descrevê-los de uma forma

computacionalmente operável”, sendo “impossível esquecer-las quando tentamos solucionar certos problemas típicos de mecanismos computacionais de reconhecimento” (AMARAL, 2009, p. 32). Ou seja, a formalização das relações entre conceitos exige a explicitação de seu significado e, assim, compreender a complexidade que envolve a valência dos verbos se torna essencial na construção de tesouros. Acredita-se mesmo que isso é essencial na construção de qualquer outro tipo de SOC.

Depois desta breve introdução, que contextualiza o assunto e estabelece o problema, o objetivo e as justificativas do estudo, o restante do presente artigo foi estruturado da seguinte maneira: na seção 2 é apresentada a Teoria da Valência Verbal, que destaca os papéis semânticos e os argumentos necessários para que o verbo tenha sentido completo; na seção 3 é discutida a identificação do Papel *Qualia* no refinamento das relações semânticas entre dois conceitos distintos, a partir da Estrutura *Qualia*, que é parte da Teoria do Léxico Gerativo, de Pustejovsky (1995; 1998); a seção 4 traz a metodologia adotada, com detalhamento dos procedimentos; na seção 5 é feita uma análise sobre os resultados obtidos; e, por fim, as considerações que apresentam as limitações e sugestões de novas pesquisas.

## **2. TEORIA DA VALENCIA VERBAL**

No Brasil, a questão valencial foi tratada pela primeira vez em 1990 por Borba, em um dicionário gramatical de verbos do português contemporâneo. Sobre a relação de dependência entre os elementos constituintes de uma frase, Borba (1996) dá o nome de coocorrência. Ou seja, para ele, a valência é, em relação ao verbo, uma rede de dependências em nível sintático e semântico. Já Perini (2003) destaca sobre a existência de uma rede de associações entre forma e significado e propõe que a análise não fique restrita a quantidade de entidades ou itens léxicos, mas a de uma abordagem de uma “rede complexa de ligações, em especial no que diz respeito à participação das formas em paradigmas morfológicos sistemáticos” (PERINI, 2003, p. 299).

Diante da perspectiva cognitiva na realização argumental, Cançado (2012) enfatiza a questão dos papéis semânticos que são definidos por uma função semântica para argumentos e adjuntos, estabelecendo-se gramaticalmente uma relação de significados mentais e conceituais para a interpretação das cenas verbais e das entidades nelas envolvidas. Ou seja, esses papéis formam a estrutura que permite a compreensão e a associação entre o aspecto sintático e as significações associadas. Perini (2015) complementa que os papéis semânticos expressam a relação semântica entre os principais constituintes de uma sentença e seus verbos. Assim, em uma sentença, pode-se ter um Agente, um Paciente, um Instrumento, um

Lugar, e assim por diante. O autor leva em conta tanto a forma quanto o papel semântico para a determinação da valência.

Para entender o comportamento gramatical dos verbos é necessário compreender a diferença entre construções, diáteses e esquemas/*frames*, que serão explicadas a partir do exemplo:

a) Pedro quebrou a vidraça.

SUJ	V	SN
Agente		Paciente

Na primeira linha tem-se a representação da estrutura sintática, em que o sujeito é representado por SUJ, o objeto (a vidraça) por SN, e o verbo (quebrou) por V. O verbo ativa o esquema cognitivo “quebrar”, incluindo outros argumentos, como o Agente (quebrador) e o Paciente (coisa quebrada). Desse modo, na segunda linha está o aspecto do significado, com o destaque para os participantes da ação. Ou seja, não é o verbo “quebrar” que está sendo analisado, mas o “esquema quebrar”, que corresponde à noção do *frame*, da gramática dos casos de Fillmore (1982). O autor diz que para compreender a estrutura semântica de um verbo, é necessário conhecer as situações caracterizadas pela junção dos casos de seus argumentos, formando uma cena esquematizada.

Com essa explanação, pode-se dizer então que as construções e as diáteses são estruturas que relacionam informação de natureza sintática (forma) e de natureza semântica (significado), em que há uma representação estabelecida pelos papéis semânticos. Para discorrer a respeito desses papéis é necessário os distinguir das Relações *Conceptuais* Elaboradas, as RCEs. Perini (2008) faz essa distinção ao falar que os papéis semânticos são semanticamente esquemáticos, enquanto as RCEs são as unidades semânticas que derivam do significado, ou são as relações cognitivas. Em linhas gerais, pode-se dizer que as RCEs são as relações cognitivas presentes nos esquemas ou *frames*, percebidas imediatamente e que fazem parte do entendimento pelo indivíduo da mensagem a partir de conhecimentos do mundo. Os papéis semânticos, por sua vez, são as relações gramaticais.

Perini (2008; 2015; em elaboração) explica que a diferenciação entre as RCEs e os papéis semânticos é que uma representa as relações cognitivas elaboradas (RCEs), utilizadas para efetuar a conexão entre a estrutura gramatical e a compreensão, gerando o que é chamado de paisagem mental; a outra se refere a um conjunto de relações esquemáticas. Já os papéis semânticos são os feixes de várias RCEs, pois relaciona a forma linguística (função sintática) e conceitos (esquemas).

Diante disso, percebe-se que os verbos possuem uma multiplicidade de significados que exprimem operações diversas e que a valência é mais complexa do que uma determinação exercida pelo verbo. Desse modo, considera-se a Estrutura *Qualia* como outro aporte teórico na contribuição do refinamento do relacionamento semântico.

### 3. ESTRUTURA QUALIA

Como recurso teórico complementar à esquematicidade semântica, a pesquisa abordou também a Estrutura *Qualia*, que é um dos níveis de representação da Teoria do Léxico Gerativo, de Pustejovsky (1995; 1998). O autor propõe um modelo de decomposição formal da linguagem com o objetivo de fornecer uma descrição mais adequada do conteúdo e de um modo que esse conteúdo possa se adaptar em diferentes contextos.

O autor fala de “léxico gerativo” como se tratasse de palavras, morfemas, entre outros, mas na verdade podemos chamá-lo de unidade cognitiva, ou seja, o conceito de elaboração introduzido por Langacker (1991), com a expressão de um nível de esquematicidade em que “uma das expressões componentes elabora uma estrutura esquemática dentro da outra” (LANGACKER, 1991, p. 7). Essas relações são importantes no processo de estabelecimento da relação entre a forma linguística e a representação cognitiva, que é o produto final do processo da compreensão.

O *Qualia* pode ser entendido como uma tentativa de codificação das formas de elaboração que são lexicalmente definíveis. O autor subdividiu a Estrutura *Qualia* em quatro aspectos: Papel Constitutivo, que expressa a relação entre a palavra e suas partes constituintes; Papel Formal, que distingue a palavra em um domínio maior; Papel Télico, que indica a função da palavra; e Papel Agentivo, que é o modo como a palavra se originou.

De acordo com Pustejovsky (1995), esses aspectos guiam a interpretação básica de um objeto ou de uma relação no mundo. O autor afirma que, a princípio, a impressão é de que a Estrutura *Qualia* pode parecer uma simples listagem de papéis causais associados a itens lexicais, no entanto ela representa o ponto de partida para operações de reconstruções semânticas, contribuindo para a caracterização de uma língua natural polimórfica, sugerindo interpretações das palavras em determinado contexto. Para melhor explicitação, pega-se como exemplo os papéis *Qualia* evocados pelo esquema “jornal”.

Jornal

Qualia=            CONST = folhas  
                      FORMAL = objeto físico escrito  
                      TELIC = ser lido  
                      AGENT = escrito por jornalistas

Em uma interpretação de sentença, a análise desses papéis poderia ser feita da seguinte maneira, considerando o exemplo da frase abaixo:

b) Joaquim trabalha no jornal, mas não sabe nem escrever uma frase correta.

Essa sentença só faz sentido pois jornal é um tipo de material escrito e feito para ser lido, o que autoriza o raciocínio de que Joaquim trabalha em um, e, por isso, precisa ter domínio da linguagem escrita, no entanto, ele não tem. Desse modo, a conjunção “mas” se justifica, introduzindo assim uma expectativa contrariada.

#### 4. METODOLOGIA

Os procedimentos metodológicos adotados foram de natureza exploratória, havendo a busca sobre informações acerca da Teoria da Valência Verbal e da Estrutura *Qualia*, tendo em vista trazer insumos para a compreensão de seus fundamentos. A pesquisa bibliográfica se valeu da revisão narrativa da literatura, quando não há definição de critérios e de sistematização da busca. Segundo Rother (2007, p. 1), nessas revisões “não [se] informam as fontes de informação utilizadas, a metodologia para busca das referências, nem os critérios utilizados na avaliação e seleção dos trabalhos”, visto que se norteiam na “interpretação e análise crítica pessoal do autor”, que, no caso desta pesquisa, tem formação em Letras/Linguística. Apoiada em Gil (1994), considera-se a presente uma pesquisa básica, pois se vislumbra a geração de novos conhecimentos sem a previsão de construção de uma aplicação específica. Quanto à abordagem, o estudo é qualitativo, visto que houve interpretação das informações exploradas, sem o uso de instrumentos estatísticos na análise das teorias.

Como insumo terminológico, trabalhou-se com os resultados de Maculan (2015), que realizou o refinamento semântico do tesouro THESAGRO brasileiro, no recorte da Intensificação Agropecuária. Aplicaram-se um total de 44 tipos diferentes de relações semânticas explícitas entre conceitos, para depois fazer análises semânticas em três relações

evocadas pelos esquemas: <afetar>, <influenciar> e <causar>. A seleção dos três esquemas teve por base o fato de se tratar de “relação[ões] associativa[s], que são comuns aos três verbetes, a presença do fator agentivo e os sinônimos comuns” (NUNES, 2019, p. 67). Neste artigo, apresentam-se os resultados da análise do esquema <afetar>, sendo que o restante das análises pode ser encontrado em Nunes (2019). Ademais, para a análise de identificação de sentidos no recorte da Intensificação Agropecuária, trabalhou-se com o mesmo *corpus* do domínio da Agricultura da pesquisa de Maculan (2015).

Para as análises, utilizou-se o método da Análise de Conteúdo e a abordagem da Linguística de *Corpus*, conforme descrito a seguir:

1- Método da Análise de Conteúdo: abordagem que tem como objetivo a busca do sentido ou dos sentidos dos conteúdos, numa análise de dados com a aplicação de inferências (BARDIN, 2006). Foram aplicadas as etapas elencadas pela autora, em três fases: 1) pré-análise (sistematização das ideias e organização do material coletado), 2) exploração do material e 3) tratamento dos resultados, inferência e interpretação (condensação e significado aos dados, com sua análise reflexiva e crítica).

2- Abordagem da Linguística de *Corpus*: técnica empregada para a coleta e exploração de corpora, que são conjuntos linguísticos textuais construídos a partir de objetivos específicos. Sardinha (2004) destaca que a definição do conteúdo e o propósito dos corpora podem seguir diferentes critérios, como modo (falado ou escrito), tempo (sincrônico, diacrônico, contemporâneo, histórico), seleção (amostragem, monitor, dinâmico, estático, equilibrado), conteúdo (especializado, regional, multilíngue), entre outros.

## 5. ANÁLISES E RESULTADOS

Nesta seção estão apresentados os resultados da aplicação da metodologia.

Quadro 1. Aplicação da metodologia

PROCEDIMENTOS METODOLÓGICOS			
FASE	PASSOS	FERRAMENTAS	RESULTADOS
(1) Pré-análise: organização dos dados	Pesquisa teórica na literatura.	Referências teóricas sobre Sociedade da Informação, Semântica verbal, Relacionamentos em SOCs, Valência Verbal e Estrutura Qualia.	Descrição dos dados coletados no Referencial Teórico.

<b>(2) Exploração do material</b>	Identificação de unidades de contexto (UC) e unidades de registro (UR).	Referencial Teórico.	UC: (1) Teoria da Valência Verbal; (2) Estrutura <i>Qualia</i> ; (3) relações semânticas. UR: A) verbo; (B) verbete; (C) construções/diátese; (D) papel semântico; (E) Relações Conceptuais Elaboradas; (F) esquemas/frames; (G) relação equivalente; (H) relação hierárquica; (I) relação associativa; (J) hiponímia/hiperonímia; (K) papel Télico; (L) papel Constitutivo; (M) papel Formal; (N) papel Agentivo.
<b>(3) Tratamento dos resultados - Abordagem linguística de corpus</b>	1º - Recorte Temático	Intensificação Agropecuária - Mesmo recorte temático definido na pesquisa de Maculan (2015).	Proposta de metodologia de restrição de sentidos de relações verbais dentro de um tesouro.
	2º - Definição de critérios para seleção dos corpora.		
	3º - Coleta dos textos	<i>Software Publish or Perish</i> ; Plataforma de busca <i>Google Scholar</i> , <i>Boserup (1965)</i> e repositório pessoal do professor Ivo Pierozzi Júnior.	
	4º - Manipulação do corpus	Software <i>Sketch Engine</i>	
	5º - Compilação do corpus	Software <i>Sketch Engine</i>	
	6º - Identificação das valências dos esquemas analisados	Dicionário de Valências de Borba (1996), Projeto ADESSE e Projeto VVP (em construção).	
	7º - Verificação dos sentidos dentro do corpus	Processo manual/intelectual.	

Fonte: Nunes, (2019, p.68-69).

Especificamente, a terceira fase possibilitou verificar a contribuição da Teoria da Valência Verbal e da Estrutura *Qualia* no refinamento dos relacionamentos presentes nos tesouros. Com as análises, foi proposta uma metodologia de restrição no domínio da Intensificação Agropecuária, com objetivo de testar a aplicabilidade das teorias como ferramentas de refinamento semânticos das relações em tesouros.

Como exemplo da identificação das valências de um verbo, apresentam-se os significados para o verbo <afetar>. Primeiramente, no projeto ADESSE, no espanhol, <afetar> é traduzido por <afectar> e assim ele é classificado quanto aos seus sentidos:

AFECTAR	
De este verbo hay más de una entrada en la base de datos	
<b>AFECTAR I</b> (Modificación) ?	Producir un efecto sobre algo. Concernir, atañer, incumbir;
<b>AFECTAR II</b> (Sensación) ?	Impresionar, emocionar, originar un efecto generalmente desagradable o negativo
<b>AFECTAR III</b> (Actividad)	Simular, fingir

Figura 1 - Sentidos evocados pelo esquema <afectar>, segundo o ADESSE.

Fonte: Nunes (2019, p. 78), sobre o Projeto ADESSE (disponível em:

<http://adesse.uvigo.es/data/verbos.php?verbo=afectar>)

O esquema <afetar> no sentido de “produzir efeito sobre algo” possui na base de dados do ADESSE 82 exemplos, sendo que em 100% dos casos são presentes os argumentos de “coisa afetada” e “afetador”. Já o papel de beneficiário aparece em dois dos exemplos. Por sua vez, a construção passiva de <afeta>, <afetado por>, está em 14 exemplos do sentido em questão.

No sentido de “lesar”, o esquema <afetar> está presente no ADESSE em 12 exemplos. Em 100% dos casos há a presença dos argumentos chamados de “experimentador” e “estímulo”. Já <afetado por> está presente em dois exemplos, mostrando a classificação semântica e potencial valencial de Afectar II. O terceiro sentido é o de “simular ou fingir”, que foi levantado em dois exemplos do banco de dados. Nos dois casos os argumentos de atuação e atividade estão presentes.

Por sua vez, no dicionário e valências, Borba (1996) levanta os sentidos para o verbo <afetar>, como demonstrado no Quadro 2.

Quadro 2. <afetar>, segundo Borba (1996)

Afetar	Classificação: estado	Complemento: nome humano
Significado 1	ser do interesse de; dizer respeito a	Que é que tenho a ver com questões que <i>afetam</i> diretamente ou que só afetam aos Soares?
Afetar	Classificação: ação	Complemento: nome abstrato ou oração infinitiva

Significado 2	fingir; simular; aparentar	Nestor Benício <i>afetou</i> alheamento. Rogério baixou o olhar <i>afetando</i> constrangimento.
<b>Afetar</b>	<b>Classificação: ação-processo</b>	<b>Complemento: nome</b>
Significado 3	lesar; molestar; atingir	Mas o problema continua <i>afetando</i> a saúde dos moradores. A cianose é uma doença que <i>afeta</i> os cães jovens. O frio <i>afeta</i> muito a senhora.

Fonte: BORBA (1996, p.36).

Observando o Quadro 2, percebe-se que <afetar>, com classificação estado, indica situações nas quais os envolvidos no processo verbal não são transformados ou alterados em sua condição inicial (BORBA, 1996). Nesses casos, o significado não aponta um objeto Agente ou Paciente ou Causativo, e, sim, um Experimentador ou Beneficiário da ação verbal. Em seguida, <afetar> na classificação ação, indica um sentido de uma atividade expressa por um objeto Agente, que aponta um fazer por parte desse objeto (BORBA, 1996). Destaca-se que, quando os verbos de ação possuem dois ou três argumentos, estes serão complementos e não serão afetados pela ação do objeto.

Como terceira e última alternativa, tem-se <afetar> na classificação ação-processo, que apresenta o sentido de uma ação que é realizada por um objeto agente e/ou causação ocorrida por um objeto causativo, cuja ação afeta o complemento. Isso está alinhado à definição estipulada por Maculan (2015), de que a relação <afeta> no domínio da Intensificação da Agropecuária, pode ser representada de duas maneiras: X <afeta> Y e Y <é afetado por> X; em que X é um Agente que altera a condição de Y. Com isso, os verbos de ação-processo têm, pelo menos, dois argumentos: um Agente ou Causativo e, outro, Afetado ou Efetuado (BORBA, 1996).

Por fim, no Projeto Valências Verbais do Português (VVP, em elaboração), o verbo <afetar> no português e os sentidos evocados por ele, há o esquema conforme a Figura 2.

C1 VSubj>Agent V NP>Patient

A qualidade da água afeta a saúde da população.  
'the quality of water affects the health of the population'

A idade afetou a memória do velhinho.  
'age affected the old man's memory'

Seu comentário não afeta o argumento.  
'your comment does not bear on the argument'

.....  
C73 VSubj>Stimulus V NP>Experiencer

A cerimônia me afetou *profundamente*. 'the ceremony  
affected me deeply'

.....  
C268 Syntax VSubj V NP

Ev.1 Agent Counterfact: [Ev. 2]

Ev.2 Qfd.thing Quality

'o irmão afetava tristeza.  
'the brother feigned sad'

In this reading, *fingir* is more usual than *afetar*; the latter sounds somewhat bookish.

Figura 2. Valência verbal de <afetar>, segundo o Projeto VVP (PERINI, em construção).

É possível verificar que na valência de <afetar> há a presença de três ocorrências de diáteses, que são representadas por uma letra e um número (C1, C73 e C268). Ao contrário de Borba (1990), o projeto VVP não especifica o sentido de “ser de interesse de” e “dizer respeito a”. No entanto, há duas construções (C1 e C73) em que o objeto assume papéis diferentes, no primeiro ele é Paciente e no segundo um Experienciador. A construção C268, como Borba (1996), tem o sentido remetido à ideia de fingimento e simulação.

Para a verificação do sentido mais adequado ao domínio, foi preciso voltar ao *corpus* com a finalidade de fazer um filtro cognitivo, em que as representações malformadas do esquema <afetar> são excluídas do domínio. Para isso, foi utilizado no programa *Sketch Engine* a funcionalidade *Concordance*, que localiza as linhas de concordância a partir da palavra pesquisada, conforme Figura 3.

The screenshot displays the 'CONCORDANCE' software interface. At the top, the search term is 'afetar' with 110 results (145 per million) in the domain 'Intensificação Agropecuária'. The interface includes a 'SUBSCRIBE' button and various tool icons. Below the search bar, there are tabs for 'Details', 'Left context', 'KWIC', and 'Right context'. The main area shows a list of 14 concordance lines, each with a unique ID and a snippet of text where the verb 'afetar' is used. The KWIC column highlights the verb in red. A 'Back to the original interface' button is visible at the bottom right of the results list.

ID	Text Snippet	KWIC	Text Snippet
#6455	zação trouxe mudanças sobre o emprego agrícola de forma diferenciada,	afetando	algumas regiões mais do que outras. Em áreas mais organizadas e capit
#8420	emática social e até do seu nível cultural. Um dos recursos naturais mais	afetados	pela agricultura é sua base física, isto é, o solo. A falta de conhecimento c
#9277	ntagens para o consumidor? A primeira geração de plantas transgênicas	afeta	o custo de produção e, portanto, beneficia mais os produtores. Apesar dis
#9808	do da chamada "moderna agricultura", trazendo inúmeros problemas que	afetam	o meio ambiente, a qualidade de vida e o processo de produção, colocanc
#13236	e com uso considerável de insumos. São aproximadamente 100.000km2	afetados	por essa dinâmica relativamente recente de uso das terras, cuja repartiçã
#21859	ialização), pode dar uma aproximação sobre como este processo poderá	afetar	a permanência ou a exclusão destes produtores na atividade nos próximc
#28372	e cultivo. Alguns autores acreditam que o recuo das florestas acabou por	afetar	o clima, fazendo-o mais seco e facilitando a expansão de desertos. Outro
#31465	mportante, pois, indagar em que direção o produto por homem-hora será	afetado	por alterações do sistema de uso da terra. Não há resposta a priori a esse
#32960	o produto por homem-hora cairá. O declínio da produtividade do trabalho	afetará	todos os trabalhadores, não apenas os recentemente adicionados, como
#54299	ecuo na densidade populacional nas regiões da África mais severamente	afetadas	. Foi provavelmente por essa razão que as florestas e capoeiras puderam
#61899	válida de como um aumento na taxa de crescimento populacional poderá	afetar	o investimento e a poupança nas comunidades rurais. A análise deverá s
#67439	senhorial como tal, o investimento rural dificilmente poderia deixar de ser	afetado	. Tais regiões onde a irrigação foi negligenciada são, com frequência, exit
#78449	99. Embora seja claro que essa mudança de política econômica deve ter	afetado	muito favoravelmente o setor agrícola, não há, na literatura, uma análise r
#87682	lugar, procurou-se mostrar de que maneira esse crescimento agrícola foi	afetado	pela mudança da política cambial de jane

Figura 3. Concordância de <afetar>.

Fonte: Nunes (2019, p. 90), tela capturada do software *Sketch Engine* (2018).

A análise foi feita manual e intelectualmente, linha por linha, pois a ferramenta apresenta as linhas de concordância do verbete <afetar>, bem como de seus *tokens* (afetado, afetará, afetou, afetados etc.) dentro do domínio da Intensificação Agropecuária. Foram constatadas 110 concordâncias, mas, após verificação individual, 3 (três) foram excluídas, pois apresentavam inconsistências para análise. Desse modo, foram consideradas 107 concordâncias com a presença da relação <afeta>, analisando-se a indicação da ocorrência em cada significado do verbo, e determinando, assim, qual valência de <afetar> de fato representa a relação no domínio em questão. Para verificação foi considerado como referência de análise os sentidos apontados por Borba (1996), resultando no conteúdo do Quadro 3.

Quadro 3. Verificação dos sentidos presentes da relação <afeta>

Sentido	Exemplos	Total no corpus
Significado 1: ser do interesse de; dizer a respeito a	1- O declínio da produtividade do trabalho afetará todos os trabalhadores, não apenas os recentemente adicionados como seria o caso se a terra por último cultivada... Relação: produtividade do trabalho <afeta> trabalhadores 2 - Dentro do marco conceitual que está sendo proposto, é preciso destacar que políticas que afetam todo o setor devem ser manejadas para criar condições gerais favoráveis para o setor agropecuário. Relação: política <afeta> setor agropecuário	05
Significado 2: fingir; simular; aparentar	Não se aplica.	00
Significado 3: lesar, molestar, atingir	1. Quais as vantagens para o consumidor? A primeira geração de plantas transgênicas afeta custo de produção e, portanto, beneficia mais os produtores. Relação: plantas transgênicas <afeta> custo de produção 2. ...ocorrência de pragas e doenças que afetam o crescimento e a persistências das espécies forrageiras... Relação: pragas <afeta> espécies forrageiras doenças <afeta> espécies forrageiras	107

Fonte: Nunes (2019, p. 90-91).

Com a análise do quadro 3, foi possível concluir que o terceiro significado (lesar, molestar, atingir) é o mais utilizado para a representação da relação <afeta> no domínio estudado, pois é passível de entendimento em todas as 107 concordâncias. Verificou-se também que em cinco das ocorrências o primeiro sentido poderia ser considerado, já que não há uma clareza sobre a modificação do estado ou localização do objeto afetado. O segundo significado (fingir ou simular), não foi encontrado em qualquer ocorrência.

Para confirmar a hipótese de que o sentido 3 seria o mais adequado para a representação da relação, verificou-se em Maculan (2015) o levantamento das relações presentes no mesmo domínio e as definições do vocabulário *Agrontology* do tesouro AGROVOC. O SOC diferencia <afeta> de outras relações pelo fato de que “X causa uma mudança de estado ou de localização de Y”, como é apresentado no sentido 3 e está presente na construção C1 do Projeto VVP (PERINI, em elaboração). Dessa maneira, o terceiro sentido apresentado por Borba (1996) e corroborado na construção C1 do Projeto VVP, seria o aceitável para a representação da relação <afeta> dentro do domínio temático da Intensificação Agropecuária.

Devido à possibilidade de uma multiplicidade de sentidos para um mesmo verbo, este estudo evidencia a necessidade de que o sentido de uma relação em tesouros deve ser restringido, como forma de controlar a ambiguidade da linguagem natural. Nessa perspectiva, considera-se a necessidade de elaboração de definições para clarear e delimitar o sentido de um dado tipo de relação semântica nos tesouros, além do que já é feito para a restrição do significado dos conceitos. Dessa maneira, o estudo da valência pode contribuir no amparo à evocação de relações mais relevantes e específicas para um dado domínio.

## 5. CONSIDERAÇÕES FINAIS

Foi testado o uso da valência verbal e dos papéis *Qualia* como argumentos para o estabelecimento de relações semânticas entre conceitos no domínio da Intensificação Agropecuária. Baseou-se em estudos prévios de Maculan (2015), que utilizou o modelo de reengenharia de tesouros de Soergel *et al.* (2004) e Lauser *et al.* (2006) para fazer o refinamento das relações semânticas no tesouro THESAGRO Brasileiro. Considerou-se que a representação dos relacionamentos semânticos entre os conceitos presentes em tesouros possui os verbos ou as expressões verbais preposicionadas como elementos fundamentais para a evocação da ideia ou da ação da relação estabelecida. Dessa forma, ficou claro que a questão do papel semântico dos verbos na representação dos relacionamentos tem um caráter teórico que precisa ser elucidado (DAHLBERG, 1978; SOERGEL *et al.*, 2004; MACULAN; LIMA, 2014; MACULAN, 2015).

Verificou-se, a partir do levantamento teórico, que as relações em um tesouro são formadas pelos esquemas ou *frames* cognitivos evocados pelos verbos ou expressões verbais, sendo essencial aprofundar na compreensão do conjunto de construções verbais (diáteses) que combinam a representação sintática a um papel semântico. Esse resultado confirmou a necessidade da elaboração de definições também para as relações semânticas em tesouros.

Com os insumos da Teoria da Valência Verbal verificou-se que a identificação dos papéis semânticos ligados a cada verbo descreve uma parte importante do seu significado. Um verbo que tem um complemento “Agente” é necessariamente um verbo de ação, por exemplo. O entendimento sobre os princípios da Estrutura *Qualia* também contribui para o processo de identificação dos sentidos dos verbos. O que Pustejovsky (1995; 1998) chama de *Qualia* é, na verdade, o esquema desse verbo, uma vez que a significação está na relação estabelecida entre os termos (que representam conceitos). Ou seja, outro aspecto percebido é que os princípios da Estrutura *Qualia* são fundamentais para estabelecer a relação entre a

forma linguística e a representação cognitiva, que é o produto final do processo de compreensão.

Conclui-se, então, que a linguagem natural é um fenômeno extremamente complexo, onde se destaca a definição dos sentidos dos verbos quando do uso para compor uma relação em um tesouro. Os diferentes tipos de classificação das fontes e a interpretação pessoal da pessoa que está verificando o sentido podem influir diretamente no resultado, uma vez que os esquemas cognitivos tratam também da percepção de mundo do usuário. No entanto, como se o contexto envolve uma ferramenta em que o vocabulário é controlado e o sentido da relação deve ser único dentro do domínio, foi proposta uma metodologia de restrição dos significados dos verbos dentro da mesma temática. Essa metodologia contou com aportes da Linguística de Corpus, que trouxe insumos para fazer o recolhimento dos corpora, com a escolha do *software Sketch Engine*, a compilação do *corpus*, além da identificação dos verbetes e seus tokens em linhas de concordância.

Considera-se que a metodologia proposta é satisfatória para um estudo teórico inicial. As limitações dizem respeito a: a) não ter alcançado um procedimento computacionalmente implementável; b) não ter trabalhado com os resultados completos de Maculan (2015). A automatização ou semiautomatização desse processo ainda parece ser um projeto distante.

Como investigações futuras, sugere-se o desenvolvimento de modelos mais elaborados para sistemas computacionais e que possam ser replicados para outros contextos. Assim, espera-se adaptar os resultados para outros domínios, além de realizar análises semânticas para o conjunto de verbos que expressam as relações aplicadas em Maculan (2015).

Acredita-se que todas as valências dos verbos que expressam relações na estrutura conceitual de um tesouro deveriam ser identificadas e determinada a definição de cada relação dentro do domínio modelado, obtendo algo similar ao que Perini (2008) denomina de “paisagem mental”, a saber, o resultado do processo de interpretação cognitiva do significado evocado pelo verbo.

## REFERÊNCIAS

- AMARAL, L. A. M. Um modelo de restrições semântico-selecionais para sistemas de processamento automático de linguagem natural. **Veredas, Revista de Estudos Linguísticos**, Juiz de Fora, v. 5, n. 2, p. 31-45, 2009.
- BARDIN, L. **Análise de conteúdo**. Tradução de L. de A. Rego e A. Pinheiro. Lisboa: Edições 70, 2006. (Obra original publicada em 1977).
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, New York, v. 284, n. 5, p. 34-43, May 2001.

- BORBA, F. S. *et al.* **Dicionário gramatical de verbos do Português contemporâneo**. São Paulo: UNESP, 1990.
- BORBA, F. S. **Uma gramática de valências para o português**. São Paulo: Ática, 1996.
- BOSERUP, E. **The conditions of agricultural growth: the economics of agrarian change under population pressure**. Chicago: Aldine, 1965.
- CANÇADO, M. **Manual de Semântica: noções básicas e exercícios**. São Paulo: Contexto, 2012.
- DAHLBERG, I. Teoria do conceito. Tradução Astério Tavares Campos. **Ci. Inf.**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.
- FARINELLI, F.; MELO, S.; ALMEIDA, M. B. Interoperabilidade semântica em sistemas de informação de saúde por meio de ontologias formais e informais: um estudo da norma OPENEHR. In: XVI ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 20 1 25 de novembro de 2016, Universidade Federal da Bahia. **Anais...** Salvador: ENANCIB, 2016.
- FILLMORE, C. J. Frame semantics. In: LINGUISTICS IN THE MORNING CALM, selected papers from SICOL-198. **Proceedings...** Soeul: Linguistic Society of Korea, 1982. p. 111-137.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 4. ed. São Paulo: Atlas, 1994.
- GOMES, H. E. (Coord.). **Manual de elaboração de tesouros monolíngues**. Brasília: Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990.
- GREEN, R. Overview of relationships in knowledge organization. In: BEAN, C. A.; GREEN, R. (Ed.). **Relationship in knowledge organization**. Dordrecht: Kluwer, 2001. Chapter 1, p. 3-18.
- LANGACKER, R. **Foundations of cognitive grammar: descriptive application**. Stanford, CA, USA: Stanford University Press, 1991. v. 2.
- LAUSER, B. *et al.* From Agrovoc to the agricultural ontology service: concept server an OWL model for creating ontologies in the agricultural domain. In: INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS, 2006, Colima, Mexico. **Proceedings...** México: DCMI, 2006.
- MACULAN, B. C. M. S.; LIMA, G. Â. B. O. Relacionamentos em tesouros: o valor semântico dos verbos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 4, p. 182-201, out./dez. 2014.
- MACULAN, B. C. M. S. **Estudo e aplicação de metodologia para reengenharia de Tesouro: Remodelagem do Thesagro**, 2015. 345f. Tese (Doutorado em Ciência da Informação), Programa de Pós-Graduação em Ciência da Informação, Escola de Ciência da Informação, Universidade Federal de Minas Gerais - UFMG, Belo Horizonte, 2015.
- MACULAN, B. C. M. S. *et al.* **Refinamento de relações em tesouros: reengenharia do THESAGRO**. In: Encontro Nacional de Pesquisa em Ciência da Informação, 16., 2015, João Pessoa. **Anais...** João Pessoa: UFPB, 2015.
- NUNES, F. R. E. **Proposta de filtros cognitivos a partir dos aportes da Teoria da Valência Verbal e dos Papéis Qualia para mapear os sentidos dos verbos que expressam as relações semânticas em Tesouros**. 2019. 153f. Dissertação (Mestrado, área de concentração Ciência da Informação) – Programa de Pós-Graduação em Gestão & Organização do Conhecimento, Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Belo Horizonte: UFMG, 2019.
- PERINI, M. A. 2003. Sobre o conceito de 'item léxico': uma proposta radical. In: ALBANO, E. *et al.* (Org.). **Saudades da Língua**. Campinas: Mercado de Letras, 2003. p. 285-310.
- PERINI, M. A. **Estudos de gramática descritiva: as valências verbais**. São Paulo: Parábola Editorial, 2008.
- PERINI, M. A. **Describing verb valency: practical and theoretical issues**. New York: Springer Cham Heildeberg, 2015.
- PERINI, M. A. **Projeto Valências Verbais do Português (VVP)**. (Em elaboração).
- PROJETO ADESSE. Disponível em: [www.adesse.uvigo.es](http://www.adesse.uvigo.es). Acesso em: 4 jun. 2018.

PUSTEJOVSKY, J. **The generative lexicon**. Cambridge: The MIT, 1995.

PUSTEJOVSKY, J. Generativity and explanation in semantics: a reply to Fodor and Lepore. **Linguistic Inquiry**, Cambridge, v. 29, n. 2, p. 289-311, primavera de 1998.

ROTHER, E. T. Revisão sistemática X revisão narrativa. **Acta Paulista de Enfermagem**, São Paulo, v. 20, n. 2, jun. 2007.

SOERGEL, D. *et al.* Reengineering thesauri for new applications: the AGROVOC example. **Journal of Digital Information**, v. 4, n. 4, 2004.

