

relações e mais amparadas por restrições convencionais, mais fortes serão as evidências que elas apresentarão de autonomia e antagonismo nos sentidos em análise. Entre as relações mencionadas pelos autores, estão a antonímia, a hiponímia/hiperonímia e a meronímia/holonímia. Os autores destacam que, quando as relações em questão são as duas últimas, para que os sentidos sejam considerados autônomos, é importante que os sentidos pertençam a campos lexicais distintos. Os dois sentidos de *bank* ('instituição financeira' e 'margem do rio') ilustram bem a situação: *bank* em *bank of the river* (margem do rio) tem *source* (fonte), *bed* (leito) e *mouth* (boca) como co-merônimos, mas não há um hipônimo ou hiperônimo evidente para todos eles; já para *bank* em *the bank in the High Street* (o banco da High Street), co-hipônimos como *building society* (edifício de uma sociedade), *insurance company* (companhia de seguros) e *financial institution* podem ser identificados como um hiperônimo, e diferentes departamentos do *banco* seriam os merônimos.

(c) *Autonomia composicional*

O critério da autonomia composicional assumirá maior importância em sentidos com níveis mais baixos de autonomia, mas ele também implica claramente unidades delimitadas por fronteiras bem definidas. Trata-se de um critério para análise semasiológica com base em fatores onomasiológicos. Esse tipo de autonomia caracteriza situações em que, em um processo composicional – combinação sintagmática entre nome e verbo, nome e adjetivo –, um dos elementos seleciona somente uma porção do sentido do outro elemento (CROFT e CRUSE, 2004). Os autores explicam que o estabelecimento de fronteiras entre os sentidos sensíveis a esse tipo de autonomia se dará em torno da porção selecionada pelo conteúdo semântico do item em questão, separando-o do restante. Como exemplo para o PB, podemos citar *banco confortável*, em que o adjetivo ignora completamente o sentido 'instituição financeira'. Por outro lado, para uma situação como *banco prático*, o critério de autonomia composicional não iria ser útil, tendo em vista os múltiplos sentidos também do adjetivo *prático*.

Os seguintes critérios podem ser tomados como os *fatores prototípicos de antagonismo*, causando grau insignificante de similaridade entre os sentidos (CROFT e CRUSE, 2004):

- (i) Dificuldade para inclusão dos sentidos como co-hipônimos em uma organização hierárquica;

- (ii) Dificuldade para a aplicação da propriedade do *englobamento*, isto é, para a assimilação dos sentidos como partes de um mesmo todo relacionadas a uma *noção unificadora* (Gestalt global) válida. O máximo que é possível é encontrarmos como unificador um elemento altamente genérico, como *coisa, entidade, local*, noções que não são suficientemente distintivas para delimitarmos o que é e o que não é, por exemplo, *banco*. Nos termos de Langacker (1987), dizemos que o conteúdo esquemático nessas situações é pouco informativo ou distante.

4.3.2.2 Grau relevante de similaridade entre sentidos plenos

Sentidos com *grau relevante de similaridade* são sentidos que devem ser considerados (a) antagônicos e (b) autônomos, porém que não se encontram em graus máximos em nenhum desses *continua* (cf. CRUSE, 1995). Por ser uma situação intermediária de antagonismo e de autonomia, é inevitável percebermos tanto fatores de autonomia e disretude entre os sentidos quanto fatores de coerência.

Dizer que dois sentidos são parcialmente antagônicos e autônomos significa dizer que entre eles há certo grau de competição, fazendo com que o leitor precise selecionar um deles para a interpretação do enunciado, mas significa dizer também que entre eles há fatores de coerência (similaridade). Sendo assim, em razão do seu potencial de autonomia, esse tipo de sentido pode ser considerado também um *sentido completo* (*full sense*, nos termos de CROFT e CRUSE, 2004), porém tem propriedades de antagonismo menores do que o que vimos na seção anterior.

Como principal fator de coerência, que nega e/ou neutraliza certas propriedades de antagonismo, destacamos a possibilidade de identificação de alguma “noção geral” (CROFT e CRUSE, 2004, p. 110). No PB, podemos considerar como exemplo novamente o item *banco*, porém agora explorando outras porções do conteúdo conceitual a ele associado. Há vários usos desse item relacionados à noção de ‘coleção ou custódia’ de dinheiro ou de alguma outra ‘commodity’, como ocorre em contextos como *o banco da rua principal, Banco Mundial, o banco de sangue, o banco de olhos, o banco de dados*, etc.⁴

⁴ Exemplos traduzidos dos originalmente apresentados em Croft e Cruse (2004, p. 110).

Esse fator de coerência justifica entendermos ser possível formar um grupo (*cluster*, cf. CROFT e CRUSE, 2004) entre esses usos de *banco*, isto é, uma categoria de sentidos. A coerência entre os membros desse grupo, no entanto, não é total: percebe-se a existência de uma marcante descontinuidade (fator de autonomia) delimitando cada um dos elementos desse grupo ‘custódia’, que formarão *subgrupos* (ex. ‘custódia de dinheiro’, ‘armazenamento e custódia de órgãos’). Para explicar esse tipo de delimitação de fronteiras entre sentidos é que Croft e Cruse (2004) fazem uso da noção de *linha falha*, que facilita a construção de fronteira/delimitação e justifica percebermos esses sentidos como autônomos ou discretos em maior ou menor grau, podendo ser percebidos como membros de uma mesma categoria ou de categorias distintas, conforme a distância entre eles. Como evidência dessa discretude, Cruse (1989) destaca sua típica habilidade para participação em domínios altamente distintos, explicada a seguir.

Segundo Cruse (1989), embora não sejam completamente distintos, eles não geram problemas no uso porque fazem parte de domínios diferentes, exigindo que um dos sentidos seja selecionado a partir do contexto linguístico imediato. A noção de domínio é entendida aqui meramente como um conjunto de itens lexicais inter-relacionados por determinada relação semântica, tal como antonímia, hiponímia, meronímia (CRUSE, 1989). Dentro do seu próprio domínio, cada sentido local funciona como um sentido plenamente delimitado. Para os sentidos associados a *banco* que mencionamos acima, isso ficou claro ao observarmos que *banco* responsável pela custódia de dinheiro pode ser considerado um tipo de ‘instituição financeira’ (relação de hiponímia), já *banco* responsável pela custódia de sangue pode ser considerado um tipo de ‘instituição de saúde’ e, ainda, *banco* de dados pode ser considerado um tipo de recurso computacional.

Outro tipo de leitura parcialmente distinta é o que ocorre em *boca*, nos contextos *a boca do cavalo*, *a boca da garrafa*, *a boca de uma caverna*, *a boca do rio*, etc. (cf. CRUSE, 1995; 1989)⁵, que estão unidos pelo fator de coerência da similaridade. Esse tipo de variação de sentido é considerado o efeito sincrônico típico de extensões metafóricas a partir de um sentido comum com base na relação de similaridade. Como evidência dessas contiguidades, Cruse (1995) apresenta o critério da falta de efeito zeugmático:

⁵ Exemplo traduzido de Cruse (1995, p. 338).

sentidos não completamente distintos são contíguos em um *continuum*, portanto, tipicamente não geram zeugma, conforme exemplo 9.

9. A boca da caverna lembra a da garrafa⁶.

O exemplo (9) sugere a existência de algum elemento esquemático unindo os dois usos de *boca* em questão. Usos mais separados por pontos mais distantes no *continuum* de similaridade são ilustrados em (10); entretanto, também não são zeugmáticos.

10. O chocolate envenenado entrou na boca da Condessa exatamente quando seu iate entrou na do rio.⁷

Do mesmo modo que os usos de *banco* mencionados, os sentidos associados a *mouth* também fazem parte de domínios diferentes e exigem seleção a partir do contexto linguístico imediato – esses são os fatores de autonomia entre esse tipo de sentidos parcialmente distintos e significativamente similares. Por exemplo, o sentido básico, ou “literal”, de *boca* é ‘boca de um animal/humano’ (a julgar por sua relativa liberdade de contexto), conforme argumenta Cruse (1989; 1995). As extensões desse sentido são relativamente frágeis porque requerem condições precisas para seu aparecimento. Somente o sentido, de acordo com o autor, tem potencial para aparecer em um contexto neutro tal como *Fizemos um projeto sobre bocas*. O sentido ‘boca do rio’ é dificilmente possível, a menos que o item lexical *rio* ocorra na vizinhança próxima em uma porção anterior do discurso. O autor entende que, mesmo quando é montado o contexto para uma viagem em um rio, ele julga mais plausível entender (11) como mais básico ou comum do que (12) se nenhuma menção recente de *rio* tiver ocorrido.

11. Vamos descer a boca do rio?

12. Vamos descer a boca?

Vimos dois tipos de sentidos que potencialmente geram leituras parcialmente ambíguas em razão da possibilidade de percebermos efeitos de coerência (negando efeitos de antagonismo, mas não totalmente) e efeitos de autonomia (negando efeitos de coerência, mas não totalmente). O tipo de multiplicidade de sentidos a ser tratado aqui é

⁶ Exemplo traduzido de Cruse (1995, p. 39).

⁷ Exemplo traduzido de Cruse (1995, p. 39).

aquele a que nos referimos como *polissemia irregular*. O exemplo mais clássico desse tipo de fenômeno é o de *posição*, quando usado em sentidos ‘local ocupado’ ou ‘atitude mental’. Uma possibilidade de explicar a similaridade que pode ser percebida entre os sentidos é dizer, fazendo associação a mecanismos cognitivos (cf. LAKOFF e JOHNSON, 1980), que ocorre em razão do empréstimo de certas propriedades semânticas de um domínio como *lugar* a um domínio como *opinião*, entendendo-se haver a seguinte associação mental: “*declarar opinião*” é “*ocupar um lugar*”.

Os seguintes critérios podem ser tomados como *fatores prototípicos de coerência* entre sentidos parcialmente distintos:

- (i) Possibilidade de aplicar a propriedade do *englobamento*, sendo possível encontrarmos uma *noção unificadora* válida ou um *conteúdo esquemático*;
- (ii) Falta de efeito zeugmático;
- (iii) Grau médio de antagonismo;
- (iv) Percepção da existência de um sentido mais neutro (“básico” ou literal) e de um sentido aplicável a contextos mais restritos (“extensão”)
- (v) Possibilidade de reconhecer fatores de continuidade revelados pela *relação de similaridade* entre os sentidos.

Os seguintes critérios podem ser tomados como *fatores prototípicos de autonomia* entre os sentidos parcialmente distintos:

- (i) Há diferentes graus de saliência entre os sentidos (ex.: sentido mais básico e sentido menos básico);
- (ii) Os sentidos são altamente discretos;
- (iii) Autonomia relacional e associação a diferentes campos semânticos;
- (iv) A relação entre os sentidos é assimétrica, configura-se pela transferência de certas propriedades de um sentido (que pertence a um domínio A) para outro sentido (pertencente a um domínio B); trata-se de uma relação de A para B, que não se aplica, portanto, de B para A, não havendo empréstimo mútuo.
- (v) Delimitação do sentido a partir do contexto linguístico imediato;

4.3.2.3 A similaridade entre facetas de sentidos

O foco agora são os *subsentidos* do tipo *faceta*. Suas principais características são (a) certo grau de autonomia, mas (b) inexistência de antagonismo. As facetas são percebidas como *componentes distintos* (subsentidos do tipo faceta) de um único sentido global (sentido formado pela cooperação entre as facetas). A principal evidência de coerência entre as facetas é o fato de que o leitor não precisa selecionar apenas uma delas para a interpretação de um enunciado, apenas colocar uma em evidência. A principal evidência de autonomia entre as facetas é o fato de que elas não são capazes de estabelecer relação de hiponímia com um único elemento.

Esse tipo de variação de leitura gera o que referimos como *polissemia regular*. O rótulo *regular* dado a esse tipo de polissemia justifica-se por diferentes razões. A primeira razão é o fato de que os sentidos relacionados, embora tenham superordenados distintos em uma rede hierárquica, são elementos que podem ser considerados “perspectivas” de um mesmo sentido. Trata-se do acesso ou do enfoque de um aspecto particular de um sentido, conforme vimos em *livro* em 4.1, que pode ser usado com foco em sua parte física ou em sua parte textual. A segunda razão para o rótulo *regular* é o fato de que o mesmo tipo de alternância semântica é recorrente do mesmo modo no léxico. Como exemplo, além da alternância bidimensional ‘objeto físico’/‘conteúdo’, uma nova dimensão pode ser percebida, que é a faceta ‘texto físico’ (ex.: para tratar do tipo de impressão ou do *lay-out*), caracterizando um tipo de alternância de tridimensional de facetas percebida também em *agenda*, *polígrafo*, *apostila*, etc. Croft e Cruse (2004) enfatizam que, “embora as facetas associadas a um mesmo item sofram fortes restrições de convenção, em princípio, não há razão para acreditar que as facetas de sentidos formam um conjunto fechado” (p. 117). Esse tipo de relação entre facetas foi mencionado na Seção 2 deste trabalho sob o rótulo primo (*cousin*), evidenciando ser possível encontrarmos uma classe de sentidos que apresenta o mesmo tipo de alternância - não podemos, no entanto, deixar de observar que há exceções (ex.: *novel* ou *romance*).

Os seguintes critérios podem ser tomados como *fatores prototípicos de autonomia* entre sentidos parcialmente distintos, conforme os autores:

(a) *Autonomia relacional*

As facetas mostram completa independência relacional, de modo que cada faceta pode participar de relações de sentidos próprias, independentemente das demais. Por exemplo, em muitos contextos, o hipônimo de *livro* forma duas taxonomias paralelas, que não se relacionam:

13. Alguns dos *livros* que nós lemos foram romances, e outros, biografias.⁸

Em (13), *livro* é construído como a faceta ‘conteúdo’, que se apresenta em relação hiperonímica com os hipônimos incompatíveis *romance* e *biografia*.

14. Alguns dos *livros* eram de capa dura outros brochura.⁹

Em (14), *livro* é construído como a faceta ‘objeto físico’, que se apresenta em relação hiperonímica com os hipônimos incompatíveis *brochura* e *capa dura*.

15. *Alguns dos *livros* eram de capa dura outros brochura.

O contexto (15) ilustra que *romance* e *capa dura* não são combinados em um contexto que espera que os itens sejam construídos como incompatíveis em um único domínio.

16. Um romance é um tipo de *livro*.

17. Um *livro* de capa dura é um tipo de *livro*.

O fato de ser normal que sentenças descontextualizadas como (16) e (17) mostram que a potencialidade para a construção das facetas mencionadas para *livro* é bem suportada por restrições convencionais.

(b) *Autonomia composicional*

Os predicados podem ser aplicados às facetas independentemente:

18. *Livro* vermelho; *livro* empoeirado.

⁸ Exemplo traduzido de Croft e Cruse (2004, p. 117).

⁹ Exemplo traduzido de Croft e Cruse (2004, p. 117).

Nas construções em (18), o adjetivo modifica somente a faceta ‘objeto físico’ de *livro*. Em (19), no entanto, o adjetivo modifica somente a faceta ‘conteúdo’, propriedade considerada por Croft e Cruse (2004) como indicação relativamente fraca de autonomia.

19. *Livro* interessante; *livro* difícil.

Em um contexto apropriado, como (20), uma pergunta de *sim* e *não* contendo *livro* pode ter potencial para ser verdadeiramente respondida de ambas as maneiras, o que significa que elas remetem, alternativamente, a uma das facetas de cada vez.

20. Você gostou do *livro*?

Fora de contexto, é provável que a pergunta seja respondida com foco na faceta ‘conteúdo’ somente, mas não é difícil visualizar uma situação em que se possam construir as facetas ‘objeto físico’ ou ‘texto físico’, por exemplo.

(c) “Núcleos” autônomos

Outra manifestação de autonomia entre as facetas apresentada pelos autores pode ser percebida em (21) e (22).

21. Não estou interessado no conteúdo, etc., estou interessado no *livro* em si.

22. Não estou interessado na encadernação, etc., estou interessado no *livro* em si.

Os contextos acima podem ser normalmente interpretados e envolvem diferentes leituras da expressão *do livro em si*: em (21), ‘objeto físico’, em (22), ‘conteúdo’. Esse fato é tomado como indicação de que *book* é associado a dois “núcleos” semânticos autônomos (CROFT e CRUSE, p. 119).

(d) Outros fatores de autonomia

O item *livro* pode ser ativado mesmo se uma de suas facetas for deixada de fora.

23. Eu ganhei um *livro* para escrever durante os minutos em que eu estiver na reunião.

24. A: Como está o seu *livro*?

B: Ah, está todo aqui [apontando para a cabeça], mas eu não escrevi nada ainda.

Por fim, conforme mostra (25), cada faceta pode ser alvo independente para extensão metafórica:

25. Um *livro* de cruzadinhas.

Vimos que somente a faceta ‘objeto físico’ foi selecionada. Por fim, os autores destacam o fato de que cada faceta pode receber nome próprio independentemente: *O pequeno príncipe* é o nome do conteúdo, não do objeto físico, nem de ambas as facetas texto e objeto físico combinadas.

Passaremos agora a tratar dos critérios que podem ser tomados como *fatores prototípicos de unificação* das facetas. Assim como Croft e Cruse (2004) mostram ocorrer com *book*, percebemos que há usos de *livro* que podem ser unificados em um sentido global, o que justifica o argumento de que as facetas não têm o *status* de sentido pleno. Eles enfatizam que, por vezes, a autonomia das facetas é superestimada, e a evidência principal é o fato de que as facetas não podem ser unificadas como hipônimos de uma única categoria: não há categoria que agrupe ‘conteúdo’ e ‘objeto físico’; além disso, elas não são partes de um todo em sentido normal. Seguem, na sequência, os critérios para unificação apresentados por Croft e Cruse (2004):

(a) *A coocorrência é a leitura prototípica*

Os autores defendem que a leitura prototípica de *livro* é a leitura combinada com ambas as facetas, embora uma possa ser selecionada sem a outra, fato que, quando acontece, é considerado periférico.

(b) *Propriedades composicionais unificadoras*

Um aspecto para unir as propriedades composicionais é a existência de predicados que combinam com ambas as facetas simultaneamente:

26. Publicar um *livro*

Em (26), temos uma situação que impossibilita separar as facetas, visto que não é possível publicar um *livro* sem haver ‘conteúdo’ e ‘objeto físico’.

Outra situação de unificação composicional das facetas é ilustrada em (27):

27. Este é um *livro* muito interessante, mas é horrível de carregar.

Observa-se que a situação de zeugma não se aplica a contextos como esses, evidenciando a inexistência de antagonismo, conforme Croft e Cruse (2004). *Interessante* modifica a faceta 'texto', e *pesado de carregar* modifica a faceta *objeto físico*.

Por fim, para justificar o status de subsentidos às facetas, Croft e Cruse (2004) retomam o que ocorre em *banco* e *teacher*. No caso de *banco*, o que ocorre é que a representação conceitual provê forte pressão para separar as leituras; no caso de *criança*, em (28) e (29), ocorre exatamente o contrário; não havendo sintomas de autonomia, a representação conceitual provê forte pressão para unificar as leituras.

28. O nome da *criança* será Izabela.

29. O nome da *criança* será Antônio.

Vemos que *criança* em (28) é interpretada como feminino e em (29) como masculino. Nos exemplos de *livro*, os autores salientam que as restrições de estabilidade não nos induzem veementemente a seguir em nenhuma das direções, então, somos relativamente livres para construir as facetas tanto de maneira autônoma quanto como unificadas em resposta a outras restrições, tais como contextuais e comunicativas. E Croft e Cruse enfatizam que isso não ocorre por alguma falha das restrições conceituais, mas por elas agirem em direções contrárias, efetivamente cancelando umas às outras.

As seguintes propriedades podem ser tomadas como *fatores prototípicos de coerência* entre as facetas de sentidos:

- (i) Ausência de antagonismo;
- (ii) Destaque de apenas uma das facetas;
- (iii) Detecção de um conjunto de sentidos que apresenta o mesmo tipo de alternância (ex.: 'objeto físico' e 'conteúdo').
- (iv) Possibilidade de identificação de um Gestalt para as facetas;

As seguintes propriedades podem ser tomadas como *fatores prototípicos de autonomia* entre as facetas:

- (i) Certo grau de autonomia entre as facetas, sendo elas percebidas como componentes distintos de um único sentido;

- (ii) Não existência de um único hiperônimo para as duas facetas;
- (iii) Seleção de outras relações paradigmáticas distintas pelas facetas.

4.3.2.4 A similaridade entre os microsentidos

Os microsentidos estão em relação regular se um for uma especialização do outro, o que acarreta, evidentemente, que o último seja uma generalização do primeiro (CRUSE, 2000a). O autor explica que é possível distinguir especialização de generalização se reconhecermos um dos sentidos como mais básico do que outro: se A é mais básico do que B, e B é mais especializado do que A, então B é uma especialização de A.

Na seção anterior, tratamos do primeiro tipo de subsentido apresentado por Croft e Cruse (2004). Passaremos agora a tratar dos subsentidos classificados como *microsentidos* pelos autores. São unidades de sentido distintas de um item que ocorrem em diferentes contextos e cujas interpretações se estabelecem em relação de incompatibilidade mútua no mesmo nível hierárquico, assim como, por exemplo, nomes de animais, como *gato*, *cachorro*, *ovelha*, *vaca*, *porco*, *cavalo*. Os microsentidos são unidades de sentido que apresentam (a) significativo grau de autonomia, embora possam ser unificados como hipônimos de uma mesma categoria e (b) não há antagonismo, visto que funcionam como *co-autohipônimos* de um sentido geral. Além dessas características, é marcante o fato de que o hiperônimo não tem *status default*. Croft e Cruse (2004) apresentam *knife* como exemplo típico. Em razão de seu comportamento muito semelhante ao de *faca*, apresentaremos o item do PB para ilustrar, visto que ele pode ser usado como uma arma, como um instrumento, como um talher, sendo cada um desses usos um microsentido do sentido global de *faca*.

Croft e Cruse destacam diferentes fontes para as evidências de autonomia dos microsentidos:

(a) *Autonomia relacional*

Cada microsentido constrói seu próprio conjunto de relações de sentido independente. Os microsentidos de *faca* (30), por exemplo, geram interpretações que

pertencem a diferentes taxonomias e que, portanto, têm diferentes *hiperônimos*, *hipônimos* e *co-hipônimos*.

30. Talher: faca, garfo, colher

Arma: faca, revólver, granada

(b) *Autonomia de condições de verdade*

Os microsentidos exibem autonomia de condições de verdade. O teste das condições de verdade através de perguntas do tipo sim/não que apresentam respostas independentes para cada microsentido. Como exemplo, podemos citar (31)

31. Mãe (na mesa; Eduardo está partindo uma carne com seus dedos): Usa a faca para cortar a carne, Eduardo.

Eduardo (está com uma faca em seu bolso, mas não com uma faca do tipo que corte): Eu não tenho.

(c) *Autonomia com base na restrição de identidade*

O teste de restrição de identidade é também considerado importante evidência de autonomia entre os microsentidos, visto que, tipicamente, há forte pressão para que anaforizante e anaforizado tenham a mesma leitura, conforme mostra (32)

32. Cícero precisa de uma *faca*; Daniel também.

Cícero enviou uma *carta*; Daniel também.

Cícero tem um *equipamento*; Daniel também.

A restrição parece ser mais forte em certas situações, o que é visto pelos autores como reflexo de graus de autonomia, e isso fica mais claro quando contrastamos (32) com situações em que a inexistência de contraste é clara, como em (33).

33. Cícero tem um *carro*; Daniel também.

Cícero tem um *primo*; Daniel também.

Para encerrar a caracterização dos *microsentidos*, consideremos as seguintes questões:

- (i) *O que faz, então, os microsentidos não poderem ser caracterizados como sentidos plenos?* Os microsentidos não são sentidos distintos porque eles são unificáveis sob um mesmo hiperônimo e não são antagônicos, de acordo com os critérios usuais (CROFT e CRUSE, 2004).
- (ii) *Por que os microsentidos não são facetas?* A razão pela qual os microsentidos não são facetas é que as facetas são de tipos ontológicos distintos (ex.: concreto e abstrato) e não podem ser unificadas sob um mesmo hiperônimo (CROFT e CRUSE, 2004).
- (iii) *Por que os microsentidos não são simples enriquecimentos contextuais de sentidos superordenados?* Há diversas razões (CRUSE, 1995): (a) têm grau maior de disretude do que simples especificações contextuais (ex. *criança* – feminino ou masculino); (b) microsentidos, ao contrário de enriquecimentos contextuais, têm suas próprias relações de sentido (autonomia relacional) e participam, ao contrário de enriquecimentos contextuais, de domínios independentes (autonomia de domínio); (c) dentro do seu próprio domínio, os microsentidos não precisam ser adjetivados para que seu conteúdo básico seja percebido, ao contrário do que ocorre em casos como o de *criança*; (d) o superordenado é marcado e não aparece somente por pressão contextual, embora varie de domínio para domínio (ex.: faca como talher, faca como arma); isso não ocorre com situações de modulações contextuais como a de *criança*, em que, em um contexto neutro, a leitura dominante é também neutra em relação ao sexo.

Ao longo desta seção, foram descritas uma série de propriedades que servem como evidências para o delineamento de fronteiras entre sentidos plenos, facetas e microsentidos. Essas noções são importantes para entendermos sobre o grau de antagonismo, similaridade e de autonomia existente entre os diferentes sentidos e para que possamos propor uma metodologia de análise dos sentidos polissêmicos

(lembramos que, quando fazemos uso do termo *sentido*, pretendemos nos referir tanto a sentidos plenos quanto a subsentidos) baseada em critérios mais explícitos do que a intuição do falante. Em razão da flexibilidade inerente à construção dos sentidos, não consideramos definitivos ou determinantes os resultados de quaisquer testes. Eles servirão como recursos auxiliares à delimitação do tipo de polissemia por serem representativos de usos estabelecidos construídos por cada item.

4.3.3 Tipos de similaridade: relações de polissemia

Em 4.3.2, tratamos das questões implicadas na delimitação dos sentidos, percebendo diferentes graus de similaridade revelados a partir de critérios de antagonismo e de autonomia entre os sentidos, fato importante para caracterizar o tipo de polissemia (e de multiplicidade de sentidos como um todo). Nesta seção, abordaremos os diferentes efeitos de saliência que os múltiplos sentidos associados a um mesmo item lexical podem apresentar, facilitando a ativação de um ou de outro no uso. Concluídas as reflexões sobre os aspectos funcionais implicados na descrição da semântica de um item polissêmico, nesta seção, o objetivo é caracterizar a similaridade entre sentidos em termos de tipo.

As relações que selecionamos para a apresentação tomaram como critério o viés sincrônico de similaridade entre sentidos, e procuramos, para enfatizar essa perspectiva, fazer uso de uma terminologia distinta daquela comumente empregada em SLC já marcada de conotação diacrônica, por ter como objetivo explicar a motivação cognitiva de determinada extensão de sentido. Para este inventário de tipos, levamos em consideração as informações que inventariamos na Seção 3 especialmente, aquelas apresentadas por Langacker (ex.: relações verticais e horizontais), Geeraerts (ex.: generalização, especialização, metáfora, metonímia) e Blank (ex.: polissemia metafórica, polissemia metonímica, polissemia por co-hiponímia, polissemia taxonômica). Acreditamos que o inventário de relações que apresentaremos a seguir dê conta do tipo de conteúdo que percebemos como mais relevante para a construção de uma rede sincrônica de polissemia: (i) explicitação de informações que abarcam relações de

polissemia entre sentidos institucionalizados em uma língua particular (PB)¹⁰; (ii) distinção entre relações regulares e irregulares de polissemia; (iii) explicitação de relações lineares entre elemento geral e elemento específico; (iv) explicitação de relações de similaridade lineares que deem conta de alternâncias do tipo que vimos ocorrer em *livro* ‘objeto físico’ e ‘conteúdo’, além de outras alternâncias, como *parte-todo*, *instituição-prédio*, entre outras tipicamente denominadas alternâncias de *figura e fundo*; (v) explicitação de relações horizontais irregulares baseadas em similaridade menos previsíveis e lineares porque envolve sentidos com importante grau de autonomia e que se associam a domínios altamente distintos.

A apresentação das relações de polissemia seguirá o seguinte critério: parte-se do maior grau de regularidade ou linearidade e chega-se ao menor. Iniciaremos pelas relações regulares entre microssentidos, *autohiponímia/autosuperordenação*; em seguida, trataremos da relação regular que ocorre entre unidades de sentidos nomeadas *facetas*, a *automeronímia/autosuperordenação*; por fim, trataremos do tipo clássico de relação de polissemia, que é a irregular, que se manifesta entre sentidos plenos.

4.3.3.1 Relação de Polissemia Regular

O foco aqui é o tipo de relação ilustrada nas ocorrências de *livro*, *cachorro* e *faca* nas seções anteriores. Trata-se da *relação de polissemia regular*. O rótulo *regular* é dado a essa relação de polissemia especialmente pelo fato de que o mesmo tipo de alternância semântica é recorrente do mesmo modo no léxico. Entre os tipos de polissemia reconhecidos como lineares, sistemáticos ou regulares, destacaremos, a seguir, os mais proeminentes: a polissemia por *parte* e *todo* (*automeronímia/autoholonímia*) e a polissemia por *é-um* (*autohiponímia/autosuperordenação*) (CRUSE, 2000).

¹⁰ Uma situação oposta à de sentidos fortemente estabelecidos por restrições de uso ocorre em nuances de sentido geradas no momento do uso (discurso). São sentidos que não fazem parte do potencial de sentido acessível por um item. Conforme Blank (2003), trata-se de uma distinção entre sentidos “lexicalizados” (ex. *livro* ‘objeto físico’ e ‘conteúdo’) e “não-lexicalizados” (ex. *sanduíche de presunto*, expressão considerada comum no discurso de garçons). Embora esse uso possa ser considerado comum e recorrente, seria um erro dizer que, entre os possíveis sentidos da expressão, está “pessoa que pediu sanduíche”. Essa leitura é gerada no discurso (BLANK, 2003).

A seguinte frase lógica¹¹ pode ser tomada como critério geral para a identificação da polissemia regular: a polissemia de uma unidade lexical A com o sentido a_1 e a_j é chamada regular em uma dada língua se existir pelo menos uma outra unidade lexical B com sentidos b_1 e b_j que sejam semanticamente distintos um do outro exatamente da mesma maneira (APRESJAN, 1973). O tipo de distinção/similaridade será determinado por cada uma das relações indicadas abaixo:

4.3.3.1.1 Relação de polissemia por automeronímia/autoholonímia

O tipo de polissemia regular mais explorado na literatura é o que relaciona *parte* e *todo*, considerando-se essa relação entre sentidos polissêmicos uma relação tão geral quanto é a meronímia.

Esse tipo de relação regular ocorre entre facetas de sentidos, elementos com alto grau de similaridade. Esse rótulo representa facetas de sentidos que podem ser simultaneamente acessadas em dado contexto, mas também se pode dar saliência para uma apenas. Embora as facetas de sentidos tenham superordenados distintos em uma rede hierárquica, elas são elementos que podem ser considerados regulares por serem “perspectivas” de um mesmo sentido geral¹², portanto, a similaridade é, de certo modo, previsível (CRUSE, 2000a). Como exemplo, mencionamos a alternância *objeto físico/conteúdo*, que é percebida também em itens lexicais como *agenda*, *polígrafo*, *apostila*, etc., e que, ainda, tem certo grau de universalidade, visto que ocorre em várias línguas do mesmo modo (CRUSE, 2000) (ex. *book* - inglês, *Buch* - alemão, *libro* - espanhol e italiano, *livre* - francês).

Tendo em vista que as alternâncias de sentidos relacionados por polissemia regular já foram bastante exploradas na literatura, é possível acessarmos um vasto material sobre nominais tipicamente associados a facetas de sentido que formam uma classe. A similaridade de sentidos percebida entre *automerônimos*/*autoholônimos* é equivalente à que ocorre entre itens lexicais merônimos. Desse modo, podemos estudar essa relação de polissemia fazendo uma analogia à relação paradigmática clássica. A partir dessa visão, vemos a relação como uma família de relações. Fazendo uma adaptação do que propõem Saint-Dizier e Viegas (1995) para a *meronímia*, aproveitando

¹¹ Adaptada de Apresjan (1973).

¹² Tais alternâncias semânticas são frequentemente consideradas como reflexo da metonímia, relação baseada em correlação, contiguidade ou proximidade conceptual. Dentro de um mesmo domínio (não “cross-domain”, por isso contiguidade conceptual), possibilita que uma entidade seja usada pela outra porque ambas coexistem nesse mesmo domínio” (EVANS & GREEN, 2006)

algumas das classes clássicas e gerais de polissemia regular (PUSTEJOVSKY, 1995), apresentamos, para encerrar a seção, o quadro 4, com um inventário de tipos gerais de relação de polissemia regular por *automeronímia* com exemplos em PB. Salientamos que esse quadro abarca apenas tipos gerais e, conforme mostramos na Seção 2, com a proposta do CoreLexII (quadro 2), esses tipos podem ser mais específicos, facilitando sua extração automática. Nossa intenção aqui, no entanto, é registrar uma tipologia que sirva como guia para nossa análise manual.

1. Todo/parte, todo/porção Ex.: <i>janela</i> Encomendei a <i>janela</i> para a sala./Passei pela <i>janela</i> .
2. Produto/substância, objeto/material Ex.: <i>chocolate</i> Comprei <i>chocolate</i> ./Esse bolo é de <i>chocolate</i> .
3. Continente e conteúdo, objeto/substância Ex.: <i>colher</i> Entortei a <i>colher</i> ./Quero uma colher de açúcar no meu café.
4. Objeto físico/conteúdo Ex.: <i>livro</i> Li o <i>livro</i> ./O <i>livro</i> é pesado para carregar.
5. Pessoas/funcionamento/instituição/estrutura física (tipo mais específico do que os anteriores) Ex.: <i>universidade</i> A <i>universidade</i> está em festa. A <i>universidade</i> inicia em fevereiro. A <i>universidade</i> foi fundada há 31 anos. A <i>universidade</i> está em obras.
6. Animal/alimento (tipo mais específico do que os anteriores) Ex.: <i>porco</i> Comprei um porco para o sítio./Comi porco no almoço.

Quadro 4 – Inventário de tipos de *automeronímia*

4.3.3.1.2 Relação de polissemia por autohiponímia/autossuperordenação

As relações de polissemia do tipo *autohiponímia* ou *autossuperordenação* são apresentadas por Cruse (2000a) e constituem relações verticais que ocorrem entre sentidos gerais e específicos em diferentes graus de detalhamento.

A autohiponímia/autossuperordenação envolve relação entre unidades de sentido que, seguindo Croft & Cruse (2004), denominamos *microsentidos*. Trata-se de uma relação entre um sentido construído em um grau menor de detalhamento e um sentido construído em grau maior de detalhamento. É o que ocorre, por exemplo, entre os sentidos de *cachorro* como ‘macho da raça canina’ (maior grau de detalhamento) e como ‘membro da raça canina’ (menor grau de detalhamento).

Na relação de *autohiponímia*, conforme Cruse (2000a), o sentido geral é geralmente construído como padrão e está havendo uma especificação em dado contexto. A relação entre o elemento geral e o específico pode ser aplicada repetidamente, marcando diferentes graus de detalhamento dos sentidos. Cruse (2000a) especifica essa relação entre especialização e generalização como *linear*, visto que o fato de o sentido **A** ser generalização do sentido **B** acarreta que **B** seja especialização de **A**. Para ilustrar, vamos detalhar os microsentidos tipicamente associados a *cachorro*, exemplo adaptado de *dog*, apresentado por Cruse (2000a, p. 111): o item *cachorro* é frequentemente associado ao sentido geral (padrão) “membro da raça canina”, como em *Os donos de gatos e cachorros devem registrar seus animais de estimação*, e uma leitura mais específica, como em *Aquele não é um cachorro, mas uma cadela*. Cruse (2000a) chama a atenção para o fato de que a leitura mais específica demonstra autonomia, visto que a segunda oração contradiz a leitura mais geral de *cachorro*, conforme mostra a figura 17.

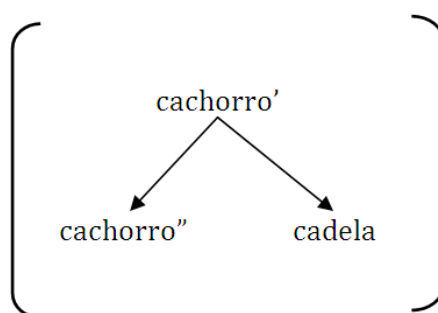


Figura 17 - Autohiponímia *cachorro*

A figura 17 mostra que *cachorro''* É-AUTOHIPÔNIMO de *cachorro'* e que *cadela* É-UM-TIPO-DE *cachorro'*, construído como ‘animal da raça canina’.

A *autossuperordenação* ocorre quando um item tem como padrão um sentido mais específico, contextualmente restrito, e está havendo uma generalização em dado contexto. Um exemplo de *autossuperordenação*, adaptado de Cruse (2000a, p. 111), é *homem*. Esse item lexical envolve a generalização do masculino, sendo usado para denotar raça humana, em contraste com o uso de *homem* denotando o ser masculino da raça humana, como em *João é um homem de caráter sem igual*, e em uma leitura mais geral, como em *O homem precisa pôr em prática alternativas ecológicas*. A figura 18 ilustra esses exemplos.

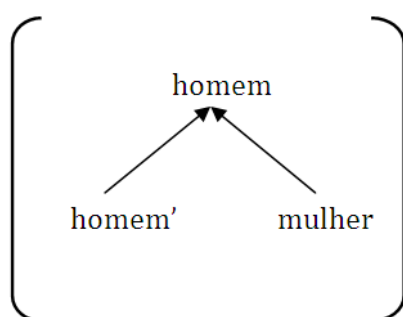


Figura 18 – Autossuperordenação *homem*

A figura 18 mostra que *homem* "É-AUTOSSUPERORDENADO-DE *homem*" e que *mulher* É-UM-TIPO-DE *homem*, como 'ser humano'.

Quando esse tipo de relação ocorre entre itens lexicais distintos, no contexto das relações semânticas nos moldes estruturalistas, o termo usualmente empregado para o elemento mais geral é *superordenado* ou *hiperônimo* e para o termo mais específico é *subordinado* ou *hipônimo*. Para identificarmos a relação hiponímica entre os nomes, podemos fazer uso do seguinte esquema: X é um Y, ou X é um tipo de Y; a mesma frase lógica é válida para sentidos polissêmicos.

A terminologia empregada aqui foi proposta em Cruse (2000a). Entretanto, essa relação entre um elemento mais geral e outro mais específico aparece na literatura também como relação de *esquema-instância* (LANGACKER, 1987; 2002) e *especialização-generalização* (denominação que tem origem na linguística diacrônica, empregada, por exemplo, por BRÉAL, e que é também utilizada por autores como Geeraerts (2000a)) e como relação de *polissemia taxonômica*, por Blank (2003). Para enfatizar nosso foco sincrônico, não faremos uso do vocabulário empregado para expressar relações de

motivação de polissemia, mas de termos cunhados para descrever sincronicamente relações entre sentidos, conforme Cruse. Outra questão importante a destacar é que olharemos para as relações de autohiponímia e autossuperordenação como relações inversas, não sendo relevante para nossa aplicação pensar em uma distinção estrita entre ambas.

4.3.3.2 Relação de Polissemia Irregular

A *relação de polissemia irregular* é caracterizada por um tipo de similaridade e pode ser aplicada a sentidos plenos que são construídos de formas diferentes ou mesmo conflitantes, porém, é possível percebermos entre os elementos alguma nuance de sentido comum, que pode ser descrita como uma “noção geral”, tal como vimos em “custódia” para *banco*. Trata-se de um princípio de agrupamento do tipo de similaridade a partir da percepção de uma *noção geral* ou como um Gestalt global. É essa noção geral que será a responsável por justificar pensarmos em certos sentidos como constituintes, mesmo com alto grau de um conjunto formado por elementos com certo grau de antagonismo e de autonomia. Essa é a relação que caracteriza o fenômeno clássico de polissemia e que foi apresentada por Langacker como uma relação gerada por extensão do tipo *similaridade* a partir do protótipo e por Blank (2003) como *polissemia metafórica*, em razão da sua motivação típica ser transferências desse tipo.

O rótulo *irregular* é aplicado porque esse tipo de relação de polissemia, embora altamente recorrente no léxico, é mais difícil de ser inferido ou previsto a partir de padrões de geração, tendo em vista que a similaridade a partir de associações com alto grau de subjetividade. Conforme temos comentado ao longo do trabalho, a similaridade que pode ser construída entre dois ou mais sentidos pode ser de diferentes tipos (forma, função, aparência, etc.), que se organizam a partir de diferentes graus de intensidade e de convencionalização. Mesmo associações objetivas, conforme já mencionamos, têm profundas marcas de similaridade. É essa percepção de similaridades baseadas em fatores subjetivos que possibilita construirmos relação entre sentidos com importante grau de discrição e antagonismo que podem ser associados a domínios discursivos diferentes ou mesmo distantes (BLANK, 2003). Além do exemplo de *cabeça* que

referimos ao longo do trabalho, podemos destacar o item *posição*, que tipicamente constrói sentidos como ‘local ocupado’ (em *Estou em posição confortável*) ou ‘atitude mental’ (em *Preciso tomar uma posição*). Uma forma de explicar a similaridade entre os sentidos é pensar em empréstimo de certas propriedades semânticas de um domínio como *lugar* a um domínio como *opinião*. Pode-se dizer que se faz a seguinte associação mental: “*declarar opinião*” é “*ocupar um lugar*”¹³; a partir daí, poderíamos dizer que é possível construir um conteúdo esquemático *localização*, englobando os dois sentidos mencionados.

A relação entre sentidos irregularmente polissêmicos será indicada em nosso trabalho pelo rótulo *é_similar_a*, entretanto, não podemos deixar de chamar a atenção para a não-especificidade do rótulo. Para demonstrar a dificuldade de encontrarmos um rótulo mais específico para a relação, vamos recuperar os exemplos de polissemia irregular que apresentamos ao longo do trabalho, procurando explicitar o tipo de similaridade que conseguimos construir em cada caso: (a) *bala* (similaridade de forma), (b) *boca* (similaridade de forma e função), (c) *cabeça* (similaridade de função e posição), (d) *banco* (no sentido de ‘custódia’, similaridade de função), (e) *posição* (similaridade por associação de processo cognitivo e processo físico). Além de haver uma variedade de associações, fica evidente que algumas das associações são mais fáceis de ser construídas do que outras e, por consequência, serão mais consensuais entre os falantes. Como exemplo de similaridade simples de ser construída, destacamos *cabeça* ‘parte do corpo humano’ e ‘quem está no comando’. Como exemplo mais difícil de ser construído, destacamos *posição*.

Os critérios para a caracterização da *polissemia irregular* são os seguintes: (i) há diferentes graus de saliência entre os sentidos, mas todos são altamente discretos; (ii) observa-se um grau médio de antagonismo entre os sentidos, já que é possível reconhecer alguma similaridade; (iii) a relação de similaridade tem grau alto de irregularidade e baixo de previsibilidade, já que a associação se dá entre sentidos

¹³ Numa perspectiva diacrônica, a polissemia assistemática é frequentemente associada ao fenômeno da metáfora (entendida como “palavras e expressões convencionais produzidas a partir de co-relações, especificadas como mapeamentos entre dois domínios”, conforme Lakoff (1987) e Lakoff & Johnson (1980)), considerada sua principal motivação. Na análise de *posição*, nota-se, de fato, que uma associação metafórica é o que está na base da similaridade existente entre os sentidos denominados irregularmente polissêmicos, visto que ela ocorre entre sentidos que mapeiam para domínios distintos.

evocadores de domínios distintos; (iv) a relação entre os sentidos é *assimétrica*¹⁴, configura-se pela transferência de certas propriedades de um sentido (que pertence a um domínio A) para outro sentido (pertencente a um domínio B); trata-se de uma relação de A para B, que não se aplica, portanto, de B para A, não havendo empréstimo mútuo.

A seguinte fórmula lógica pode ser tomada como critério geral para a identificação da polissemia irregular: uma unidade lexical A com o sentido a_1 e a_2 é chamada irregular em uma dada língua se NÃO existir outra unidade lexical B com sentidos b_1 e b_2 que sejam semanticamente distintos um do outro exatamente da mesma maneira da distinção entre a_1 e a_2 .

4.4 QUADRO SÍNTESE DA SEÇÃO

¹⁴ A propriedade de assimetria é destacada por Evans & Green (2006) ao descreverem a polissemia gerada por metáfora.

		Formam <i>polyset</i> *		
Multiplicidade de sentidos	Homonímia	Polissemia irregular	Polissemia regular	
Tipo de sentido	Sentidos plenos 1	Sentidos plenos 2	Subsentidos	
			Microsentidos	Facetas
Exemplos	Manga, <i>banco</i> (inst./assento)	Posição, cabeça, boca, <i>banco</i> (inst.=custódia)	Cachorro, faca	Livro, janela
Similaridade	Grau insignificante	Leituras semidistintas (grau relevante)	Grau funcional	Grau alto
Tipo de relação	-	Assimétrica (irregular)	Simétrica (regular - recorrente)	
		<i>tem_similaridade_com</i>	<i>é_autohipônimo_de</i>	<i>é_automerônimo_de</i>
Autonomia	Grau máximo	Grau significativo	Grau significativo - Há incompatibilidade mútua no mesmo nível hierárquico – são co-autohipônimos.	Grau mínimo
Antagonismo	Grau máximo	Grau médio	Grau mínimo	Inexistente
Capacidade de unificação	Altamente resistente	Ocorre. <i>Tipo de unificação:</i> Possibilidade de construção de uma ‘noção geral’.	Ocorre. <i>Tipo de unificação:</i> um subsentido é uma subespecificação do outro.	Ocorre. <i>Tipo de unificação:</i> os subsentidos são componentes distintos de um único sentido. Partes de um mesmo todo.
Seleção de atenção	Grau máximo	Grau máximo	Algum grau. O subsentido hiperônimo não tem <i>status</i> de atenção <i>default</i> .	Grau baixo - Seleção de perspectiva. Não é necessário delimitar apenas uma das facetas, mas é possível (núcleos autônomos).
Efeito zeugmático	Grau máximo	Difícilmente ocorre	Grau significativo.	Difícilmente ocorre
Autonomia relacional	Grau máximo. Campos semânticos distintos.	Grau máximo. Campos semânticos distintos.	Grau alto.	Grau máximo. Campos semânticos similares, tipos ontológicos distintos, mas não são unificadas sob um mesmo hiperônimo.
Autonomia composicional	Completa (não muito relevante)	Completa (não muito relevante)	Completa (não muito relevante)	Ocorre, mas também pode servir como evidência de unificação.
Outros critérios de autonomia			a. Autonomia de condição de verdade. b. A ativação do co-hipônimo não exige adjetivação.	a. “Núcleos” autônomos. b. Independência de ativação das facetas.

Quadro 5 – Síntese da seção 4.

SEÇÃO 5 - DELINEAMENTO LINGÜÍSTICO DO MODELO DE REPRESENTAÇÃO: O CONSTRUCTO *POLYSET*

As discussões empreendidas nas seções anteriores convergem neste ponto para a proposição de uma metodologia de análise e representação linguístico-computacional da polissemia de nominais que complemente as informações léxico-conceituais já codificadas em uma *wordnet*. Nesta seção, o objetivo é apresentar o delineamento linguístico desse modelo. Para isso, como já observamos na introdução, adotamos a estratégia de pesquisa *task-driven* (cf. NIRENBURG e RASKIN, 2004), de modo que procuraremos manipular os constructos teóricos propostos no âmbito da SLC, sem perder de vista o objetivo de o constructo a ser proposto contribuir para a ampliação da utilidade das *wordnets* para o desenvolvimento de sistemas de PLN. Essa meta exigiu, ao longo de todo o trabalho, um direcionamento da pesquisa para a proposição de uma representação linguística que possa ser codificada computacionalmente. Salientamos que essa orientação aplicada não exige - nem serve de justificativa para - um olhar leviano à teoria linguística e, conforme procuraremos demonstrar, essa tarefa interdisciplinar pode ser realizada com sucesso.

A metodologia geral da tese, como argumentamos na introdução, seguiu aquela proposta por Dias-da-Silva (1996, 1998; 2003; 2006), e, portanto, previu a integração dos três domínios mutuamente complementares: o linguístico, o linguístico-computacional e o computacional. Para a análise linguística, foram necessárias reflexões referentes aos dois primeiros domínios. O **domínio linguístico-computacional** tem o papel de estabelecer a ligação entre o primeiro e o terceiro domínios. Assim, a discussão que se fez no âmbito desse domínio forneceu o cenário computacional da pesquisa, ao mapear as lacunas e as exigências da aplicação-alvo. A conclusão dessa etapa, descrita na Seção 2, evidenciou que uma representação formal da polissemia de nominais em uma base de dados lexicais possibilita que um sistema de PLN manipule níveis distintos de generalidade de sentidos. Desse modo, a base se torna mais útil, por exemplo, tanto para sistemas de tradução automática e de processamento de informação, que exigem um

grau de granularidade mais fina, quanto para sistemas de classificação de textos, que necessitam de um grau de granularidade mais grosseiro.

Conhecido o cenário de aplicação, trabalhamos a segunda etapa da pesquisa: o **domínio linguístico**. Embora ele esteja, em certa medida, subordinado a necessidades computacionais, procuramos explorar o modelo escolhido em profundidade, de maneira a tornar a representação linguístico-computacional também adequada a um contexto linguístico-cognitivo. A flexibilidade e a irregularidade do significado lexical, tão enfatizadas na SLC, parecem, em um primeiro momento, colocar uma limitação à sua codificação computacional. Conforme procuramos demonstrar nas Seções 3 e 4, trata-se apenas de uma impressão. A abordagem de redes, mais especificamente, a de redes de polissemia, revelou ser a solução ideal, por elas possibilitarem uma representação que abrange dois tipos de codificação: a que codifica a representação de sentidos gerais e abstratos e a que codifica a representação de níveis de instanciação dessas abstrações, fornecendo uma estratégia para a representação do significado lexical em diferentes níveis de generalidade e de estabilidade. Considerar a polissemia um fenômeno categorial possibilitou entendermos que os sentidos podem ser decompostos em termos de “categorias” e, desse modo, são explicados como entidades complexas e a partir de uma estrutura multidimensional. Com base nas investigações relatadas na Seção 4, seguindo o paradigma da SLC, o modelo de representação deve apresentar as seguintes características: ser capaz de codificar (i) os tipos de sentido (sentidos plenos e subsentidos do tipo facetas e microsentidos), (ii) os graus distintos de saliência, (iii) os graus distintos de delimitação de sentidos (maior ou menor autonomia entre os sentidos) e (iv) os fatores de similaridade, que caracterizam as relações de sentido e os tipos de polissemia.

Tomando as *wordnets* como a aplicação alvo e a SLC como referência teórica, esta seção visa a desenvolver a primeira etapa de aplicação da teoria estudada, dedicando-se à caracterização do modelo representacional para a polissemia de nominais: os *polysets*. Propomo-nos a representar fundamentalmente a realidade linguística sincrônica dos nominais polissêmicos do PB. Entretanto, a representação proposta tem *status* múltiplo conforme descreveremos.

Esta quinta Seção está organizada em duas grandes partes: em 5.1, abordaremos *o polyset como um constructo de status múltiplo* e, em 5.2, *os componentes e a estrutura do polyset*.

5.1 O *POLYSET* COMO UM CONSTRUCTO DE *STATUS* MÚLTIPLO

O *polyset*¹⁵ deve ser entendido como um “conjunto de sentidos polissêmicos”. Trata-se de um constructo destinado à representação estruturada do conteúdo semântico associado a itens lexicais polissêmicos. Os critérios propostos para a construção dos *polysets* possibilitam a representação da polissemia de nominais porque (i) representam o sentido em diferentes níveis de generalidade; (ii) explicitam o tipo de relação existente entre os sentidos associados a uma mesma estrutura fonológica; (iii) explicitam diferentes graus de saliência entre os sentidos que formam um conjunto polissêmico. Desse modo, a partir da aplicação dos critérios propostos para a construção dos *polysets*, a base da WN.Br terá representações distintas para itens homônimos e itens polissêmicos (de diferentes tipos). Destacamos, mais uma vez, que a base teórico-metodológica para a proposição do *polyset* é a SLC e a sua estruturação interna é descrita em termos de uma rede de polissemia com enfoque sincrônico nos moldes propostos por Geeraerts (2006) e Blank (2003), por esse tipo de rede ser multidimensional e livre de formato fixo.

Os *polysets* são considerados constructos com *status* múltiplo porque são construídos em termos científico, pois ancoram-se teoricamente na SLC, e tecnológico, pois se prestam ao PLN, precisando, portanto, dar conta dos requisitos e das exigências de um modelo de representação em dois domínios: linguístico-cognitivo e linguístico-computacional.

Para dar conta desses papéis, a representação do *polyset* conforma-se a critérios. Para nos ajudar na formulação desses critérios, tomaremos os propostos por Helbig (2006),¹⁶ para a apresentação do seu robusto modelo de representação do conhecimento baseado em redes e organizado em múltiplas camadas, o MultiNet. O modelo de Helbig pretende representar o conhecimento linguístico como um todo e de maneira independente de língua (nível conceitual) e de aplicação computacional. Tendo em vista que, ao contrário do MultiNet, os *polysets* são construídos para servir a propósitos específicos, descrição linguística e aplicação computacional, nem todos os

¹⁵ A denominação *polyset* é motivada pela denominação *synset*. Enquanto este é um constructo criado para representar um conjunto de itens lexicais que compartilham num mesmo conceito, aquele é um constructo que propomos para representar um conjunto de sentidos que se aproximam por polissemia.

¹⁶ Di Felippo (2008) explorou o modelo MultiNet para a construção de ontologias aplicada à WN.Br.

critérios propostos por Helbig (2006) serão aproveitados. Feita a necessária adaptação, chegamos ao seguinte perfil a que o *polyset* deve atender:

1. Requisitos gerais:

- a) **Adequação cognitiva:** as estratégias representacionais devem possibilitar a modelagem de estruturas conceituais humanas (na medida em que são conhecidas) e sua manifestação na semântica das línguas naturais.
- b) **Interoperabilidade:** os meios de representação devem ser aplicáveis tanto a investigações linguísticas quanto linguístico-computacionais;
- c) **Homogeneidade:** os meios de representação devem ser capazes de expressar os diferentes tipos de sentido polissêmico e, além disso, devem possibilitar a integração da descrição semântica de um item específico à descrição semântica como um todo;
- d) **Reutilização:** os critérios de construção e os constructos resultantes devem ser explícitos e precisos a fim de possibilitar sua reutilização por outras equipes de pesquisadores;
- e) **Viabilidade:** os constructos devem ter potencial de tratamento técnico e implementacional;
- f) **Automaticidade:** os meios descritivos devem permitir o processamento automático do conhecimento.

2. Requisitos internos

- a) **Completeness:** os sentidos polissêmicos instanciados em uma língua natural devem ser representados pelos recursos descritivos previstos pelo modelo;
- b) **Granularidade ótima:** diferentes níveis de granularidade de sentidos devem ser inseridos no modelo;

- c) **Consistência:** deve-se cuidar para que não sejam derivadas informações inadequadas ou logicamente contraditórias a partir de outras inseridas no modelo;¹⁷
- d) **Multidimensionalidade:** os sentidos devem ser descritos de maneira a expressar sua estrutura multidimensional, que é influenciada por diferentes níveis de representação do conteúdo semântico;
- e) **Interpretabilidade local:** os construtos, embora devam ser logicamente interpretáveis por si, de maneira independente de sua inclusão em uma base de conhecimento sobre o léxico da língua, devem ser compatíveis à estrutura geral desse léxico.

É com esse perfil que os componentes e a estrutura do *polyset* serão propostos. Ao final da Seção 7, quando tivermos concluído a apresentação da codificação computacional dos dados, será possível refletirmos sobre a adequação do *polyset* a esse perfil.

5.2 OS COMPONENTES E A ESTRUTURA DO *POLYSET*

Ao longo desta seção passaremos a descrever os componentes e a estrutura do constructo *polyset*. A construção do *polyset* prevê a manipulação dos dois elementos estruturais básicos de uma rede de polissemia: os *nós* e os *arcos*. A construção da rede prevê, por sua vez, a identificação dos itens lexicais polissêmicos e a representação de seus sentidos com base em *cópus*. Os componentes do *polyset* são itens lexicais, conjuntos de sentidos e relações de polissemia, conforme descreveremos a seguir.

¹⁷ Para isso, recomenda-se a aplicação combinada dos testes e das etapas de análise dos sentidos polissêmicos.

5.2.1 Os *nós* e a caracterização dos sentidos polissêmicos

Os *nós* da rede representam os sentidos polissêmicos e os *arcos*, as relações de polissemias que os unem. O *nó* pode instanciar uma unidade simples ou complexa, podendo ser descrito por uma breve glosa, por um hiperônimo, ou por um sinônimo. Para a descrição do conteúdo dos *nós*, identificamos um item lexical polissêmico e analisamos o conteúdo veiculado por ele, com base em *cópus* e em recursos léxicos. Como *cópus*, utilizamos, preferencialmente, a *web* “como *cópus*”, através de buscas simples no sistema Google¹⁸. Em casos especiais, se necessário, recorremos ao *Cópus* NILC¹⁹. Como recursos léxicos, usamos dicionários monolíngues (WEISZFLOG, 1998, WEBSTER’S, 1997), FERREIRA, 1999, HOUAISS, 2001, BORBA, 2002; “Random House Webster’s”, 1994), bilíngues (HOUAISS, CARDIM, 1982), AULETE (2009) e *wordnets* (WN.Br, WN.Pt, WN.Pr).

A seguir, caracterizaremos cada componente da estrutura do *polyset*, são eles: *item lexical*, *constante*, *sentidos* e *conjuntos de sentidos*.

5.2.1.1 O item lexical

No modelo que estamos propondo, os itens lexicais instanciados em uma língua particular são, conforme discutimos na Seção 3, entidades simbólicas que estabelecem o pareamento entre forma (estrutura fonológica) e sentido (estrutura semântica/conceitual). Como já assinalara Saussure, os critérios dessa associação são definidos por padrões de convenção. É preciso lembrar, também, que um item lexical com múltiplos sentidos é aquele que ativa um determinado tipo de rotina cognitiva, de modo que uma mesma estrutura fonológica provê acesso a diferentes porções relacionadas do espaço conceitual. Classificar um item lexical como polissêmico pressupõe, assim, entender que seu conteúdo é formado por sentidos múltiplos que apresentam algum tipo de similaridade.

¹⁸ Disponível em <http://www.google.com.br/>

¹⁹ Disponível em <http://www.linguateca.pt/CETENFolha/>

Os itens lexicais substantivos do PB foram extraídos, fundamentalmente, dos próprios textos teóricos estudados, tanto no domínio linguístico quanto no computacional; poucos foram selecionados diretamente no *cópus*. Fizemos essa opção com o objetivo de testar a validade do modelo representacional proposto para resolver questões discutidas ao longo do trabalho e que são frequentes na literatura sobre o tema.

5.2.1.2 A constante

A denominação *constante* é empregada para descrever a *estrutura fonológica/grafêmica* de um item lexical, por esta ser a porção formal comum aos sentidos que compõem o *polyset*. A constante tem papel fundamental no modelo, visto que o reconhecimento da polissemia passa pela identificação de uma mesma estrutura fonológica/grafêmica (simples ou complexa) associada a uma estrutura semântica complexa, ou seja, a múltiplos sentidos relacionados. Em termos práticos, a constante tem aqui duas funções principais: servir como rótulo ou identificador do *polyset* e como chave de busca nas *wordnets*, nos dicionários e nos *cópus*.

5.2.1.3 Os sentidos

Sob a denominação *sentido polissêmico*, incluímos tanto sentidos plenos quanto subsentidos (facetar e micros sentidos). Como dissemos, eles são os *nós* da rede de polissemia que estrutura o *polyset*. Sem perder de vista os princípios da SLC, consideraremos os sentidos entidades cognitivas convencionalmente associadas aos itens lexicais e construídos no uso que se faz da língua. Dessa maneira, qualquer registro ou representação da semântica lexical adquire aqui o *status* de abstração a partir dos dados e, portanto, um "constructo" (KILGARRIFF, 2006, p. 29).

Conforme mencionamos, os itens lexicais polissêmicos analisados foram extraídos fundamentalmente dos textos estudados. A seleção, por sua vez, se deu em dois momentos: a análise de ocorrências mais ilustrativas de polissemia selecionadas da

literatura e a análise da frequência dessas ocorrências no *cópus* de referência. A identificação dos sentidos foi feita a partir da análise combinada de diferentes fontes. A primeira etapa previu a análise (a) do inventário de sentidos polissêmicos extraídos dos dicionários monolíngues do PB (WEISZFLOG, 1998, WEBSTER'S (1997), FERREIRA, 1999, HOUAISS, 2001, BORBA, 2002); (b) dos *synsets* da WN.Br²⁰ e (c) das informações disponíveis na WN.PT, ambos os recursos são limitados por estarem em construção. A segunda etapa previu a identificação dos itens lexicais do inglês que são equivalentes aos do português. Os itens lexicais do inglês desempenham o papel de chave de busca para a consulta à WN.Pr. Para esse processo, tomamos como referência o dicionário inglês-português português-inglês "Dicionário Eletrônico Webster's" (HOUAISS, CARDIM, 1982) e o Dicionário Inglês-Inglês Eletrônico "Random House Webster's" (1994). O terceiro passo foi a consulta à WN.Pr, incluindo a análise das glosas nela propostas para cada sentido e das relações semânticas nela incluídas. O uso dos dicionários e das bases *wordnets* como fontes de busca dos sentidos se justifica por serem esses registros de usos bem estabelecidos, fornecendo um repertório desejável e controlado de informações para os fins deste trabalho. Por fim, enfatizamos que, concomitantemente às consultas aos dicionários e às *wordnets*, consultas ao *cópus* de referência (Internet via Google e *CópusNilc*) foram feitas para verificarmos os usos de determinado item lexical, para atestarmos os sentidos registrados nos dicionários ou, ainda, para identificarmos possíveis sentidos novos.

A partir dessas fontes de busca de sentidos, identificamos informações semânticas de tipos diferentes: usos reais registrados no *cópus*, diferentes acepções registradas nos dicionários, glosas e relações semânticas codificadas nas *wordnets*. A partir dessa variedade de informações, definimos estratégias para o registro de cada sentido analisado. Nessa análise, a opção foi pela clareza, objetividade, precisão, além da praticidade. O conteúdo dos *nós*, portanto, pode ser registrado a partir de um dos seguintes elementos:

²⁰ Trata-se da base derivada do *Thesaurus eletrônico do Português*: um dicionário eletrônico de sinônimos e antônimos formado por cerca de 44 mil unidades lexicais (Dias-Da-Silva, 2003a; Dias-Da-Silva e Moraes 2003).

(a) Uma glosa reduzida

A glosa da rede de polissemia deve ser entendida como uma “adaptação reduzida” da tradução livre da glosa da WN.Pr²¹, combinada com a definição dos dicionários consultados. A construção da glosa que propomos para as unidades de sentido da rede de polissemia observa o seguinte critério: usar o número mínimo de itens lexicais necessários para expressá-la. Por exemplo, para o *polyset*, a glosa da WN.Pr exemplificada na nota 7 é reduzida para “objeto físico”. Como a informação da glosa não é manipulável pelo sistema computacional, julgamos que essa redução não compromete o desempenho do sistema e é suficiente para auxiliar o analista a identificar o sentido por ela codificado.

(b) Um hiperônimo suficientemente específico

A informação do hiperônimo de um item lexical orienta a análise de delimitação do sentido lexical. Por essa razão, ele é usado, como conteúdo do *nó* da rede, quando ele contribuir de modo mais preciso e informativo para a delimitação do sentido. Por exemplo, um dos sentidos do item lexical *base* registrado nos dicionários e na WN.Pr é assim especificado: “tudo o que serve de apoio ou suporte”; o seu hiperônimo direto na WN.Pr, em PB, é “suporte”. Notamos que o hiperônimo identifica, em apenas uma palavra, o que a especificação identifica por meio de uma frase. Logo, nesse exemplo, a opção é pelo hiperônimo.

(c) Um sinônimo representativo

Um sinônimo é usado para preencher o *nó* da rede de polissemia quando ele puder ser substituído, com o mínimo de alteração semântica, em um número grande de contextos²². Por exemplo, uma das glosas da WN.Pr associadas ao item lexical *posição* é “the act of putting something in a certain place or location” e o *synset*

²¹ Nas *wordnets*, a glosa é entendida como uma definição²¹ informal que representa o sentido codificado no *synset*, ou seja, o *nó* da rede. Um exemplo de glosa da WN.Pr é, para o *synset* {*book, volume*}, “physical objects consisting of a number of pages bound together”.

²² É importante que se diga que não se pretende com isso considerar que o sinônimo escolhido é absoluto, fenômeno raro ou inexistente.

correspondente é {placement, location, locating, position, positioning, emplacement}. Nesse exemplo, usar *localização*, tradução de *location*, componente do *synset*, é suficiente para descrever o conteúdo do *polyset*.

(d) Uma noção geral

Um elemento que usamos para representar o registro do sentido de um item lexical que exhibe polissemia irregular é uma *noção geral*, que codifica um conteúdo semântico esquemático e não tem o mesmo *status* dos mencionados acima, pois serve como uma noção globalizante que representa a similaridade entre os sentidos e não os sentidos propriamente ditos. Lembramos que esse tipo de noção, embora generalizante, não tem *status* de hiperônimo e deve permitir a delimitação de grupos de sentidos. Para a identificação da noção geral, consideraremos as informações fornecidas pelos dicionários, pelas *wordnets*, pelo corpus, além do conhecimento do analista sobre usos convencionalmente associados ao item em análise. Esse tipo de informação tem a função de rotular um agrupamento de sentidos, explicitando o tipo de similaridade percebida entre eles e serve, portanto, como elemento auxiliar na análise e estruturação do *polyset*. O exemplo citado ao longo do trabalho foi a noção de “custódia”. Essa noção indica um grupo de sentidos associados a *banco*. Já as noções de “entidade, coisa ou local”, visto que não são capazes de distinguir o que é um *banco*, não são adequadas.

Para avaliarmos o impacto quantitativo do modelo de *polysets* na organização da WN.Br, fez-se uma filtragem automática dos dados disponíveis na base²³. A estatística mostrou que de um total de 16.925 entradas para nominais, 5.957 são codificados como polissêmicos, segundo a noção grosseira de polissemia adotada pelos idealizadores da base. Esses números são importantes para revelar o impacto quantitativo do modelo proposto para a organização geral da base.

Este nível inicial de análise permitiu que construíssemos o primeiro inventário dos múltiplos sentidos associados a um item lexical. Nesta fase, situações de monosssemia (ex.: criança, tio, meia-idade) foram desconsideradas. Esse primeiro

²³ A filtragem dos dados da base da WN.Br tem sido feita pela equipe do NILC-SC, grupo parceiro do projeto de construção da WN.Br.

inventário exigiu outra filtragem: eliminação dos itens lexicais homônimos, tarefa que seguiu esta análise inicial.

5.2.1.4 Um conjunto de sentidos

Identificados os sentidos que marcam os *nós* da rede de polissemia, o próximo passo é construir o *conjunto de sentidos* de acordo com a percepção de alguma similaridade. Os múltiplos sentidos delimitados na análise inicial constituem apenas uma lista (estrutura homonímica). A partir dessa lista, tendo em vista que a estratégia polissêmica é a preferencial em um *polyset*, em razão das diversas vantagens computacionais e cognitivas observadas ao longo do trabalho, o objetivo é especificar os fatores de coerência que se estabelecem entre os sentidos. Sempre que eles forem especificados, constrói-se um conjunto de sentidos, o *polyset* propriamente dito. A partir desse segundo nível de construção do *polyset*, temos a indicação de alguma similaridade, mas ainda não se sabe com precisão de que tipo ela é.

Podemos assim dizer que os *polysets* complementam as informações associadas aos *synsets*, refinando-os semanticamente e são, ao mesmo tempo, mais livres em sua estrutura, uma vez que o sentido pode ser composto de elementos que fazem parte de níveis distintos de generalidade: podemos ter uma glosa, um hiperônimo ou um sinônimo. Um exemplo de diferenças entre os componentes do *polyset* é dado na comparação dos conjuntos de sentidos que identificamos para os itens lexicais *livro*, enquanto ‘objeto físico’ e ‘conteúdo’ (glosa simplificada) e *posição*, enquanto ‘localização’ e ‘opinião’ (sinônimos). O processo de construção dos agrupamentos de *nós* da rede previu a consulta a todas as fontes de sentidos mencionadas acima e levou, também, em consideração os fatores de coerência e de autonomia, além do tipo de relação de sentido e dos efeitos de saliência.

5.2.1.5 A construção dos *polysets*: a delimitação dos tipos sentidos

A tarefa de delimitação dos tipos sentidos constitui a primeira fase específica na construção do *polyset*. Nas duas etapas anteriores, a tarefa foi a de organizar e registrar as informações disponibilizadas em diferentes fontes. Assim, a partir deste ponto é que passamos a apresentar a proposta de representação dos diferentes tipos de sentido identificados nas seções 3 e 4.

Entre os sentidos polissêmicos, na seção 4, identificamos sentidos semi-distintos e sentidos cooperativos, conforme o grau de delimitação. Os sentidos semi-distintos, os irregulamente polissêmicos, são os sentidos plenos que têm um grau médio de antagonismo e um grau significativo de autonomia. Os sentidos cooperativos apresentam certo grau de autonomia, com um grau de antagonismo menor ou inexistente. Conforme são delineadas as fronteiras entre os diferentes sentido, as diferentes leituras podem ser classificadas como *sentidos plenos* ou subsentidos do tipo *facetar* e *microsentidos*. Sob o rótulo *sentido pleno*, incluímos aqueles que apresentam distinções marcantes, bem estabelecidos, incrustados no léxico da língua e fortemente restritos por restrições convencionais estáveis (CROFT e CRUSE, 2004). Sob o rótulo de *subsentidos*, incluímos aqueles que apresentam grau de antagonismo baixo ou inexistente. Há, portanto, a possibilidade de unificação, que pode ser de diferentes tipos: enquanto as facetar são entendidas como partes de um mesmo sentido geral, os microsentidos são entendidos como especializações desse sentido geral.

Feitas essas contextualizações, passamos ao foco da seção: inventariar, no quadro 6, os fatores de coerência, ou seja, as evidências de similaridade, e os fatores de discretude, ou seja, as evidências de autonomia entre os sentidos. Para o inventário apresentado aqui, selecionamos aqueles critérios da Seção 4 que julgamos os mais relevantes para a construção de um *polyset*.

TIPOS DE SENTIDO	EVIDÊNCIAS DE COERÊNCIA	EVIDÊNCIAS DE AUTONOMIA
Sentidos plenos	1. Presença da noção geral 2. Ocorrência marginal do efeito zeugmático	(1) Grau de seleção de atenção máximo (Como o leitor é obrigado a selecionar apenas um dos sentidos no âmbito do enunciado, a delimitação entre eles ocorre já no contexto linguístico sentencial) (2) Possibilidade de identificação de um sentido mais neutro e outro menos neutro, aplicável a contextos mais restritos (3) Presença de autonomia relacional

		(4) Presença de autonomia composicional
Subsentidos do tipo facetas	(1) Capacidade de unificação (as facetas do sentido são partes de um mesmo sentido) (2) Inexistência de antagonismo (3) Grau de autonomia marginal. (4) Grau de seleção de atenção marginal. (5) Ocorrência marginal do efeito zeugmático. (6) Coocorrência de leitura prototípica. (7) Presença de propriedades composicionais unificadoras	(1) Presença de autonomia relacional (as facetas de sentido têm hiperônimos distintos a observação sobre o tipo de hiperônimo possibilita a percepção de classes regulares de polissemia) (2) Presença de núcleos autônomos, isto é, ocorrência de independência na ativação das facetas (3) Grau de seleção de atenção marginal (não há necessidade de seleção de apenas uma das facetas)
Subsentidos do tipo microsentidos	(1) Presença de hiperônimo comum (importante fator unificador) (2) Grau mínimo de antagonismo (os microsentidos são co-autohipônimos)	(1) Microsentidos em relação de incompatibilidade mútua (co-autohipônimos). (2) Presença de autonomia relacional (3) Presença da autonomia de condições de verdade. (4) Presença de autonomia de domínio.

Quadro 6 - Evidências de coerência e de autonomia

5.2.1.6 Efeitos de saliência entre os sentidos

A análise dos efeitos de saliência entre os sentidos constitui a segunda fase de construção de um *polyset*. Sua identificação precisa, no entanto, somente é concluída a partir da terceira atividade, quando se os rotulam arcos. Aqui, destacamos duas questões fundamentais: a identificação do núcleo da categoria e a identificação dos efeitos de saliência, tanto a psicológica como a estrutural.

No contexto das discussões teóricas das seções 3 e 4, entende-se que, de acordo com os critérios semânticos sincrônicos de delimitação do núcleo de uma categoria de sentidos, podemos ter os seguintes cenários típicos: (a) sentido único – quando há um sentido saliente que se sobrepõe aos outros, caracterizando a polissemia irregular; (b) sentidos duplos e independentes – quando a independência é parcial, caracteriza-se a polissemia irregular ou a polissemia regular do tipo autohiponímia; (c) sentidos duplos e sobrepostos – quando há a ocorrência de um modo independente e caracterizando a polissemia regular do tipo automeronímia.

Para a identificação do núcleo de uma categoria de sentidos representada em termos de um *polyset*, consideraremos dois critérios fundamentais: a predominância na rede de polissemia e a frequência de ocorrência em *cópus*, conforme descreveremos.

5.2.1.4.1. Saliência psicológica

A detecção do efeito da saliência psicológica funcional para a delimitação do núcleo de uma categoria de sentido pressupõe a análise do(s) sentido(s) que ocorrem(m) com um número maior de ligações. O critério é considerado válido por duas razões. Entende-se que ele é o sentido cognitivamente mais saliente e, portanto, o que tem maior vantagem psicológica, por facilitar o acesso aos demais. Embora a primeira razão seja suficiente para os propósitos da análise sincrônica, é relevante mencionar que o critério é também válido para a identificação dos sentidos primários em abordagens que fazem interface entre a sincronia e a diacronia, mais comuns na SLC. O fato de o sentido ocorrer com um número maior de relações (os *links*) pode ser indicativo de que os seus componentes estão mais frequentemente presentes em outros sentidos.

5.2.1.4.2 Saliência estrutural

O efeito de saliência estrutural, ou de distinção, está relacionado à frequência de ocorrência de um sentido em relação a outro. Para a contagem do peso estrutural de um sentido integrante de um *polyset*, é feita uma busca pelo item lexical sob análise na web como cópula, analisando-se as 10 primeiras ocorrências trazidas pelo motor de busca do Google™. Esse exercício tem o objetivo de ilustrar a metodologia, visto que o número de ocorrências analisadas para cada item é pequeno. É importante notar também que o resultado das buscas é subordinado à ocorrência do sentido entre as páginas mais consultadas.

5.2.2 Os arcos da rede de sentidos polissêmicos

Os *arcos* de uma rede de sentidos polissêmicos representam as *relações* entre os diferentes sentidos polissêmicos. São as relações as responsáveis por dar forma e por justificar o agrupamento dos sentidos em um *polyset*. É somente após a conclusão dessa fase de análise que o *polyset* é caracterizado por completo. De maneira análoga ao que fizemos na apresentação da teoria na Seção 4, o processo de especificação das relações na modelagem dos dados prevê duas fases. A primeira fase, descrita anteriormente, é mais geral, e consiste na detecção da existência ou não de relação entre os sentidos para

a composição do conjunto de sentidos. A segunda fase é mais específica, e consiste na explicitação do tipo de relação que rotula os arcos. O tipo de codificação da relação determina (a) o tipo de grau de generalidade ou de especificidade entre os sentidos, (b) o sentido mais saliente da rede (isto é, aquele que tem maior peso estrutural por ocorrer com um número maior de *links*) e (c) o grau de produtividade da combinação de sentidos no léxico como um todo. Nas subseções seguintes, descreveremos as propriedades essenciais de cada uma das relações já discutidas na Seção 4.

5.2.2.1 A relação de polissemia regular

Há sentidos que são relacionados pela polissemia regular do tipo *automeronímia/autoholonímia* (relação de PARTE-TODO). A relação ocorre quando os subsentidos (as facetas) associam-se por relação de *parte-todo* (ex.: porção/substância, objeto físico/conteúdo). A similaridade entre as facetas de sentido é alta, de maneira que ambas as facetas são componente distintos de um mesmo sentido geral. Como a principal característica das facetas de sentido é a de elas apresentam grau mínimo de autonomia, não é necessário delimitar apenas uma delas em um contexto, embora seja comum que a cada contexto uma esteja em relevo. Dos exemplos que referimos ao longo do trabalho, destacamos *livro* ‘objeto físico’ e *livro* ‘conteúdo’. O rótulo da relação é *é_automerônimo_de* e o do seu reverso é *é_autoholônimo_de*. Para a identificação da relação regular de polissemia do tipo 2, a automeronímia, a aplicação da seguinte fórmula lógica precisa ser verdadeira: A é *parte de* B, e A e B são facetas de sentido. A regularidade da relação se dá na medida em que podemos perceber outros pares de sentidos relacionados também por automeronímia. Lembramos que a automeronímia abarca uma família de relações que não constituem partes em sentido estrito, tais como *objeto-substância*, *continente-conteúdo*, *objeto físico-conteúdo*, etc. O rótulo da relação é *é_automeronímia_de* e o do seu reverso é *é_autoholonímia_de*.

Há os sentidos que são relacionados pela polissemia regular do tipo *autohiponímia/autosuperordenação* (relação de *É-UM*). Essa relação ocorre quando um dos subsentidos (microsentidos) é construído de modo mais específico em relação aos outros. A similaridade entre os microsentidos sentidos é funcional, de maneira que o

microsentido subordinado é um subtipo do superordenado e herda todas as suas características, acrescentando, porém, outras que o distinguem dos seus co-autohipônimos. Como ocorre em relações hierárquicas, os co-autohipônimos em um mesmo nível hierárquico são mutuamente incompatíveis. Dos exemplos que mencionamos ao longo do trabalho, destacamos *faca* como ‘utensílio doméstico’, ‘arma branca’ e ‘instrumento cirúrgico’. O rótulo da relação é *é_automerônimo_de* o do seu reverso é *é_autosuperordenado_de*. Para a identificação da relação regular de polissemia do tipo 1, a autohiponímia, a aplicação da seguinte frase lógica precisa ser verdadeira: *A é um B*, ou *A é um tipo de B*, e *A* e *B* são microsentidos. A regularidade da relação se dá na medida em que detectamos outros pares de sentidos relacionados também por autohiponímia.

5.2.2.2 A relação de polissemia irregular

Há sentidos com alto grau de independência uns dos outros, mas, mesmo assim, é possível que se perceba alguma relação de similaridade (objetiva ou subjetiva) entre eles. Ao contrário das relações que se detectam na polissemia regular, trata-se de uma relação horizontal, não havendo herança entre os sentidos. O que ocorre é uma comparação entre os sentidos de modo que podemos dizer que há algum grau de similaridade entre eles. Esse tipo de similaridade caracteriza a polissemia irregular, situação em que dois ou mais sentidos, com grau médio de antagonismo e significativo grau de autonomia, associam-se de algum modo a uma mesma “noção geral”. Entre os exemplos apresentados ao longo do trabalho, destacamos *cabeça* ‘membro do corpo humano’ e *cabeça* ‘chefe’ e *posição* ‘local ocupado’ e *posição* ‘opinião’. O rótulo para essa relação é *é-similar-a*. A aplicação da seguinte frase lógica precisa ser verdadeira: Um item lexical *A* com os sentidos a_1 e a_2 apresenta polissemia irregular se **NÃO** existir nenhuma outra unidade lexical *B* com sentidos b_1 e b_2 que sejam semanticamente distintos um do outro exatamente de maneira diferente da distinção entre a_1 e a_2 .

Para sintetizar a seção 5.2.2, apresentamos o quadro 7, que resume as propriedades centrais de cada relação.

PROPRIEDADES	POLISSEMIA REGULAR		POLISSEMIA IRREGULAR
	Rótulo		Rótulo
	<i>é automerônimo_de</i> ou <i>é autoholônimo_de</i>	<i>é autohipônimo_de</i> ou <i>é autosuperordenado_de</i>	<i>é similar_a</i>
Tipo de sentido	Faceta	Microsentido	Pleno
Tipo de Relação	Regular	Regular	Irregular
Tipo de similaridade	Hierárquica Relação entre parte-todo e variantes	Hierárquica Relação entre geral e específico	Horizontal Relação objetiva ou subjetiva a uma noção geral
Grau de similaridade	Alto	Funcional	Intermediário
Grau de autonomia	Mínimo	Incompatibilidade mútua de microsentidos (co-autohipônimos)	Alto
Grau de antagonismo	Inexistente	Mínimo	Médio
Fórmula lógica	[A <i>é parte de</i> B, e A e B são facetas de sentido]	[A <i>é um</i> B, ou A <i>é um tipo de</i> B, e A e B são microsentidos]	Um item lexical A com os sentidos a1 e a2 apresenta polissemia irregular se NÃO existir nenhuma outra unidade lexical B com sentidos b1 e b2 que sejam semanticamente distintos um do outro exatamente de maneira diferente da distinção entre a1 e a2

Quadro 7 – Propriedades centrais da relação de polissemia.

5.2.3 Síntese da seção

Ao longo desta seção, caracterizamos cada componente de um *polyset*, apresentando sua função e funcionamento na ao longo dos três etapas gerais de construção do modelo. O quadro 8 resume as principais etapas de sua construção

A CONSTRUÇÃO DO <i>POLYSET</i>			
	Etapas preliminares		Etapas de construção
	Definição dos <i>nós</i> da rede	Indicação dos conjuntos de sentidos	Construção da rede de polissemia
1 Item lexical polissêmico	Identificar	x	x
1.2 Constante	Identificar	x	x
1.3 Sentidos polissêmicos	Delimitar potencial de sentidos e <i>status</i>	Analisar	Analisar
1.3.1 glosas	Construir	Analisar	Analisar
1.3.2 noção geral	Identificar	Analisar	Analisar
1.3.3 <i>cópus</i>	Selecionar e analisar	Analisar	Analisar
1.3.4 <i>wordnets</i>	Consultar	Consultar	Analisar
1.3.5 dicionários	Consultar	Consultar	Analisar
2. Grupos de sentidos	x	Formar	Rotular
3. <i>Status</i> de delimitação	Analisar	Analisar	Especificar
3.1 Sentido pleno	x	Analisar	Identificar
3.2 Subsentido	x	x	x
3.2.1 Faceta	x	Analisar	Identificar
3.2.1 Microsentido	x	Analisar	Identificar
4. Efeitos de saliência	x	Analisar	x
4.1 Psicológica	x	x	Identificar
4.2 Estrutural	x	x	Identificar
5. Tipo de relação de sentido	x	Identificar a existência	Especificar
REGULAR	x	x	Indicar
5.1 automeronímia	x	x	Indicar
5.2 autohiponímia	x	x	Indicar
IRREGULAR	x	x	Indicar
5.3 similaridade	x	x	Indicar

Quadro 8 – Níveis gerais de estruturação de um *polyset*

Conforme descreve o quadro 8, um *polyset* é construído em dois grandes níveis: etapas preliminares de construção do *polyset* e etapas de construção do *polyset*. É importante destacar que ambos os momentos têm um funcionamento integrado e mutuamente cooperativo. Em síntese, as tarefas desenvolvem-se conforme a descrição que segue.

As **etapas preliminares de construção do *polyset*** têm dois objetivos gerais: definir os *nós* da rede e indicar os conjuntos de sentidos. Para determinar os *nós* que comporão a rede de polissemia que estrutura o *polyset*, é necessária a realização das seguintes tarefas: (i) a identificação do item lexical polissêmico no PB; (ii) a delimitação do seu potencial de sentido (os *nós* da rede). Como exemplo, para o item lexical *livro*, selecionamos os sentidos *livro* ‘objeto físico’ e *livro* ‘conteúdo’. Para indicar os conjuntos de sentidos relacionados, com base nas informações delimitadas no nível anterior e nas informações fornecidas pela WN.Pr, são construídos, para cada sentido identificado,

conjuntos de sentidos sem a indicação ainda do tipo de relação. Cada conjunto de sentidos é identificado por uma constante que representa o item lexical polissêmico em análise. Essa análise é útil não só para a tarefa interna ao processo de construção de *polyset*, mas também para a indicação de uma possível integração entre a análise semasiológica do *polyset* e a análise onomasiológica do *synset*, pois o sentido do item lexical em questão precisará ser expresso linguisticamente através de um sinônimo, de um hiperônimo ou de uma glosa simplificada. Voltando ao exemplo de *livro*, o conjunto de sentidos pode ser representado, do seguinte modo $\{\{\text{livro, conteúdo}\}, \{\text{livro, objeto físico}\}, \}$ _{livro}.

As **etapas de construção do *polyset***. Quando chegamos nesse nível, já construímos os *nós* (sentidos) da rede e os agrupamentos entre os sentidos (conjuntos de sentidos). Falta-nos identificar as seguintes informações: (i) o *status* dos sentidos, percebendo (a) a ocorrência de sentido pleno, de faceta ou de microsentido e (b) as diferenças de saliência, e (ii) os rótulos das relações, explicitando o motivo dos agrupamentos. Será neste nível que a rede de polissemia tomará forma, sendo enriquecida com os *nós* e *arcos* rotulados. Os rótulos com os quais trabalhamos são *é-similar-a*, *é_autohipônimo_de*, *é-autosuperordenado-de*, *é-automerônimo-de*, *é-autoholônimo-de*. Além disso, são também indicadas as propriedades regular ou irregular para cada relação. Em formato gráfico, o exemplo de *livro* é ilustrado na figura 19.

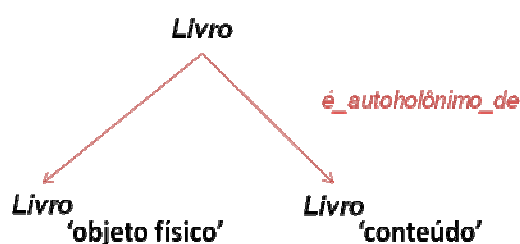


Figura 19 - Representação gráfica dos automerônimos de *livro*

SEÇÃO 6 – CONSTRUÇÃO E IMPLEMENTAÇÃO DO *POLYSET*

Na seção anterior, apresentamos o constructo *polyset* como um modelo de análise e representação linguístico computacional da polissemia de nominais que tem a finalidade de servir à complementação das informações léxico-conceituais disponíveis em uma *wordnet*. O objetivo desta seção é apresentar um exercício de aplicação dessa proposta, relatando os processos desenvolvidos para a construção e implementação do *polyset*. Sendo assim, dividimos a seção em três momentos principais: no primeiro, apresentamos um exercício de análise e representação linguística; no segundo, um exercício computacional; e, no terceiro, destacaremos as principais contribuições do modelo *polyset* para as *wordnets*, em especial, para a WN.Br. Lembramos que a descrição proposta tem *status* múltiplo: o de representação linguística, pois caracteriza um fenômeno que se manifesta na superfície da língua; o de representação linguístico-computacional, pois se configura como uma descrição formalizável das informações linguísticas; e, ainda, o de representação linguístico-cognitiva, pois tem o *status* de “isomorfismo funcional”, entendendo-se que a realidade psicológica de uma representação não é imediata, mas pode ser representativa do que temos na mente (SANDRA e RICE, 1995, p. 102).

Os itens lexicais em análise serão *fonte*, *banco* e *café*. Os exemplos foram escolhidos por (a) serem recorrente na literatura, recebendo tratamentos diversos, (b) por ilustrarem as relações de polissemia estudadas e, sobretudo, (c) por gerarem diferentes configurações para a rede que estrutura o *polyset*. Os sentidos que recheiam os nós da rede de polissemia proposta para cada item lexical não têm o objetivo de esgotar suas possibilidades de uso, mas o de ilustrar o potencial representacional do *polyset*, a partir de sentidos registrados nos dicionários e *wordnets* pesquisados que tiveram ocorrência no corpus. A partir desse critério, entendemos que o *polyset* registrará sentidos convencionalizados e relevantes sincronicamente. Para a delimitação e registro dos sentidos, conforme observamos na Seção 5, levamos em consideração as acepções dos recursos léxicos analisados. A descrição dos resultados da etapa linguística foi organizada em quadros síntese, registrando as *etapas preliminares*, e em *redes de polissemia*, representando a configuração do *polyset* para cada um dos itens lexicais

investigados. Os resultados da etapa de implementação computacional serão apresentados a partir das telas do editor Protégé-OWL.

A seção está organizada do seguinte modo: em 6.1, apresentaremos a *face linguística de construção do polyset*; em 6.2, o foco é a *a face computacional de construção do polyset*; em 6.3, são sintetizadas as *contribuições do modelo polyset para a WN.Br*; por fim, em 6.4, sintetizaremos a seção.

6.1 A FACE LINGUÍSTICA DE CONSTRUÇÃO DO *POLYSET*

O objetivo desta seção é descrever, de acordo com o ferramental teórico delimitado nas Seções 3 e 4, as ocorrências de polissemia em nominais do português. Trata-se da segunda etapa do domínio linguístico desta tese responsável pela análise e representação dos dados. Essa tarefa é realizada em dois momentos, conforme referimos anteriormente. O primeiro deles é explicitado no quadro 9, que serve como modelo para a descrição das etapas preliminares de construção do *polyset*, que buscam a definição dos *nós* da rede de polissemia e a indicação dos conjuntos de sentidos.

DEFINIÇÃO DO NÓ DE UM <i>POLYSET</i>	
1. Constante	Indicar a constante
2. Exemplo	Indicar um exemplo para cada sentido em análise
3. Chave de busca em inglês	Indicar a chave de busca correspondente a cada sentido
4. Glosa reduzida	Indicar (fonte: <i>wordnets</i> e dicionários)
5. Hiperônimo relevante	Entre as possíveis traduções dos hiperônimos fornecidos pela WN.Pr, indicar o hiperônimo mais representativo, quando relevante. O hiperônimo representará o nó da rede somente se for considerado o componente mais relevante para descrever o sentido em análise e suficientemente específico.
6. Sinônimo relevante	Entre as possíveis traduções dos sinônimos fornecidos pela WN.Pr, indicar o sinônimo mais representativo, quando relevante. O sinônimo representará o nó da rede somente se for considerado o componente mais representativo do sentido em análise.
7. Nó do <i>polyset</i>	O conteúdo do nó da rede de polissemia deve servir como referência ao analista humano, que contará também com a frase-exemplo para facilitar/complementar sua delimitação. Seu conteúdo pode ser composto por: a. um sinônimo representativo b. um hiperônimo suficientemente específico c. uma glosa reduzida
8. Conjunto de sentidos	Agrupar os sentidos em conjuntos rotuladas pela <i>constante</i> .

Quadro 9 - Resultado das etapas preliminares de construção do *polyset*

Concluídas as etapas preliminares, passamos ao segundo momento da análise e representação linguística mencionado anteriormente. Elaboraremos, como resultado das etapas de construção do *polyset*, uma rede de sentidos polissêmicos, representando graficamente as diferentes relações de polissemia e os efeitos de saliência entre os sentidos. Lembramos que o formato da rede não segue nenhum padrão pré-definido (tal como as redes radiais e esquemáticas analisadas na Seção 3), pois, conforme mencionamos na seção anterior, ela é inspirada nos modelos de Geeraerts (2006) e Blank (2003). O quadro 10 apresenta uma síntese das tarefas dessa fase.

CONSTRUÇÃO DA REDE DE POLISSEMIA	
1. Organização da rede	Nesta etapa é construída uma rede manual que ilustra a distribuição dos sentidos.
2. Rótulo das relações	Os rótulos com os quais trabalharemos são <i>é-similar-a</i> , <i>é-autohipônimo-de</i> , <i>é-autosuperordenado-de</i> , <i>é-automerônimo-de</i> , <i>é-autoholônimo-de</i> .
3. Noção geral	*Componente identificado somente em itens irregularmente polissêmicos. Ele serve como especificação do tipo de similaridade.

4. Status dos sentidos I	Os tipos são: sentido pleno, faceta, microsentido. O tipo de sentido ficará evidente no modelo a partir da organização hierárquica e das relações entre sentidos.
5. Status dos sentidos II	
a. saliência psicológica	Determinada pela predominância.
b. saliência estrutural	Frequência no <i>corpus</i> . Lembramos que, para essa análise, fizemos um consulta aos primeiros 10 resultados de uma busca no Google. Tarefa que serve como ilustração da saliência, mas que é quantitativamente limitada.

Quadro 10 - Resultado das etapas de construção do *polyset*

6.1.1 Análise do item lexical fonte

A DEFINIÇÃO DOS NÓS	
1. Constante	Fonte
2. Exemplos	1. O aproveitamento de uma fonte de água requer o estudo geológico completo de toda a área, no qual se deve levar em conta as condições de poluição dos arredores, da área de captação e da área de infiltração, e o isolamento da fonte, para evitar qualquer contaminação bacteriana.²⁴ 2. A presente fonte ornamental cibernética é provida de um número adequado de bocais variados de saída de água constando de meios ejetores de água, meios reguladores de caudal e meios de pilotagem e ligação elétrica com um microprocessador de controle individual dos bocais e complementada por uma série de elementos decorativos e espetaculares complementares coordenados entre si, facilmente aplicável às fontes instaladas em praças, parques e outros sítios com fins decorativos.²⁵ 3. As formigas invadem o videocassete e vão fazer ninho na fonte, peça que engloba o transformador de voltagem.²⁶ 4. Os exemplos mostram que a multiplicidade de fontes científicas que a web oferece representa uma dificuldade para os usuários, quando eles têm de avaliar e selecionar as informações.²⁷ 5. "A experiência brasileira com biocombustíveis e fontes renováveis de energia mais limpas e baratas abre a perspectiva de mais emprego e renda no campo e é nossa resposta ao desafio da mudança do clima", afirmou Lula a Koroma, primeiro líder de Serra Leoa a visitar o Brasil²⁸. 6. A televisão continua a ser a principal fonte de informação para a maioria dos europeus, mesmo tendo se tornado menos informativa e mais sensacionalista.²⁹
1. Chaves de busca em inglês	1-2 <i>fountain</i> 3-6. <i>source</i> *7. <i>font (typeface)</i>
2. Hiperônimos relevantes	1. formação geológica 2. estrutura, construção 3. - 4. publicação 5. -

²⁴ Fonte: <http://www.emdiv.com.br/pt/mundo/asmaravilhas/1364-fonte-nascente.html>

²⁵ Fonte: www.patentesonline.com.br/fonte-ornamental-cibernetica-77535.html

²⁶ Fonte: CórpusNilc.

²⁷ Fonte: www.comciencia.br/reportagens/cultura/cultura10.shtml

²⁸ Fonte: <http://g1.globo.com/Noticias/Mundo/0,MUL1273371-5602,00-LULA+OFERECE+AJUDA+A+SERRA+LEOA+PARA+PRODUZIR+BIOCOMBUSTIVEIS.html>

²⁹ Fonte: <http://www.observatoriodaimprensa.com.br/artigos.asp?cod=351MON005>

	6. comunicador
3.Sinônimos relevantes	1. nascente 2. chafariz 3. cabo de energia 4. bibliografia, documentação 5. origem, causa 6. informante
4.Glosas WN.Pr/acepções de dicionários (relevantes)	1. Local em que brota a água proveniente do solo (AULETE, 2009) 2. Instalação para aproveitar a água de nascente natural, ou mesmo água encanada. (HOUAISS, 2001) 3. Circuito de eletricidade (BORBA, 2002) 4. Texto de apoio ou referência. (BORBA, 2002) 5. Aquilo que dá origem, matriz, procedência. (combinação de sentidos apresentados no dicionário Houaiss, 2001) 6. Pessoa que fornece informações a repórteres. (HOUAISS, 2001)
5.Nós do <i>polyset</i>_{fonte}	1. Nascente de água 2.Instalação para aproveitar a água de nascente natural, ou mesmo água encanada 3. Circuito de eletricidade 4. Texto de referência 5. Origem, causa, procedência 6. Que ou quem transmite informações
6.Conjuntos de sentidos	$\{ \text{['nascente de água', 'instalação para aproveitar a água de nascente natural, ou mesmo água encanada']}_{1-2}, \text{['origem', causa, procedência, 'circuito de eletricidade', 'texto de referência']}_{3-6} \}_{\text{fonte}}$

Quadro 11 – Resultado da etapa preliminar de construção do *polyset*_{fonte}

6.1.1.2 ETAPAS DE CONSTRUÇÃO DO *POLYSET*_{fonte}:

A rede de polissemia para os sentidos associados a *fonte* está representada na figura 20.

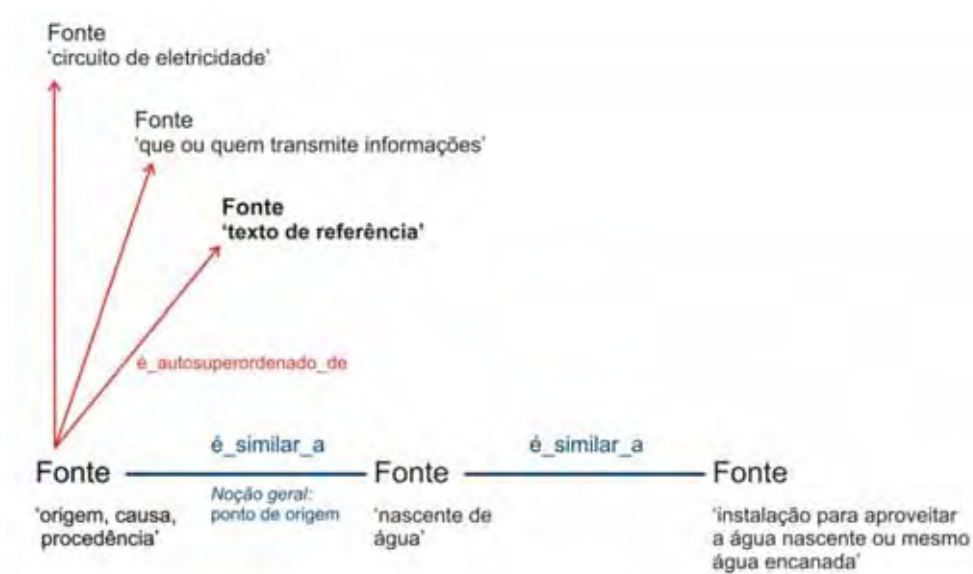


Figura 20 - Polysset fonte

Os sentidos associados a *fonte* selecionados pelos exemplos (1-6) ilustram uma situação em que uma mesma unidade fonológica provê acesso a um conteúdo conceitual delimitado de formas distintas. Conforme representado na figura 20, os sentidos estão associados tanto por polissemia irregular (em azul) quanto e por polissemia regular (em vermelho). A relação de polissemia irregular, expressa através do rótulo *é_similar_a*, caracteriza uma aproximação subjetiva entre 'nascente de água', 'origem, causa, procedência' e 'instalação para aproveitar a água de nascente natural, ou mesmo água encanada'. A similaridade que percebemos entre esses sentidos plenos é subjetiva e pode ser descrita por meio da noção geral *ponto de origem*: no exemplo (1), tem-se ilustrado um ponto natural de origem de água, no (2), o foco é um ponto artificial de origem de água e, no (5), salienta-se a origem de energias renováveis. Ilustrando a relação de polissemia regular, vemos o nó 'origem, causa, procedência' que pode funcionar como conteúdo esquemático para uma série de leituras percebidas como microsentidos por se encaixarem nos critérios de Croft & Cruse (2004) para a identificação de uma relação entre sentido especializado e sentido geral. Entre os critérios, conforme mostramos no quadro 5, destacamos o grau significativo de autonomia, a seleção de atenção e a possibilidade de unificação dos subsentidos como subespecificação de outro. Nos exemplos em análise, em (3), o foco é a origem da energia elétrica, em (4), a origem do conteúdo científico; em (6), a origem de informações sobre

atualidade. Sendo assim, o nó que representa o exemplo (5) estabelece com os nós (3), (4) e (6) uma relação de autosuperordenação.

Em negrito na figura estão os sentidos de *fonte* como ‘origem, causa, procedência’ e ‘texto de referência’. O primeiro é o que apresenta maior predominância na rede, portanto, o que tem maior saliência psicológica. O segundo é o que tem maior saliência estrutural, considerando os resultados da busca no Google.

Além dos sentidos compõem o *polyset*_{fonte}, entre os sentidos mais frequentes no corpus associados à unidade fonológica *fonte*, está ‘conjunto de caracteres do mesmo tipo e tamanho’ (BORBA, 2002), em contextos como *preparamos uma coleção de fontes de computador para fazer convite de casamento*³⁰. Conforme demonstra a figura 20, no entanto, esse sentido ficou de fora do *polyset* por não conseguirmos construir nenhuma relação de similaridade com os demais.

6.1.2 Análise do item lexical *banco*

A DEFINIÇÃO DOS NÓS	
1. Constante	Banco
2. Exemplos	1a. “Ao digitar sua senha do <i>banco</i> para confirmar a operação, o Itaú Shoptline se comunica com a operadora do cartão que autoriza o pagamento ao estabelecimento comercial onde a compra foi feita.”³¹ 1b. “Nos demais Estados do país a greve nos <i>bancos</i> privados será suspensão a partir desta quinta, segundo a Contraf”.³² 1c. “Nos dias 25, 31 e 1º de janeiro, os <i>bancos</i> ficarão fechados.”³³ 1d. “Fácil comodidade para fazer consultas, efetuar pagamentos, transferências e outras transações sem precisar ir ao <i>banco</i>.”³⁴ 2a. “Voluntários fazem campanha por <i>banco</i> de órgãos do ABC.”³⁵ 2b. “O <i>banco</i> de sangue de São Paulo obedece normas nacionais e internacionais de segurança.”³⁶

³⁰ Fonte: <http://fonte.gratuita.com.br/>

³¹ Fonte: http://www.itaubr.com.br/conveniencia/serv_comprasnet_pgto.htm

³² Fonte: www1.folha.uol.com.br/folha/dinheiro/ult91u333681.shtml

³³ Fonte: http://g1.globo.com/Noticias/Economia_Negocios/0,,MUL931724-9356,00.html

³⁴ Fonte: <http://www.bb.com.br/portalbb/page22,136,3524,0,0,2,8.bb?codigoNoticia=3261>

³⁵ Fonte: <http://www.metodista.br/rroonline/cidades/voluntarios-querem-opo-no-abc>

³⁶ Fonte: <http://www.bssp.com.br/doacao.html>

	2c. “O banco de olhos não escolhe e nem tem preferência de qualquer espécie, pois a pessoa que irá receber os olhos entrará numa lista de espera seguindo uma ordem cronológica de inscrição.”³⁷ 3a. O objetivo principal de um sistema de banco de dados é possibilitar um ambiente que seja adequado e eficiente para uso na recuperação e armazenamento de informações.³⁸ 3b. O banco de imagens <i>on-line</i> que disponibiliza o trabalho de fotógrafos de todo o mundo, e que se destaca pelos preços que pratica para imagens de elevadíssima qualidade: entre um e cinco dólares por imagem.³⁹ 4. Jovem versátil sai do banco e garante vitória no fim do jogo.⁴⁰ 5. Sua ergonomia é ótima, pois o controle de altura do banco e a regulação da direção proporcionam uma boa posição para dirigir.⁴¹ As doações diárias de sangue não completam as necessidades do banco.⁴² 6. As andorinhas vêm das ilhas costeiras da costa da América do Norte e Europa e utilizam o banco como parada obrigatória todos os anos, entre setembro e março.⁴³
7. Chave de busca em inglês	<i>bank</i>
8. Hiperônimos relevantes	1a. instituição financeira 6. formação geológica
9. Sinônimos relevantes	1a. companhia bancária
10. Glosas WN.Pr/acepções de dicionários (relevantes)	1a-d. Estabelecimento de crédito, para transação de fundos públicos ou particulares (BORBA, 2002) 2. Local onde se armazenam componentes orgânicos de doadores para tratamento ou uso de outras pessoas (<i>banco</i> de sangue/ leite/ olhos/ esperma etc.) (AULETE, 2009) 3. Depósito ou conjunto de coisas armazenadas para uso ou distribuição futuros, certos serviços ou operações, etc. (<i>banco</i> de dados, <i>banco</i> de imagens). (AULETE, 2009) 4. (não consta em nenhum dos recursos léxicos analisados, mas é frequente no corpus) 5. Móvel usado como assento com ou sem encosto, rústico, feito de madeira, pedra ou concreto. (BORBA, 2002) Qualquer assento improvisado. (BORBA, 2002) 6. Elevação especialmente dentro do mar. (WN)

³⁷ Fonte: <http://boasaude.uol.com.br/lib/ShowDoc.cfm?LibDocID=3611&ReturnCatID=1776>

³⁸ Fonte: <ftp://ftp.unicamp.br/pub/apoio/treinamentos/bancodados/cursodb.pdf>.

³⁹ Fonte: www.sobresites.com/design/imagens.htm

⁴⁰ Fonte:

<http://www.sidneyrezende.com/noticia/49932+lateral+improvisado+sai+do+banco+e+garante+vitória+no+fim+do+jogo>

⁴¹ Fonte: <http://www.parana-online.com.br/canal/automoveis/news/35531/>

⁴² Fonte: <http://3planassiste.pgr.mpf.gov.br/informes/pr-banco-de-sangue-necessita-de-doadores>

⁴³ Fonte: <http://www.nectonsub.com.br/item/845/>

11. Nós do <i>polyset</i>_{café}	1. Instituição 1a. Instituição financeira 1b. Pessoas 1c. Funcionamento 1d. Sede física onde funciona a instituição financeira 2. Local onde se armazenam componentes orgânicos de doadores para tratamento ou uso de outras pessoas. 2a <i>banco</i> de sangue 2b <i>banco</i> de olhos 3. Depósito ou conjunto de coisas armazenadas para uso ou distribuição futuros, certos serviços ou operações, etc. 3.1 coleção de dados 3.2 coleção de imagens 4. Reserva, suplente 5. Assento 6. Aglomeração de areia, conchas, fósseis e detritos de rochas
12. Conjuntos de sentidos	<i>Polyset</i>¹ { ['instituição financeira', 'pessoas', 'funcionamento', 'sede física onde funciona a instituição financeira']_{a-d}, ['local onde se armazenam componentes orgânicos de doadores para tratamento ou uso de outras pessoas' (<i>banco</i> de sangue)_{2a} (<i>banco</i> de olhos)_{2b}]₂, ['depósito ou conjunto de coisas armazenadas para uso ou distribuição futuros, certos serviços ou operações, etc. (<i>banco</i> de dados, <i>banco</i> de imagens).'] (<i>coleção</i> de dados)_{3a} (<i>coleção</i> de imagens)_{3b}]₃, ('reserva, suplente')₄ }_{banco1} <i>Polyset</i>² { ('assento')₅, ('aglomeração de areia, conchas, fósseis e detritos de rochas')₆ }_{banco2}

Quadro 12 – Resultados das etapas preliminares de construção do *polyset*_{banco}

6.1.2.1 ETAPAS DE CONSTRUÇÃO DO *POLYSET*_{banco}:

O item lexical *banco* é frequentemente citado na literatura como exemplo de homonímia, conforme mencionamos ao longo deste trabalho, por não poder ser percebida qualquer similaridade entre *banco* ‘instituição financeira’, sentido (1), e ‘assento’, sentido (8). Esses não são, no entanto, os únicos sentidos convencionalmente associados ao item, de maneira que diferentes agrupamentos podem ser percebidos em seu conteúdo semântico. Conforme indicamos no quadro 13, *banco* é associado a dois *polysets* independentes construídos a partir da análise dos sentidos referentes aos exemplos (1-6): o primeiro é formado pelos sentidos (1-4) e o segundo pelos sentidos (5 e 6).

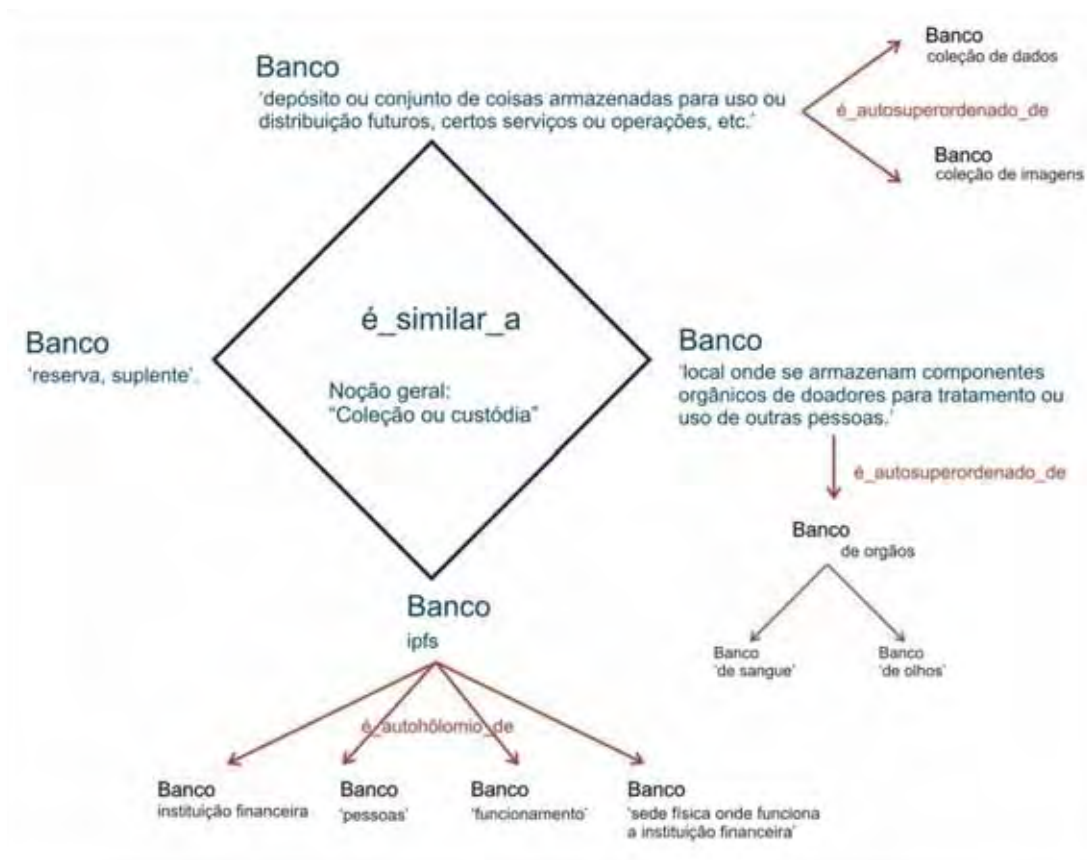


Figura 21 - Polyset1 banco

A figura 21 representa o *polyset1* associado a *banco*. A rede é estruturada em termos de polissemia regular e irregular. Detalhando o que apresentamos na Seção 4, optamos por organizar o conjunto de sentidos lá referidos como associados à 'coleção ou custódia' de algo de valor em diferentes níveis. O primeiro nível relaciona quatro sentidos por similaridade subjetiva (em azul), que é baseada na função de armazenamento de algo de valor, são eles: (a) *banco* como 'instituição' (representado na figura pelas iniciais das facetas de sentido 'instituição financeira', 'pessoas', 'funcionamento' e 'sede física onde funciona a instituição financeira'), (b) 'local onde se armazenam componentes orgânicos de doadores para tratamento ou uso de outras pessoas', (c) 'depósito ou conjunto de coisas armazenadas para uso ou distribuição futuros, certos serviços ou operações, etc.' e (d) 'reserva, suplente'. Trata-se de uma similaridade subjetiva entre quatro sentidos plenos (a-d). Um segundo nível de relações, e está aqui o nosso acréscimo, associa os sentidos por polissemia regular, em que o

sentido pleno (2), ‘local onde se armazenam componentes orgânicos de doadores para tratamento ou uso de outras pessoas’, serve como conteúdo esquemático para as leituras codificadas como microsentidos ‘*banco* de sangue’ e ‘*banco* de córnea’, relação de autohiponímia/autosuperordenação.

Entre os sentidos exemplificados em (1a-d), respectivamente, ‘instituição financeira’, ‘pessoas’, ‘funcionamento’ e ‘sede física onde funciona a instituição financeira’, pode ser percebida uma relação de polissemia regular do tipo *automeronímia*. Trata-se da classe mencionada em Buitelaar (2000) e que retomamos no quadro de fechamento da Seção 4. A regularidade é percebida por podermos encontrar outros itens lexicais que apresentam o mesmo tipo de variação entre ‘instituição’/ ‘pessoas’/ ‘funcionamento’/ ‘estrutura física’, entre eles, *escola*, *creche*, *universidade*.

A figura 22 ilustra o *polyset2* associado ao rótulo *banco*, relacionando os sentidos ‘móvel usado como assento’ e ‘aglomeração de areia, conchas, fósseis e detritos de rochas’ por similaridade.



Figura 22 - Polyset2 *banco*

A similaridade que construímos entre os sentidos (8), *banco* como ‘assento’ e (9), ‘aglomeração de areia, conchas, fósseis e detritos’, foi motivada especialmente pela combinação entre o exemplo (9) e a acepção geral apresentada em (8), em que *banco* é definido como ‘qualquer assento improvisado’ (BORBA, 2002). A partir da combinação dessas informações, entendemos que *banco* nos sentidos (8) e (9) associam-se em termos da similaridade subjetiva relacionada à noção geral *função de apoio* ou da similaridade objetiva baseada na forma de um *banco* do tipo móvel e um *banco* de areia ou fósseis e detritos. Não percebemos, no entanto, os sentidos exemplificados em (8) e (9) como microsentidos de um sentido mais geral em razão do fato de que o uso de um *banco* de areia, por exemplo, como apoio ou assento constrói-se apenas em situações especiais, como foi o que ocorreu em (9); fato que impossibilita preencheremos um dos

critérios principais da relação de *autohiponímia* que é a capacidade de unificação do tipo *é-um*. Há alto grau de autonomia e antagonismo entre esses sentidos, configurando-se como sentidos plenos relacionados por similaridade.

6.1.3 Café

A DEFINIÇÃO DOS NÓS	
1. Constante	café
2. Exemplos	1a. É toda essa composição que faz do café uma bebida natural e saudável.⁴⁴ 1b. Tomei uma água, um café. Não demorou muito. Dali eu saí e fui para a sala da ministra conduzida pela Erenice. Nós conversamos amenidades.⁴⁵ 2. Estas são as bases do cultivo orgânico do café que pode, entretanto, variar dependendo se o cultivo é mais ou menos intensivo.⁴⁶ 3. O café não é remédio, mas a comunidade médico-científica já considera a planta como funcional (previne doenças mantendo a saúde) ou mesmo nutracêutica (nutricional e farmacêutico).⁴⁷ 4. Durante a torra do café, esses ácidos clorogênicos formam novos compostos bioativos: os quinídeos.⁴⁸ 5. O uso da cor café na pintura provoca efeitos especiais e personaliza ambientes com estilo.⁴⁹ 6. Você é da turma que pula o café da manhã pensando em economizar calorias e acha que está fazendo um belíssimo negócio?⁵⁰ 7. Vamos ao café, fazer um lanchinho para conseguir ir até o final do expediente?⁵¹
3. Chaves de busca em inglês	(1-5) <i>coffee</i> (6) <i>bar</i> (7) <i>breakfast</i>
4. Hiperônimos relevantes	1a. alimento > bebida alimento > bebida > porção

⁴⁴ Fonte: http://www.abic.com.br/sabor_cafe.html

⁴⁵ Fonte: <http://www1.folha.uol.com.br/folha/brasil/ult96u612562.shtml>

⁴⁶ Fonte: <http://www.cnpab.embrapa.br/publicacoes/sistemasdeproducao/cafe/cafe.htm>

⁴⁷ Fonte: http://www.abic.com.br/sabor_cafe.html

⁴⁸ Fonte: http://www.abic.com.br/sabor_cafe.html

⁴⁹ Fonte: <http://www.revistacafeicultura.com.br/index.php?tipo=ler&mat=17104>

⁵⁰ Fonte: http://boaforma.abril.com.br/edicoes/189/fechado/Dieta/conteudo_130.shtml

⁵¹ Fonte: <http://twitter.com/ArantesGM/status/3181969279>

	2. planta > árvore 3. objeto natural > semente 4. substância 5. cor 6. refeição 7. local
5. Sinônimos relevantes	6. bar, boteco
6. Glosas WN.Pr/acepções de dicionários relevantes	1.a Bebida feita com a infusão de grãos de café (WN.Pr). b Porção individual dessa bebida (HOUAISS, 2001). 2. Árvore cafeeiro (HOUAISS, 2001). 3. Fruto do cafeeiro (HOUAISS, 2001). 4. Produto resultante do beneficiamento, torrefação, resfriamento, descanso e moagem dos grãos de café (AULETE, 2009). Pó de café. (BORBA, 2002). 5. A cor marrom do café ('grão') seco ou torrado, de matiz amarelado ou avermelhado e tonalidade entre o pardo ou um tom muito escuro, quase preto, a que ger. tb. corresponde a cor da bebida (HOUAISS, 2001). 6. Refeição rápida onde se toma a bebida (HOUAISS, 2001). 7. Estabelecimento comercial onde se vende essa bebida (BORBA, 2002).
7. Nós do <i>polyset</i>_{café}	1a. bebida b. porção da bebida 2. cafeeiro 3. fruto do cafeeiro 4. pó de café 5. cor 6. refeição rápida 7. bar
8. Conjuntos de sentidos	$\{('bebida', 'porção da bebida', 'cafeeiro', 'fruto do cafeeiro', 'pó de café')_{1-4}, ('cor')_5, ('refeição rápida')_6, ('bar')_7\}_{café}$

Quadro 13 – Resultado das etapas preliminares de construção do *polyset*_{café}

6.1.3.1 ETAPAS DE CONSTRUÇÃO DO *POLYSET*_{café}:

A rede de polissemia para os sentidos associados a *café* está representada na figura 23.

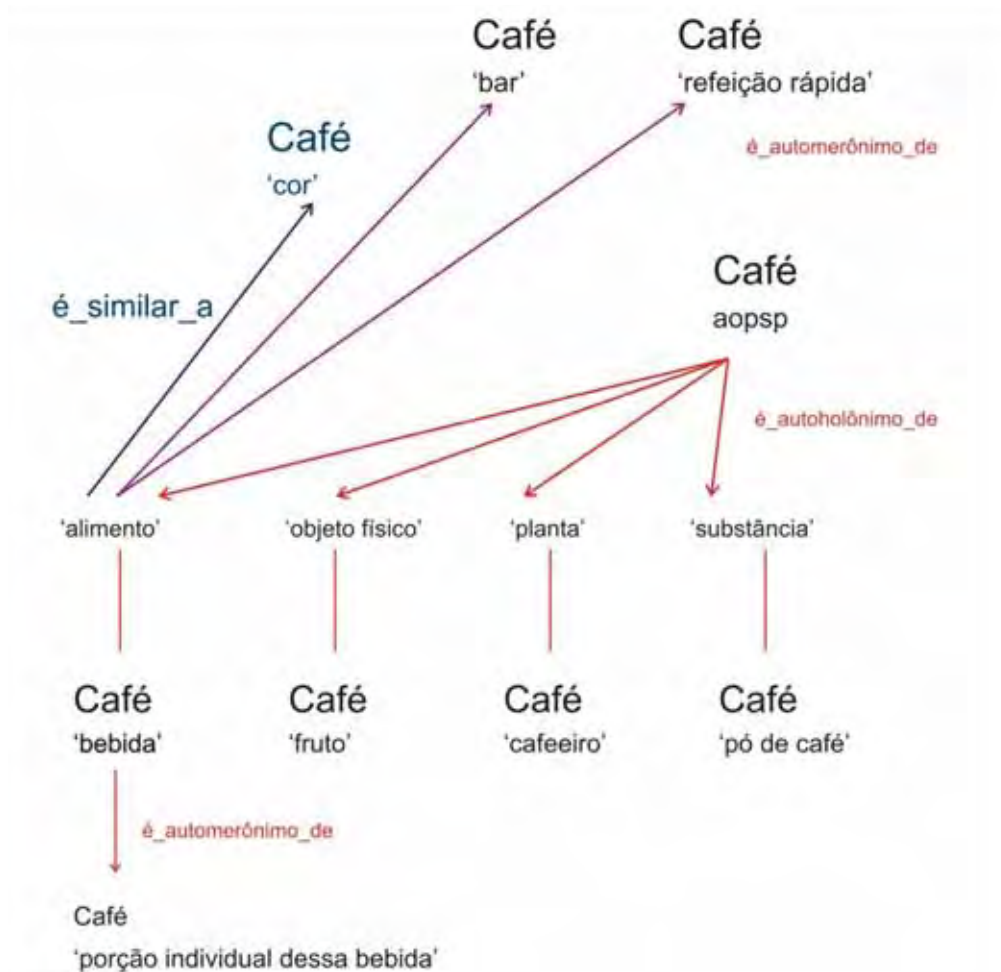


Figura 23 - Poliset *Café*

Conforme ilustra a figura 23, os sentidos associados a *café* podem ser agrupados em dois grandes conjuntos: polissemia irregular (em azul) e polissemia regular (em vermelho). A relação de polissemia irregular, expressa através do rótulo *é_similar_a*, caracteriza um conjunto de sentidos plenos relacionados por similaridade: entre *café* no sentido de 'cor', exemplo (5), e *café* no sentido de 'alimento', exemplo (1), percebemos um tipo de similaridade objetiva, baseada na percepção visual que associa a cor do café com a cor de objetos em geral. A relação de polissemia irregular agrupa diferentes conjuntos de facetas de sentido. O primeiro grupo que podemos identificar é aquele que relaciona uma classe regular de polissemia descrita em Buitelaar (1998), a classe que abreviamos, seguindo os moldes do autor, em "aopsp", para representar um sentido composto pelas facetas 'alimento', 'fruto', 'cafeeiro', 'pó_de_café'. Esse é o tipo clássico de polissemia regular rotulado como autoholonímia. A regularidade, nesse caso, justifica-se

por podermos perceber itens como *chocolate*, *erva-doce*, *anis*, *amendoim* que podem delimitar o mesmo tipo de relação. O segundo grupo de sentidos relacionados por polissemia regular é entre *café* ‘bebida’ (exemplo 1), ‘bar’ (exemplo 7) e ‘refeição rápida’ (6). Rotulamos a relação como automeronímia, visto que a bebida café é tipicamente parte do cenário de um bar, por ser vendida no bar, e é parte também típica do evento refeição, como é o café da manhã, ilustrado em (6). A distribuição das conexões entre os sentidos mostra que o sentido com maior saliência psicológica é *café* como ‘bebida’ e os resultados da busca no Google mostraram também esse como o sentido mais recorrente, portanto, com maior saliência estrutural.

6.2 A FACE COMPUTACIONAL DE CONSTRUÇÃO DO *POLYSET*

Nesta seção, o objetivo é equacionar o domínio computacional da tese, tarefa que compreende a apresentação de uma estratégia para a codificação do *polyset*. Para a realização de tal tarefa, utilizaremos o editor Protégé-OWL, que possibilita a criação da rede de polissemia, sua visualização e pesquisa de termos, além da codificação e exportação da base de conhecimento para arquivos em diferentes formatos (OLW, RDF, CLIPS, HTML, N-TRIPLE, N3, TURTLE). Para incluir no editor o conteúdo descrito em 6.1, adaptaremos, conforme as necessidades representacionais do fenômeno da polissemia, as recomendações da World Wide Web Consórcio (W3C) para a codificação da base WN.Pr em Ontology Web Language (OWL). Conforme mencionamos na introdução, o *status* que pretendemos dar à representação proposta no domínio computacional é o que Nirenburg e Raskin (2004, p. 14) denominam “tese fraca de IA”. Esse tipo de representação de conhecimento nega que um sistema computacional modele diretamente o conhecimento humano sobre semântica. Ao invés disso, sugere haver equivalência funcional, isto é, sugere que os programas computacionais podem alcançar resultados com qualidade análoga à humana, apesar de não usarem exatamente os mesmos métodos. Passaremos à explicitação das estratégias para a codificação de um *polyset* no editor Protégé-OWL⁵².

⁵² A mesma versão do Protégé-OWL foi testada e aprovada por em Di Felippo (2008) para o alinhamento de conceitos das *wordnets* WN.Pr e WN.Br, fato que demonstra a facilidade de integração das informações representadas no modelo *polyset* às demais informações em construção para a WN.Br.

6.2.1 A CODIFICAÇÃO DE UM *POLYSET* EM OWL ATRAVÉS DO PROTÉGÉ-OWL

Reconhecendo a importância das *wordnets* para sistemas de PLN, no contexto da W3C, Gangemi (2004)⁵³ propôs padrões para a codificação da WN.Pr no formato Ontology Web Language (OWL) com o objetivo de criar um esquema representacional comum aos diferentes projetos, facilitando o intercâmbio de dados. A opção pela linguagem OWL deve-se ao fato de ela ser a linguagem recomendada pela W3C para a representação de ontologias⁵⁴ em ambientes de Web Semântica (do Inglês: *Semantic Web*)⁵⁵. A OWL é recomendada por não ser uma simples linguagem de marcação de hipertexto com informações estruturais apenas (como é o HTML), mas por possuir semântica e sintaxe próprias para a modelação do conteúdo semântico de textos, contribuindo para a realização de inferências (ALVES, 2005). Tendo em vista as vantagens da codificação de conteúdo semântico em OWL, faremos uso do editor Protégé-OWL que, além de possibilitar a modelagem dos dados nessa linguagem, possibilita que eles sejam exportados para outros formatos, mostrando-se um recurso de grande utilidade (ex. CLIPS, HTML, N-TRIPLE, N3, TURTLE).

Seguindo as especificações propostas por Gangemi (2004), e sua equipe, para a codificação dos constructos da WN.Pr (ex.: formas lexicais, *synset*, sentidos e relações) em OWL, apresentaremos a seguir os critérios para representação dos *polysets* no Protégé-OWL, recurso programado para abarcar todos os constructos previstos pela linguagem. Seguiremos o mesmo esquema de codificação das informações semânticas para facilitar a integração do *polyset* à estrutura geral das *wordnets*. Seguindo a proposta da W3C (GANGEMI, 2004), os componentes que estruturam um *polysets* em OWL são representados como a figura 24⁵⁶.

⁵³ Disponível em <http://www.w3.org/2001/sw/BestPractices/WNET/tf.html>

⁵⁴ Uma definição clássica de ontologia é a de Gruber (1993, p.199), que afirma que “uma ontologia é uma especificação explícita de uma conceitualização (CHISHMAN, et. al). A definição do termo não é consensual na literatura, mas, para os fins desta tese, ela é suficiente.

⁵⁵ Padrão de representação de dados na web a fim de possibilitar que sistemas de PLN processem o conteúdo semântico de bases de dados lexicais criado pela W3C. O conteúdo semântico de documentos disponíveis na Internet estão codificados fundamentalmente em na linguagem padrão para a marcação de hipertexto, HTML (HyperText Markup Language) (ALVES, 2005).

⁵⁶ Figura adaptada da proposta de codificação da *wordnet* em OWL disponível em <http://www.w3.org/2001/sw/BestPractices/WNET/wordnet-sw-20040713.html>

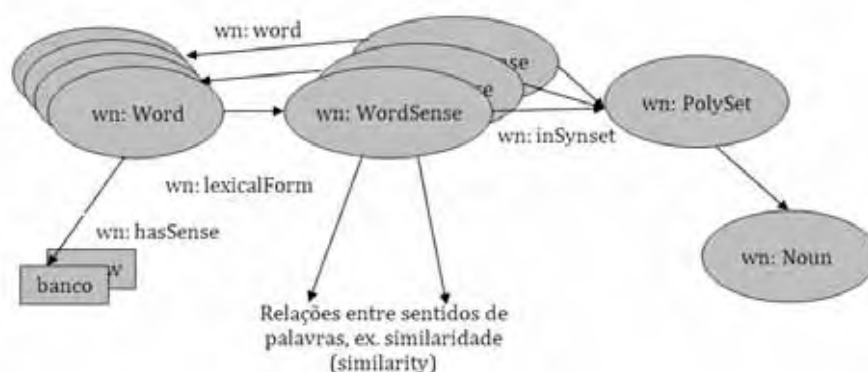


Figura 24 - Os componentes do *polyset* em OWL

Conforme retratamos na figura 24, há dois elementos que estruturam uma representação em OWL que são centrais ao modelo de *polyset*, são eles: *classes* e *propriedades*. As classes estão representadas nas esferas cinza e as propriedades nos rótulos das setas que ligam cada esfera. No esquema padrão do Protegé-OWL, as classes são responsáveis pela definição das entidades codificadas (ex.: *polyset*, *word*, *banco*). As classes genéricas do esquema de Gangemi (2004) para uma *wordnet* são (i) *Synsets*, que podem ser do tipo *synsets* para adjetivos, advérbios, nomes e verbos, (ii) *Word*, que pode ter como subclasse *collocation*, para incluir unidades complexas (na definição do autor, *string* composta por duas ou mais palavras conectadas por espaços ou hífen, por exemplo *man-eating*), (iii) *WordSense*, que pode se referir ao significado de adjetivos, advérbios, nomes ou verbos. Por recomendação de Knublauch et al. (2004), as classes são grafadas com iniciais maiúsculas e sem espaço entre as palavras. A figura 25 mostra esses elementos em estrutura hierárquica no Protégé-OWL.

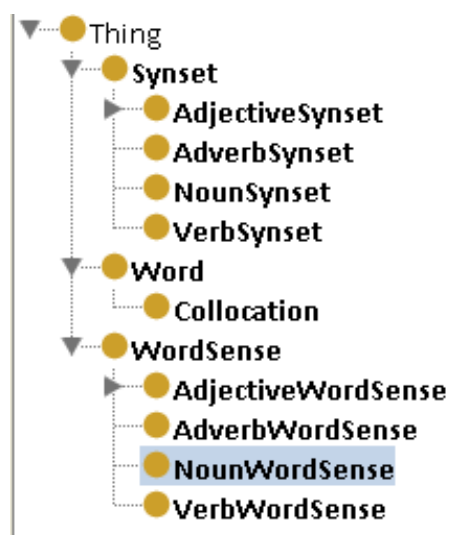


Figura 25 - Classes genéricas de uma *wordnet* em OWL

Adaptadas ao modelo *polyset*, essas classes genéricas são *polyset*, *Noun_Polyset*, *Word* e *Word_Sense*, conforme mostra a figura 26.

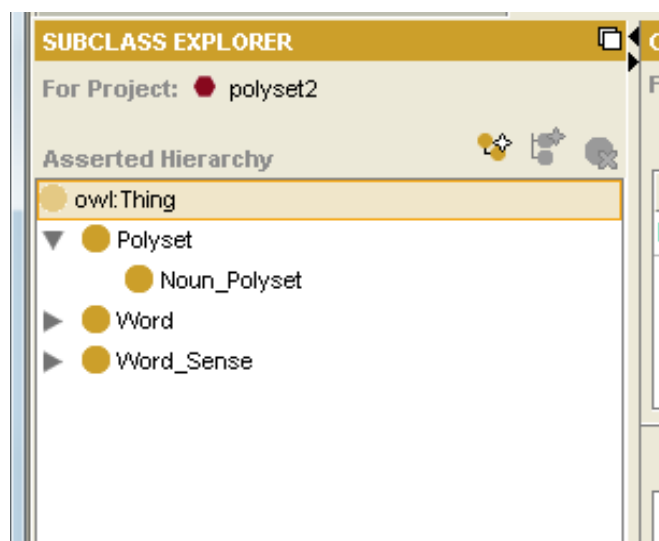


Figura 26 - Classes genéricas do modelo *polyset*

Dois tipos de propriedades codificadas pelo editor interessam ao nosso modelo: *object property*, para definir as relações de sentido, e *datatype property*, para definir outros tipos de características entre os sentidos, tais como *regular* e *noção geral*. Para as propriedades, a recomendação é inicial minúscula e as demais palavras com iniciais

maiúsculas, também sem espaço. A seguir, descreveremos como cada um dos componentes do *polyset* são codificados no Protégé-OWL, seguindo Gangemi (2004).

Os **nós das redes de polissemia** são os sentidos e estão representados do mesmo modo que os nós da rede *wordnet*: em termos de *classe* e estruturados hierarquicamente, conforme figura 27.

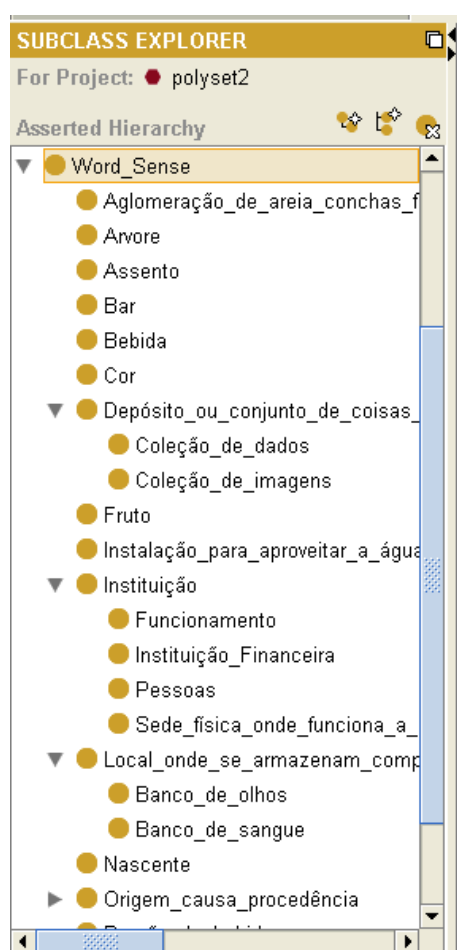


Figura 27 - Os nós como classes em estrutura hierárquica

Conforme apresentamos na Seção 5, os **nós** que compõem um *polyset* são os sentidos polissêmicos, e sua identificação pressupõe perceber sua associação convencional com uma mesma estrutura fonológica. Esses dois polos que integram um item lexical ficam evidentes no modelo de codificação proposto por Gangemi (2004), pois são inseridos como classes separadas. Essa estratégia faz com que não seja

necessário inserir mais de uma vez um mesmo item lexical com índices que os diferenciem e tornem possível a representação das relações que estruturam o modelo *polyset* do modo com descrevemos na seção teórica desta tese. Conforme indicamos na figura 28, o polo formal (Word) é representado pela classe *word* e recebe a propriedade, do tipo *datatype*, *lexicalForm*, e o polo semântico (Word Sense) é representado pela classe *WordSense*, e é especificado por diferentes propriedades do tipo *object property*. A conexão entre os polos se dá por meio da propriedade *hasSense* (*object property*), conforme a figura 28.

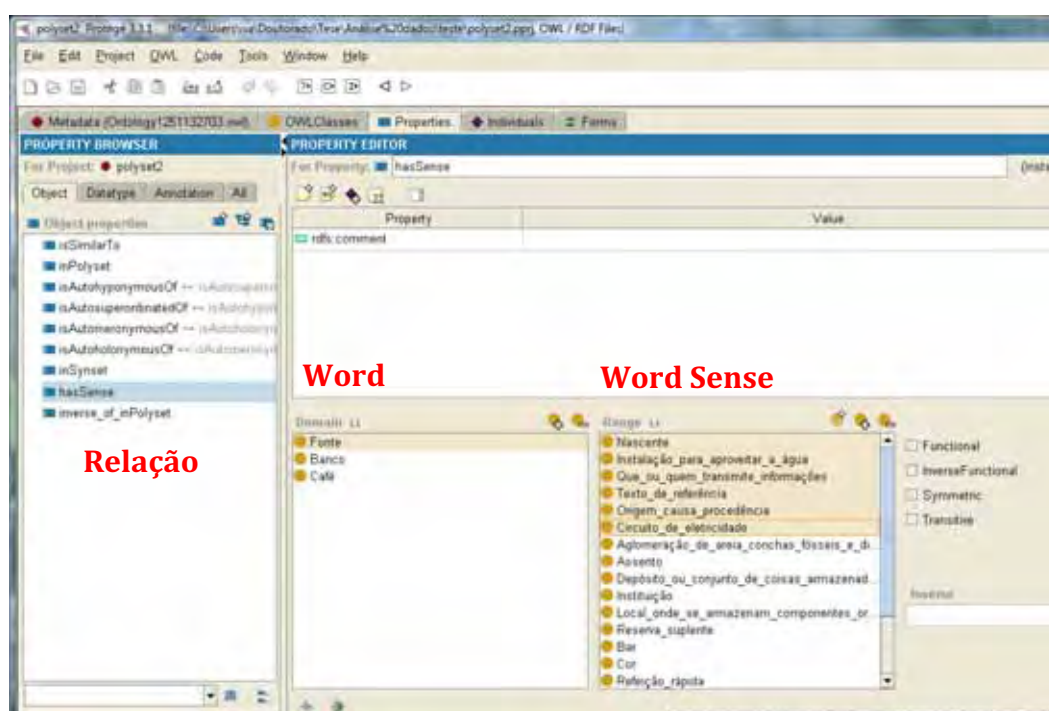


Figura 28 - A relação entre estrutura fonológica e semântica como uma *object property* do tipo *hasSense*

A codificação da relação entre uma mesma *word* e mais de uma *WordSense* significa que determinada *word* tem mais de um sentido. Na WN.Pr, conforme vimos na Seção 2, não se indica a existência ou não de conexões entre os diferentes sentidos associados a uma mesma *word*. O que as *wordnets* tipicamente codificam é quando *words* distintas associam-se a um mesmo *WordSense* (sinonímia), informação que, na proposta de Gangemi (2004) é explicitada por meio da relação *inSynset*. A proposta do modelo *polyset* é, justamente, codificar essas relações que têm ficado de fora das

wordnets. Desse modo, quando dois ou mais sentidos estiverem associados a uma mesma *word* receberão a propriedade (*object property*) ***inPolyset***, por analogia à relação *inSynset*. Um *polyset*, portanto, deve ser entendido, conforme descrito em 5, como um **conjunto de sentidos polissêmicos**. Essa informação, ilustrada na figura 29, é um importante acréscimo do modelo *polyset* ao paradigma *wordnet*. Na figura vemos que os sentidos de *fonte* como ‘nascente’, ‘instalação_para_aproveitar_água’, ‘origem_causa_procedência’ fazem parte de um *polyset* de nominais (Noun_Polyset).

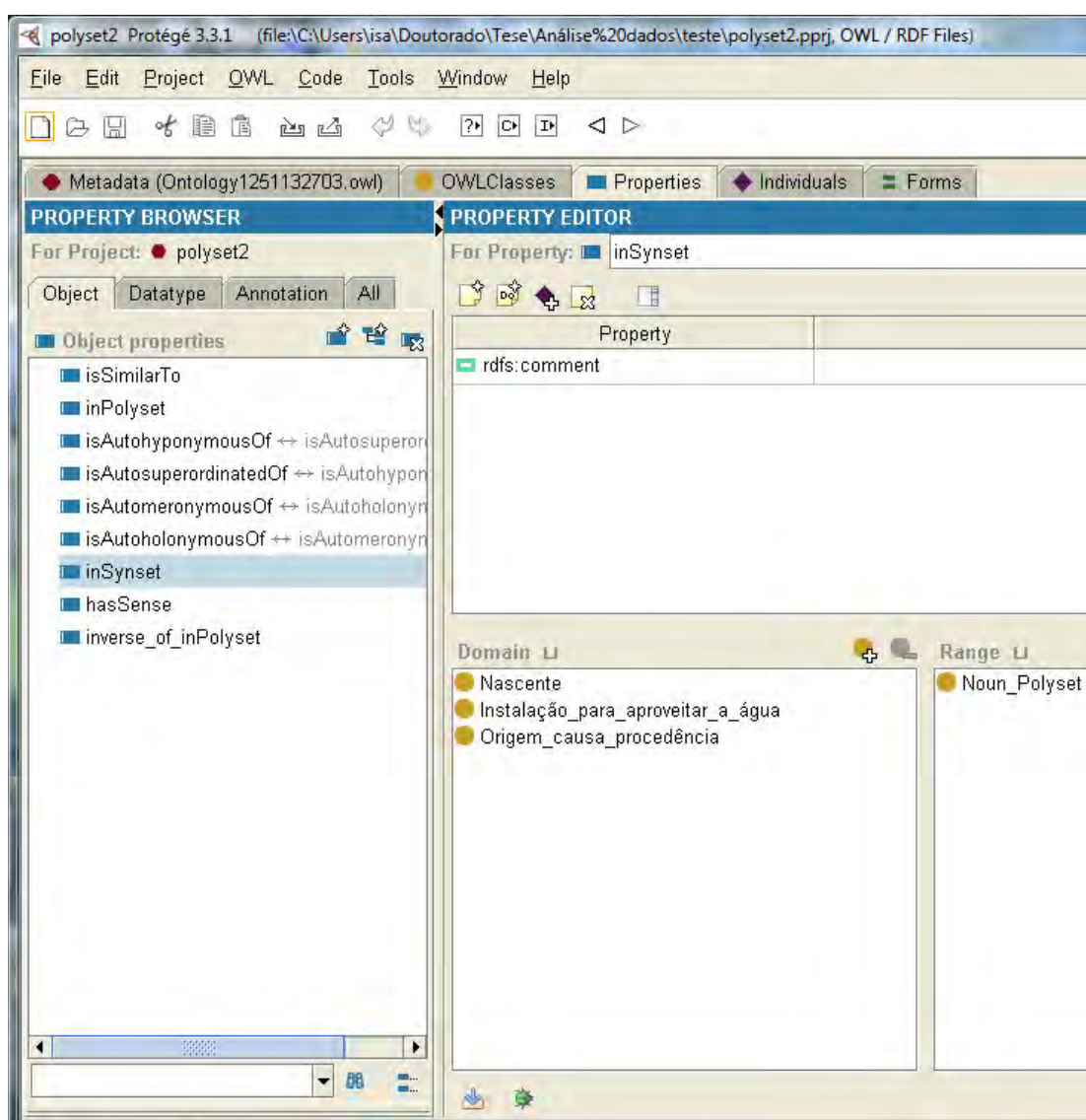


Figura 29 - Conjunto de sentidos polissêmicos como *object property* do tipo *inSynset*

Os **arcos da rede de polissemia** que estruturam um *polyset* são, conforme caracterizamos na Seção 5, as relações que estruturam a rede. Defendemos, nas seções anteriores, uma representação dos sentidos polissêmicos em diferentes níveis de generalidade e especificade. É, conforme descrevemos, a partir dessa estruturação de sentidos que os *polysets* estão organizados. Vimos também, na Seção 2, que essa organização hierárquica entre os sentidos polissêmicos não tem sido, até onde temos notícias, codificada sistematicamente na WN.Pr, de modo que não há indicação no esquema representacional de Gangemi (2004). Fazendo uso da mesma lógica de Gangemi para a codificação das relações léxico-conceituais das *wordnets*, representaremos as relações entre sentidos polissêmicos por meio de propriedades do tipo *property object*, conforme a figura 30.

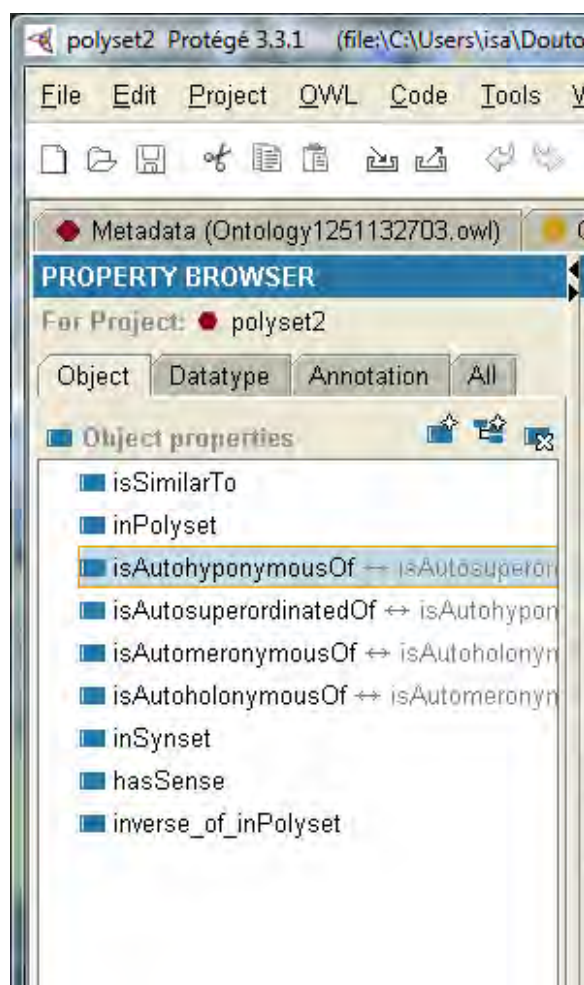


Figura 30 - Relações como *object properties*

Para codificar cada relação de sentido na aba de *object property*, é necessário definir um sentido de base (*domain*, no Protégé-OWL) e um alvo (*range*, no Protégé-OWL). O quadro 14 lista as diferentes relações codificadas no modelo *polyset* e indica o domínio e o alvo de cada relação

<i>Property</i> (propriedade)	<i>Domain</i> (domínio)	<i>Range</i> (alvo)
OBJECT PROPERTY		
<i>inPolyset</i>	WordSense	WordSense
<i>hasWordSense</i>	Polyset	WordSense
<i>Word</i>	WordSense	Word
<i>isSimilar</i>	WordSense	WordSense
<i>isAutohiponimousOf</i>	WordSense	WordSense
<i>isAutosuperordinatedOf</i>	WordSense	WordSense
<i>isAutomeronymousOf</i>	WordSense	WordSense
<i>isAutoholonymousOf</i>	WordSense	WordSense
DATATYPE PROPERTY		
globalNotion	WordSense	

Quadro 14 – Domínio e alvo nas relações de sentido

Para a estruturação dos sentidos plenos e dos subsentidos que compoem um *polyset*, informação que não está presente no padrão para *wordnets*, seguiremos a lógica de estruturação de informações do Protégé-OWL: hierarquização dos sentidos e indicação de *object properties*. A relação de autohiponímia (e seu inverso, a autosuperordenação) é expressa na figura 31.

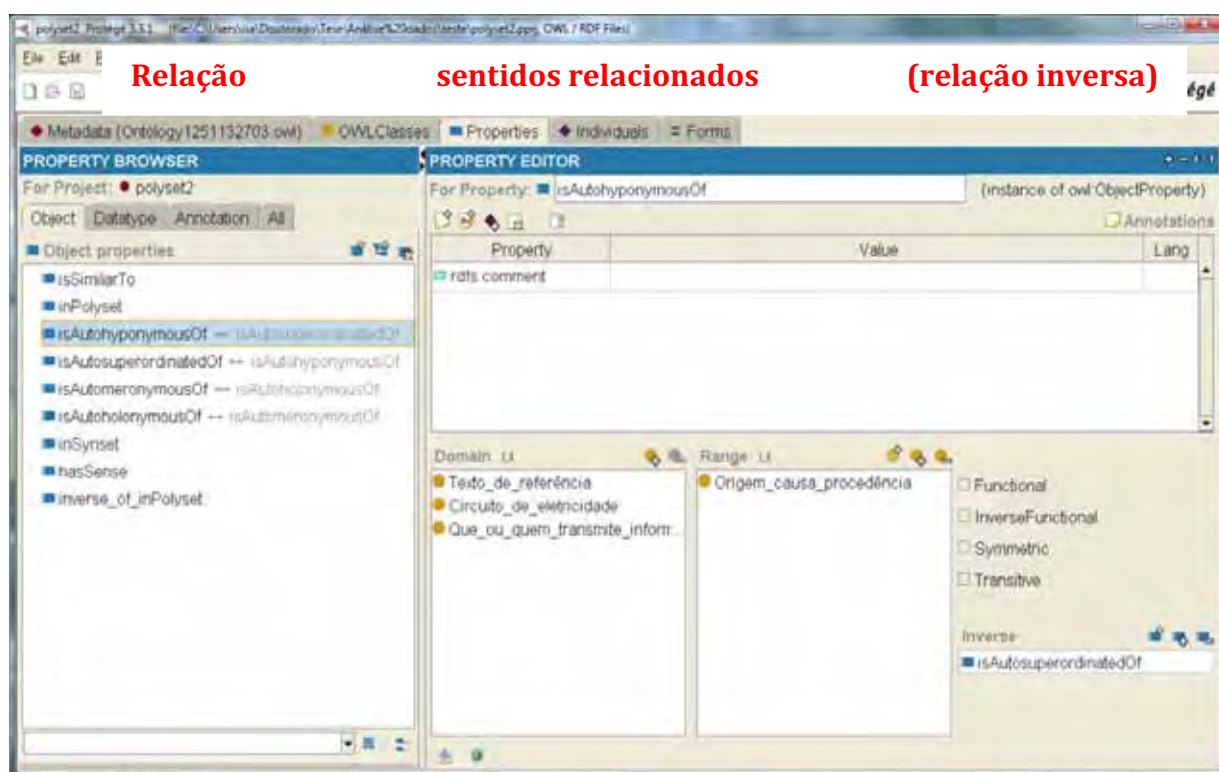


Figura 31 - Relação de autohiponímia/autosuperordenação

A relação de autoholonímia (e seu inverso, a automeronímia) é representada do mesmo modo, através da codificação de uma *object property*, conforme mostra a figura 32.

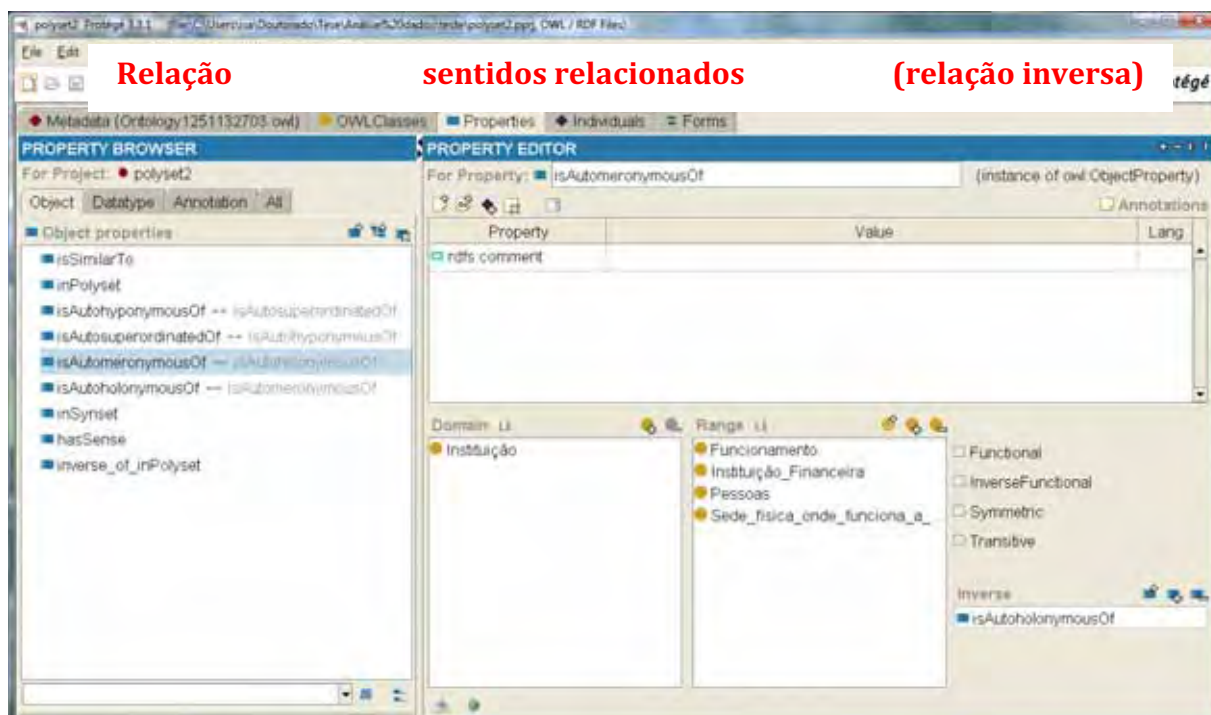


Figura 32 - Relação de automeronímia/autoholonímia

Para concluir a codificação de um *polyset*, é preciso ainda explicitar distintos graus de saliência entre os sentidos. Essa é mais uma informação que não está prevista no esquema de Gangemi (2004) por não constar na WN, configurando-se como um acréscimo do modelo *polyset*. Os efeitos de saliência entre os sentido são codificados no através de relações do tipo *datatype property*: para indicar o sentido com maior saliência psicológica, o rótulo da relação é *psycSalience*; para indicar o sentido com maior saliência estrutural ou de frequência, o rótulo da relação é *freqSalience* – conforme mostra a figura 33.

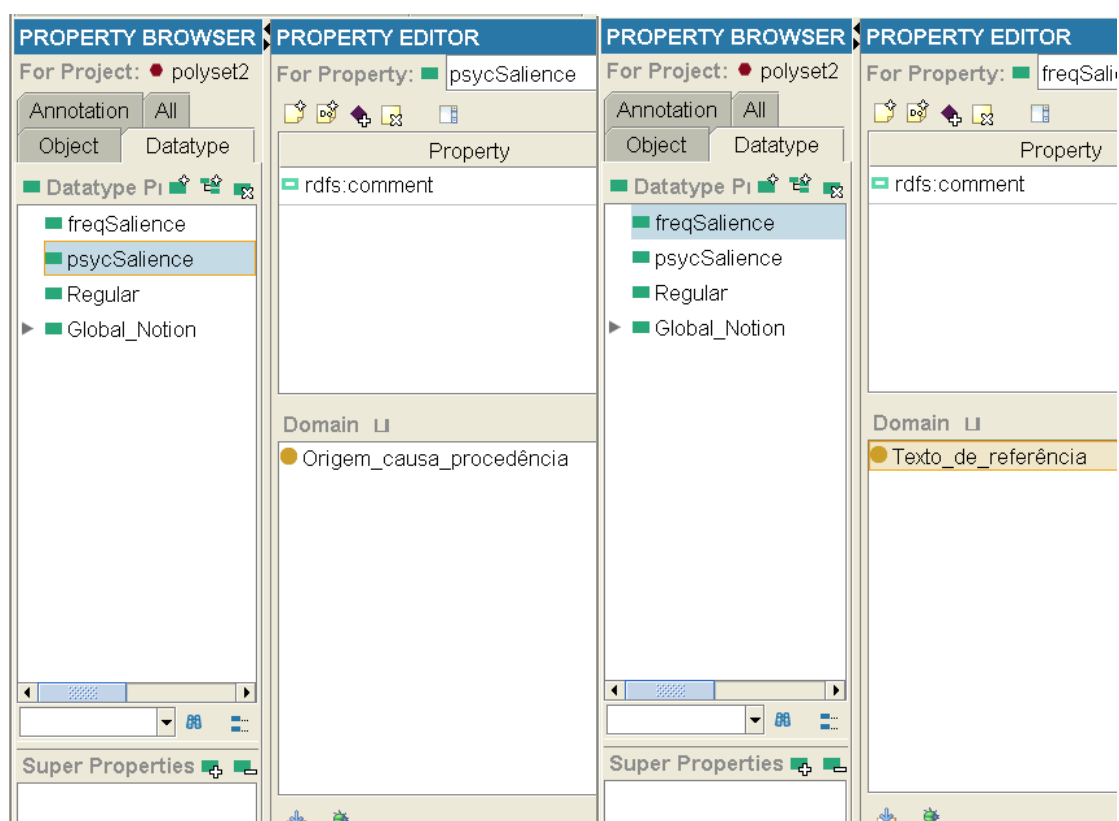


Figura 33 - Graus de saliência

Para concluir a seção, apresentamos, na figura 34, uma visualização completa do exercício realizado no formato de rede construído a partir de um *plug-in* que pode ser habilitado no Protégé-OWL, o OWL-Viz.

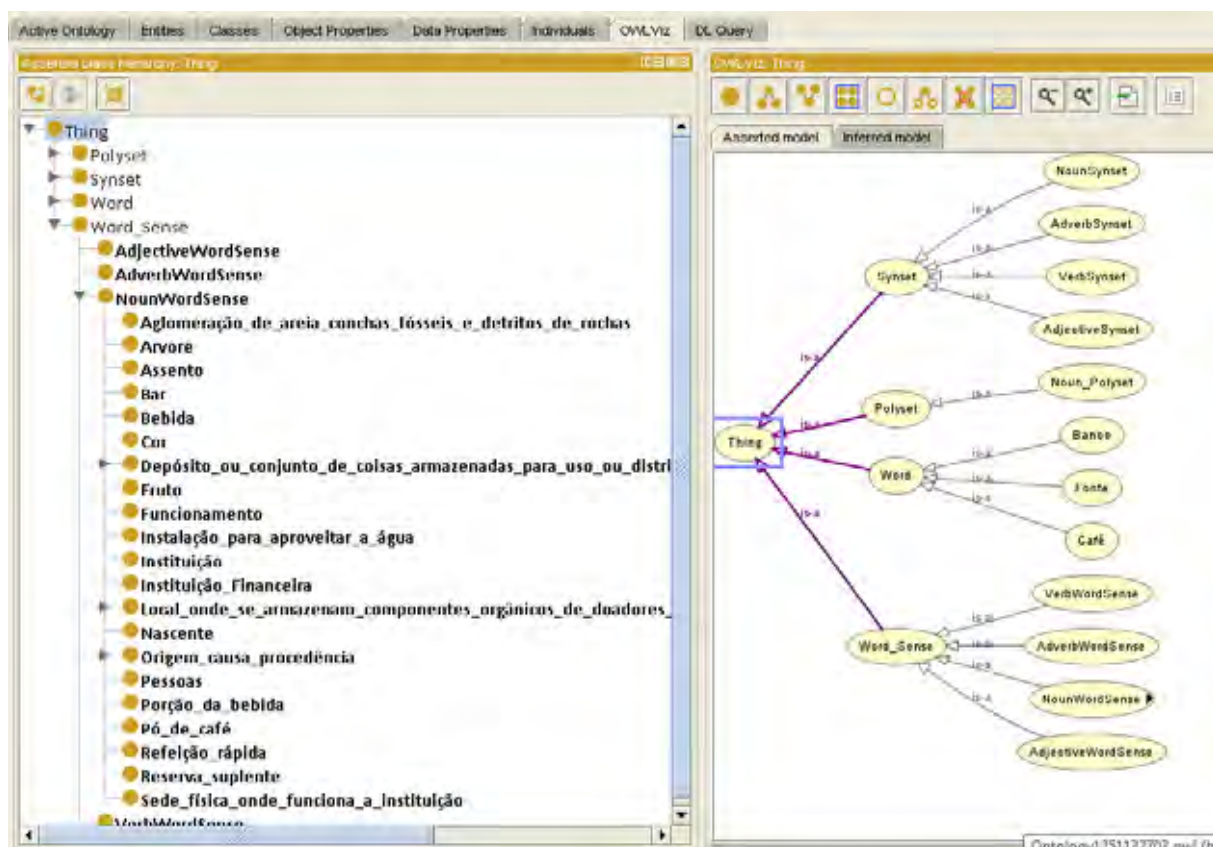


Figura 34 - Visão geral dos *polysets* codificados no Protégé-OWL

6.3 SÍNTESE DAS CONTRIBUIÇÕES DO MODELO *POLYSET* PARA A WORDNET.BR

A WordNet.Br (WN.Br) (DIAS-DA-SILVA, 2005; 2006), o foco de aplicação desta tese, está em desenvolvimento. Desde 2002, e, em sua situação atual, contém mais de 20 mil unidades lexicais analisadas, somando mais de 20 mil *synsets*, entre verbos, substantivos, advérbios e adjetivos. As relações identificadas até o momento para as formas verbais, adverbiais e adjetivais foram *sinonímia* e *antonímia*; além disso, foi feita a indexação de parte desses *synsets* com seus equivalentes na WN2.0. A metodologia de desenvolvimento adotada pelo grupo é por abrangência, e não por profundidade.

A inclusão dos nomes, foco desta tese, está em fase embrionária, e o tratamento da polissemia segue, por enquanto, a regra geral da WN.Br, apresentando apenas estatísticas sobre itens com sentido múltiplo. Trata-se de uma quantificação do número

de *synsets* de que um item lexical participa. Por exemplo, se uma mesma forma ocorre em três *synsets*, estipula-se que seu grau de polissemia é 3, independentemente do tipo de relação que se estabelece entre os sentidos. Nesse cenário, temos já uma base numerosa e com uma grande quantidade de substantivos polissêmicos. Conforme mencionamos na Seção 5, de um total de 16.925 entradas para nominais, 5.957 são codificados como polissêmicos. A organização desses conjuntos de sinônimos em termos de *polyset* terá, portanto, grande impacto na estruturação da WN.Br.

Destacamos duas como as contribuições centrais do modelo proposto para uma *wordnet* e, em especial, à WN.Br. A primeira delas é a proposição de uma metodologia para a representação estruturada para a polissemia de nominais, complementando as informações léxico-conceituais já disponibilizadas pela base. A segunda contribuição é a proposição de uma metodologia para a codificação computacional de tais informações, de maneira compatível com a proposta de codificação da WN.Br em OWL apresentada pela W3C. Para demonstrar visualmente essa compatibilidade, integramos a representação de *polysets* à base WN em OWL disponibilizada pela W3C⁵⁷ no Protégé-OWL, conforme a figura 35.

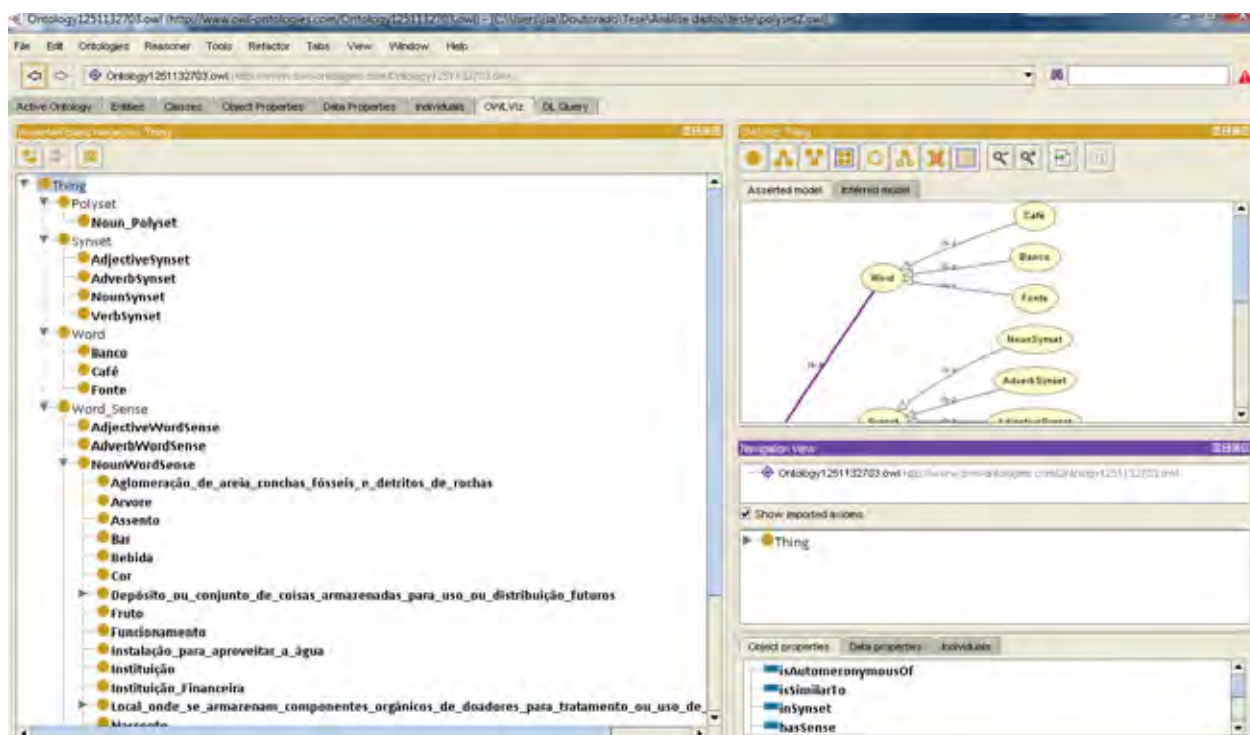


Figura 35 - wordnet e polysets em OLW no Protégé-OWL

⁵⁷ Disponível em <http://www.w3.org/2006/03/wn/wn20/download/>

Vimos na Seção 5 que há uma série de requisitos (cf. HELBIG, 2006) importantes para uma representação linguística que pretende ter *status* cognitivo e computacional. Para encerrar esta seção de apresentação das contribuições do modelo ao desenvolvimento da WN.Br, enfatizamos que incluir uma descrição de nominais polissêmicos nos moldes propostos fará com que a base esteja de acordo com os seguintes requisitos:

- (a) *Adequação cognitiva* – a representação da polissemia em termos de rede, conforme descrevemos na Seção 3 atribui o *status* cognitivo ao modelo;
- (b) *Interoperabilidade* – conforme ficou evidenciado na Seção 6, todas as propriedades previstas para o modelo *polyset* puderam ser representadas computacionalmente com auxílio do Protégé-OWL e do esquema de representação de Grangemi (2004);
- (c) *Homogeneidade e interpretabilidade local* – Estruturar os *polysets* em termos de rede de polissemia possibilitou representar as diferentes relações de sentido, conforme vimos em 6.1, e integrar essa representação à estrutura do léxico como um todo, conforme vimos em 6.3;
- (d) *Reutilização* – A adequação do modelo a esse requisito geral evidencia-se por dois motivos. O primeiro deles deve-se ao fato de as informações estarem codificadas em um editor que possibilita a exportação de sua base de conhecimentos para diferentes formatos. O segundo motivo deve-se ao fato de os *polysets* terem sido estruturados em OWL seguindo o padrão proposto pela W3C;
- (e) *Viabilidade e automaticidade* – Os *polysets* têm potencial para tratamento técnico e implementacional, conforme vimos evidenciado na Seção 6;
- (f) *Compleitude e consistência* - a possibilidade de representar polissemia regular e irregular a partir de um mesmo constructo, sem que sejam derivadas informações inadequadas;
- (g) *Granularidade ótima* - a possibilidade de representar sentidos polissêmicos em diferentes níveis de generalidade e estabilidade;
- (h) *Multidimensionalidade* – o constructo *polyset*, em sua codificação no Protégé-OWL está apto a representar informações provenientes de diferentes níveis de generalidade e de diferentes tipos, conforme vimos nos tipos de

propriedades que podem ser evidenciada a partir da combinação das diferentes propriedades codificadas (*object property* e *datatype property*).

6.4 SÍNTESE DA SEÇÃO

Para finalizar a Seção 6, com o objetivo de sintetizar as etapas de construção de um *polyset*, elaboramos o que chamaremos de regras gerais para construção de um *polyset*. Para sua aplicação, é necessário estar de posse das informações elencadas na Seção 5, pois aqui apresentamos apenas um guia geral que registra os tipos de relação de polissemia inventariados nesta tese. Tais regras foram construídas procurando seguir estilo similar ao algoritmo apresentado em Buitelaar (2000), na tentativa de evidenciar possíveis momentos para automatização dos processos de construção de um *polyset*. Tendo em vista esse objetivo de automatização, indicaremos somente uma *wordnet* como referência para os sentidos. Para a aplicação de tais critérios ao PB, não se pode esquecer que, enquanto a WordNet.Br estiver em processo de construção, é necessário seguir as etapas preliminares sintetizadas no quadro 9, realizando estudo de dicionários, *wordnets* e *corpus*. As regras gerais são as seguintes:

Regra geral A – Um item lexical é considerado com sentido múltiplo sempre que participar de mais de um *synset*.

Regra geral B* - Por *default*, cada novo *synset* (sentido) deve ser listado na base de dados lexicais de maneira independente.

* Esse valor *default* é anulado sempre que possível aplicar a regra C.

Regra geral C - Sempre que alguma relação de polissemia for identificada, o padrão *default* é quebrado e uma das seguintes regras é aplicada:

Regra 1 – Se forem identificados sentidos plenos ligados por alguma *noção geral unificadora*, indicar a relação *é_similar_a* para rotular a relação entre eles.

Regra 2 – Se a relação de polissemia for *regular*, aplicar um dos seguintes rótulos:

2.1 Se forem identificadas as relações do tipo *é_um* entre os sentidos (microsentidos), aplicar o rótulo *é_autohipônimo_de* (a relação reversa é *é_autosuperordenado_de*).

Regra 3 – Se forem identificadas as relações do tipo *é_parte_de* entre os sentidos (facetas), aplicar o rótulo *é_automerônimo_de* (a relação reversa é *é_autosuperordenado_de*).

SEÇÃO 7 - CONCLUSÃO

O objetivo geral desta tese foi propor uma estratégia de representação linguístico-computacional da semântica de nominais polissêmicos que complemente as informações léxico-conceituais codificadas em *wordnets*, com vistas a ampliar a utilidade das bases para sistemas de PLN. O estudo foi desenvolvido seguindo uma metodologia de pesquisa que integra três domínios mutuamente complementares: o linguístico-computacional, o linguístico e o computacional, conforme propõe Dias-da-Silva (1996, 1998, 2003, 2006). O domínio linguístico-computacional ocupou a posição de “pivô” entre os dois outros domínios. Foi ele o responsável por fornecer o tema para esta pesquisa e por articular a relação entre os domínios linguístico e computacional. Respeitando essa estratégia de pesquisa, o conteúdo desta tese foi distribuído em cinco seções, cujos resultados serão sintetizados a seguir.

A Seção 2 ficou responsável pelo equacionamento da primeira etapa desenvolvida no domínio linguístico-computacional: a investigação sobre a relação polissemia, PLN e *wordnets*. A partir das reflexões empreendidas nessa seção, foi possível perceber que a falta de representação explícita da polissemia em bases do tipo *wordnet* tem sido apontada como uma das principais limitações para a maior utilização do recurso em sistemas de PLN. Vimos que diversas técnicas automáticas têm sido propostas para extrair similaridades entre os sentidos, e as técnicas que fazem uso das *wordnets* têm sido as que apresentam melhores resultados. Por não haver nesses recursos informações para as relações de polissemia, relações do tipo *sinonímia* e *hiponímia* têm sido exploradas para fins de desambiguação. Ao estudarmos o potencial descritivo do paradigma *wordnet*, ficou evidente que, em sua base, a polissemia não está prevista; isso porque, o significado lexical é representado sob a ótica onomasiológica, tendo a *sinonímia* como a relação estruturadora fundamental. A investigação sobre a WN.Pr, a WN.Pt e a GN comprovou essa dificuldade e demonstrou que, embora os projetos reconheçam a importância de representar o fenômeno, não apresentam uma estratégia adequada à descrição da polissemia como um todo. Vimos que há a preocupação com o fenômeno, porém, dada as dificuldades de sua delimitação por exigir um minucioso estudo linguístico, ele tem sido deixado em segundo plano. Concluímos a

Seção 2 cientes da importância da codificação da polissemia e informados sobre a situação atual de desenvolvimento das diferentes *wordnets*, portanto, informados sobre as possibilidades e as necessidades das *wordnets*.

As Seções 3 e 4 tiveram como foco o desenvolvimento da fase teórica do domínio linguístico. Na Seção 3, caracterizamos a polissemia sob o viés da SLC, tornando possível lançarmos mão de um arcabouço de constructos adequados à descrição das principais questões envolvidas na representação da polissemia. Olhar para o fenômeno do significado múltiplo sob o viés cognitivo, possibilitou descrever os sentidos como uma entidade complexa, estruturada em termos de redes. As redes de polissemia sincrônicas, em sua configuração livre e multidimensional, conforme propõem Blank (2003) e Geeraerts (2006), demonstraram ser a estratégia descritiva mais adequada à representação da flexibilidade do sentido para os propósitos desta tese. Nessa configuração de rede, há espaço para, em uma mesma estrutura, representar sentidos mais específicos e instáveis - particulares a determinada situação de comunicação -, e sentidos mais gerais, convencionais e estáveis - comuns a diversas situações de comunicação. Essa flexibilidade é especialmente interessante para a construção de uma base de dados lexicais que pretende ser útil a sistemas computacionais diversos como é uma *wordnet*.

A Seção 4 foi importante para compreendermos os processos envolvidos na organização interna de uma categoria de sentidos polissêmicos, investigando questões como delimitação de sentidos, graus de saliência e relações de sentido. No contexto linguístico-cognitivo, as relações de polissemia representam rotinas cognitivas que o falante realiza para delimitar sentidos compreendendo a língua em suas diversas situações. De acordo com o paradigma cognitivo da semântica, entende-se que essas rotinas cognitivas realizam-se baseadas em habilidades de percepção e em padrões de convenção. No contexto linguístico-computacional, a representação de diferentes relações entre os sentidos possibilita que o sistema perceba diferentes graus de similaridade, de saliência e de estabilidade entre os sentidos.

A Seção 5 foi dedicada à apresentação do delineamento linguístico do modelo de representação proposto: os *polysets*. Defendemos que a construção de um *polyset* seja realizada em dois momentos principais. O primeiro momento, denominado *etapas preliminares*, é dedicado à definição dos *nós* da rede de polissemia e à indicação dos agrupamentos entre os sentidos. Entre as atividades realizadas nessa etapa estão a

consulta ao *córpus*, a consulta aos recursos léxicos e a indicação de um conteúdo para cada um dos *nós* do *polyset*. O segundo momento, denominado *etapas de construção do polyset*, é dedicado à definição propriamente dita da rede de polissemia. É nessas fases que são rotulados os *arcos* da rede de polissemia, identificado o *status* de delimitação dos sentidos e seus efeitos de saliência. Do ponto de vista linguístico, o *polyset* é considerado construído quando se tem estruturada uma rede manual que representa as múltiplas relações entre os sentidos polissêmicos associados a um item lexical. Do ponto de vista linguístico-computacional e computacional, a construção do *polyset* se dá quando é finalizado o processo de inclusão dos dados na ferramenta de edição, etapa realizada na Seção 6.

Reservamos para a seção final da tese, a Seção 6, o espaço para a construção e implementação do *polyset*, respondendo à etapa aplicada dos domínios linguístico, linguístico-computacional e computacional desta pesquisa. Do ponto de vista linguístico, destacamos que os exemplos analisados revelaram a riqueza do fenômeno da polissemia. O item lexical *fonte* é tipicamente tratado na literatura como um exemplo de polissemia irregular, *banco* é exemplo típico de homonímia e *café* é exemplo de polissemia regular. Conforme ficou evidente nas redes de polissemia construídas para cada um dos itens lexicais, o conjunto de sentidos associados a um item é, de fato, múltiplo e precisa ser organizado em termos de uma estrutura multidimensional. Somente com esse tipo de configuração de rede é possível representar as diferentes relações, os diferentes níveis de saliência, de especificidade e de sensibilidade ao contexto situacional que são percebidos no espaço semântico acessível a uma unidade do léxico. Do ponto de vista computacional, destacamos a importância do esquema de codificação que seguimos para a inclusão do *polyset* no Protégé-OWL. Seguir um modelo padrão facilita a reutilização da representação e a interação entre grupos de pesquisa. A possibilidade e facilidade de interação entre representações que seguem os mesmos princípios ficou evidente ao apresentarmos a tela do Protégé-OWL que integra os *polysets* à estrutura de uma *wordnet* que contém informações clássicas.

Conforme mencionamos ao longo do trabalho, não há consenso sobre qual é a melhor abordagem para a análise e a representação da polissemia. A definição se dá em termos de sua adequação aos objetivos da pesquisa. Sem perder de vista esse fato, entendemos que o modelo adotado aqui é suficiente para os nossos propósitos por representar o significado lexical de um modo flexível, podendo expressar as diferentes

conexões entre eles, bem como diferentes níveis de especificidade e de generalidade em uma estrutura hierárquica.

Como trabalhos futuros, prevemos a expansão da aplicação do modelo a outras classes de palavras. Considerando as reflexões teóricas apresentadas, ficou evidente que o modelo de *polysets* foi proposto para substantivos, entretanto, sua estrutura de base, as redes de polissemia, não é restrita a eles, de modo que a abrangência do estudo pode ser expandida em trabalhos futuros. Uma outra expansão prevista para este trabalho, para facilitar a replicação da análise para um número maior de dados, é a automatização de algumas das etapas, por exemplo, a busca nos dicionários e no *corpus*.

REFERÊNCIAS BIBLIOGRÁFICAS

ALANI, H. TGVizTab: An Ontology Visualisation Extension for Protégé. In: **Knowledge Capture (K-Cap'03), Workshop on Visualization Information in Knowledge Engineering**, October 26, Sanibel Island, Florida, USA, 2003.

ALLWOOD, J. Meaning potentials and context: Some consequences for the analysis of variation in meaning. In: CUYCKENS, H.; DIRVEN, R.; TAYLOR, J. (Ed.). **Cognitive approaches to lexical semantics**. Berlin: Mouton de Gruyter, p. 29-66, 2003.

ALVES, I. M. da R. **O uso da Semântica Verbal em Sistemas de Extração de Informação: a Construção de uma Ontologia de Domínio Jurídico**. Dissertação de Mestrado em Lingüística e Língua Portuguesa – Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, 2005.

ALVES, I. M. da R. **Os Tradutores Automáticos e a Ambigüidade: Uma Abordagem Crítica**. Trabalho de Conclusão de Curso de Graduação em Letras (Português-Inglês) – Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, 2003.

APRESJAN, J. **Regular Polysemy**. *Linguistics*, p. 142, 5-32, 1973.

AULETE, C. **Dicionário contemporâneo da língua portuguesa**. Edição brasileira original: Hamílcar de Garcia. Versão Eletrônica. Atualizada em 2009.

BLANK, A. Polysemy in the lexicon and discourse. In: NERLICH, B. et al (Ed.). **Polysemy: flexible patterns of meaning in mind and language**. Hawthorne, N.Y.: Mouton de Gruyter, 2003. p. 267-293.

BORBA, F. S. (Coord.). **Dicionário de usos do português do Brasil**. São Paulo: Ed. da UNESP, 2002.

BOYD-GRABER, J.; FELLBAUM, C.; OSHERSON, D.; SCHAPIRE, R. Adding dense, weighted connections to WordNet. In: **Proceedings of the 3rd International WordNet Conference**, Jeju Island, Korea. 2006.

BRÉAL, M. **Ensaio de Semântica: Ciência das Significações**. São Paulo: EDUC, 1992.

BRISARD, F. The English tense-system as an epistemic category: the case futurity. In: VERSPOOR, M.; LEE, K. D.; SWEETSER, E. (Eds.). **Lexical and syntactical constructions and the construction of meaning**. Amsterdam: John Benjamins, p. 271-286, 1997.

BROCCIAS, C. Approachs to grammar. In: KRISTIANSEN, G.; ACHARD, M.; DIRVEN, R.; IBÁÑEZ, F. J. R. M. (Eds.) **Cognitive Linguistics: Current Application and Future Perspectives**. Mouton de Gruyter: Berlin, p. 81-118, 2006.

BROWN, S. W. Polysemy in the mental lexicon. **Colorado Research in Linguistics**, Boulder, v. 21, p. 4, June 2008. Disponível em: <http://www.colorado.edu/ling/CRIL/Volume21_Issue1/paper_BROWN.pdf>. Acesso em: 15 mar. 2009.

BRUGMAN, C. **Story of over**. M.A. thesis--University of California, Berkeley, 1981.

BRUGMAN, C.; LAKOFF, G. Cognitive typology and lexical networks. In.: GEERAERTS, D. (Ed.). **Cognitive Linguistics: Basic Readings**, Berlin: Mouton de Gruyter, p. 109-139, 2006.

BUITELAAR, P. **Corelex: systematic polysemy and underspecification**. PhD Thesis--Department of Computer Science, Brandeis University, Waltham, 137p., 1998.

BUITELAAR, P. Reducing lexical semantic complexity with systematic polysemous classes and underspecification. In: **Proceedings of the ANLP2000 Workshop on Syntactic and Semantic Complexity in Natural Language Processing Systems**, Seattle, USA, p. 14-19, 2000.

CANÇADO, M. **Manual de Semântica – Noções Básicas e Exercícios**. Belo Horizonte: Editora UFMG, p. 183. 2005.

CHISHMAN, R. L. de O. **A teoria do léxico gerativo: uma abordagem crítica**. 2000. 185 f. Tese (Doutorado em Lingüística e Letras)--Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.

CHISHMAN, R. L. de O; BERTOLDI, A.; ALVES, I. M. da R. **Integrando léxicos semânticos e ontologias: uma aproximação a favor da web semântica** (*no prelo*)

COPESTAKE, A.; BRISCOE, T. Semi-productive polysemy and sense extension. **Journal of Semantics**, Nijmegen, v. 12, n. 1, p. 15-67, 1995. Disponível em: <<http://www.cl.cam.ac.uk/~aac10/papers/jsem.ps.gz>>. Acesso em: 20 jan. 2006.

CROFT, W.; CRUSE, D. A. **Cognitive linguistics**. Cambridge: Cambridge University Press, 2004.

CRUSE, D. A. **Lexical semantics**. Cambridge: Cambridge University Press, 1986.

CRUSE, D.A. Aspects of the micro-structure of word meanings. In: RAVIN, Y.; LEACOCK, C. (Ed.). **Polysemy: theoretical and computational approaches**. Oxford: Oxford University Press, p. 30-51, 2000.

CRUSE, D.A. **Meaning in language: an introduction to semantics and pragmatics**. Cambridge: Cambridge University Press, 2000.

CRUSE, D.A. Polysemy and related phenomena from a cognitive linguistic viewpoint. In: SAINT-DIZIER, P.; VIEGAS, E. (Eds.). **Computational lexical semantics**. Cambridge: Cambridge University Press, p. 33-49, 1995.

CUYCKENS, H.; ZAWADA, B. (Ed.). **Polysemy in cognitive linguistics**: selected papers from the Fifth International Cognitive Linguistics Association, Amsterdam, 1997. Amsterdam: Benjamins, 2001.

DEANE, P. D. Polysemy and cognition. *Língua*, 75, p. 325-361, 1988.

Di FELIPPO, A. **Delimitação e alinhamento de conceitos lexicalizados no inglês norte-americano e no português brasileiro**. Tese de Doutorado em Linguística e Língua Portuguesa. Faculdade de Ciências e Letras (FCL), Universidade Estadual Paulista (UNESP). Araraquara, São Paulo, 253p., 2008.

DIAS-DA-SILVA, B. C. A construção da base da Wordnet.Br: conquistas e desafios. In. **III WORKSHOP em Tecnologia da Informação e da Linguagem Humana. XXV Congresso da Sociedade Brasileira de Computação**. São Leopoldo: Universidade Federal do Rio Grande do Sul, 2005.

DIAS-DA-SILVA, B. C. **A face tecnológica dos estudos da linguagem: o processamento automático das línguas naturais**. 272 f. Tese (Doutorado em Linguística e Língua Portuguesa)—Faculdade de Ciências e Letras, Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 1996.

DIAS-DA-SILVA, B. C. A rede WordNet e a compilação de um thesaurus eletrônico. **Fórum Lingüístico**, Florianópolis, v. 3, n. 2, p. 157-176, 2003.

DIAS-DA-SILVA, B. C. Bridging the gap between linguistic theory and natural language processing. In: INTERNATIONAL CONGRESS OF LINGUISTICS, 16., 1997, Paris. **Proceedings...** Oxford: Elsevier Sciences, 1998. p. 1-10.

DIAS-DA-SILVA, B. C. Human language technology research and the development of the brazilian portuguese WordNet. In: HAJCOVÁ, E.; KOTESOVCOVÁ, A.; MÍROVSKY, J. (Eds.): **Proceedings of the 17th International Congress of Linguists**. Prague: Matfyzpress, MFF UK, 12p., 2003.

DIAS-DA-SILVA, B. C. O estudo lingüístico-computacional da linguagem. **Letras de Hoje**, Porto Alegre, v. 41, p.103-138, 2006.

EVANS, V.; GREEN M. **Cognitive Linguistics: An Introduction**. Mahwah, NJ: Lawrence Erlbaum Associates. 2006.

EVENS, M. W. **Relational Models of the Lexicon**. Cambridge: Cambridge University Press, 1988.

FAUCONNIER, G.; TURNER, M. Blending as a central process of grammar. In: GOLDBERG, A. (Ed.). **Conceptual structure, discourse and language**. Stanford: CSLI Publications, p.113-130. 1996.

FAUCONNIER, G.; TURNER, M. Polysemy and conceptual blending. In: NERLICK, B.; TODD, Z.; HERMAN, V.; CLARKE, D.D. (Eds.). **Polysemy and Flexible Patterns of Meaning in Mind and Language**, p. 79-94, 2003.

FELLBAUM, C.; MILLER, G.A. Whither WordNet. Power Point Presentation, 2007.

FERREIRA, A. B. de H. **Dicionário Aurélio eletrônico século XXI (Versão 3.0)**. São Paulo: LexiKon Informática, 1999. 1 CD-ROM.

FILLMORE, C. J.; JOHNSON, C. R; PETRUCK, M. Background to FrameNet. **International Journal of Lexicography**. v.16, n.3, p.235-250, 2003.

GEERAERTS, D. (Ed.). **Cognitive linguistics: basic readings**. Berlin: Mouton de Gruyter, 2006.

GEERAERTS, D. **Words and other wonders: papers on lexical and semantic topics**. Berlin: Mouton de Gruyter, 2006.

GOLDBERG, A. **Constructions: a construction grammar approach to argument structure**. Chicago: The University of Chicago Press, 1995.

GUARINO, N. Understanding, Building and Using Ontologies. **KAW96 Knowledge Acquisition Workshop**. Banff, Alberta, Canada, 1997.

HAMP, B.; FELDWEG, H. GermaNet: a lexical-semantic net for German. In: **PROCEEDINGS of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications**. Madrid, 1997.

HANDKE, J. **The structure of the lexicon: human versus machine**. Berlin: Mouton de Gruyter, 1995.

HIRSCH, E. **Dividing reality**. New York: Oxford University Press, 1997.

HOUAISS, A. **Dicionário eletrônico Houaiss da língua portuguesa**. Rio de Janeiro: FL Gama Design, 2001. 1 CD-ROM.

HOUAISS, A.; VILLAR, M. de S. **Dicionário eletrônico Houaiss da língua portuguesa (versão 1.0)**. Rio de Janeiro: Objetiva, 2001. 1 CD-ROM.

JANSSEN, T. Monosemy versus polysemy. In: CUYCKENS, H.; DIRVEN, R.; TAYLOR, J. (Ed.). **Cognitive approaches to lexical semantics**. Berlin: Mouton de Gruyter, p. 93-122, 2003.

KATZ, J., FODOR, J. A. The structure of a semantic theory. **Language**, 39, p. 170-210, 1963.

KAY, P.; FILLMORE, C. Grammatical constructions and linguistic generalizations: the what's X doing Y? construction. **Language**, 75(1), p. 1-33, 1999.

KILGARRIFF, A. **Polisemy**. PhD thesis--University of Sussex, Brighton, UK, 1992. Disponível em: <<http://www.kilgarriff.co.uk/Publications/1992-K-thesis.pdf>>. Acesso em: 15 mar. 2009.

KLEIN, D. E.; MURPHY, G. L. Paper has been my ruin: Conceptual relations of polysemous senses. **Journal of Memory and Language**, 47, p. 548-570. 2002.

KLEPOUSNIOTOU, E. Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. **Journal of Neurolinguistics** v. 20:1, p. 1-24. 2007.

KNUBLAUCH, H.; FERGERSON, R. W.; NOY, N. F.; MUSEN, M. A. The protégé OWL plugin: an open development environment for Semantic Web Applications. In: **Third International Semantic Web Conference – ISWC 2004**, Hiroshima, Japan, p. 1-15, 2004.

KROVETZ R; CROFT, W.B. Lexical Acquisition and Information Retrieval. In: **Fifth Conference on Applied natural language processing**, p.380-387. Disponível em <<http://portal.acm.org/citation.cfm?id=146802.146810>>. Acesso em 30 maio 2006.

KRYMOLOWSKI, Y.; ROTH, D. Incorporating Knowledge in Natural Language Learning: A Case Study. In: **COLING-ACL'98 workshop on the Usage of WordNet in Natural Language Processing Systems**. 1998.

LAKOFF, G. **Women, fire, and dangerous things**: what categories reveal about the mind. Chicago: University of Chicago Press, 1987.

LAKOFF, G.; JOHNSON, M. **Metaphors We live by**. Chicago: University of Chicago Press, 1980.

LANGACKER, R. W. **The cognitive basis of grammar**. Berlin: Mouton de Gruyter, ed. 2, 395p., 2002.

LANGACKER, R.W. **Concept, image, and symbol**: the cognitive basis of grammar. Berlin: Mouton de Gruyter. 1990.

LANGACKER, R.W. **Foundation of cognitive grammar**. Stanford: Stanford University Press, v.1: Theoretical prerequisites, 516p., 1987.

LANGACKER, R.W. **Grammar and Conceptualization**. Berlin: Mouton de Gruyter, 2000.

LEHRER, A. Homonymy and polysemy: measuring similarity of meaning. **Language Sciences**, 3, p. 33-39, 1974.

LEHRER, A. Polysemy in derivational affixes. In: NERLICH, B. et al (Ed.). **Polysemy: flexible patterns of meaning in mind and language**. Berlin: Mouton de Gruyter, p. 217-232, 2003.

LEHRER, A. **Semantic fields and lexical structure**. Amsterdam : North Holland, 1974.

LEVELT, W. J. M. **Speaking: to intention to articulation**. Cambridge, Mass.: The MIT Press, 1992.

LYONS, J. **Semantics**. Cambridge: Cambridge University Press, 1977.

MARRAFA, P. **WordNet do Português: uma base de dados de conhecimento lingüístico**. Lisboa: Instituto Camões, 2001.

MENDES, S.; CHAVES, R. P. Enriching WordNet with qualia information. In: **Proceedings of the NAACL 2001 Workshop on WordNet and Other Lexical Resources**, Pittsburgh, June 2001. Disponível em : <<http://www.seas.smu.edu/~rada/mwnw/papers/WNW-NAACL-212.pdf.gz>>. Acesso em: 15 mar. 2009.

MICHAELIS moderno dicionário da língua portuguesa. Disponível em: <<http://michaelis.uol.com.br/moderno/portugues/index.php>>. Acesso em: 15 mar. 2009.

MILLER, G.A. Nouns in WordNet. In: FELLBAUM, C. (Ed.). **WordNet: an electronic lexical database**. Cambridge: The MIT Press, p. 23-46, 1998.

MILLER, G.A. WordNet: an online lexical database. Special issue of **International Journal of Lexicography**, 3(4), 1990.

MILLER, J.; JOHNSON-LAIRD, P. N. **Language and perception**. Cambridge, Mass.: Belknap Press of Harvard University Press, 1976.

NIRENBURG, S. e RASKIN V. **Ontological Semantics**. Mass., Cambridge: The MIT Press, 2004.

NIRENBURG, S.; RASKIN, V. **Ontological semantics**. Cambridge, Mass.: MIT Press, 2004.

NIRENBURG, S.; RASKIN, V. Ten Choices for Lexical Semantics. In. **Computer and Cognitive Science** MCC-96-304, New Mexico State University: Computing Research Laboratory, 1996.

PETERS, W.; PETERS, I.; VOSSSEN, P. Automatic sense clustering in Euro WordNet. In: **Proc. First Intl Conf on Language Resources and Evaluation**. Granada, Spain, p. 409-416, 1998.

PINKER, S. **Do que é feito o pensamento**. São Paulo: Companhia das Letras, 561p., 2008.

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

RAVIN, Y.; LEACOCK, C. (Eds.). **Polysemy: Theoretical and Computational Approaches**. Oxford: O.U.P. 2000.

RAVIN, Y.; LEACOCK, C. (Ed.). **Polysemy: theoretical and computational approaches**. Oxford: Oxford University Press, 2000.

RICE, S. Growth of a lexical network: Nine English preposition in acquisition. In: Cuyckens, H., Dirven, R. e Taylor J.R. (Eds.) **Cognitive Approaches to Lexical Semantics**. Berlin/New York: Mouton de Gruyter, p. 243-280, 2003.

RICE, S. Growth of a lexical network: nine English prepositions in acquisition. In: CUYCKENS, H.; DIRVEN, R.; TAYLOR, J. R. (Ed.). **Cognitive approaches to lexical semantics**. Berlin: Mouton de Gruyter, p. 243-280, 2003.

ROSCH, Eleanor. Cognitive representations of semantic categories. *Jornal of Experimental Psychology: General* 104, p. 192-233, 1975.

SAINT-DIZIER, P.; VIEGAS, E. **Computational lexical semantics**. Cambridge: Cambridge University Press, 1995.

SANDRA, D.; RICE, S. Network analyses of prepositional meaning: mirroring whose mind – the linguist's or the language user's?. **Cognitive Linguistics**, v. 6, n. 1, p. 89-130, 1995.

SILVA, A.S. **A Semântica de Deixar: Uma Contribuição para a Abordagem Cognitiva em Semântica Lexical**. Braga: Empresa do Diário do Minho, Ltda., 1999.

SILVA, A.S. **O Mundo dos Sentidos em Português: Polissemia, Semântica e Cognição**. Coimbra: Edições Alamedina, p.392, 2006.

SILVA, A.S. O sentido múltiplo: polissemia, semântica e cognição. In: Feltes, H.P.M. (Ed.), **Produção de Sentido. Estudos transdisciplinares**. Caxias do Sul: Editora da Universidade de Caxias de Sul, p.91-115, 2003.

TALMY, L. Lexicalization patterns: semantic structure in lexical forms. In: SHOPEN, T. (Ed.) **Language typology and syntactic description: grammatical categories and the lexicon**. (v.3). Cambridge: Cambridge University Press, p. 57-149, 1985.

TALMY, L. **Toward a cognitive semantics**. Cambridge, Mass.: MIT Press, 2 v., 2000.

TAYLOR, J. R. **Cognitive grammar**. Oxford: Oxford University Press, 2002.

TAYLOR, J. R. Cognitive models of polysemy. In: NERLICH, B. et al (Ed.). **Polysemy: flexible patterns of meaning in mind and language**. Berlin: Mouton de Gruyter, p. 31-47, 2003b.

TAYLOR, J. R. **Linguistic categorization: prototypes in linguistic theory**. London: Clarendon Press; New York: Oxford University Press, 1995.

TAYLOR, J. R. Polysemy's paradoxes. **Language Sciences**, v. 25, p. 637-655, 2003a.

TAYLOR, J.R. **Polysemy and the lexicon**. In: G Kristiansen, M. Achard, R. Dirven, & F. J.R. de Mendonza Ibanez (Eds.), *Cognitive Linguistics: Current Application and Future Perspectives*. Berlin- New York: Mouton de Gruyter. 2006, p. 51-80.

TAYLOR, J.R. Prototypes in cognitive linguistics. In: Robinson, P.; Ellis, N.C. **Handbook of cognitive linguistics and second language acquisition**. New York: Routledge, p. 39-65, 2008.

TUGGY, D. Ambiguity, polysemy, and vagueness. In: Dirk G. **Cognitive Linguistics: Basic Readings**. 2006. P. 167-184.

TYLER, A; EVANS, V. Reconsidering prepositional polysemy networks: the case of over. In: NERLICH, B. et al. (Ed.). **Polysemy: flexible patterns of meaning**. Berlin: Mouton de Gruyter, p. 95-160, 2003b.

ULLMAN, S. **Semantics: an introduction to the science of meaning**. Oxford: Blackwell, 1962.

VANDELOISE, C. Representation, prototypes, and centrality. In: TSOHATZIDIS, S. (Ed.). **Meanings and prototypes: studies in linguistic categorization**. London: Routledge, p. 403-437, 1990.

VOSSEN, P et al. Compatibility in interpretation of relations in EuroWordNet. **Computers and the Humanities**, Dordrecht: Kluwer Academic Publishers, v. 32, p. 153-184, 1998.

VOSSEN, P. EuroWordNet: a multilingual database with semantic networks. **Computers and the Humanities**, Dordrecht, v. 32., n. 2-3, p. 73-89, 1998.

VOSSEN, P.; FELLBAUM C. Universals and idiosyncrasies in multilingual WordNet. In: BOAS, H.C. (ed). **Trends in linguistics: multilingual FrameNets in computational lexicography**. Berlin: Mouton de Gruyter, P. 319-346, 2009.

WEINREICH, U. Webster's third: a critique of its semantics. **International Journal of American Linguistics**, New York, v. 30, p. 405-409, 1964.

WEISZFLOG, W. (Ed.). **Michaelis português: moderno dicionário da língua portuguesa: versão 1.0**. São Paulo: DTS Software Brasil, 1998. 1 CD-ROM.

ZLATEV, J. Polysemy or generality? Mu. In: CUYCKENS, H.; DIRVEN, R.; TAYLOR, J. R. (Ed.). **Cognitive approaches to lexical semantics**. Berlin: Mouton de Gruyter, p. 447-494, 2003.