

Validade empírica das redes de polissemia para o significado preposicional

Empirical validity of polysemy networks for prepositional meaning

Aparecida de Araújo Oliveira

Universidade Federal de Viçosa, Viçosa, Minas Gerais / Brasil
cidaaraujo007@gmail.com

Pedro Ivo Vieira GoodGod

Universidade Federal de Viçosa, Rio Paranaíba, Minas Gerais / Brasil
pedro.god@ufv.br

Resumo: O presente artigo avalia a validade empírica de uma rede de polissemia proposta para a preposição *em* do português do Brasil, comparando as perspectivas do linguista e de usuários leigos da língua sobre categorização. A rede baseou-se no modelo cognitivo de Rede Esquemática, de R. Langacker (1987, 2008), com relações de esquema/instância e de extensão semântica, e envolveu 48 frases da internet, que representavam 24 padrões de uso de *em* definidos pelo linguista. Participaram como sujeitos em um experimento psicolinguístico 32 estudantes de graduação, sem formação em linguística, todos eles falantes nativos de português do Brasil. Esses sujeitos classificaram as frases do *corpus* segundo semelhanças percebidas no significado da preposição. Em entrevistas individuais realizadas após a tarefa, os participantes relataram suas estratégias de categorização. Uma matriz de dissimilaridade criada com base na classificação feita pelos informantes foi submetida a análises de agrupamentos pelos métodos de Ward (1963) e de Tocher (*apud* RAO,

1952).¹ Este revelou sete categorias criadas pelos informantes, quatro das quais apresentaram motivação semântica mais forte que as outras três. O número de grupos formados variou entre os participantes, assim como variou o número de frases nos grupos criados. Além da coerência nas ligações entre as categorias, os resultados revelaram um nível relevante de isomorfismo entre a rede proposta originalmente e a avaliação dos informantes, corroborando a estrutura relacional dessa rede e uma parte significativa de sua granularidade. Em geral, este estudo também comprovou a maior saliência do ESPAÇO sobre os demais domínios, assim como a saliência do TEMPO entre domínios não espaciais.

Palavras-chave: validade empírica; categorização; rede de polissemia; preposição; experimento psicolinguístico.

Abstract: This paper evaluates the empirical validity of a polysemy network proposed for the preposition *em* in Brazilian Portuguese, by comparing the linguist's and the non-specialist language user's views of categorization. The network was based on the Schematic Network cognitive model of R. Langacker (2008, 1987), with scheme/instance and semantic extension relationships, and involved 48 sentences taken from the internet, which represented 24 usage patterns of *em* as defined by the linguist. Thirty-two undergraduate students with no training in linguistics, all native speakers of Brazilian Portuguese, took part as subjects in a psycholinguistic experiment in which they sorted these sentences according to similarities perceived in the meaning of *em*. After completing this task, the participants reported their categorization strategies in individual interviews. A dissimilarity matrix based on the participants' classification was subjected to cluster analyses by the Ward (1963) and Tocher (cited in RAO, 1952) methods. The latter revealed seven categories created by the participants, four of which exhibited stronger semantic motivation than the other three. Participants differed in the number of groups they formed, and each participant created groups of different sizes. As well as coherent connections among categories, the results showed an important level of isomorphism between the originally proposed network and the participants' evaluation, corroborating the relational structure of the network and a significant part of its granularity. Overall, the study also confirmed the greater predominance of SPACE over the other domains and of TIME among non-spatial domains.

¹ Rao (1952) é considerado referência original do método de Tocher. Ele expõe a proposta de K. D. Tocher; porém, não informa a fonte consultada.

Keywords: empirical validity; categorization; polysemy network; preposition; psycholinguistic experiment.

Recebido em: 20 de maio de 2016.

Aprovado em: 23 de setembro de 2016.

1 Introdução

Além de explicações coerentes para os processos cognitivos relacionados ao uso da linguagem, a Linguística Cognitiva procura oferecer confirmações do modelo “baseado no uso”, o qual faz parte de sua essência. Como mostram Gries *et al.* (2005) e Gries *et al.* (2010), nos últimos anos, pesquisadores desse quadro teórico têm-se empenhado em validar empiricamente os muitos construtos e princípios que norteiam esse ramo da Linguística, por meio da metodologia de *corpus* ou da experimentação psicolinguística, ou pela conjugação de ambas.

Em particular, os experimentos psicolinguísticos podem revelar semelhanças e diferenças entre usuários da língua, no acesso ao conteúdo conceitual. Essas diferenças podem ser descritas como variações entre os aspectos linguísticos que recebem maior atenção consciente e entre os níveis diferenciados de prática com a introspecção linguística (TALMY, 2005, p. 2, 11). O presente artigo descreve um experimento realizado com o objetivo de comparar a análise introspectiva do linguista sobre a polissemia da preposição *em* no uso corrente do português do Brasil e a análise intuitiva de falantes nativos leigos sobre os mesmos dados. O construto cuja validade se pretende testar é o modelo de Rede Esquemática de polissemia (LANGACKER, 2008, 1987), descrito a seguir, na seção 3.

As preposições têm sido tema de inúmeras investigações em Semântica Cognitiva, apoiadas em um ou outro modelo de rede de polissemia, que, por sua vez, baseiam-se no conceito de categorias prototípicas (ROSCH, 1978). De acordo com essa visão, o polo semântico de uma palavra ou construção geralmente é formado por um conjunto de sentidos interligados, dos quais um ou mais se destacam por sua saliência cognitiva. Entre esses estudos, encontram-se Brugman (1981), Lakoff (1987), Deane (1992), Cuyckens (1993), Dewell (1994), Vandeloise (1991), Kreitzer (1997), Teixeira (2001), Tyler e Evans (2003) e Evans e Tyler (2004).

Embora várias dessas pesquisas tenham se baseado no uso real da língua e empreguem modelos considerados coerentes, Sandra e Rice (1995) criticam a grande variedade desses modelos e, principalmente, questionam sua presumida natureza cognitiva. De modo particular, esses autores chamam a atenção para a pretensa compatibilidade entre a visão do pesquisador e a do falante leigo. Seu trabalho inspirou a presente investigação.

2 A polêmica sobre as representações mentais

Sandra e Rice (1995, p.100-102) discutem dois componentes distintos da realidade psicológica dos modelos de rede. O primeiro diz respeito aos processos mentais envolvidos na categorização como mecanismo de mudança da língua, os quais podem ser verificados diacronicamente tanto por meio das relações entre usos mais antigos e novos usos de um mesmo item quanto pelos estudos sobre a aquisição da língua. O segundo componente cognitivo refere-se à realidade das representações mentais. Trata-se do objeto central do referido artigo, que questiona e avalia o grau de correspondência entre as redes de polissemia propostas por diferentes pesquisadores e o léxico mental do usuário da língua.

Uma versão forte de isomorfismo assume uma correspondência total entre os nódulos e ligações na rede e uma representação dessa estrutura semântica armazenada na memória do falante, hipótese considerada de difícil comprovação. Entretanto, Sandra e Rice avaliam que a abundância de conceitos da psicologia na literatura linguística parece corroborar a lógica das ligações propostas nos modelos de rede mais comuns. Tyler e Evans (2003), por exemplo, consideram que seu modelo de rede coincide com o léxico mental dos falantes, porque não inclui certas distinções contextuais, como ocorre na descrição de *over* (LAKOFF, 1987):

[...] nem todos os padrões de uso estão contidos na rede semântica. Enquanto parte da variação em usos de uma palavra é *necessariamente* instanciada na memória de longo prazo e, dessa forma, persiste na rede semântica, alguns usos são criados *on-line*, no desenrolar da interpretação normal dos enunciados.² (TYLER; EVANS, 2003, p. 7, ênfase e tradução nossas).

² “[...] not all usages are contained within the semantic network. While some of the variation in uses of a word must be instantiated in long-term memory, and hence

Outros teóricos, entre os quais se encontra Langacker (1987), rejeitam o isomorfismo por acreditarem que o nível de aprofundamento descritivo proposto para as redes pelos linguistas pode não ser percebido de igual maneira pelo falante leigo. Langacker pondera que algumas distinções e algumas semelhanças em níveis mais abstratos podem não ser percebidas pelo falante leigo, porque “alguns domínios são intrinsecamente salientes e abundantes no processamento cognitivo a exemplo daqueles pertencentes ao espaço e à visão.” (LANGACKER, 1987, p. 380, tradução nossa).³ Além disso, estruturas pertencentes a domínios muito abstratos são menos salientes que aquelas pertencentes a domínios ligados à percepção sensorial. Acredita-se que tal fato se reflita na força e amplitude dos agrupamentos formados pelos informantes do presente estudo.

Entre os pesquisadores que negam o isomorfismo entre rede e léxico mental, Sandra e Rice (1995, p. 103) identificam dois grupos, com opiniões distintas. Por um lado, alguns reconhecem o caráter psicológico da rede nos processos cognitivos envolvidos no surgimento de novos usos, mas rejeitam a ideia de que padrões de uso sejam representados no léxico mental dos falantes.

Por outro lado, há os que afirmam que algumas características das redes, tais como seu alto nível de refinamento ou “granularidade”, estão presentes na memória de longo prazo e que os falantes *seriam capazes de perceber distinções* entre categorias mais amplas e entre categorias mais finas. Entretanto, não necessariamente perceberiam *semelhanças* entre grandes categorias, como as que ocorrem entre domínios, porque a transparência dessas relações se perdeu no tempo. No presente trabalho, ao se propor a comparação entre o processo de categorização informado por conhecimentos teóricos e aquele realizado intuitivamente, essa última visão de isomorfismo foi assumida como premissa.

3 Ponto de partida: a rede de polissemia

A polissemia é entendida como um fenômeno no qual dois ou mais sentidos de uma palavra, relacionados entre si, são associados a

persist in the semantic network, some uses are created on-line in the course of regular interpretation of utterances”.

³ “Some domains are intrinsically salient and pervasive in cognitive processing (e.g. those pertaining to space and vision)”.

uma única forma linguística (SILVA, 2006, p.10) em um processo de categorização. Por exemplo, o termo “papel” pode designar diferentes conceitos, como ‘matéria fabricada com fibras vegetais’ ‘documento’, ‘ação, função’. O primeiro desses sentidos é considerado a origem de extensões semânticas que geram os demais por meio de processos cognitivos, como a metonímia e a metáfora. Na Linguística Cognitiva, os diversos sentidos de uma palavra constituem uma categoria prototípica (ROSCH, 1978), com elementos cognitivamente mais salientes – o(s) protótipo(s) – e outros mais periféricos. As categorias prototípicas se organizam por relações de semelhanças de família, à maneira de diferentes tipos de “jogos” que formam uma única categoria com esse nome, sem apresentarem uma característica que seja comum a todos (WITTGENSTEIN, 1999 [1953]). Como Teixeira (2001, p. 61) salienta, o protótipo é a “estrutura conceptual modelar que, para os falantes, corresponde prioritariamente a uma determinada conceptualização”. Não se restringe, portanto, a entidades do mundo físico; explica também o significado de verbos, preposições, advérbios, enfim, “todas as organizações mentais linguisticamente traduzíveis”.

A polissemia contrasta, primeiramente, com a monosseμία, que ocorre quando uma palavra ou expressão linguística tem apenas um sentido, e as diferenças de significado, se houver, são elaboradas pela interpretação do contexto. Esses dois fenômenos distinguem-se da homonímia, caracterizada pela existência de dois signos linguísticos com a mesma forma fonológica e/ou gráfica, com significados distintos não relacionados. A homonímia ocorre com “calo” (substantivo) e “calo” (verbo), os quais têm pronúncia e grafia idênticas, mas significam, respectivamente, uma ‘região da camada exterior da pele que se encontra mais espessa e endurecida’ e ‘faço silêncio’. A homonímia pode ser resultado de uma coincidência histórica, a qual tenha levado duas palavras de etimologias distintas a assumirem a mesma forma com o passar do tempo. Sincronicamente, pode ser causada pelo esvaziamento da relação percebida entre os sentidos de um termo com uma origem comum (DUBOIS *et al.* 2009, p. 326-327).

Langacker (2008, p. 37; 1987, p. 74) interpreta o conceito de categoria prototípica por meio do seu Modelo de Rede Esquemática de polissemia, o qual é empregado no presente estudo. A Rede Esquemática se organiza em termos de esquema/instância (padrão mais abstrato/conceito mais específico) e em termos de sancionamento de novos usos, alguns

dos quais podem se tornar unidades linguísticas convencionalizadas. A esquematização implica a categorização dos usos linguísticos de forma hierárquica, visto que o usuário da língua percebe padrões de uso e forma esquemas, que ele organiza em níveis distintos de abstração. Esses mesmos padrões são usados para incluir (sancionar) novos usos em uma ou outra categoria da língua.

A natureza prototípica do modelo de Langacker decorre dos tipos de sancionamento previstos: total ou parcial. Quando o sancionamento é total, os membros da categoria semântica incorporam perfeitamente a descrição do esquema ou noção de nível superior que instanciam (ou elaboram). A noção mais abstrata de localização é compartilhada por várias outras mais específicas, tais como os esquemas imagéticos [EM CIMA DE], [DENTRO DE] e [ATRÁS DE].

Entretanto, essa mesma noção de localização pode ser estendida para outro domínio conceitual, como o TEMPO. Nesse caso, ocorre extensão semântica com sancionamento parcial, para um domínio sem as mesmas características do ESPAÇO. Do mesmo modo, os conceitos de especificação e de localização fazem parte da rede de polissemia da preposição *em*, e a relação entre eles é também de sancionamento parcial. Na perspectiva assumida no presente estudo, e em consonância com a Teoria da Metáfora Conceitual (LAKOFF; JOHNSON, 1980) e a das Metáforas Primárias (GRADY, 1997), a relação de especificação que *em* evoca em expressões como “especialista *em* direito trabalhista” é uma extensão metafórica de um uso espacial dessa preposição, oriunda de certa consequência funcional de alguns usos locativos de *em*. Com a repetição do uso, a preposição passou a ser empregada com o novo significado mesmo na ausência da configuração espacial motivadora.

Em termos de categorização prototípica, na cultura brasileira, o esquema de [LARANJA] é um exemplo mais saliente de [FRUTA] que o de [TOMATE]. O primeiro desses conceitos – e não o segundo – pode ser então o protótipo da categoria fruta. Igualmente, o sentido espacial de localização é considerado protótipo na rede de polissemia de *em*. Isso ocorre porque esse sentido espacial tem maior saliência cognitiva. Por sua vez, especificação constitui um significado mais periférico na categoria por ser mais abstrato e, conseqüentemente, menos saliente, como explica Langacker na seção anterior.

A polissemia da preposição *em* foi analisada originalmente com base nos seguintes critérios semânticos: domínio conceitual, esquema

imagético evocado nos usos espaciais ([CONTENTOR] e [TRAJETO]) e suas subespecificações em noções topológicas (inclusão, contato, adjacência). Foram identificados sentidos esquemáticos – especificação e localização – correspondentes a conceitos com um nível de abstração superior, mais frequentes na rede, que perpassam o significado da preposição em diferentes usos e que são derivados do esquema imagético de [CONTENTOR]. Abaixo desses dois sentidos mais amplos, a maior parte das distinções é de natureza contextual.

Nessa análise introspectiva, foram considerados locativos metafóricos os usos temporais de *em* e aqueles tendo eventos, atividades e estados emocionais como complemento da preposição, nos quais os elementos relacionados não são objetos concretos. Nesses casos, a noção de localização emerge porque os conceitos ligados pela preposição *em*, ainda que abstratos, teoricamente são conceitualizados como objetos metafóricos, o que permite que se estabeleça entre eles alguma relação locativa (LAKOFF; JOHNSON, 1980). Por outro lado, usos de especificação – portanto, não locativos – também são considerados metafóricos pelo linguista, mesmo que digam respeito à cor, ao material ou à forma de objetos concretos, tendo em vista que o conceito de especificação, em si, não é espacial, mas reflete uma possível restrição ao significado dos objetos especificados. Como já mencionado, segundo a concepção teórica adotada, especificação emerge de um efeito funcional (GRADY, 1997) do esquema de [CONTENTOR], como uma restrição imposta ao conteúdo.

Na descrição proposta, o antecedente da preposição representa o trajeto, e o conseqüente, o marco, conforme Langacker (1987), representados no exemplo “placa_{TRAJETOR} no acostamento_{MARCO}”. Essa terminologia está associada à assimetria figura e fundo, da Teoria da Gestalt, e implica uma série de distinções na conceitualização dessas entidades. Talmy (2000a, p. 183-185) define a figura (trajeto) como a entidade “potencialmente móvel ou em movimento”, cuja localização ou trajeto percorrido se discute. O fundo (marco) é a “moldura de referência ou entidade estática dentro de uma moldura de referência”. Em virtude dessa função, o marco aparece mais cedo na cena/na memória, é percebido mais rapidamente, e sua geometria tende a ser mais elaborada na conceitualização. O marco também tende a ser mais estático e mais conhecido do falante. Por essas razões, suas propriedades espaciais tendem a ser mais relevantes para o significado da preposição.

4 Análise prévia das frases

As frases empregadas no experimento psicolinguístico foram adaptadas de usos retirados da internet e representam 24 categorias semânticas pré-definidas para o significado da preposição *em*. Das categorias e subcategorias usadas no experimento, dezessete são locativas, nove das quais pertencem ao domínio espacial, três, ao temporal, e cinco, a outros domínios concretos e abstratos. Entre os usos locativos, foram incluídos dois pares de frases em contexto dinâmico (com trajetória): um par representando o domínio espacial e outro, um domínio não espacial. A proporção entre usos estáticos e dinâmicos é explicada adiante e tem a ver com as categorias obtidas em pesquisa de *corpus* realizada anteriormente (OLIVEIRA, 2009). Com o sentido amplo de especificação, são testadas sete subcategorias de domínios concretos e abstratos, consideradas como usos metafóricos emergindo do efeito pragmático de controle. A maior parte da variação em contextos espaciais pôde ser explicada pela interpretação (*construal*, em LANGACKER, 1987) e pela relação funcional Contentor/conteúdo (VANDELOISE, 1991), e os demais, por meio de processos metafóricos e metonímicos.

A escolha das categorias de análise segue certos princípios teóricos resumidos na sequência. Oliveira (2011) apresenta uma explicação mais detalhada da rede de polissemia de *em* no português do Brasil, por meio do esquema de [CONTENTOR].

Sentido esquemático de localização:

1.1 Domínio espacial:

1.1.1 Localização estática no interior de um contentor (o marco da preposição é conceitualizado como um objeto tridimensional que envolve o trajetor):

1.1.1.1 Inclusão total:

- a) Carregava um pão fresquíssimo no saco de papel pardo.
- b) Bebê nasce num carro por falta de ambulância.

1.1.1.2 Inclusão parcial (marco envolve parte do trajetor):

- a) O gelo no uísque destruiu seu teor alcoólico.
- b) Ainda havia muitos icebergs no mar.

1.1.1.3 Inclusão de um trajetor vazio:

- a) O buraco na parede foi feito com dois pedaços de ferro.
- b) A miniatura possui um pequeno trincado no parabrisa.

1.1.1.4 Inclusão em um contendor não canônico:

- a) Entenda os protestos do setor rural na Argentina.
- b) Os alfabetos orientais atrapalham a ciência nesses países.

1.1.2 Inclusão em um contendor ao final de um movimento (inclui a noção de trajetória):

- a) Joguei os livros na bolsa.
- b) Você deve colocar pouco ar no balão.

1.1.3 Contato em posição canônica no eixo vertical (trajetor fica acima do marco, gerando um efeito funcional de suporte):

- a) Havia faixas estendidas no gramado.
- b) O governo dá incentivo para quem plantar no telhado.

1.1.4 Contato com mudança da ordem dos objetos no eixo vertical ou da orientação do eixo:

- a) Quantos funcionários públicos são necessários para se trocar uma lâmpada no teto da repartição?
- b) O jornalista não pode considerar o Brasil e Minas Gerais como apenas um retrato na parede.

1.1.5 Localização pontual (trajetor se localiza em um ponto determinado do marco):

- a) O senhor não viu a placa no acostamento?
- b) Use um grampo na ponta da linha para facilitar a troca de isca.

1.1.6 Adjacência:

- a) Ela estava sentada numa mesa do Café Capricieux.
- b) A Pousada Villa das Pedras fica na estrada que liga Brasília a Pirenópolis.

1.2 Localização metafórica (trajetor e marco abstratos são conceitualizados como objetos):

1.2.1 Domínio temporal:

1.2.1.1 Coincidência com o intervalo de tempo:

- a) Trabalhei nesses 15 anos de pesquisa sem parar.
- b) Nesses tempos de globalização, Turquia e Brasil fazem parte da categoria de países emergentes.

1.2.1.2 Localização pontual no tempo:

- a) O Overmundo completou um ano no último dia 7.
- b) A chegada da família real no Brasil em 1808 aprimorou o serviço postal da então colônia.

1.2.1.3 Localização no final de um intervalo de tempo:

- a) Quero te ver em meia hora.
- b) Você será redirecionado para a página do Instituto em 5 segundos.

1.2.2 Estados emocionais conceitualizados como contentores:

- a) Na angústia, o homem experimenta a finitude de sua existência humana.
- b) Eu me sentia bem naquela solidão.

1.2.3 Metáfora do conduto (as palavras são contentores com informação semântica):

- a) Um pouco sobre civilização nas palavras de Bertrand Russell.
- b) Se quiser continuar comigo é nesses termos.

1.2.4 Atividades são conceitualizadas como contentores:

- a) Não há muita perspectiva do Brasil [sic] participar nesses grandes projetos.
- b) Dejetos usados na fertilização degradam microbacias.

1.2.5 Eventos são pontos no tempo:

- a) Há esperança de contar com o meia Roger no clássico contra o Santos.
- b) Ayrton enfrenta Alemão no paredão do BBB.

1.2.6 Mudança de estado (a passagem de um estado para outro é um trajeto metafórico saindo de um contentor para outro):

- a) O uso de drogas pode constituir-se em um caso de dependência.
- b) Jonas transformou-se no primeiro desaparecido político brasileiro.

2 O sentido esquemático de especificação (metafóricos não locativos):

2.1 Tipo:

- a) É cirurgião-dentista, especialista em Dentística Restauradora.
- b) Bem-vindo ao Programa de Pós-graduação em Direito da UFRGS.

2.2 Meio ou instrumento:

- a) Esta página é melhor visualizada em Internet Explorer 5 ou superior.
- b) Fala e escreve bem em inglês e francês.

2.3 Material:

- a) Abajur decorado em porcelana fria.
- b) Cadeira de couro com estrutura em tubo de aço.

2.4 Estrutura interna:

- a) Vendo som completo ou em partes.
- b) Os anúncios das seções Imóveis são publicados em subseções.

2.5 Forma do objeto:

- a) Robô em forma de lagarta simula os movimentos do bicho.
- b) Pesquisador inventou uma bateria que carrega energia em estado sólido.

2.6 Forma da atividade:

- a) Rodou em círculos, tropeçou e proferiu blasfêmias.
- b) O ECA pressupõe um atendimento em rede.

2.7 Cor:

- a) A chita ou o algodãozinho estampados nessas cores desbotam facilmente.
- b) Os designers trabalharam para criar um vestido em vermelho vibrante.

Os exemplos 1.1.1.4.a e 1.1.1.4.b apresentam marcos concretos bidimensionais (*Argentina; esses países*) que, portanto, não são contedores canônicos. Entretanto, o fato de apresentarem uma região interior e um limite que a separa do exterior possibilita sua conceitualização como marcos delimitados. Tyler e Evans (2003, p. 184) assim justificam o uso de *in* tendo campos, desertos, países, cidades, entre outros como marcos. Sandra e Rice (1995, p. 110) classificam esses locais como marcos de “dimensionalidade desconhecida”.

São apresentados dois tipos de localização espacial, sendo uma estática e outra dinâmica. Essa classificação se baseia em Talmy (2000a, p. 180-191). Segundo ele, na conceitualização de cenas espaciais, um objeto pode se encontrar em duas “disposições espaciais” básicas: ele pode ficar estacionário ou se deslocar em relação ao fundo (movimento translativo). Isso coloca os conceitos de localização e movimento no centro da representação linguística do espaço. Como mencionado anteriormente, na descrição de uma cena espacial, os elementos de classes fechadas e a estrutura das orações distinguem aquilo que é figura daquilo que é fundo. Nessa mesma obra, Talmy (2000a, p. 182) associa a disposição da figura parada à sua localização e a da figura em movimento ao seu trajeto. Em ambos os casos, é possível que a língua também distinga sua orientação. Dessa forma, os usos espaciais representados anteriormente se enquadram em uma ou outra categoria.

Segundo tipologia proposta por Talmy (2000b, p. 221) para a representação dos eventos de movimento, as línguas “aparentemente se dividem em [...] duas categorias, com base no padrão característico no qual a estrutura conceitual do macroevento é mapeada para a estrutura sintática”. (Tradução nossa)⁴

⁴ “The world’s languages generally seem to divide into a two-category typology on the basis of the characteristic pattern in which the conceptual structure of the macro-event is mapped onto syntactic structure.”

Em linhas bastante gerais, os idiomas que codificam, no verbo, o esquema central do movimento (portanto, o trajeto) são chamados “de moldura de verbo”; aqueles que codificam o trajeto do movimento fora do verbo, através de um satélite, são chamados “de moldura de satélite”.

De acordo com essa tipologia, a língua portuguesa compartilha com as demais neolatinas um padrão de lexicalização predominante do primeiro tipo (BATORÉO, 2000), como demonstram os exemplos abaixo, de Silva e Batoréo (2010, p. 242, Figura 13), que comparam os paradigmas das línguas neolatinas com o das germânicas:

Português

1. O João atravessou o rio **a nado**. ?? * O João **nadou** através do rio.
 MOVIMENTO+PERCURSO MODO MOV+MODO PERCURSO

Inglês

2. ??(*) John crossed the river **swimming**. (4) John **swam** across the river.
 MOVIMENTO+PERCURSO MODO MOV+MODO PERCURSO

Dessa forma, nos exemplos 1.1.2.a e 1.1.2.b, os verbos (*joguei; colocar*) expressam o trajeto do movimento do trajetor em direção ao marco e a preposição *em* codifica a localização dessa figura ao final desse movimento. O emprego de *em* é predominante estático no português do Brasil, mas compete com certos usos da preposição *a* em contextos dinâmicos. Em comparação, no português europeu, pode-se exprimir movimento apenas com *a* (“ida ao Brasil”), que também codifica localização estática – ‘junto a’ (“o fio ao pescoço”). Segundo Batoréo (2000, p. 444), variedades africanas do português também utilizam *em* no lugar de *a*.

As classes 1.1.3 e 1.1.4 representam variações que são significativas para outras preposições do português. Contato é uma noção frequentemente associada a *sobre* e a *em cima de*. Entretanto, apenas nos exemplos 1.1.3, com objetos dispostos na ordem canônica no eixo vertical (*faixas estendidas no gramado; plantar no telhado*), pode haver troca de *em* por *sobre* ou *em cima de*. A inclusão desses exemplos visava testar se os informantes levariam em conta essa variação na configuração espacial. Vandeloise (1991) discute a relevância dessas mudanças na semântica das preposições francesas *sur* e *sous*.

Por fim, alguns autores, a exemplo de Tyler e Evans (2003, p. 178-179), sugerem que, na interpretação de determinadas cenas espaciais, a conceitualização do marco como um ponto unidimensional, servindo

como referência de localização, justificaria certos usos locativos de preposições como *at* no inglês. Vandeloise (1991, p. 5-6 e capítulo 11) prefere uma definição funcional de localização para um marco adimensional de *à* do francês, associada à noção de ritual. O objetivo neste ponto do presente estudo foi testar a possibilidade de a localização ser interpretada como pontual em função da configuração do trajetor (*placa*) em 1.1.5.a, e em função das configurações do trajetor (*grampo*) e do marco (*ponta*) em 1.1.5.b.

5 O experimento psicolinguístico

O experimento psicolinguístico baseia-se naquele relatado por Sandra e Rice (1995) sobre as preposições *in*, *on* e *at* do inglês, e, do mesmo modo, envolve a leitura de enunciados originalmente produzidos na modalidade escrita. Porém, diferentemente do experimento de Sandra e Rice, o presente estudo inclui entrevistas individuais realizadas logo após a execução da tarefa de classificação.

5.1 Objetivos, raciocínio e hipóteses

Visou-se descobrir (i) o nível de granularidade que os informantes pudessem perceber entre sentidos na rede e (ii) a possibilidade de os informantes perceberem as relações entre categorias de domínios.

Em tarefas de classificação como a empregada no presente estudo, os participantes são submetidos a um determinado número de estímulos (neste caso, frases) para serem agrupados com base em algum princípio pré-determinado (neste caso, semelhança de significado). O raciocínio por trás do experimento é o de que as eventuais distinções e associações entre usos, feitas pelos informantes, refletem diferenças contextuais e algumas representações existentes em seu léxico mental, tais como a distinção entre o *em* locativo e o de especificação e as associações entre grupos maiores. Entretanto, o experimento revela, principalmente, as estratégias de categorização empregadas pelos sujeitos.

A hipótese nula admitia duas possibilidades:

1. Os informantes não perceberiam qualquer diferença de significado, no nível das representações mentais ou no das distinções contextuais, demonstrando uma atitude monossêmica em relação à preposição. Isso implicaria um único sentido mais vago, porém mais saliente, que pudesse ser especificado por outros itens lexicais presentes na frase, ou mesmo, por elementos extralinguísticos.

2. A maioria dos agrupamentos seria aleatória, visto que os informantes apenas tentariam cumprir a tarefa de agrupar, sem se preocupar com o critério definido. Essa hipótese é testada na entrevista.

Como hipótese de pesquisa, e em consonância com a rede proposta, acreditava-se que os informantes formassem grupos de maneira coerente, demonstrando sua capacidade de discriminação, e que haveria grupos grandes e pequenos. Ainda de acordo com a premissa das redes de polissemia, esperava-se que as distinções mais finas fossem mais fortes, ou seja, que mais informantes fizessem distinções finas, e menos informantes fizessem distinções entre categorias maiores. Como consequência, os grandes grupos e os pequenos grupos se formariam em pontos diferentes da escala de dissimilaridade que constitui o gráfico da Figura 1. Também se esperava que muitos informantes fossem capazes de perceber algum tipo de semelhança entre grupos menores (estrutura relacional da rede) e que, provavelmente, poucos percebessem relações entre grupos maiores.

5.2 Participantes, método e conteúdo das frases

A tarefa *off-line* de classificação foi proposta a 32 falantes nativos de português do Brasil, de ambos os sexos, que concordaram formalmente em participar da pesquisa, autorizando a divulgação dos resultados. Todos eles eram calouros universitários de cursos variados, sem formação prévia em Letras, com idade média de 22 anos na ocasião da coleta. As variáveis extralinguísticas não são consideradas.

Em sessões individuais, cada informante recebeu 48 frases numeradas aleatoriamente (Figura 1), em tiras individuais de 4 cm X 19 cm, com a palavra *em* (*no, na, naquela*, etc.) sublinhadas. Todos foram informados de que essas frases deveriam ser agrupadas por critério de semelhança de sentido da palavra sublinhada e de que o número de grupos e o de palavras em cada grupo seriam de livre escolha do participante, podendo inclusive haver um só grupo com todas as frases ou grupos com apenas um elemento. A partir desse ponto, cada um deles dispôs de tempo livre para executar a tarefa, o que levou entre 30 e 45 minutos.

Para testar o maior refinamento da rede, foram incluídos pares de sentenças que diferiam de outros por características tais como o tipo de inclusão (parcial ou total), a natureza do contato entre entidades, o tipo de localização no tempo e, ainda, pela distinção aspectual pontual/durativo entre eventos e atividades. Acredita-se que essa última diferença seja bem menos saliente para o falante leigo que a distinção entre tempo e espaço, por exemplo.

Por sugestão da Profa. Heliana R. Mello (comunicação pessoal), foram evitados certos enunciados nos quais a motivação semântica para o uso sincrônico de *em* fosse pouco transparente ou a preposição ocorresse com verbos cuja regência muitas vezes só se consolida na linguagem do falante por meio do ensino formal.⁵

5.3 Tratamento dos dados

As respostas de todos os informantes foram tabuladas em matrizes simétricas individuais com valores binários (0 e 1), nas quais o valor “zero” significa diferença, e o valor “um” representa semelhança entre frases. A partir da soma das trinta e duas matrizes, obteve-se uma matriz de coincidências (o número de vezes que determinado par foi formado pelos informantes). A divisão da primeira matriz pelo total de informantes produziu uma segunda matriz, de frequências relativas, cujo complemento aritmético possibilitou o cálculo da última matriz, de dissimilaridade, que foi então submetida a uma análise de agrupamentos, pelo método de Ward (1963). Esse método de classificação foi sugerido pelo Prof. Stephan Gries, em comunicação pessoal.

Análise de agrupamentos é um termo genérico para se referir a procedimentos estatísticos que possibilitam a formação de grupos ou categorias de objetos e, nesse caso, de frases. Tipicamente, cada objeto pertence a um único grupo e o conjunto de todos os agrupamentos contém todos os objetos (EVERITT *et al.*, 2001), como ocorrido no presente estudo.

O método proposto por Ward (1963, p. 236-238) forma grupos hierárquicos de conjuntos mutualmente exclusivos, segundo o princípio de se obter o máximo de similaridade possível entre os membros, no tocante às características medidas. Esse método de agrupamento parte de n conjuntos, cada um contendo um indivíduo apenas. Considera-se que o máximo de informação esteja disponível nesse estágio e, portanto, o desvio dentro de cada conjunto seja igual a zero. À medida que são formados pares de indivíduos, mantém-se o desvio padrão dentro de cada grupo menor que aquele entre os grupos. Assim, reduz-se o número de conjuntos ($n-1$) enquanto se mantém a perda de informação no valor mínimo possível. Seguindo esse princípio, esses novos subconjuntos podem receber um novo membro, ou outro pareamento ($n-2$) pode ser feito, preservando-se o mesmo princípio de semelhança ótima dentro de cada grupo. Esse procedimento

⁵ Exemplos de usos que não foram aproveitados: *Luciano não confiava na esposa. A declaração do motorista o implicou no caso.*

pode ser repetido até que todos os membros, separados individualmente no início, fiquem juntos em um só grupo ($n, n-1, \dots, 1$).

Ward denomina o procedimento “agrupamento hierárquico”, em função dos inúmeros estágios, os quais podem ser representados por um diagrama bidimensional conhecido como dendrograma, ou diagrama de árvore, contendo todas as fusões ou divisões ocorridas em cada etapa. O método também possibilita que seja observada a perda de informação a cada nova fusão. No presente estudo, pode-se observar a distância entre agrupamentos, tomando-se por referência a escala na parte inferior da Figura 1.

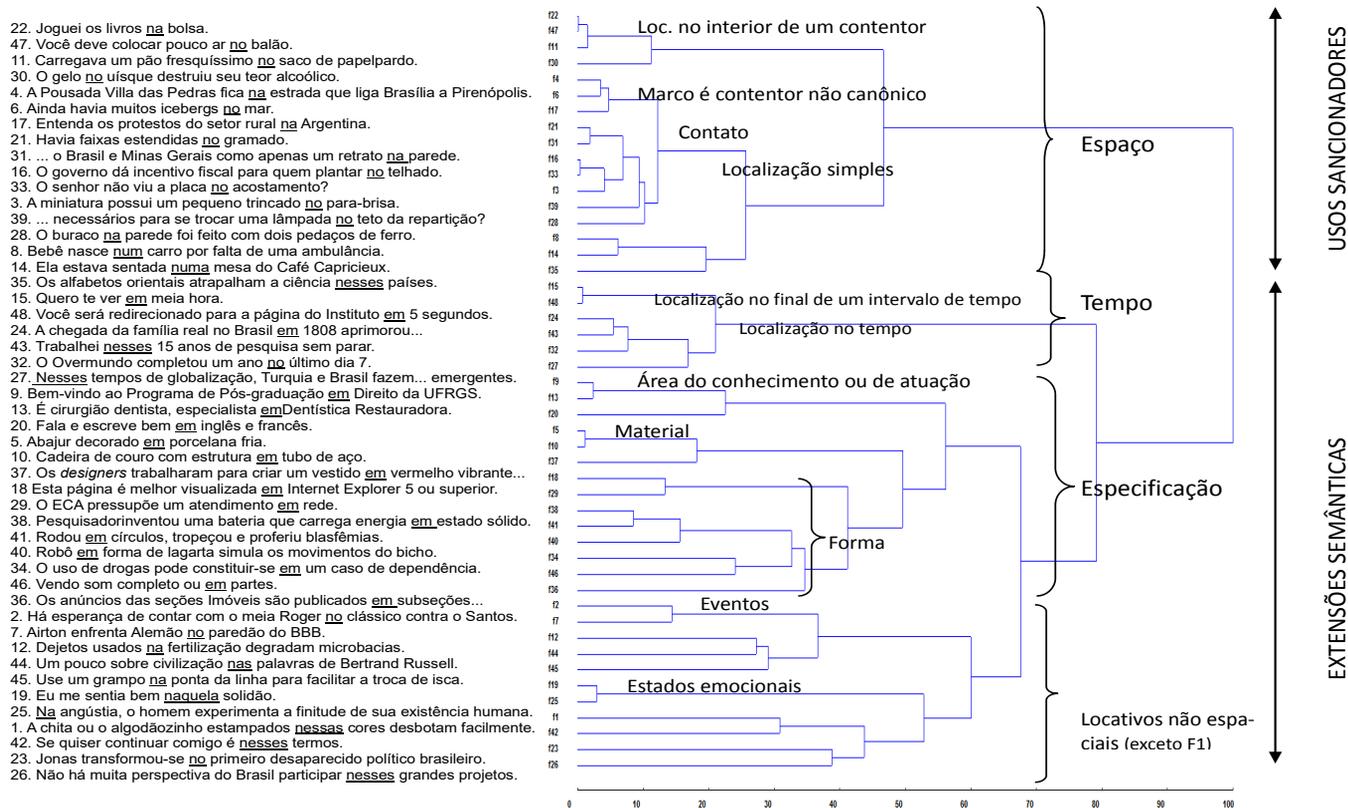
Na sequência, submeteu-se a mesma matriz de dissimilaridade ao método de otimização de Tocher (*apud* RAO, 1952), e o resultado é apresentado na Tabela 1 e na Figura 2. Conforme descrito por Rao (1952, p. 363), o método sugerido por K. D. Tocher inicia o processo de agrupamento com dois objetos muito próximos em termos de semelhança (que neste estudo, são pares de frases associadas por muitos participantes). A seguir, o método associa outro objeto com a menor distância média em relação aos dois primeiros objetos. Depois disso, um quarto objeto é comparado aos três anteriores de acordo com o mesmo princípio de menor distância, e assim por diante. Sempre que um novo objeto entra para o grupo, a distância média de dissimilaridade entre os membros desse grupo é recalculada. Um novo grupo é formado quando um objeto apresenta uma distância média maior que a média de distância entre os membros do grupo pré-existente.

Esse método foi útil na obtenção de um ponto de corte para definir o nível de granularidade das escolhas dos informantes.

6 Análise

O dendrograma gerado pelo *software* Genes (CRUZ, 2006), apresentado na Figura 1, é o resultado da análise hierárquica sobre os dados coletados. Sua estrutura reflete a força das relações entre as sentenças, como percebidas pelos informantes, e a escala de valores (0-100) abaixo da árvore demonstra a proporção dos participantes que agruparam dois usos pelo critério de semelhança. Cada nóculo representa um agrupamento e o comprimento dos ramos representa as distâncias nas quais os grupos são formados. A análise partiu de uma matriz de dissimilaridade e, portanto, os valores mínimos na escala de distâncias refletem o grau máximo de semelhança obtido: um grupo que se forma próximo de zero apresenta um nível muito alto de semelhança. Os rótulos dos objetos analisados (f1=frase (1); f2=frase (2), etc.) aparecem junto aos nóculos terminais.

FIGURA 1 – Dendrograma gerado pelo Método Ward de agrupamento, com os níveis de semelhança obtidos para os usos de *em*



6.1 A coerência das categorizações

Levando-se em conta as definições de polissemia, monossemia e homonímia apresentadas na introdução deste artigo, uma visão geral do dendrograma mostra que a hipótese nula não se confirma em nenhum de seus aspectos. Primeiramente, a possibilidade de ocorrer monossemia no sentido forte foi descartada, uma vez que vários grupos coerentes se formaram em diferentes níveis, isto é, no geral, os informantes não consideraram que os usos da preposição tivessem um só significado. Essa visão seria confirmada caso a maioria dos nódulos tivesse se formado próximo ao ponto “zero” da escala (semelhança máxima), na Figura 1, e se a Figura 2 e a Tabela 1 apresentassem um só grupo.

Nosso resultado é similar àquele observado por Sandra e Rice (1995, p. 108). Segundo esses autores, seus “sujeitos não produziram agrupamentos monolíticos”, e, em nosso experimento com *em*, apenas uma participante reuniu todas as frases em um único grande grupo. Essa decisão não foi aleatória, entretanto, e o sentido esquemático encontrado não foi localização, como pode ser visto na transcrição da entrevista.

Sj15: Olha, eu agrupei todas as frases em um só grande grupo porque eu acho que toda vez que as... essas palavras sublinhadas aparecem, *acontece uma especificação na frase*. Como se os elementos da frase estiverem dentro de uma categoria. [...] Por exemplo, na frase “Joguei os livros na bolsa”. “Na bolsa”. Eu poderia ter jogado os livros no chão ou jogado os livros na mesa. [...] “Robô em forma de lagarta simula os movimentos do bicho.” É uma especificação do robô: “em forma de lagarta”. O robô poderia ter outra forma. [...] “Se quiser continuar comigo é nesses termos”. Se continuar comigo nesses... poderia ser em outros termos. [...] “Cirurgião dentista especialista em dentística restauradora.” Ele poderia ser especialista em... em outra coisa.

Aquilo que a informante denomina especificação é um valor semântico único, mais vago, instanciado por usos locativos ou não, nos diferentes domínios representados no experimento. Portanto, todas as distinções (localização, especificação – por forma, material, cor, etc.) abaixo desse nível são menos salientes para ela.

A possibilidade de os grupos terem sido constituídos aleatoriamente também não se confirmou, visto que os agrupamentos foram coerentes, e todas as justificativas apresentadas nas entrevistas foram igualmente coerentes, baseadas em critérios semânticos. Por exemplo, Sj6 e Sj19

usam palavras distintas (sublinhadas) para se referirem à ideia de localização, mas ambos perceberam esse sentido como uma categoria.

Sj6: O grupo um, é, eu *reuni por*; *por local*, com *um sentido de, de lugar* que... têm essas frases. Você gostaria que eu lesse essas frases?

Sj 19: Eu usei o critério de... é... qual... é... qual a outra palavra que o “em” poderia ser substituído. [...] no grupo quatro, por exemplo, o “em” *pode ser substituído pela palavra “onde”*. [...] E no dois são... *sempre vem com características*. É sempre características do... da... do que foi falado antes. Ouve só: Abajur decorado em porcelana fria. É uma característica da... da decoração.

Abaixo, a informante Sj7 descreve usos não espaciais da preposição que ela agrupou em uma só categoria diferente de duas outras formadas antes. Ela emprega o critério de domínio evocado (por exemplo, o TEMPO), mas reserva a noção de localização para usos espaciais. Para os demais casos, ela procura uma categoria “guarda-chuva”, que possa incorporar usos em domínios menos salientes. Além disso, Sj7 apoia-se no marco para decidir sobre o tipo de relação criada a cada vez.

Sj7: A meu ver, *esse aqui já não é nem uma coisa [lugar] nem outra [tempo]*. Não tem assim... uma característica que me chama mais a atenção. [...] Olha, *eles dão mesmo coisas mais vagas*. [...] Por exemplo... igual a esse. “Em rede”. Rede é uma coisa muito ampla. Muito... Não é uma coisa específica. Então, *eu fiz esse grupo assim, de coisas não específicas*. [...] é uma coisa assim... um caso é uma coisa vaga, igual essa outra frase. A 34. Um caso é uma coisa vaga. Entendeu? Então, assim... coisas mais vagas. [...] “Em partes”, “nas palavras”, “termos”, “círculos”, são todas coisas mais vagas. Não é uma coisa precisa.

Também se observa que os nódulos do dendrograma distribuem-se ao longo da escala na Figura 1, ou seja, nenhum dos 32 informantes devolveu as 48 sentenças separadas. O mesmo foi relatado por Sandra e Rice (1995, p. 108) para as três preposições estudadas.⁶ Entretanto, apesar

⁶ Em seu artigo (p.108), Sandra e Rice discutem os resultados obtidos para a preposição *in*. Entretanto, segundo eles, “exactly the same conclusions apply to the other two prepositions [*at, on*] as well”.

das indicações do gráfico, não se pode eliminar com certeza a ideia de homonímia máxima, tampouco afirmar que o resultado obtido seja uma resposta a um componente da tarefa de classificação. Foi solicitado aos participantes que agrupassem as frases “de acordo com SEMELHANÇA DE SENTIDO”, embora as instruções também incluíssem a informação “sem limite máximo ou mínimo para o número de grupos ou para o número de frases em cada grupo”.

No dendrograma da Figura 1, a coerência se manifesta na organização dos grupos. Logo de início, observa-se uma distinção principal entre usos espaciais e não espaciais. Além disso, a grande maioria dos usos espaciais está reunida no alto dessa figura, indo de f22 a f35. Esse grupo é seguido por outro menor, de f15 a f27, que incorpora todos os usos temporais. Mais abaixo, vê-se um grande grupo de usos não locativos de especificação, de f9 a f36, com uma única exceção no uso de estado final em f34. Essa frase talvez não seja exceção de fato, tendo em vista que foi associada a grupos de usos com o sentido de forma, e adquirir uma nova forma é um significado possível para “constituir-se em...”. Por fim, surge outro grupo de locativos abstratos, que vai de f2 a f26, com apenas uma exceção: um uso da categoria cor em f1.

As entrevistas e o gráfico na Figura 1 confirmam, assim, a hipótese de haver coerência na categorização dos significados da preposição pelos informantes. Essa coerência entre os grupos também é demonstrada pelo método de otimização de Tocher (Tabela 1 e Figura 2). A análise multidimensional revelou sete diferentes grupos, quase todos com base em alguma dimensão semântica. O grupo <1> representa majoritariamente localização espacial; o grupo <2> contém majoritariamente usos de especificação; o grupo <3> somente apresenta usos temporais; o grupo <4> representa majoritariamente estados emocionais. O grupo <5> contém dois usos locativos abstratos – estado final e um denominado previamente metáfora do conduto. Como se observa na Figura 1, os grupos <4> e <5> se distinguem pela grande saliência da relação entre f19 e f25 em comparação com f42, que completa o grupo <4> na Tabela 1. Essa saliência inexistente no grupo <5>. É possível que a formação dos três últimos grupos (<5>, com dois objetos, e <6> e <7>, com um único objeto cada) tenha sido causada pela maior esquematicidade da relação expressa pela preposição, incluindo um uso não locativo de cor. Dessa forma, não fica muito clara a motivação semântica por trás dos três últimos grupos formados pelo método de Tocher. Contudo, a parte inferior do dendrograma na Figura 1 fornece certa lógica para esses agrupamentos, que são majoritariamente locativos não espaciais e menos transparentes que outros locativos.

TABELA 1 – Formação dos agrupamentos pelo método de Tocher

GRUPO	ACESSOS
<1>	22 47 11 6 33 21 16 31 3 4 28 17 14 39 8 30 7 45 2 12
<2>	5 10 37 40 38 41 29 46 36 13 9 20 34 18
<3>	15 48 24 32 43 27
<4>	19 25 42
<5>	23 44
<6>	26
<7>	1

Fonte: A pesquisa.

6.2 Níveis de granularidade

Após se confirmar a capacidade de discriminação e de esquematização dos informantes, discute-se agora o nível de granularidade nas suas classificações e a força de coesão entre os grupos. Como se pôde observar na Figura 1, os participantes também diferenciaram inicialmente duas grandes categorias: uma de usos espaciais (que denominamos “sancionadores”) e outra de não espaciais (que denominamos “extensões de sentido”). Além disso, dezesseis deles mencionaram o termo “especificação” em suas entrevistas.

Entre os grandes domínios mencionados na seção anterior, como esperado, o ESPAÇO foi o domínio mais saliente, seguido pelo TEMPO, formando grupos muito nítidos nos dois gráficos. Nas entrevistas, vinte dos trinta e dois sujeitos fizeram menção clara ao domínio espacial, e dezenove, ao temporal, em contraste com onze que mencionaram os estados emocionais, e seis, a categoria cor. Na entrevista, alguns chegaram a assumir a diferença de nível de saliência, como no excerto abaixo:

Sj 29: O grupo quatro eu achei mais fácil que é mais... *no local, em algum lugar.*

Quanto às distinções mais finas, alguns grupos menores foram confirmados por um número maior de informantes e aparecem perto do zero da escala de dissimilaridade no dendrograma. Esses grupos pequenos, porém altamente coesos, representam subcategorias do domínio espacial, do domínio temporal e da categoria especificação.

O primeiro desses subgrupos pode ser identificado bem no alto do dendrograma, separado dos demais usos espaciais. Ele contém exemplos de inclusão ou contenção, começando com f22 e f47, os quais representam ações de movimento para dentro de um CONTENTOR (31/32 associações), recebendo depois o uso estático f11 (28/32 associações) e, por fim, f30 (22/32 associações). Os marcos tridimensionais bem delimitados – *bolsa*, *balão* e *saco de papel* – e os trajetores tangíveis resultaram na grande saliência da noção de contenção. Essa saliência é menor no exemplo f30, porque o marco – *uísqe* – não apresenta limites claros.

f22. Joguei os livros na bolsa.

f47. Você deve colocar pouco ar no balão.

f11. Carregava um pão fresquíssimo no saco de papelpardo.

f30. O gelo no uísque destruiu seu teor alcoólico.

Embora a geometria do trajetor tenda a ser caracterizada de forma mais simplificada que a do marco, distinções mais finas no domínio espacial foram relacionadas por Sandra e Rice (1995, p. 109-110) não apenas ao número de dimensões do marco, mas também à natureza tangível do trajetor. Esse é um resultado que corrobora a premissa das redes de polissemia sobre a influência das propriedades dessas entidades na formação dos grupos. Em seu experimento com a preposição *in*, as frases 3-5 abaixo formaram um grupo separado de (6), em que um trajetor “intangível” ou vazio (*hole*) ocorre com um marco bidimensional (*sweater*), e (7), em que as dimensões do marco (*Japan*) são “desconhecidas”.

3. *Are you putting onions in the stew?* [Você vai colocar/está colocando cebolas no ensopado?]

4. *Don't put that in your mouth.* [Não ponha isso na sua boca.]

5. *My pen is in the drawer.* [Minha caneta está na minha gaveta.]

6. *There's a hole in your sweater.* [Tem um buraco no seu blusão.]

7. *In Japan, they eat raw fish.* [No Japão, eles comem peixe cru.]

Em nosso experimento com *em*, trajetores vazios, ou intangíveis, não foram encontrados entre as instâncias prototípicas do esquema de [CONTENTOR]. F3 – *trincado no para-brisa*– e f28 – *buraco na parede* – foram tratados genericamente como outros usos apenas locativos, em grupos com menor coesão, como se vê no dendrograma. A separação entre f28 e f31 (*retrato na parede*) resultou da concretude do trajetor *retrato*, já que ambas as frases apresentam o mesmo marco.

Devido à maior esquematicidade de *em* comparado às preposições inglesas *in* e *on*, o trajetor concreto de f31 e os marcos bidimensionais *parede* e *gramado* contribuíram para a formação de um segundo subgrupo com f21 com base nas noções de contato e suporte, com 27/32 associações. Esse subgrupo demonstra que os participantes não levaram em conta as orientações distintas do eixo sobre o qual se alinham os objetos, como proposto na análise prévia.

f21. Havia faixas estendidas no gramado.

f31. [...] o Brasil e Minas Gerais como apenas um retrato na parede.

Os exemplos f16 – *plantar no telhado* – e f33 – *a placa no acostamento* – também foram considerados altamente semelhantes (29/32 associações) e se juntaram ao par f21-f31 no estágio seguinte. Embora fatores como a centralidade da localização dos objetos no significado de f16 e de f33 possam ter reforçado a ligação entre essas frases, a proximidade entre os quatro usos na perspectiva dos falantes demonstra coerência que pode ser atribuída à presença de um marco bidimensional.

Os demais usos espaciais se agruparam próximos do início da escala, em torno do ponto 10. O alto grau de coesão entre os subgrupos espaciais e de semelhança entre seus elementos confirma, mais uma vez, a saliência desse domínio básico em nossa experiência.

Fora do domínio espacial, também foram observados pares com grande força coesiva. Ainda que não tenha sido mencionada nas entrevistas, a categoria localização no final de um período de tempo (f15 – *Quero te ver em meia hora* – e f48 – *Você será redirecionado para a página do Instituto em 5 segundos*) formou-se em um nódulo muito próximo de zero na escala de distâncias, com 28/32 associações. Já era prevista essa grande saliência do TEMPO, e esse par, especificamente, é formado por usos que apresentam a maior semelhança entre si no tocante ao conteúdo temporal, já que ambas as frases evocam um tempo futuro quase imediato.

Fora dos domínios do ESPAÇO e do TEMPO, é relativamente saliente (26/32 associações) a semelhança percebida entre f19 – *naquela solidão* – e f25 – *na angústia* –, cujos marcos são estados emocionais. A associação entre esses usos era esperada porque a EMOÇÃO constitui um domínio abstrato que se destaca no cotidiano das pessoas.

No grande grupo de especificação, duas distinções mais finas podem ser observadas. A primeira delas é um subgrupo de especificação

no sentido restrito (f9 – *pós-graduação em direito* – e f13 – *especialista em dentística restauradora* –, com 27/32 associações). Essa categoria, chamada “tipo” na análise prévia, pode ser justificada pela estabilidade da construção *em* + “área de conhecimento” modificando substantivos (e verbos) de um único campo lexical. A segunda distinção inclui as especificações em f5 – *abajur decorado em porcelana fria* – e f10 – *estrutura em tubo de aço* –, consideradas semelhantes pela maioria dos participantes (28/32 associações). Como pode ser observado no dendrograma, esse último é o subgrupo com a maior coesão dentro de especificação. Sem dúvida, mesmo não se tratando de uma relação espacial, a preposição *em* introduz uma propriedade muito concreta e saliente desses objetos, que é o material do qual eles são fabricados.

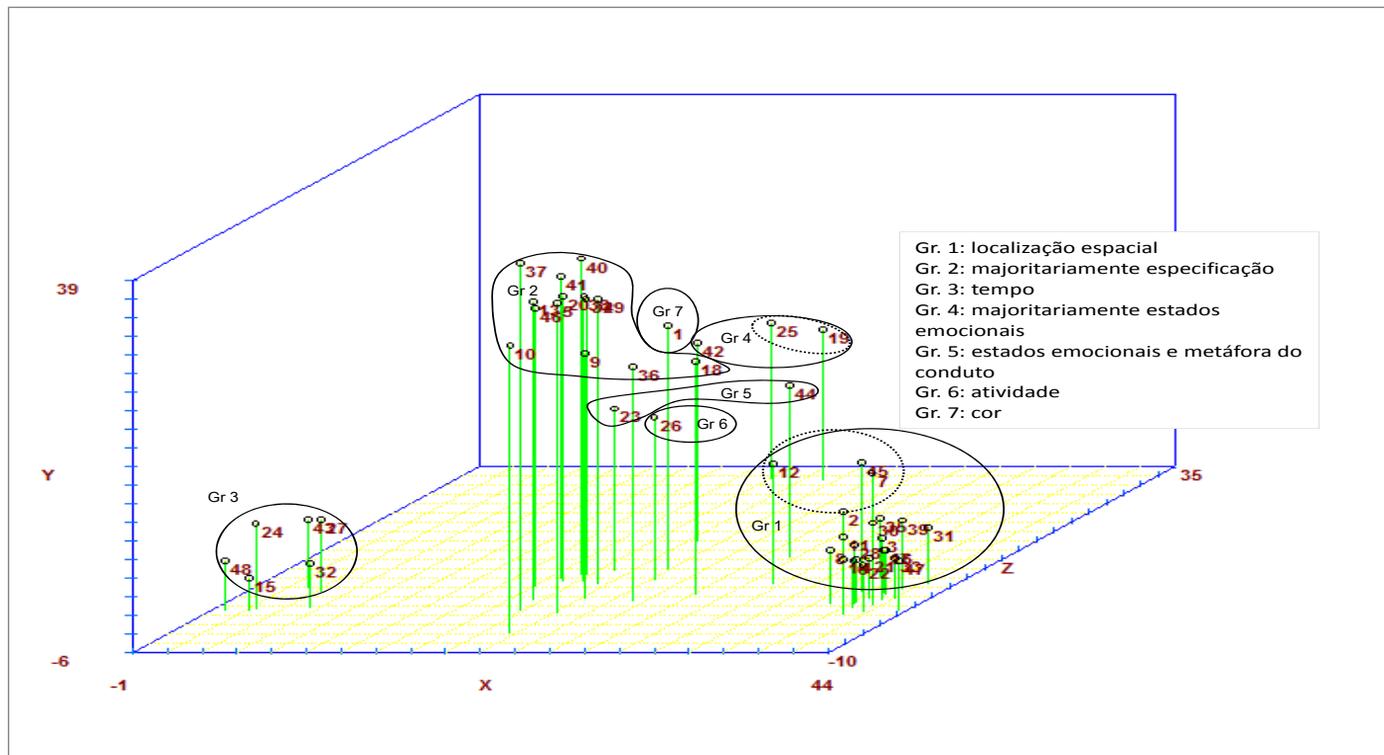
Por fim, os informantes também estabeleceram relações menos fortes de semelhança dentro de grupos maiores. Um exemplo disso é o par com nível intermediário de coesão (20/32 associações), formado por f2 – *o meia Roger no clássico contra o Santos* – e f7 – *Alemão no paredão do BBB* –, cujos marcos são eventos. Outro caso menos saliente (14/32) é a associação entre os usos f38 a f36, relacionados à forma de entidades e de atividades. Com força de coesão menor que a existente entre usos nos domínios espacial e temporal, ainda assim, essas associações se destacam entre categorias não locativas. Também se observou que certos grupos não locativos com muita força interna ficaram nitidamente separados de outros subgrupos do mesmo grupo maior. Certamente, esses são domínios com grande saliência cognitiva, mesmo não representando experiências primárias.

Corroborar-se, assim, a hipótese de que um maior número de informantes perceberia distinções mais finas entre os grandes grupos. Da mesma maneira que nos resultados obtidos por Sandra e Rice (1995, p. 108), confirma-se, no presente estudo, a premissa de que algumas categorias de usos são mais facilmente distinguíveis que outras.

6.3 A estrutura relacional

A capacidade de os participantes perceberem a estrutura relacional refere-se à coerência da ligação entre usos de categorias distintas. Menos intensamente que a relação entre domínios, essa capacidade pôde ser comprovada pela formação de um grande grupo de locativos e não locativos, separado dos usos espaciais, por volta do ponto 70 na escala do dendrograma (Figura 1). Outro exemplo foi a associação de locativos espaciais e não espaciais no grupo <1> do gráfico do método de Tocher (Figura 2).

FIGURA 2 – Representação 3D dos grupos de frases formados pelo método de otimização de Tocher



Fonte: A pesquisa.

A elipse pontilhada dentro do grupo <1> contém os usos não espaciais f2 – *no clássico contra o Santos* –, f7 – *no paredão do BBB* – e f12 – *na fertilização* –, além do uso espacial f45 – *na ponta da linha*. Levando-se em conta que se trata da união de quatro usos, 9/32 é um número razoável de associações do sentido esquemático de localização às categorias de atividade e de eventos e, ainda, a localização pontual no espaço. Essa associação também pôde ser observada no discurso de alguns informantes:

Sj 19: “Vendo som completo ou em partes”. “Dejetos usados na fertilização...”, “um pouco sobre civilização nas palavras...”. *Eu caracterizaria como “referencial” do físico. Um local abstrato, não físico.*

Sj30: Aqui é *lugar*. Igual “na parede”, “no uísque”, “na mesa”, em vários lugares. E aqui, “visualizada em Internet”. É o lugar que ela é visualizada. Eu achei que estava referindo a algum lugar. “nesses países”, “no paredão do *big brother*”, “no para-brisa”. *Em todas eu achava que estava relacionada a um lugar.*

Outra evidência de que os falantes perceberam relações entre domínios diferentes é o fato de eles terem produzido um grupo com vários tipos de especificação (16/32). Eles associaram instrumento (f20 – *em francês*) a dois usos bastante claros de especificação propriamente dita (f9 – *pós-graduação em direito* – e f13 – *especialista em dentística restauradora* – com 27/32 associações), e, ainda, f5 e f10 (material, 28/32) a f37 (cor), a associação entre esses três últimos tendo ficado em 19/32.

Por outro lado, não se pode afirmar com certeza que os falantes tenham percebido alguma relação entre as categorias maiores ou se houve influência do método estatístico. Tome-se como exemplo a separação total entre os usos do ESPAÇO e os demais usos em geral (extensões de sentido na Figura 1). Os métodos hierárquicos começam sempre com todos os elementos em um só grupo. A partir daí, inicia-se uma série de subdivisões, ou então, ao contrário, todos os objetos são considerados diferentes no início do cálculo e vão sendo agrupados até formarem um só grupo contendo todos eles. Essa poderia ser a razão para a ausência de um nóculo unificador entre grupos maiores dentro dos limites da escala no dendrograma.

6.4 Comparação com o modelo de rede proposto

Além dos grandes grupos e dos pequenos subgrupos altamente coesos (seção 6.2) discutidos até este ponto do texto, que também foram previstos pelo linguista, as escolhas dos informantes revelaram diferentes níveis de granularidade, coerência na formação de grupos e identificação de uma estrutura relacional. Assim, em grande parte, as respostas no experimento coincidiram com a primeira proposta de Rede Esquemática, ainda que certas distinções mais finas não tenham sido reproduzidas pelos participantes, como já previsto pelo próprio Langacker (2008, p. 37). Essas distinções/não distinções mais finas são discutidas a seguir.

De fato, os informantes não marcaram claramente a diferença entre inclusão total e parcial. Pouco mais da metade deles (18/32) associaram f30 – *gelo no uísque* – a f6 – *icebergs no mar* –, consideradas usos de inclusão parcial na análise prévia. A primeira foi associada a usos típicos de inclusão em um marco definido, e a segunda ficou em um grupo com marcos que não são contentores prototípicos. O mesmo pode ser dito sobre f4 – *a pousada fica na estrada* – e f17 – *os protestos na Argentina* –, que apresentam marcos bidimensionais não canônicos (ver seção 4 acima).

Além disso, contraria a análise prévia a associação discutida na seção 6.2, dos dois usos com trajetores vazios (f3 – *o trincado no parabrisa* – e f28 – *o buraco na parede*) a outros tipos de localização espacial mais vaga. Como consequência, na rede de polissemia, é possível que os usos 1.1.1.3 e 1.1.1.4 devam ser excluídos da classe de localização estática em um contentor.

As associações explicadas na seção 6.2, sobre tipos diferentes de inclusão, de contato e de localização pontual, refletem a tendência dos informantes de classificarem esses usos como localização não específica, indo de encontro ao nível de refinamento da análise prévia. Essa conclusão também se apoia em trechos das entrevistas como este:

P: No mar, na estrada, numa mesa, no paredão. Você entende todas essas aí como dando uma ideia de localização.

Sj21: É isso aí. Com certeza. Então, “O senhor não viu a placa no acostamento?” Local, lugar. “Airtton enfrenta Alemão no paredão”. “Ela estava sentada numa mesa de café”. “A pousada fica numa estrada que liga Brasília a Pirenópolis”. “Os alfabetos orientais atrapalham a ciência *nesses* países”.

Para efeito de descrição linguística, a classificação feita pelos participantes sugere que os pares 1.1.3, 1.1.4 e 1.1.5, propostos na análise prévia, possam ser fundidos em uma categoria de usos espaciais de localização com contato. De fato, nem a língua nem os sujeitos envolvidos fazem distinção no emprego de *em* quanto ao eixo de orientação ou quanto ao tipo de localização (pontual e não pontual).

Não houve associação entre mudanças de estado e alguma ideia de movimento abstrato. F23 – *Jonas transformou-se no primeiro desaparecido político* – e f34 – *o uso de drogas pode constituir-se em um caso de dependência* – tiveram apenas 9/32 associações e aparecem em grupos distintos de acordo com a análise pelo método de Tocher.

Por outro lado, no que diz respeito à distinção aspectual entre eventos e atividades, 20/32 associações foram feitas entre f2 – *no clássico* – e f7 – *no paredão do BBB* –, que são eventos bem definidos. Ao contrário, as atividades, que são durativas e sem limites claros, apareceram dispersas na parte inferior do dendrograma (f12 – *na fertilização* – e f26 – *nesses grandes projetos*), formando grupos menos coesos com exemplos de metáfora do conduto (f44) e estado final (f26). Dessa forma, também fora do domínio espacial, o mais delimitado demonstra ser cognitivamente mais saliente.

Como mostra a Figura 2, o domínio temporal formou um grupo fortemente coeso e também apresentou subdivisões, tais como o subgrupo de usos temporais previsto na primeira rede, descrevendo eventos em um futuro próximo, diferenciados de outros usos temporais. Porém, não houve discriminação entre pontualidade e duração, como o comprova o considerável grau de semelhança obtido (24/32), neste domínio, entre f24 – *a chegada da família real em 1808* – e f43 – *trabalhei nesses quinze anos*. A pontualidade no tempo foi tratada por alguns sujeitos como uma instância de especificação geral, como se lê a seguir.

Sj21: ... pra mim, *essas palavras dão o sentido de estar especificando as coisas*. Por exemplo, “Bem vindo ao programa de pós-graduação em direito da UFRGS.” Então, aqui está especificando. Pós-graduação em. Está especificando. “Cirurgião dentista especialista em dentística restauradora”. “Trabalhei nesses quinze anos de pesquisa”. Está especificando. Não foi neste nem naquele. [...] O mesmo caso aqui. “O overmundo completou um ano no último dia sete”, “Cadeira de couro com estrutura em tubo de aço”. Está especificando, qual tubo? De aço. Fica um pouco confuso.

7 Comentários finais

O presente artigo apresentou os resultados de um experimento proposto como segunda abordagem empírica a uma rede de polissemia. Nele se demonstrou a interpretação intuitiva de 32 falantes leigos nativos do português do Brasil, em relação aos usos da preposição *em*. Embora a negação da hipótese de homonímia possa ter sido enfraquecida pela inclusão da expressão “semelhança de sentidos” na instrução dada aos participantes, foi possível descartar a hipótese de monossemia, tendo sido obtido, neste caso, um resultado muito parecido com aquele relatado por Sandra e Rice (1995). No todo, observou-se que os participantes classificaram as frases de forma coerente.

Levando-se em conta o objetivo (i) – identificar o nível de granularidade que os falantes são capazes de perceber entre sentidos na rede –, observou-se que foi mais forte a distinção entre usos espaciais (sancionadores) e não espaciais (extensões).

Com relação aos primeiros, ficou comprovada a importância das propriedades físicas dos objetos relacionados, em particular os marcos, como critérios de classificação dos usos, mas também dos trajetores em alguns casos, como observaram Sandra e Rice em seu experimento. Entidades concretas mais delimitadas refletem mais facilmente efeitos pragmáticos como contenção e suporte, reforçando as hipóteses de classificação dos participantes. Essa informação é ainda mais relevante no caso da preposição *em*, cuja semântica é mais vaga que a das preposições da língua inglesa analisadas por esses pesquisadores. Com relação às extensões semânticas, os participantes diferenciaram, com maior nitidez, usos temporais, usos envolvendo estados emocionais e usos não locativos relacionados à percepção sensorial.

Também ficou demonstrada a capacidade dos participantes para fazer distinções finas, dentro e fora do domínio espacial, embora em menor grau que na rede proposta originalmente. Entre as distinções feitas pelo linguista, que não se concretizaram na tarefa, estão inclusão total versus parcial no domínio do ESPAÇO, e aspecto pontual versus durativo, NO TEMPO.

Com relação ao objetivo (ii) – verificar se as relações entre categorias de domínios seriam realmente percebidas –, demonstrou-se também que, embora uma noção comum de localização tenha sido percebida pelos participantes, essa noção foi suplantada pela distinção entre domínios, especialmente o ESPAÇO, o TEMPO e os usos de especificação.

Essas observações confirmam um isomorfismo parcial, em particular quanto ao padrão de estrutura e das relações típicas do modelo de Rede Esquemática de polissemia.

Finalmente, acerca dos procedimentos de análise, o ponto de corte na escala de distâncias do dendrograma é subjetivo, conforme relatam Everitt *et al.* (2001). Para obtê-lo, foi de crucial importância o emprego combinado do método hierárquico de Ward (1963) e da análise multidimensional de Tocher (*apud* RAO, 1952), o que possibilitou a divisão de grupos nos moldes descritos na seção 5.3.

Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por bolsas de doutoramento e de estágio sanduíche para a primeira autora.

Aos informantes da pesquisa, alunos da Universidade Federal de Viçosa (Rio Paranaíba, MG) e da Funedi-UEMG (Divinópolis, MG), e ao Prof. Maurício J. Faria (Funedi-UEMG), pois, sem eles, nossa pesquisa não poderia ser realizada.

Aos professores Heliana R. Mello (UFMG) e Stephan Gries (UCSB) por suas sugestões mencionadas ao longo deste texto. Vale ressaltar que quaisquer inadequações percebidas pelo leitor nesses pontos são de inteira responsabilidade dos autores deste artigo.

Aos revisores anônimos, cujas sugestões contribuíram imensamente para o aprimoramento de nosso texto.

Referências

BATORÉO, H. J. *Expressão do Espaço no português europeu*: contributo psicolinguístico para o estudo da linguagem e cognição. Lisboa: Fundação Calouste Gulbenkian; Fundação para a Ciência e Tecnologia, 2000.

BRUGMAN, C. *The story of “over”*: polysemy, semantics and the structure of the lexicon. 1981. 116 f. Dissertação (Mestrado em Linguística). University of California at Berkeley, 1981.

CRUZ, C. D. *Programa GENES*: análise multivariada e simulação. Viçosa: Editora da UFV, 2006.

CUYCKENS, H. The Dutch preposition *in*: a cognitive-semantic analysis. In: ZELINSKY-WIBBELT, C. (Ed.). *The semantics of prepositions: from mental processing to natural language processing*. Berlin/New York: Mouton de Gruyter, 1993. p. 27-71. DOI: <https://doi.org/10.1515/9783110872576.27>.

DEANE, P. D. Polysemy as the consequence of internal conceptual complexity: The case of *over*. *Proceedings of the Eastern States Conference on Linguistics (ESCOL)*, v. 9, Ohio State University, 1992.

DEWELL, R. B. *Over again: image-schema transformations in semantic analysis*. *Cognitive Linguistics*, ICLA – International Cognitive Linguistics Association, v. 5, n. 4, p. 351-380, 1994.

DUBOIS, J. *et al. Dicionário de Linguística*. Tradução de Izidoro Blikstein *et al.* 11. ed. São Paulo: Cultrix, 2009.

EVANS, V.; TYLER, A. Spatial experience, lexical structure and motivation: the case of *in*. In: RADDEN, G.; PANTHER, K. (Ed.). *Studies in linguistic motivation*. Berlin/New York: Mouton de Gruyter, 2004. p. 157-192.

EVERITT, B.; LANDAU, S.; LEESE, M. *Cluster analysis*. 4. ed. London: Arnold, 2001. DOI: <https://doi.org/10.1002/9781118887486.ch6>.

GRADY, J. E. *Foundations of meaning: primary metaphors and primary scenes*. 1997. 299 f. Tese (Doutorado) – University of California at Berkeley, 1997.

GRIES, S. T.; HAMPE, B.; SHÖNEFELD, D. Converging evidence: bringing together experimental and corpus data to the association of verbs and constructions. *Cognitive Linguistics*, ICLA – International Cognitive Linguistics Association, v. 16, n. 4, p. 635-676, 2005.

GRIES, S. T.; HAMPE, B.; SHÖNEFELD, D. Converging evidence II: more on the association of verbs and constructions. In: NEWMAN, J.; RICE, S. (Ed.). *Empirical and experimental methods in cognitive/functional research*. Stanford, CA: CSLI, 2010. p. 59-72.

KREITZER, A. Multiple levels of schematization. A study in the conceptualization of space. *Cognitive Linguistics*, ICLA – International Cognitive Linguistics Association, v. 6, n. 4, p. 291-325, 1997.

LAKOFF, G. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: The University of Chicago Press, 1987. DOI: <https://doi.org/10.7208/chicago/9780226471013.001.0001>.

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

LANGACKER, R. W. *Cognitive Grammar: a basic introduction*. Oxford: Oxford University Press, 2008.

LANGACKER, R. W. *Foundations of cognitive grammar: theoretical prerequisites*. Stanford: Stanford University Press, 1987. v. 1.

OLIVEIRA, A. A. Mapeamentos metafóricos a partir do esquema imagético de contentor: o caso de ‘em’. *Revista da Associação Brasileira de Linguística*, Niterói, v. 10, n. 2, p. 43-65, 2011.

OLIVEIRA, A. A. *Relações semântico-cognitivas no uso da preposição ‘em’ no português do Brasil*. 2009. 309 f. Tese (Doutorado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

RAO, C. R. *Advanced statistical methods in biometric research*. New York: J. Wiley, 1952.

ROSCH, E. Principles of categorization. In: ROSCH, E. L.; LLOYD, B. (Ed.). *Cognition and categorization*. Hillsdale, N.J.: Erlbaum, 1978. p. 27-48.

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

SILVA, A. S. *O mundo dos sentidos em português: polissemia, semântica e cognição*. Coimbra: Almedina, 2006.

SILVA, A. S.; BATORÉO, H. J. Gramática Cognitiva: estruturação conceptual, arquitectura e aplicações. In: BRITO, A. M. (Org.). *Gramática: história, teorias, aplicações*. Porto: Faculdade de Letras da Universidade do Porto, 2010. p. 229-251.

TALMY, L. Foreword. In: GONZALEZ-MARQUEZ, M.; COULSON, S.; SPIVEY, M. (Ed.). *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, 2005. p. xi-xxi.

TALMY, L. *Toward a cognitive semantics: concept structuring systems*. Cambridge, London: The MIT Press, 2000a. v. 1.

TALMY, L. *Toward a cognitive semantics: typology and process in concept structuring*. Cambridge, London: The MIT Press, 2000b. v. 2.

TEIXEIRA, J. *A verbalização do espaço: modelos mentais de frente/trás*. Braga: Universidade do Minho. Centro de Estudos Humanísticos, 2001. Disponível em: <<http://repositorium.sdum.uminho.pt/handle/1822/4517>>. Acesso em: 7 maio 2015.

TYLER, A.; EVANS, V. *The semantics of English Prepositions: spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press, 2003. DOI: <https://doi.org/10.1017/CBO9780511486517>.

VANDELOISE, C. *Spatial Prepositions: a case study from French*. Tradução de A. R. K. Bosch. Chicago; London: The University of Chicago Press, 1991.

WARD, J.H. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, Taylor & Francis on-line, v. 58, p. 236-244, 1963.

WITTGENSTEIN, L. *Investigações filosóficas*. Tradução de José Carlos Bruni. São Paulo: Nova Cultural, 1999 [1953]. (Coleção Os Pensadores).