



Como montar uma rede construcional? Uma proposta metodológica aplicada às construções de complementação sentencial do português brasileiro

Diogo Pinheiro

UFRJ

<https://orcid.org/0000-0003-2403-5040>

diogopinheiro@letras.ufrj.br

Dayanne Ximenes

UFRJ

<https://orcid.org/0000-0002-3848-1474>

dayannecolosso@gmail.com

RESUMO

Redes hierárquicas baseadas em classes semânticas são centrais à Gramática de Construções Baseada no Uso. Como consequência, um desafio metodológico emerge: como identificar as classes semânticas que definem construções intermediárias de uma rede construcional? A fim de contribuir para responder a essa questão, este estudo desenvolve e aplica uma proposta metodológica à Construção de Complementação Sentencial do português brasileiro. O método proposto prevê o mapeamento de construções com verbo especificado à luz da Teoria dos Exemplos, a postulação de categorias semânticas abstratas por meio da observação dos atributos semânticos latentes em *clusters* de exemplares e o emprego de Escalonamento Multidimensional para verificar a realidade psicológica da rede postulada.

PALAVRAS-CHAVE: Complementação sentencial. Rede construcional. Escalonamento multidimensional.

How to build a construcional network: a methodological proposal applied to Brazilian Portuguese clausal complementation

ABSTRACT

Semantic class-based hierarchical networks lie at the heart of Usage-based Construction Grammar. As a result, a methodological challenge emerges: how to identify semantic classes that define middle-level constructions in a constructional network? This paper aims at demonstrating a specific method designed to meet this challenge. In order to do so, the method is applied to Brazilian Portuguese clausal complementation constructions. The proposed method involves the mapping of lexically specific constructions in an Exemplar Theory fashion, the identification of emerging semantic class and the verification of the resulting network by means of Multidimensional Scaling.

KEYWORDS: Clausal complementation. Constructional network. Multidimensional scaling.



1. Primeiras palavras

Na análise de construções gramaticais de nível sintático, redes hierárquicas baseadas em classes semânticas têm se revelado centrais e recorrentes em análises ancoradas na Gramática de Construções (GC), e em especial na Gramática de Construções Baseada no Uso (GCBU). Croft (2012, cap. 9), em particular, propõe explicitamente que o tratamento goldbergiano em termos de polissemia construcional (GOLDBERG, 1995) deve ser substituído por uma análise baseada em redes hierárquicas, nas quais um conjunto de construções monossêmicas se associa a formas específicas por meio de relações taxonômicas e horizontais.

Mais especificamente, o autor sugere que construções altamente gerais (como, por exemplo, a construção bitransitiva do inglês, que tem a forma *SUJ V OBJ1 OBJ2*) se vinculam a um conjunto de *verb class specific constructions*, isto é, construções semanticamente mais específicas e marcadas quanto à classe semântica admitida no *slot* verbal (por exemplo, *SUJEITO + VERBO DE DOAÇÃO + OBJ1 + OBJ2*, *SUJEITO + VERBO DE MOVIMENTO BALÍSTICO + OBJ1 + OBJ2*, etc.). Adicionalmente, o autor propõe que cada uma dessas construções está, por sua vez, vinculada a uma *verb specific construction*, isto é, uma construção especificada quanto a um item verbal particular (por exemplo, *SUJEITO + GIVE + OBJ1 + OBJ2*). Dessa maneira, a postulação de uma única construção altamente polissêmica dá lugar a um amplo conjunto de construções com diferentes níveis de esquematicidade e todas elas, em si mesmas, monossêmicas.

Ao mesmo tempo em que esse novo tratamento é certamente vantajoso em relação à abordagem goldbergiana¹, ele levanta uma questão prática: como identificar as classes semânticas relevantes em cada caso particular? Essa questão é ainda mais desafiadora se levarmos em conta que o modelo não limita, em princípio, a quantidade de níveis de classes semânticas, isto é, entre a construção mais genérica e a mais concreta (entendendo-se a mais concreta como a *verb specific construction*) pode existir mais de um nível intermediário com classes especificadas em diferentes graus de generalidade.

Este artigo se configura, fundamentalmente, como uma tentativa de responder a essa questão. Especificamente, busca-se apresentar uma proposta de caminho metodológico que permita definir essas classes por meio de procedimentos sistemáticos e rigorosos.

Para demonstrar o caminho metodológico desenvolvido, recorreremos a um objeto gramatical do português brasileiro (PB) que, até onde sabemos, carece de descrição construcionista: a estrutura morfossintática formada por sujeito seguido de verbo complementado por uma oração introduzida pelo complementizador “que” (por exemplo, *João contou que a festa foi boa*). Neste trabalho, dada a sua inserção no quadro teórico da GC, referimo-nos a essa estrutura como Construção de Complementação Sentencial (CCS).

Nosso interesse pela exploração desse padrão em particular não advém apenas das necessidades deste trabalho. Em estudo anterior (XIMENES; PINHEIRO, 2020), investigamos a natureza do conhecimento que “bloqueia” a combinação de certos verbos com essa construção, impedindo que ela licencie supergeneralizações. Em outras palavras, buscávamos explicações para o

¹ Em função de limitações de espaço, não é possível discutir essas vantagens aqui. Remetemos o leitor para Croft (2012).

fato de que os falantes sabem que certas formulações com complemento sentencial, ainda que pareçam sintática e semanticamente plausíveis, não são possíveis no português brasileiro.

Naquele momento, notamos que os verbos verificados na CCS são aqueles que podem ser complementados por sintagmas que denotam proposições². Logo, um verbo como *comprar* certamente não seria verificado nessa construção, já que seu complemento sempre se configura como uma entidade, e não uma proposição. O intrigante, no entanto, é que verbos como *criticar* e *elogiar*, apesar de admitirem complementos proposicionais, não parecem ser possíveis nessa moldura sintática (?? *Ele criticou que o filme estava entediante*).

A fim de compreender essa aparente contradição, desenvolvemos, para aquela pesquisa, um experimento de produção induzida em que os participantes deveriam oferecer descrições orais de cenas representadas em vídeos curtos. Para isso, eles deveriam usar ora verbos reais, ora verbos inventados pertencentes a dois grupos semânticos: anúncio e crítica. Os resultados mostraram que, independentemente de o verbo ser real ou inventado, sua classe semântica era um fator relevante: os participantes tinham maior probabilidade de usar a estrutura de complementação sentencial com verbos de anúncio do que com verbos de crítica, independentemente de o verbo ser real ou inventado. Isso permitiu demonstrar que o fato de o falante evitar a combinação de certos verbos do PB com a CCS não se deve somente a uma postura conservadora (o fato de não experienciarem essa combinação em particular), mas também a um conhecimento de natureza semântica.

Esse experimento reforça a convicção de que a organização dos verbos em categorias é gramaticalmente relevante para o fenômeno da complementação sentencial – e justifica, assim, que a CCS seja tomada como objeto de investigação deste estudo. Em outras palavras, nossos resultados anteriores sugerem que a complementação sentencial no PB pode ser descrita nos termos de uma rede hierárquica de construções gramaticais, em que uma construção altamente esquemática – a estrutura SUJEITO + VERBO + COMPLEMENTO INTRODUZIDO PELO COMPLEMENTIZADOR “QUE” – está ligada a um conjunto de construções definidas em função de classes semânticas (as *verb class specific constructions*), as quais, por sua vez, se vinculam a construções ainda mais específicas, nas quais o *slot* verbal inclui itens particulares (as *verb specific constructions*). Sendo assim, para além de apresentar uma proposta metodológica específica, este artigo procura, ainda, contribuir para uma descrição construcionista de uma parcela específica da rede construcional do PB.

Como se verá, o método proposto aqui se baseia, fundamentalmente, na ideia de que o pesquisador deve buscar replicar o processo de montagem da rede construcional que é levado a cabo, de forma inconsciente, pela criança em fase de aquisição: um processo *bottom-up* no qual se parte de exemplares concretos para só então se chegar, gradualmente, a categorias cada vez mais abstratas. Mais especificamente, o método inclui os seguintes passos: (i) identificação de construções lexicalmente preenchidas com base em dados de *corpora*; (ii) mapeamento dessas construções segundo suas afinidades semânticas, o que resultará na identificação de *clusters* de exemplares; (iii) a partir da observação dos atributos semânticos latentes nesses *clusters*, postulação de categorias semân-

² Entendemos “proposição” como toda representação mental de um conteúdo que seja passível de ser avaliada como verdadeira ou falsa. Nesse sentido, uma sentença como “o livro é caro” denota uma proposição, mas o mesmo não se aplica a um item como “livro” (que denota uma entidade, e não uma proposição) (LAMBRECHT, 1994).

ticas abstratas; (iv) aplicação de um experimento psicolinguístico de julgamento de similaridade a fim de verificar a realidade psicológica dos *clusters* e categorias propostos; e (v) aplicação de um segundo experimento psicolinguístico, desta vez de julgamento de aceitabilidade, com o objetivo de identificar se as classes semânticas estão definidas em nível apropriado de generalidade. Neste artigo, porém, por limitações de espaço, não será possível apresentar o passo (v).

Em larga medida, o método proposto aqui é tributário do estudo de Bybee e Eddington (2006) sobre construções de mudança de estado do espanhol. Em uma análise fortemente influenciada pela Teoria dos Exemplos (TE)³, os autores identificam, por meio de busca em *corpus*, os itens que ocorrem no *slot* adjetival dessas construções. Na sequência, solicitam a uma falante nativa que os posicione em um plano bidimensional de acordo com suas afinidades semânticas e então colocam essa análise à prova por meio de um experimento de julgamento de similaridade. Como se vê, portanto, as três primeiras etapas metodológicas propostas por nós se inspiram diretamente no artigo de Bybee e Eddington (2006).

Contudo, buscamos dar um passo além em pelo menos um sentido importante: assumimos que o mapeamento de exemplos do tipo proposto por Bybee e Eddington constitui o primeiro nível de uma rede construcional hierárquica (um tipo de entendimento que não está presente no estudo desses autores). Por essa razão, buscamos dar conta também dos níveis mais abstratos, postulando a emergência de generalizações semânticas. Nos termos de Croft (2012), poderíamos dizer que Bybee e Eddington (2006) se mantêm restritos ao nível das *verb specific constructions*, ao passo que, neste estudo, nós buscamos partir dessas construções lexicalmente preenchidas para, por indução/esquematização, chegar ao nível (ou níveis) das *verb class specific constructions*. Nesse sentido, procuramos aqui promover uma integração sistemática entre a Teoria dos Exemplos e a Gramática de Construções.

Por se constituir, em grande medida, como a apresentação de uma proposta metodológica, este artigo está organizado em função do caminho metodológico a ser percorrido. Assim, começamos pela etapa de coleta de dados e análise preliminar, com vistas ao estabelecimento de um recorte metodológico (seção 2). Em seguida, passamos à etapa de construção da rede construcional (seção 3) e, por fim, chegamos à verificação empírica dessa mesma rede (seção 4). Por fim, as considerações finais apresentam reflexões sobre o método proposto, bem como possíveis desdobramentos para a descrição da Construção de Complementação Sentencial.

2. Análise preliminar: um recorte do objeto de estudo

A análise aqui proposta se inicia pela coleta de dados. Para isso, recorreremos a três bases de dados: Nilc/São Carlos, Corpus Brasileiro e Revista Piauí. As duas primeiras são *corpora* eletrônicos disponíveis no site da Linguatca (linguatca.pt); a terceira foi construída a partir de 27 notícias publicadas na revista Piauí entre 14/09/18 e 05/02/19.

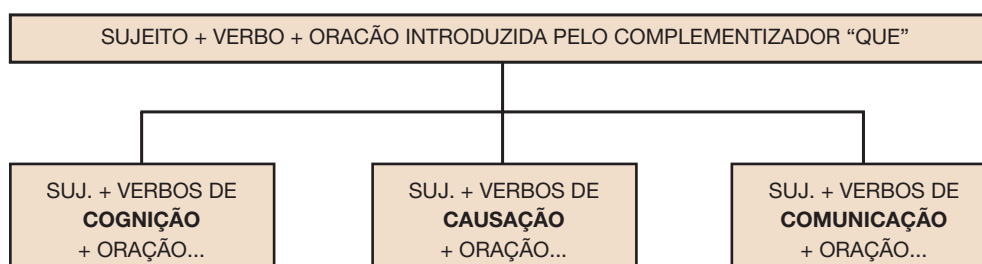
³ Oriunda da psicologia cognitiva, a TE ingressou na linguística pela via dos estudos fonético-fonológicos e posteriormente adentrou o domínio das investigações em morfossintaxe. Remetemos o leitor interessado para Bybee e Eddington (2006) e Bybee (2013).

Ao final da coleta, obtiveram-se 1.931 ocorrências da construção (frequência *token*), associadas a 170 verbos distintos (frequência *type*). Como essa quantidade é excessivamente elevada para o tipo de investigação que se pretende empreender (afinal, como se verá, este estudo prevê uma análise semântica individual de cada *type*), foi necessário limitá-la a um subconjunto dos dados obtidos. Assim, a fim de estabelecer um critério para esse recorte, submetemos a totalidade dos dados a uma análise preliminar.

Nessa análise, verificamos que os 170 verbos podem ser distribuídos em três grupos: cognição (como *aceitar*), causação (como *impedir*) e comunicação (como *dizer*)⁴. Consideramos pertencentes ao grupo de cognição aqueles verbos que indicam processos ou estados mentais; ao grupo de causação, aqueles que indicam processos que provocam, ou pelo menos induzem, a ocorrência de outras situações; e ao grupo de comunicação, os verbos que indicam situações de representação de conteúdos por meio da voz, escrita ou outras formas de expressão. Essa organização permite distribuir os 170 *types* da seguinte maneira: 57 na classe de cognição, 28 na classe de causação e 107 na classe de comunicação⁵.

Essa análise pode ser traduzida, em termos construcionistas, na postulação de três construções de complementação sentencial intermediárias que se relacionam horizontalmente entre si e, ao mesmo tempo, estabelecem uma relação taxonômica com a CCS mais esquemática:

FIGURA 1. As três categorias semânticas gerais da rede da complementação sentencial



O estabelecimento dessas três classes fornece o critério que buscávamos para proceder a um recorte nos dados: optamos por restringir nossa análise às construções com verbos de comunicação, o que permite reduzir o total de *types* para 107 (em vez dos 170 originais).

A opção por essa classe se deve a dois fatores. Em primeiro lugar, como vimos, trata-se do grupo mais amplo, razão pela qual é também o que provavelmente apresenta estrutura interna mais complexa. Em segundo lugar, os verbos que mais frequentemente são utilizados na CCS estão abrigados nessa classe. Veja-se, por exemplo, que os três verbos comunicativos mais recorrentes têm frequência *token* igual a 284 (*dizer*), 139 (*afirmar*) e 126 (*esclarecer*), enquanto

⁴ Toda vez que mencionamos o significado dos verbos encontrados nos *corpora*, referimo-nos ao seu sentido na estrutura de complementação sentencial, e não isoladamente. Em outras palavras, estamos discutindo, na prática, o significado de construções semiesquemáticas que incluem esses verbos (e não o significado dos verbos em si mesmos).

⁵ Vale destacar que muitas construções podem pertencer a mais de uma classe semântica simultaneamente, como se verá adiante.

os mais recorrentes da classe de cognição têm frequência *token* de 93 (*achar*), 60 (*saber*) e 49 (*acreditar*), e os da classe de causação têm todos menos de dez ocorrências.

A rede construcional representada na figura 1 pode constituir parte do conhecimento do falante do PB acerca da complementação sentencial; no entanto, uma descrição detalhada requer que classes ainda mais específicas estejam representadas. Um dos fatores que apontam para essa necessidade é o fato de que a existência das categorias de comunicação, causação e cognição não é suficiente para explicar o motivo pelo qual certos verbos, embora pudessem perfeitamente estar abrigados em alguma dessas classes semânticas, são, na prática, evitados na CCS. Por exemplo, os verbos *elogiar* e *criticar*, ainda que admitam complementos proposicionais e denotem eventos de comunicação, produzem sentenças de aceitabilidade marginal se inseridos na CCS (?? *Zé elogiou que a comida estava boa*). Além disso, a GCBU prevê que classes mais abstratas emergem a partir de outras mais concretas, e que estas não desaparecem da memória uma vez que as categorias mais gerais são formadas. Juntas, essas duas observações sugerem a necessidade de postular classes verbais mais específicas, a serem representadas em níveis da rede construcional inferiores àquele apresentado na figura acima.

Mas como identificar essas classes? Como já ficou dito, um dos objetivos deste artigo é propor um caminho metodológico que possibilite essa identificação. Para isso, propomos que o pesquisador replique a trajetória *bottom-up* que é percorrida pelo aprendiz, partindo da comparação entre exemplares concretos para identificar de que maneira eles se distribuem em *clusters* e, a partir daí, postular categorias progressivamente mais abstratas. Agora que já restringimos nossa base de dados a uma classe específica, reduzindo assim o total de *types* a serem analisados, podemos iniciar essa tarefa.

3. Construindo a rede das construções de complementação sentencial de comunicação

Esta seção está dividida em dois blocos: enquanto a primeira parte (2.1) descreve o primeiro nível da rede construcional (aquele composto pelas *verb specific contructions*), em um tipo de representação típico da Teoria dos Exemplares, a segunda parte (2.2) descreve a parte intermediária da rede (aquela composta pelas categorias que emergem a partir das características compartilhadas por construções do primeiro nível).

3.1 Exemplares e *clusters*: o nível mais baixo da rede construcional

Em conformidade com as premissas da TE, o primeiro passo do método aqui proposto consiste em comparar, item a item, todos os exemplares selecionados, considerando suas características funcionais. O resultado dessa comparação será um mapa de exemplares em que, seguindo a notação própria da TE, os elementos semanticamente mais similares são representados como espacialmente mais próximos.

Ao proceder a essa análise introspectiva dos 107 *types* incluídos na nossa base de dados, chegamos ao mapa apresentado na figura 2. Nessa representação, indicamos apenas o verbo por



FIGURA 2. Representação por exemplares do primeiro nível da rede



uma questão de clareza visual – é importante lembrar que cada exemplar do mapa é, na verdade, uma *verb specific construction* que especifica o item em questão no *slot* verbal.

Na figura 2, observa-se *dizer* ao centro e em destaque. Tal posicionamento decorre da percepção de que ele tem afinidade semântica com todos os itens do mapa. A observação desse comportamento nos deu uma pista de que ele poderia ser um membro central. De fato, a construção com *dizer* é, de longe, a mais frequente, com 284 ocorrências contra 139 de *afirmar*, a segunda colocada⁶.

Em função de limitação de espaço, não é possível descrever e justificar em detalhes a estruturação desse mapa de exemplares. Sobre isso, então, fazemos duas observações. Em primeiro lugar, remetemos o leitor para Ximenes (2020), onde podem ser encontradas análises e justificativas. Em segundo lugar, lembramos que, independentemente da análise proposta, o ponto fundamental aqui é a demonstração de um método empírico que permita avaliar de forma objetiva as hipóteses desenvolvidas com base em um exercício de análise semântica. Sendo assim, qualquer mapa de exemplares descrito a esta altura é, de todo modo, apenas uma hipótese, a ser verificada empiricamente (e potencialmente modificada). Voltaremos a esse ponto adiante.

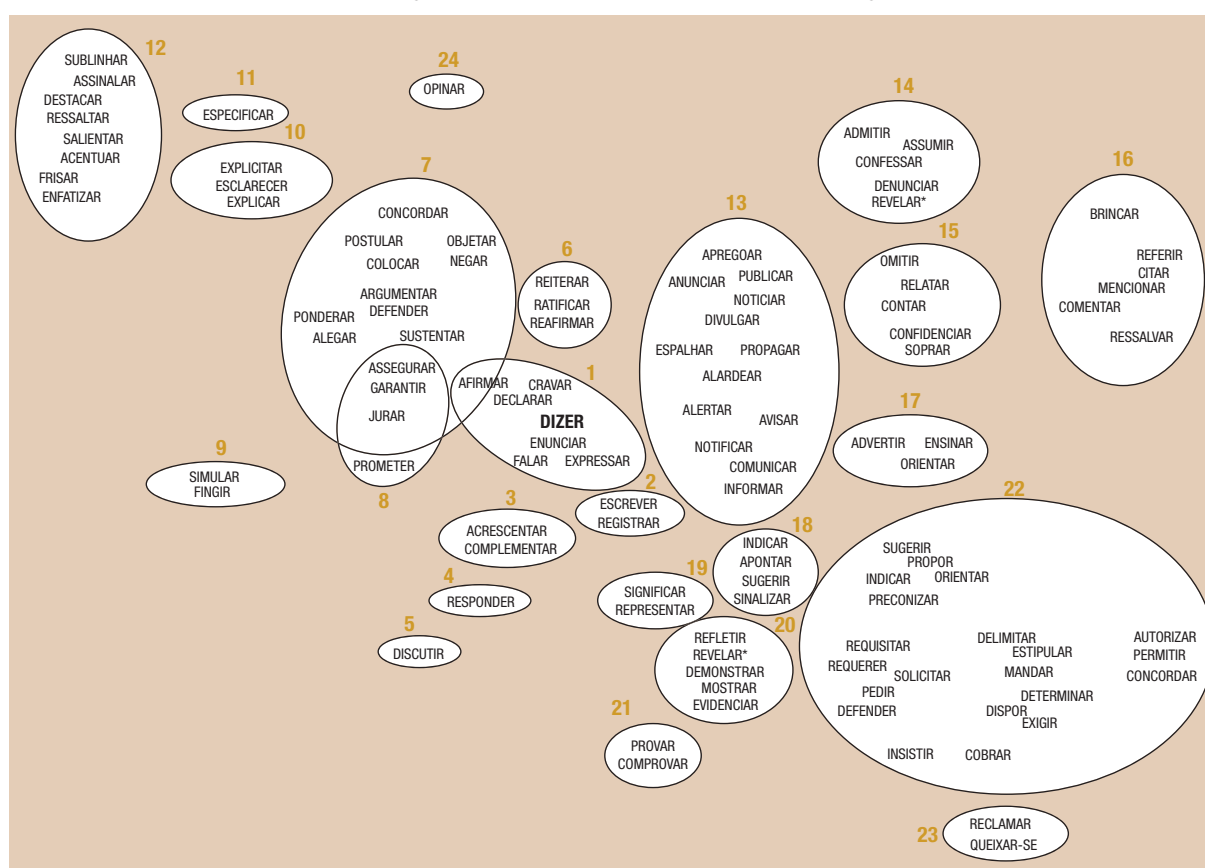
⁶ Embora haja críticas por parte dos pioneiros da Teoria dos Protótipos em relação a usar a frequência como critério para estabelecer prototipicidade (ver, por exemplo, Rosch (1978)), Bybee (2010, p. 79) argumenta que esse critério é aplicável quando se trata de objetos gramaticais. De todo modo, nosso interesse não é estabelecer o elemento mais frequente como o protótipo no sentido roschiano, mas como o elemento cognitivamente mais acessível da categoria.

3.2 De exemplares a categorias: os níveis intermediários da rede

Tomando como base os *clusters* representados visualmente no mapa de exemplares, o próximo passo do método aqui proposto consiste na identificação de categorias semânticas latentes nesse mapa – as quais deverão então ser interpretadas como *verb class specific constructions*.

Em função das limitações de espaço naturais em um artigo, não será possível descrever, justificar e exemplificar todas as (meso)construções propostas (para isso, remetemos novamente o leitor a Ximenes (2020)). Em vez disso, adotaremos aqui a seguinte estratégia: inicialmente, reapresentamos o mapa de exemplares descrito anteriormente, desta vez com a indicação de categorias semânticas gerais; em seguida, apresentamos o quadro 1, adiante, com a identificação de todas as (meso)construções propostas e a lista dos verbos (mais precisamente, *verb specific constructions*) associados a cada uma; e, finalmente, apresentamos nossa proposta de rede construcional⁷.

FIGURA 3. Representação por exemplares acrescida de explicitação dos conjuntos



Como se observa, a figura 3 identifica categorias semânticas latentes no mapa de exemplares apresentado na figura 1. Considerando que cada exemplar é uma construção semi-preenchida e cada conjunto corresponde a uma *verb class specific construction*, esse mapa traduz de forma explícita, ainda que latente, dois níveis construcionais. No entanto, é possível observar que algumas

⁷ Mas, como se verá, algumas porções da rede serão omitidas da representação em nome da clareza visual e em função de limitações de espaço.

categorias são internamente estruturadas em *(sub)clusters*. Sob uma ótica construcionista, esses *(sub)clusters* devem ser interpretados como um terceiro nível construcional. O quadro 1 identifica esses três níveis hierárquicos, todos eles situados abaixo do nível da construção mais geral de comunicação (que se encontra, por sua vez, abaixo da CCS mais geral, conforme a figura 1).

QUADRO 1. Construções definidas por classes semânticas

CONSTRUÇÃO SUPERORDENADA ⁸ (classe semântica mais geral)	CONSTRUÇÃO SUBORDINADA ⁹ (classe semântica mais específica)	ITENS VERBAIS ¹⁰
Comunicação neutra		falar, enunciar, afirmar, declarar, expressar, cravar
Representação perene de um conteúdo		escrever, registrar
Continuação do discurso		acrescentar, complementar
Resposta		responder, replicar, retrucar
Discussão		debater, discutir
Repetição		reafirmar, ratificar, reiterar
Argumentação	Neutra	argumentar, defender, sustentar, alegar, colocar, ponderar, postular, concordar
	Com alto grau de certeza	assegurar, garantir, jurar, afirmar
	Contra-argumentação	objetar, negar
Comprometimento enfático		jurar, assegurar, garantir, prometer
Fingimento		simular, fingir
Esclarecimento		explicar, esclarecer, explicitar
Especificação		especificar, discriminar
Destaque		acentuar, destacar, frisar, enfatizar, ressaltar, salientar, sublinhar e assinalar
Informação	Neutra	notificar, avisar, alertar, informar, comunicar
	Para um conjunto de pessoas	noticiar, anunciar, divulgar, espalhar, propagar, apregoar, publicar, alardear

(continua)

⁸ Esta coluna apresenta a construção que se apresenta como superordenada no conjunto dos três níveis construcionais representados no quadro 1. Ao mesmo tempo, porém, as construções identificadas nela estão subordinadas à CCS de comunicação mais geral (localizada no segundo nível da figura 1); esta, por sua vez, é subordinada à CCS mais geral; esta, por sua vez, é subordinada à construção mais geral de complementação verbal; e assim sucessivamente. Ou seja, considerados apenas os níveis hierárquicos identificados na tabela, elas correspondem ao que às vezes é referido como “macroconstruções”. Por outro lado, se incorporamos ao recorte a CCS de comunicação (ou quaisquer outros níveis acima deste, conforme listado no início desta nota), elas se tornam mesoconstruções.

⁹ As construções aqui identificadas são diretamente subordinadas àquelas indicadas na primeira coluna, o que significa que estão a dois níveis de distância da CCS de comunicação, a três níveis de distância da CCS mais geral, a quatro níveis de distância da construção mais esquemática de complementação verbal, e assim sucessivamente (ver nota 8). Sendo assim, qualquer que seja o recorte considerado, as construções representadas nesta coluna corresponderão ao que é às vezes referido como “mesoconstruções” – já que elas estarão acima das *verb specific constructions* indicadas na coluna da direita.

¹⁰ Considerado o recorte definido nesta pesquisa, as construções indicadas nesta coluna correspondem ao que é às vezes referido como “microconstruções”. Mas vale lembrar que é esperado que haja construções ainda menos esquemáticas, e portanto hierarquicamente inferiores a estas – por exemplo, aquelas a que Croft (2012) se refere como *verb and object specific constructions*. Não há nenhuma razão apriorística para supor que a existência de um slot verbal preenchido determine, por si só, o nível menos esquemático de uma rede construcional.

(continuação)

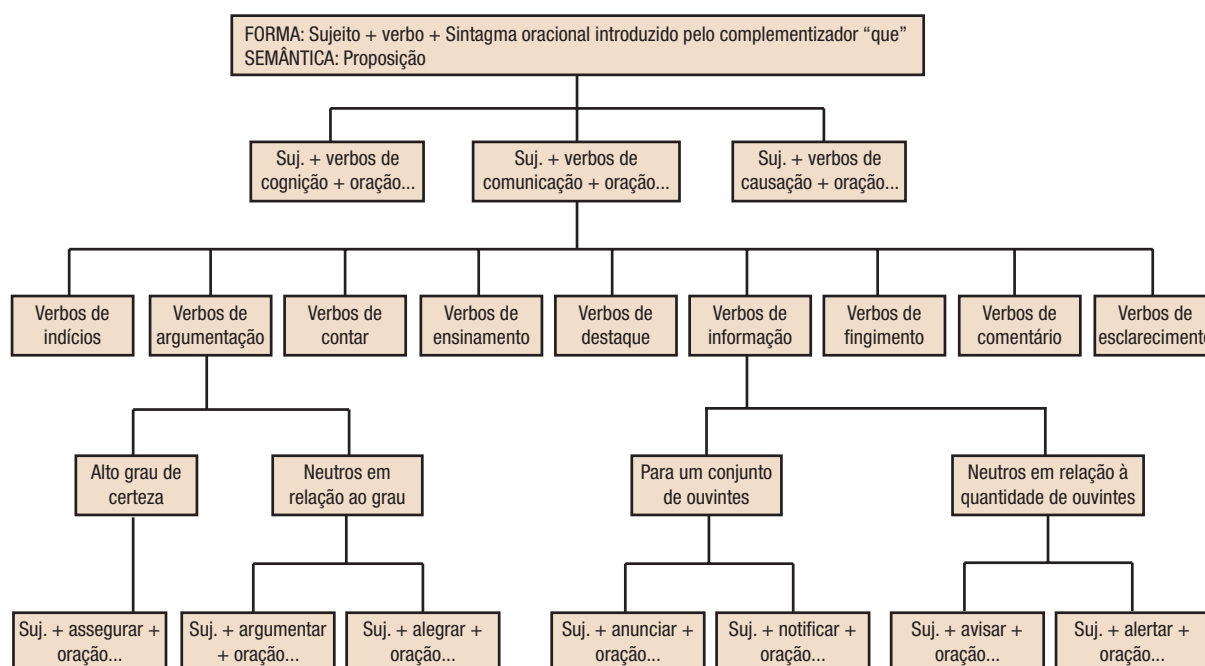
CONSTRUÇÃO SUPERORDENADA ⁸ (classe semântica mais geral)	CONSTRUÇÃO SUBORDINADA ⁹ (classe semântica mais específica)	ITENS VERBAIS ¹⁰
Revelação	Neutra	orientar, ensinar, advertir
	Com referente do sujeito negativamente afetado	admitir, assumir, revelar, denunciar, confessar
Relato	Neutro	contar, relatar, omitir
	Às escondidas	confidenciar, soprar
Comentário		citar, referir, ressaltar, comentar, referir, brincar
Ensino		orientar, ensinar, advertir
Sinalização (dar indícios)		sugerir, indicar, apontar, sinalizar
Simbolização		representar, significar
Evidenciação (tornar perceptível)		evidenciar, demonstrar, mostrar, refletir, revelar
Confirmação (por meio de provas)		provar, comprovar
Atos diretivos	Recomendação	indicar, sugerir, orientar, preconizar e propor
	Ordem	mandar, exigir, delimitar, estipular, dispor e determinar
	Pedido	defender, pedir, solicitar, requerer requisitar
	Ênfase ou insistência	Cobrar, insistir
	Concessão	autorizar, concordar e permitir
Reclamação		queixar-se, reclamar
Opinião		opinar, palpar

Sobre o quadro 1, façamos três observações. Em primeiro lugar, observe-se que alguns verbos são associados a mais de uma construção de nível intermediário. Isso, além de ser perfeitamente usual em GC, permite capturar certas diferenças semânticas sutis entre os itens verbais. Por exemplo, a diferença entre o uso de *jurar* expressando o comprometimento com a realização de uma atividade futura (“juro que não vou faltar”) e comprometimento com a validade de uma proposição (“eu juro que Paris é a capital da França”) é capturada pela postulação da associação da *verb specific construction* com *jurar* a duas construções de nível intermediário distintas. Note-se que isso permite ainda capturar a diferença semântica entre *jurar* e *prometer*, a saber, o fato de que este último exibe apenas a segunda dentre as duas acepções propostas para *jurar*. Nessa rede, isso é capturado vinculando-se *verb specific constructions* com *jurar* a apenas uma construção de nível intermediário.

Em segundo lugar, note-se que algumas construções mais gerais estão associadas a sub-construções definidas a partir de classes semânticas mais específicas, ao passo que outras estão diretamente vinculadas às construções lexicalmente preenchidas. Em terceiro lugar, e o mais importante, é evidente que muitas outras formas de organização semântica são possíveis com base nos dados analisados. Nesse sentido, voltamos a insistir em um ponto já comentado anteriormente: o que está apresentado na figura 2 e no quadro 1 é uma *hipótese*, a qual, embora se-

mantidamente fundamentada, está longe de ser o único modo de categorização justificável em termos descritivos. Com efeito, este é o ponto fundamental deste artigo: dado que, em geral, há muitas maneiras semanticamente razoáveis de agrupar exemplares, formando *clusters* de *verb specific constructions*, linguistas alinhados a esse tipo de modelo teórico necessitam de um método que permita a eles optar pela organização mais apropriada. A hipótese ilustrada pela figura 2 e pelo quadro 1 é, como se verá adiante, o ponto de partida de um método empírico cujo objetivo último é chegar a uma organização construcional psicologicamente real.

FIGURA 4. A rede hierárquica das construções de complementação sentencial



A figura 4 apresenta a rede de construções de complementação sentencial formada por cinco níveis. O nível mais baixo é formado pelas construções com verbo especificado. Na imagem, estão representadas, por uma questão de clareza visual, apenas sete construções. No entanto, entende-se que todas as 107 construções de comunicação obtidas nos *corpora* se estabelecem nesse nível.

Os três níveis acima são os níveis intermediários da rede e estão definidos pelas classes semânticas que emergem das construções mais baixas. O primeiro nível intermediário (de baixo para cima) está exemplificado na imagem por quatro construções, mas representa 14 delas, a saber, todas as que estão indicadas na coluna do meio do quadro 1. O segundo dos níveis intermediários está exemplificado por nove construções, mas representa 24 – a saber, todas as que estão indicadas na coluna da esquerda do quadro 1.

O último dos níveis intermediários é formado pelas três grandes classes mencionadas no início deste artigo: comunicação, cognição e causação. Acima delas, assumimos que pode existir, ainda que não seja requerida na prática, uma construção mais esquemática cuja única especificidade semântica, no que diz respeito ao verbo da cláusula matriz, é que ele seja do tipo que admite complementos proposicionais (ver introdução e nota 1).

4. Verificando a rede das construções de complementação sentencial comunicativas

Passamos agora à etapa de verificação empírica da rede proposta. Para isso, desenvolvemos um experimento de julgamento de similaridade, cuja tarefa consiste em avaliar o quão próximos ou distantes semanticamente são dois itens apresentados em um par. Todos os voluntários (60 sujeitos entre 17 e 38 anos) foram apresentados a duplas de sentenças completas, nas quais o verbo relevante estava sublinhado, e instruídos a indicar o grau de similaridade entre os verbos em uma escala *Likert* de 5 pontos, que ia de “verbos com sentido nada parecidos” (1) até “verbos com sentido bastante parecidos” (5).

Na presente pesquisa, caso desejássemos avaliar a realidade psicológica da organização de todos os verbos que compõem as construções do grupo de comunicação, seria necessário parear cada um dos 107 verbos a todos os demais, gerando uma quantidade de estímulos muito além do que é viável incluir em um único experimento. Por isso, optamos por selecionar um subconjunto do total de itens possíveis na construção para submeter ao teste experimental¹¹. Para esse recorte, optamos por selecionar duas subclasses que, de acordo com a nossa hipótese, exibem estrutura interna complexa: a classe dos verbos de informação e a classe dos verbos de argumentação.

Somados os verbos que compõem essas duas subclasses, chega-se a 23 verbos, dentre os quais selecionamos 16 – oito de cada um – para o experimento. Do *cluster* de anúncio, utilizamos os verbos *notificar*, *avisar*, *alertar* e *comunicar* (*subcluster* 1) e *anunciar*, *divulgar*, *espalhar* e *alardear* (*subcluster* 2). Do *cluster* de argumentação, os verbos *jurar*, *garantir*, *assegurar* e *afirmar* (*subcluster* 1) e *sustentar*, *alegar*, *argumentar* e *defender* (*subcluster* 2). Pareando cada um desses verbos com todos os outros, obtém-se o total de 120 pares de sentenças a serem avaliadas.

4.1 Análise estatística: sobre o EMD

As respostas fornecidas pelos voluntários foram analisadas por meio da ferramenta estatística conhecida como Escalonamento Multidimensional (EMD). Essa técnica permite obter estimativas de similaridade entre itens e grupos de itens por meio de uma apreciação visual. Em outras palavras, o EMD converte uma matriz numérica de distâncias (perceptuais/conceituais) em um mapa que revela, especialmente, as relações entre os itens investigados. Esse tipo de tratamento de dados tem ampla aplicabilidade em estudos de ciências cognitivas graças ao seu potencial para validar suposições sobre a organização do espaço psicológico (HOUT; PAPESH; GOLDINGER, 2013).

A tarefa de mapear itens de natureza abstrata e mental (como são as construções gramaticais) é bastante complexa, ainda mais se considerando que devem ser apreciadas as relações individuais de cada um com todos os outros. Com efeito, quando se trata de medidas experimentais, os dados que revelam as distâncias entre os pares são frequentemente conflitantes. Em EMD, um conflito é criado quando, por exemplo, dois itens A e B estão estabelecidos

¹¹ A mesma opção foi feita por Bybee e Eddington (2006). Como o fenômeno investigado por eles envolvia 161 adjetivos, avaliar a distância semântica de todos eles significaria incluir mais de 10 mil pares de sentenças no experimento; por isso, eles reduziram para 20 adjetivos. No nosso caso, o total de 107 verbos seria igualmente inexequível.

numericamente como muito próximos um do outro ao mesmo tempo em que um item C é estabelecido simultaneamente como muito próximo de A e muito distante de B. Isso pode ocorrer porque, como os participantes comparam dois itens por vez, eles podem não levar em conta suas classificações anteriores.

Apesar de o EMD buscar refletir as distâncias numéricas, traduzindo-as em distâncias espaciais do modo mais fidedigno possível, em casos como esses, em que o posicionamento dos pares é incompatível, algumas “manobras” inevitavelmente deverão ser realizadas, de modo que certas relações entre itens serão distorcidas. Para que o pesquisador saiba se, de fato, havia conflito nos dados, o *output* do tratamento estatístico fornece um valor de “*stress*”. Esse número se configura como um indicador do nível de discrepância entre as distâncias estimadas no mapa e os valores numéricos do *input*, de modo que, quanto mais baixo o valor de *stress*, melhor o ajuste entre essas duas realidades (a numérica e a visual) (HOUT; PAPESH; GOLDINGER, 2013).

4.2. Resultados e discussão

Para a montagem de um mapa de exemplares via EMD, os graus de similaridade obtidos experimentalmente devem ser invertidos ($1 > 5$; $2 > 4$; etc.), de maneira tal que os maiores valores reflitam o maior grau de distância semântica (e, portanto, maior distância espacial). Após essa inversão, é necessário somar os valores obtidos para cada par. No nosso experimento, o resultado desses procedimentos levou ao quadro a seguir:

QUADRO 2. Matriz com grau de similaridade

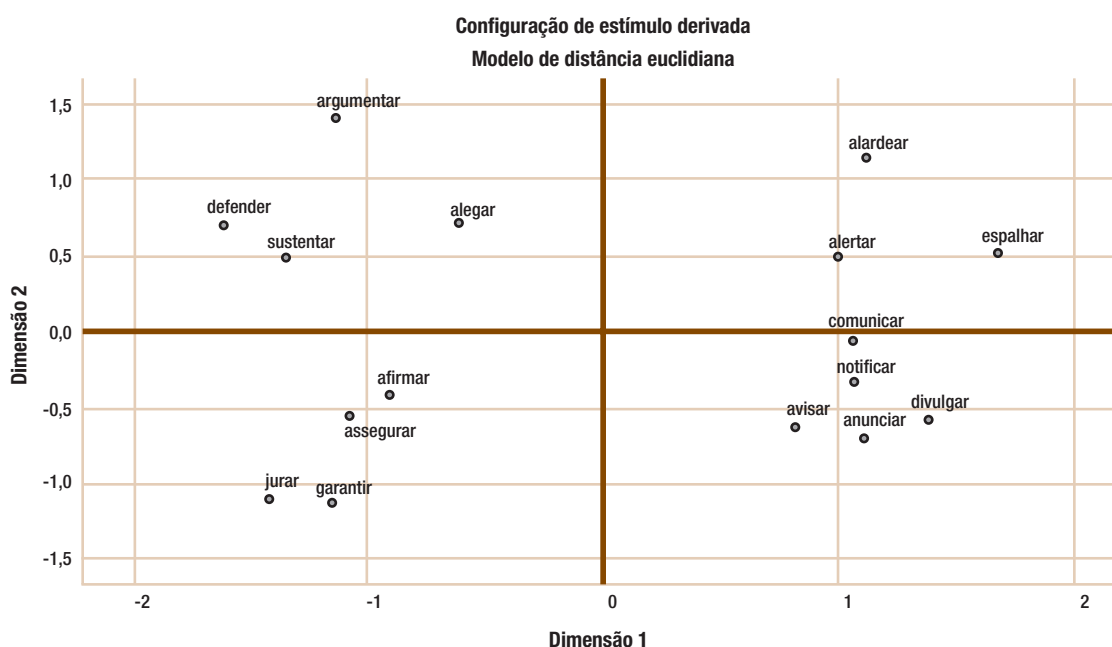
	divulgar	espalhar	notificar	alertar	comunicar	avisar	alardear	anunciar	assegurar	garantir	defender	jurar	afirmar	alegar	argumentar	sustentar
divulgar																
espalhar	85															
notificar	64	85														
alertar	76	86	79													
comunicar	55	82	48	70												
avisar	63	101	47	53	40											
alardear	90	68	96	67	104	92										
anunciar	48	84	54	94	54	44	98									
assegurar	103	121	88	106	108	101	107	117								
garantir	121	124	114	101	106	96	124	105	47							
defender	126	125	110	116	121	106	114	124	56	87						
jurar	131	124	113	116	119	122	110	107	76	56	102					
afirmar	91	122	87	113	106	84	113	85	47	54	86	88				
alegar	109	109	87	91	84	99	100	97	80	102	95	100				
argumentar	123	126	115	95	109	115	115	117	106	116	54	113	57	60		
sustentar	113	114	118	115	100	111	113	111	70	82	57	99	58	69	62	

Esse formato de apresentação dos resultados em uma matriz é exatamente o requerido para traduzir os números em informações espaciais por meio do EMD. Para que possamos conferir, a partir dela, a nota resultante dos julgamentos dos participantes em relação a um determinado par, basta identificar, na primeira coluna da tabela, a linha que comporta um dos verbos, e, em seguida, localizar qual dentre as demais colunas aquela que corresponde ao segundo verbo do par.

Em uma primeira aproximação desses resultados numéricos, vê-se que parecem apontar na direção pretendida por nossas hipóteses. Por exemplo: os valores referentes aos pares formados por verbos previamente classificados como pertencentes a classes distintas variou entre 84 e 131, enquanto aqueles hipotetizados como pertencentes à mesma classe variou entre 40 e 116. Além disso, enquanto não há nenhum valor menor que 84 nos pares de classes distintas, 30 dos 54 pares da mesma classe (55,5%) foram julgados com notas inferiores a essa. Esses números parecem confirmar a ideia de que os 16 verbos contemplados se distribuem em dois *clusters* distintos: um com semântica de argumentação e outro com semântica de anúncio.

Contudo, a observação das notas atribuídas a cada par individualmente não contribui diretamente para a criação de um panorama geral da organização desses verbos. Para isso, é necessário recorrer a uma análise estatística via EMD, cujo *output* é um mapeamento espacial de cada um dos verbos, considerando a distância de cada um deles em relação a todos os outros. Observe-se:

FIGURA 5. Resultado do Escalonamento Multidimensional



A figura 5 demonstra com clareza a divisão apontada por nós entre dois grandes *clusters* (correspondentes à construção de argumentação e à construção de informação). Com efeito, *jurar*, *afirmar*, *garantir*, *assegurar*, *argumentar*, *defender*, *sustentar* e *alegar* estão posicionados mais à

esquerda, enquanto *anunciar*, *avisar*, *divulgar*, *notificar*, *comunicar*, *alertar*, *alardear* e *espalhar* estão mais à direita do mapa.

Segundo a figura 2 e o quadro 1, também esperávamos encontrar uma divisão interna às construções de argumentação – em um agrupamento, *afirmar*, *assegurar*, *garantir* e *jurar*, como aquelas que indicam argumentação com alto grau de certeza; em outro, *argumentar*, *alegar*, *defender* e *sustentar*, como aquelas que são neutras em relação ao grau de certeza. E, com efeito, como se pode observar, as quatro construções hipotetizadas como marcadas quanto ao grau de certeza se concentram no quadrante inferior esquerdo, enquanto as quatro restantes se agrupam no canto superior esquerdo.

Ainda segundo a figura 2 e o quadro 1, esperávamos encontrar uma divisão interna entre as construções de informação – em um grupo, aquelas que indicam informação para um conjunto de ouvintes (*anunciar*, *alardear*, *divulgar* e *espalhar*); em outro, aquelas que são neutras em relação a isso (*alertar*, *comunicar*, *avisar* e *notificar*). Essa distribuição, no entanto, não é verificada no mapa.

O fato de essa hipótese não ter sido confirmada, contudo, não nos deixa sem respostas. Uma vantagem desse tipo de análise estatística é que, em casos como esse, ela não apenas refuta a suposição que tínhamos sobre a organização do espaço psicológico, mas nos apresenta uma outra possibilidade de organização. Por exemplo, observamos que os verbos *comunicar*, *notificar*, *avisar*, *divulgar* e *anunciar* formam um aglomerado bastante coeso, de forma que devemos imaginar propriedades capazes de reuni-los. De outro lado, *alertar*, *espalhar* e *alardear* também parecem formar um *cluster*, embora relativamente menos coeso. A partir da análise desse mapeamento, é possível fazer inferências acerca de quais parâmetros teriam sido levados em conta pelos participantes para organizar os exemplares dessa maneira.

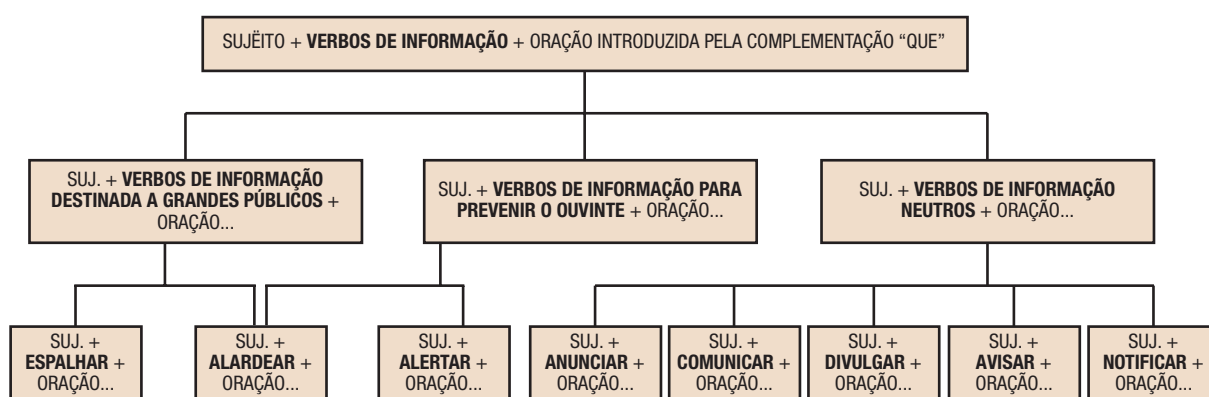
No *cluster* formado por *alertar*, *espalhar* e *alardear*, uma afinidade óbvia se nota entre *espalhar* e *alardear*: ambos denotam a ação de transmitir informação para um público amplo. Isso não explica, porém, o fato de que a distância semântica percebida entre *alardear* e *alertar* (67) é quase idêntica àquela sentida pelos falantes em relação a *alardear* e *espalhar* (68). Uma possibilidade é que os falantes, ou pelo menos parte deles, assumam que *alardear* envolve (sempre ou tipicamente) a comunicação de um evento negativo, em relação ao qual os destinatários devem se prevenir (daí a necessidade de “fazer alarde”). De fato, o mapeamento de exemplares representado nos permite considerar essa hipótese, como forma de explicar a distribuição gerada pelas respostas dos falantes.

Com base na percepção de que *alardear* compartilha uma característica particular com *alertar* e outra com *espalhar*, entendemos que esses três verbos parecem se reunir por meio de uma estrutura de semelhança de família. Como consequência dessa compreensão, verificamos que os verbos do quadrante inferior direito podem ter sido agrupados justamente por serem neutros em relação a essas duas características, de modo que os participantes os consideraram muito próximos entre si simplesmente por serem todos usados para veicular uma informação. Assim, ao se depararem com um par formado por *alardear* e *avisar*, por exemplo, esses participantes não os consideraram muito parecidos porque o uso de *alardear* conferiria um tom de *alerta* e, ao mesmo tempo, de difusão (ou “espalhamento”) da informação, que não é veiculado por *avisar*.

Essas percepções permitem reconsiderar a ideia de que *divulgar* e *anunciar* seriam marcados quanto ao fato de denotar eventos de comunicação para mais de um receptor. Em outras palavras, é possível que esses verbos sejam percebidos como neutros em relação ao número de ouvintes e, por isso, estejam armazenados, na rede construcional, em posição próxima à de *avisar*, *notificar* e *comunicar*.

Diante do mapa gerado pelo experimento e das reflexões feitas a partir dele, concluímos que as categorias que emergem a partir das construções de informação difere daquela por nós pensadas anteriormente. A figura 6 mostra, então, a rede construcional da CCS parcialmente modificada de acordo com as ponderações desenvolvidas aqui.

FIGURA 6. Construções de informação modificadas



Essas reflexões são apenas um exemplo das potencialidades do EMD para a construção de mapas de exemplares e redes construcionais. Naturalmente, elas não chegam perto de esgotar a descrição da CCS do PB; no entanto, permitem alguns avanços nesse sentido, ao mesmo tempo em que – talvez mais importante – oferecem um método que pode ser aplicado a outras porções dessa rede a fim de respaldar uma descrição empiricamente fundamentada do conhecimento linguístico do falante¹².

5. Palavras finais

Este artigo postulou e ilustrou uma proposta metodológica para se definir *verb class specific constructions*, que tipicamente correspondem aos níveis intermediários de uma rede construcional. Para isso, adotamos como objeto um subconjunto das construções de complementação sentencial do PB, a saber, as construções de comunicação.

Com isso, esperamos ter alcançado um objetivo duplo. De um lado, buscamos oferecer uma contribuição descritiva ao avançar na análise de uma porção da rede construcional do PB que, até onde sabemos, ainda não foi contemplada. De outro, buscamos oferecer uma contribuição

¹² Para uma análise mais completa, ver Ximenes (2020).

teórico-metodológica específica. Interessantemente, essa contribuição poderá ser futuramente estendida, inclusive, para as parcelas da rede construcional da CCS que, embora tenhamos podido postular neste artigo, não pudemos aqui verificar empiricamente.

REFERÊNCIAS

- BYBEE, J. **Language, usage and cognition**. Cambridge: University Press, 2010.
- BYBEE, J. Usage-based theory and exemplar representations of constructions. In: HOFMANN, T.; TROUSDALE, G. (eds.) **The Oxford Handbook of Construction Grammar**. Oxford: University Press, 2013.
- BYBEE, J.; EDDINGTON, D. A usage-based approach to Spanish verbs of ‘becoming’. **Language**, 82, 2, 2006, p. 323-355.
- CROFT, W. **Verbs: aspect and clausal structure**. Oxford: University Press, 2012.
- GOLDBERG, A. E. **Constructions: A construction grammar approach to argument structure**. Chicago: University Press, 1995.
- HOUT, M.; PAPESH, M.; GOLDINGER, S. Multidimensional scaling. **Cognitive Science**, v.4, n.1, 2013, p.93-103.
- LAMBRECHT, K. **Information structure and sentence form: Topic, focus and the menal representation of referents..** Cambridge: University Press, 1994.
- ROSCH, E. Principles of categorization. In: ROSCH, E.; LLOYD, B. (eds.) **Cognition and categorization**. Hillsdale, NJ: Lawrence Erlbaum, 1978, p. 27-48.
- XIMENES, D. **Uma proposta metodológica para o desenvolvimento de uma rede de construções: investigando a complementação sentencial do português brasileiro**. 148f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Programa de Pós-Graduação em Linguística, 2020.
- XIMENES, D.; PINHEIRO, D. Semiprodutividade construcional: uma investigação empírica da complementação sentencial do português brasileiro. **(Con)textos linguísticos**, v. 14, n. 28, 2020, p. 177-197.

