

Terceira parte: Particularidades

Capítulo 5: Idiossincrasias e Problemas

Mário A. Perini, Sigrid Fraiha
Lúcia Fulgêncio, Regina Bessa Neto

1. ANOMALIAS NA ANÁLISE DA LÍNGUA

A língua não se compõe apenas de regularidades; inclui também, em proporção desconhecida, casos particulares, exceções, irregularidades etc. Ou, para usar uma antiga nomenclatura, a língua compreende um componente **anomalístico**.

Essas anomalias são tradicionalmente colocadas no léxico; assim, um item pode ser marcado como exceção a uma regra (como o plural de **mal**, que é **males** e não o esperado * **mais**), e isso é parte da informação contida no léxico a respeito desse item.

Quando tratamos de fenômenos gramaticais dependentes de traços semânticos, é de se esperar que surja grande quantidade de anomalias — já que cada item léxico é, por assim dizer, semanticamente

único. Queremos dizer com isso que um item como, digamos, **orelha** não é exatamente igual a nenhum outro item, pois apenas ele designa aquela parte do corpo. Desse modo, a semântica encerra um grande potencial anomalístico.

Por outro lado, existem também anomalias formais (mórficas), a mais evidente das quais é o próprio potencial funcional dos itens: **aquele** pode ser determinante, mas não pré-determinante nem termo livre; **algum** pode ser determinante ou termo livre, mas não pré-determinante etc.

Na verdade, em princípio pode-se esperar que haja anomalias em qualquer parte da estrutura da língua. Se há ou não, e em que proporção, é algo que só a investigação poderá responder. Neste capítulo apontamos alguns casos particulares de interesse, que por um lado estabelecem limites ao nosso estudo (e, acreditamos, a qualquer estudo generalizante), e por outro lado sugerem direções promissoras de pesquisa.

2. MARCAS IDIOSSINCRÁTICAS

O sintagma seguinte é bem formado:

(1) Animal mamífero

e sabemos que sua inversão, * **mamífero animal**, não é, por causa da Condição de Restritividade Crescente (CRC). O mesmo acontece com

(2) Animal anfíbio

Agora, há itens que são semanticamente semelhantes a **mamífero** e **anfíbio**, mas que não podem ocorrer em sintagmas como (1):

(3) * Animal réptil

(4) * Animal peixe

(5) * Animal ave

Qual será a melhor explicação para essa diferença?

Uma explicação semântica generalizada terá de enfrentar certas dificuldades, justamente por causa da semelhança de significado entre **mamífero**, **réptil** e **peixe**: todos esses itens denotam subdivisões da classe maior dos animais. Não nos parece impossível que eventualmente se encontre uma explicação semântica para o comportamento excepcional de **mamífero** frente a **peixe** e **réptil**. No momento a solução não é de todo clara.

A possibilidade que nos parece mais plausível é atribuir esses fatos a uma diferença semântica entre os itens em questão — mas uma diferença que não afeta seu potencial referencial, que é muito semelhante em todos os casos (subconjuntos dos animais). Essa diferença, mais abstrata, reside na possibilidade ou não de cada item exprimir uma qualidade. Assim, **mamífero** e **anfíbio** serão marcados <+Q>: além de poderem referir-se a uma entidade (como em um mamífero), podem igualmente ser usados para exprimir uma qualidade (**animal mamífero**, **características mamíferas**). Já **réptil**, **inseto**, **peixe**, **ave** serão marcados <-Q>, não podendo portanto ser usados para exprimir uma qualidade — daí a inaceitabilidade de * **características répteis** etc. Nestes casos a função qualificativa é preenchida por itens semanticamente relacionados, marcados <+Q> (e, aliás, <-R>) como **reptiliano**, **písceo**, **ornitológico** etc.

Desse modo, a má formação de

(3) * Animal réptil

vem do fato de que **réptil** está aí em uma posição (após o indicador) onde é forçoso que assuma acepção qualificativa; mas o item léxico **réptil** é marcado <-Q>, sendo portanto inadequado para receber essa acepção. Em

(1) Animal mamífero

a má formação não surge, porque **mamífero** é marcado <+Q>.

3. LEXIFICAÇÃO

3.1. EXPRESSÕES FEITAS

O sistema desenvolvido neste livro dá conta da maior parte das restrições de ordenação observadas entre os termos do SN; e temos esperança de encontrar outras condições que expliquem boa parte dos casos ainda pendentes. Resta, entretanto, um resíduo, e pelo menos parte dele terá de ser atribuída ao processo de lexificação.

Entendemos por lexificação a fixação de uma ordem de termos (que, a se seguir estritamente a gramática, deveria ser livre) em um sintagma de alta frequência no discurso. O processo pode ser entendido como o de conversão de uma estrutura sintática (construída segundo regras gerais) em um item léxico (construído idiossincraticamente).

O processo geral pode ser exemplificado nas construções seguintes, onde se encontra lexificada uma seqüência de dois itens coordenados por meio de *e*:

(6) a. Filme em preto e branco

b. ?? Filme em branco e preto

(7) a. Exame de alto a baixo

b. * Exame de baixo a alto

(8) a. Roupa de cama e mesa

b. * Roupa de mesa e cama

(9) a. Empregada de forno e fogão

b. * Empregada de fogão e forno

E' evidente que os mecanismos gramaticais (mórficos ou sêmicos) não darão conta da ordenação estrita nesses casos. Consideramos, portanto, as seqüências **em preto e branco, de alto a baixo e cama e mesa** como itens léxicos — ou, mais exatamente, construções lexificadas (“frases feitas”). Não nos parece possível dar uma explicação estrutural para a preferência por uma ordenação sobre a outra: trata-se do resultado de acidentes históricos, e se têm ou tiveram alguma motivação esta não é lingüística.

Acreditamos que certos SNs nos quais a ordenação dos termos escapa às explicações usuais devem ser também colocados nessa categoria. Estamos conscientes dos perigos dessa decisão: aceitar a lexificação como análise implica o risco de criar um *deus ex machina* capaz de resolver todos os problemas em princípio, sem fornecer nenhuma explicação real. Por isso propomos que só se analisem como lexificações sintagmas de uso particularmente freqüente. São muitas vezes termos pertencentes a linguagens técnicas, e nesses casos a freqüência se verifica apenas no âmbito da linguagem técnica em questão.

Não nos parece conveniente prescindir de todo dessa noção porque estamos convencidos da existência na língua do processo de lexificação, talvez identificável em mais de um estágio (de incipiente a total). E' o caso de

(10) a. Parede ventricular direita

b. * Parede direita ventricular

que não nos parece passível de explicação gramatical.

A idiosincrasia da formação de um sintagma como (10) aparece mesmo na semântica: **parede ventricular direita** não se refere à parede direita do ventrículo, mas antes à parede (de todos os lados) do ventrículo direito. A formação mórfica e semântica é irregular, e parece simples efeito de um acidente histórico ligado ao surgimento do termo.

3.2. SELEÇÃO IDIOSSINCRÁTICA

Os casos em que se observa a acepção classificadora (“C”) de um TL que vem após o indicador parecem ser estágios incipientes de lexificação.

Sabemos que o TL é classificador em **engenheiro eletrônico** (uma pessoa que lida com eletrônica), mas não em **professor eletrônico** (isto é, uma máquina de ensinar, e não uma pessoa que ensina eletrônica). Não vemos nenhuma motivação semântica para essa diferença de interpretação, já que nada impediria, em princípio, de chamarmos de **professor eletrônico** uma pessoa que ensinasse eletrônica. A seleção nesses casos é idiossincrática: a acepção classificadora do adjetivo determina (“seleciona”) determinado grupo de itens compatíveis com ela; no caso de **eletrônico**, é selecionado o item **engenheiro**, mas não o item **professor**. Essa escolha parece ser independente de quaisquer propriedades semânticas do item, assim como de suas propriedades mórficas.

Por outro lado, não se observa esse tipo de seleção idiossincrática funcionando para a acepção não-classificadora. Em outras palavras, qualquer coisa pode ser “eletrônica” na acepção não-classificadora, mas apenas algumas (**engenheiro**, **engenharia**, **técnico** ...) podem ser “eletrônicas” na acepção classificadora.

Esse fenômeno pode ser interpretado como um sinal incipiente de lexificação; de certo modo, seria adequado encarar **engenheiro eletrônico** como um item único, possivelmente armazenado no léxico como um bloco. Sob outros pontos de vista, contudo, essas seqüências se comportam como formações mórficas normais, e não como expressões idiomáticas: por exemplo, os elementos de **engenheiro eletrônico** podem, na fala, ser separados por pausa, hesitação ou interjeição, o que não se dá com as expressões idiomáticas (James, 1972). Teríamos, portanto, em **engenheiro eletrônico** um caso intermediário entre as expressões idiomáticas e as construções sintáticas normais.

4. ALGUNS FATOS E SUA INTERPRETAÇÃO

4.1. EFEITO DA VÍRGULA

Tradicionalmente, o acréscimo de uma vírgula separando o indicador de um TL seguinte acarreta mudança de função: o TL (“adjunto adnominal”) passa a ser “aposto”. Em termos da língua escrita essa análise faz sentido, e será mantida aqui. Vamos agora examinar brevemente as conseqüências semânticas dessa mudança de funções.

Vimos acima que, como conseqüência da ação da Condição x-C, um TL só pode ter a acepção classificadora se vier imediatamente depois do indicador (ou então imediatamente depois de outro TL também com acepção classificadora, estando este último contíguo ao indicador). Podemos exemplificar com

(11) O engenheiro mecânico

(12) Um ataque cardíaco

Em ambos os casos o TL depois do indicador é claramente de acepção classificadora.

No entanto, se introduzirmos uma vírgula, não só mudaremos a função de TL para aposto, como forçaremos uma interpretação não-classificadora (provavelmente qualificativa):

(13) O engenheiro, mecânico, começou logo o trabalho.

(14) ?? O ataque, cardíaco, ocorreu na quinta feira.

Note-se como (13) nos força a entender o engenheiro como um robô, ou coisa parecida. Nessa acepção, **mecânico** aceita **muito** e **modalizadores**, o que indica que se trata de uma acepção qualificativa. Igualmente, em (14), **cardíaco** precisa ser interpretado qualificativamente, o que resulta em baixa de aceitabilidade; já não está disponível a interpretação “ataque do coração”; ficamos com “ataque de

características cardíacas”, que é difícil de processar.

Esse efeito da vírgula, e da concomitante mudança de função mórfica, parece ser geral, e aponta para a possibilidade de que a Condição x-C seja sensível à função mórfica do elemento examinado, e não apenas a sua ordenação. Isso já está incluído na definição da condição, dada na seção 2.1: ela se aplica apenas a TLs. Em algum lugar da gramática deverá ser incluída também a informação de que um aposto não pode ter interpretação classificadora.

Se encaramos a formulação das condições em termos de facilidade de emblocamento, vemos que o efeito da diferença de funções (TL / aposto) é inteiramente esperado. O emblocamento semântico é dependente da estrutura mórfica, e as funções mórficas estão tradicionalmente ligadas à estrutura de constituintes; por exemplo, é geralmente aceito que um “adjunto adnominal” se liga ao “núcleo do sintagma” de maneira mais estreita do que um aposto.

4.2. ACEITAÇÃO DE COMPLEMENTO

O complemento do sintagma adjetivo (CSA) é o termo que se segue ao adjetivo em

(15) Um artista ansioso *por reconhecimento*

(16) Uma pesquisadora digna *de um Nobel*

(O CSA é uma função mórfica proposta e definida em Perini, 1989; 1995).

A ocorrência de CSA é incompatível com a interpretação classificadora; parece ser exclusiva de TLs com interpretação qualificativa.

A testagem ampla das possibilidades de ocorrência do CSA é dificultada pelo fato de que a maioria dos adjetivos não o aceita. No entanto, é preciso observar que nos casos em que ocorre o CSA a interpretação do adjetivo é invariavelmente qualificativa. Este é um

fato que precisa ser mais investigado e integrado no conjunto da análise.

4.3. COORDENAÇÃO DE TERMOS LIVRES

A coordenação de TLs está sujeita a restrições que acreditamos serem semânticas. Assim, temos

(17) Uma moça alta e gorda

com dois TLs em acepção qualificativa, mas não

(18) * Um engenheiro mecânico e brasileiro

com um TL em acepção classificadora e outro em acepção não-classificadora. Já em

(19) Um engenheiro mecânico e nuclear

a aceitabilidade parece ser bem mais alta (ambos os TLs são classificadores); se (19) parece um pouco estranho, isso se deve provavelmente à raridade da situação descrita.

Ao que tudo indica, a coordenação só se faz se ambos os TLs tiverem interpretações paralelas. Em (18), pois, o problema está provavelmente na dificuldade em interpretar **mecânico** como qualificativo, ou então **brasileiro** como classificador.

Em (17) o problema não se coloca, porque tanto **alta** quanto **gorda** podem ser entendidos facilmente como qualificativos.