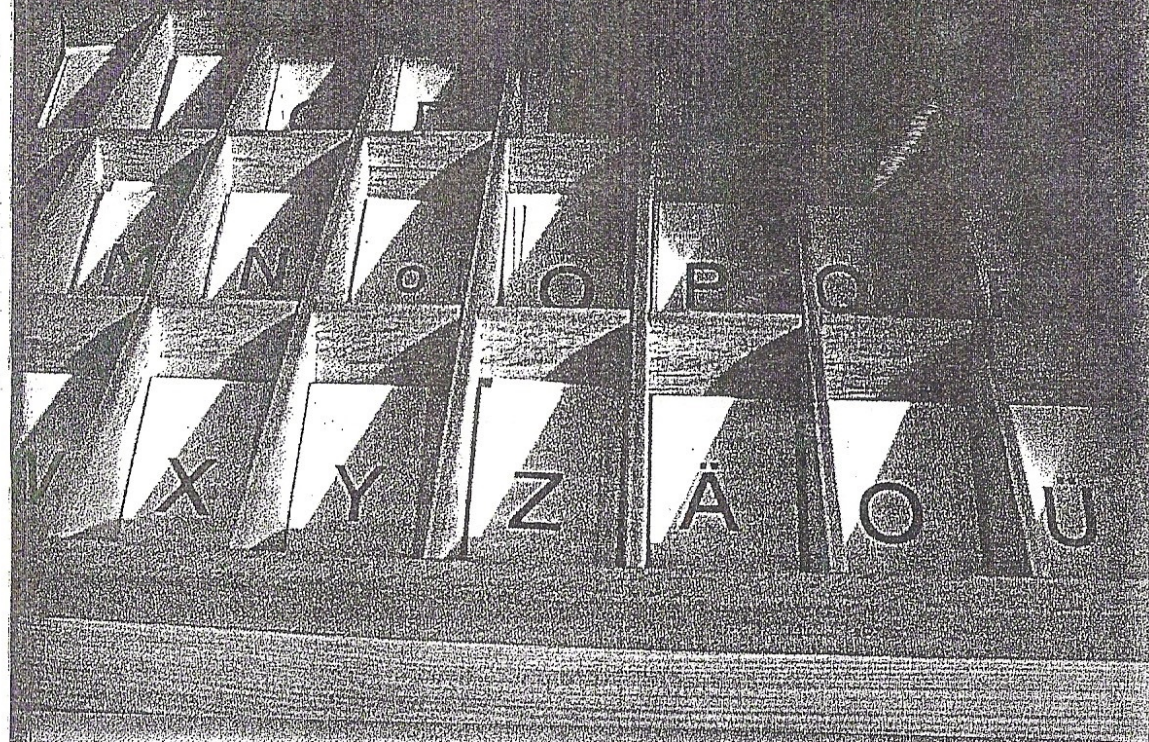


Lingua[gem]

1. *Português ou brasileiro? Um convite à pesquisa*, 6ª ed.
Marcos Bagno
2. *Linguagem & comunicação social — visões da lingüística moderna*
Manoel Luiz Gonçalves Corrêa
3. *Por uma lingüística crítica*, 2ª ed.
Kanavillil Rajagopalan
4. *Educação em língua materna: a sociolingüística na sala de aula*, 4ª ed.
Stella Maris Bortoni-Ricardo
5. *Sistema, mudança e linguagem — um percurso pela história da lingüística moderna*
Dante Lucchesi
6. "O português são dois" — *novas fronteiras, velhos problemas*, 2ª ed.
Rosa Virgínia Mattos e Silva
7. *Ensaio para uma sócio-história do português brasileiro*
Rosa Virgínia Mattos e Silva
8. *A lingüística que nos faz falhar*
Fábio Lopes da Silva, Kanavillil Rajagopalan [orgs.]
9. *Do signo ao discurso — Introdução à filosofia da linguagem*
Inês Lacerda Araújo
10. *Ensaio de filosofia da lingüística*
José Borges Neto
11. *Nós chegemos na escola. E agora?*, 2ª ed.
Stella Maris Bortoni-Ricardo
12. *Doa-se lindos filhotes de poodle — Variação lingüística, mídia e preconceito*
Maria Marta Pereira Scherre
13. *A geopolítica do inglês*
Yves Lacoste [org.] e Kanavillil Rajagopalan
14. *Gêneros — teorias, métodos, debates*, 2ª ed.
J. L. Meurer, Adair Bonini, Désirée Motta-Roth [orgs.]
15. *O tempo nos verbos do português: uma introdução à sua interpretação semântica*
Maria Luiza Monteiro Sales Corôa
16. *Considerações sobre a fala e a escrita — fonologia em nova chave*
Darcília Simões
17. *Princípios de lingüística descritiva — introdução ao pensamento gramatical*, 2ª ed.
M. A. Perini
18. *Fundamentos empíricos para uma teoria da mudança lingüística*
U. Weinreich, W. Labov, M. I. Herzog
19. *Por uma lingüística aplicada INdisciplinar*
Luiz Paulo da Moita Lopes (org.)
20. *Origens do português brasileiro*
Anthony Julius Naro | Maria Marta Pereira Scherre

MÁRIO A. PERINI



Princípios de lingüística descritiva

INTRODUÇÃO AO PENSAMENTO GRAMATICAL

parábola

EDITOR: Marcos Marcionilo

CAPA E PROJETO GRÁFICO: Andréia Custódio

CONSELHO EDITORIAL: Ana Stahl Zilles [Unisinos]

Carlos Alberto Faraco [UFPR]

Egon de Oliveira Rangel [PUCSP]

Gilyan Müller de Oliveira [UFSC, Ipol]

Henrique Monteagudo [Universidade de Santiago de Compostela]

José Carlos Sebe Bom Meihy [NEHO/USP]

Kanavillil Rajagopalan [Unicamp]

Marcos Araújo Bagno [UnB]

Maria Marta Pereira Scherre [UFRJ, UnB]

Rachel Gazolla de Andrade [PUC-SP]

Salma Tannus Muchail [PUC-SP]

Stella Maris Bortoni-Ricardo [UnB]

CIP-BRASIL. CATALOGAÇÃO NA FONTE
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ

P522p

Perini, Mário A. (Mário Alberto)

Princípios de lingüística descritiva: introdução ao
pensamento gramatical / Mário A. Perini. - São Paulo :
Parábola Editorial, 2006.

(Lingua[gem]; 17)

Inclui bibliografia

ISBN 978-85-88456-52-4

1. Língua portuguesa - Gramática. 2. Língua portuguesa -
Brasil. 3. Lingüística. I. Título. II. Série.
06-1085

CDD 469.5

CDU 811.134.3'36

Direitos reservados à

PARÁBOLA EDITORIAL

Rua Clemente Pereira, 327 - Ipiranga

04216-060 São Paulo, SP

Fone: [11] 6914-4932 | Fax: [11] 6215-2636

home page: www.parabolaeditorial.com.br

e-mail: parabola@parabolaeditorial.com.br

Todos os direitos reservados. Nenhuma parte desta obra pode ser reproduzida
ou transmitida por qualquer forma e/ou quaisquer meios (eletrônico ou me-
cânico, incluindo fotocópia e gravação) ou arquivada em qualquer sistema
ou banco de dados sem permissão por escrito da Parábola Editorial Ltda.

ISBN: 978-85-88456-52-5 (antigo 85-88456-52-4)

2ª REIMPRESSÃO: junho de 2007

© do texto: Mário A. Perini, 2006

© da edição: Parábola Editorial, São Paulo, maio de 2006

ÍNDICE

Prólogo	9
Apresentação	11
Objetivos	11
Focalizando a gramática	13
O que incluir?	13
Sobre o treinamento do lingüista	14
Sugestões: como usar o livro em aula	15
Agradecimentos	15
0. Introdução: Gramática e lingüística	17
0.1. De que trata este livro	17
0.2. Descrição e prescrição	21
0.3. O que é uma "gramática"?	23

1ª PARTE: NOÇÕES BÁSICAS

1. Fatos e hipóteses	27
1.1. De aves e hipóteses	27
1.2. Classificando palavras	28
1.3. A gramática é feita de hipóteses	29
1.4. O que é que é tão difícil?	32
1.5. As hipóteses e a estrutura da língua	32
2. De onde vêm os dados?	35
2.1. Dados e hipóteses	35
2.2. Coletando dados	36
2.3. Vantagens e perigos	37
2.3.1. O <i>corpus</i>	37
2.3.2. <i>Introspecção</i>	39
2.4. Agarre-se aos fatos!	40
3. Forma e significado	41
3.1. De onde tiramos o significado?	41
3.2. Uma teoria errada sobre as palavras e o contexto	43
3.3. O significado da estrutura	45

3.3.1. <i>Ordem dos termos</i>	45
3.3.2. <i>Constituintes</i>	46
3.4. <i>Ambigüidade</i>	47
4. <i>O objetivo da descrição</i>	49
4.1. <i>Selecionando o que é relevante</i>	49
4.2. <i>O caráter explícito da análise</i>	51
5. <i>Componentes do significado</i>	55
5.1. <i>Estratégias de interpretação</i>	55
5.2. <i>Os dois componentes</i>	57
6. <i>Onde entra a gramática</i>	59
6.1. <i>Regras gramaticais</i>	59
6.2. <i>Definições</i>	60
6.3. <i>Regras para quê?</i>	61
6.4. <i>Sintagmas nominais e verbos</i>	64
7. <i>Princípios de análise</i>	67
7.1. <i>O que é uma língua?</i>	67
7.2. <i>Afirmações formais, semânticas e simbólicas</i>	68
8. <i>Regras descritivas</i>	71
8.1. <i>Descrivendo a realidade da língua</i>	71
8.2. <i>Agentes e pacientes antes do verbo</i>	71
9. <i>Sintaxe e semântica</i>	75
9.1. <i>Formas sem significado</i>	75
9.2. <i>Formas com significado</i>	76
9.3. <i>A sintaxe residual</i>	77
10. <i>Relações simbólicas na morfologia e na sintaxe</i>	83
10.1. <i>Nomeação e falação</i>	83
10.2. <i>Bons relógios e relógios japoneses</i>	87

2ª PARTE: INSTRUMENTOS DE ANÁLISE

11. <i>Níveis e unidades</i>	93
11.1. <i>Palavra, lexema, sintagma</i>	93
11.2. <i>Tipos de sintagmas</i>	95
11.3. <i>Associação dos sintagmas</i>	99
11.4. <i>Sintagmas: unidades de forma e de significado</i>	100
11.5. <i>Estrutura sintagmática e a montagem do significado</i>	101
11.6. <i>Propriedades dos constituintes</i>	102
11.6.1. <i>Constituintes, uns dentro dos outros</i>	102
11.6.2. <i>Constituintes e o contexto sintático</i>	103

12. <i>Funções sintáticas</i>	105
12.1. <i>Funções</i>	105
12.2. <i>Relações de ordem</i>	105
12.2.1. <i>Sujeito e objeto direto</i>	105
12.2.2. <i>Complicações</i>	109
12.2.3. <i>Orações sem sujeito</i>	109
12.2.4. <i>Predicativo do sujeito</i>	111
12.3. <i>Relações de concordância</i>	113
13. <i>Funções sintáticas: mais exemplos</i>	115
13.1. <i>O núcleo do predicado</i>	115
13.2. <i>Predicativo do objeto</i>	116
13.3. <i>Adverbiais</i>	117
13.4. <i>Termos do sintagma nominal</i>	118
13.5. <i>Preposições e seus complementos</i>	120
13.6. <i>Resumo: funções sintáticas</i>	120
14. <i>Funções semânticas</i>	121
14.1. <i>Agente, paciente</i>	121
14.2. <i>Lista dos papéis temáticos</i>	122
14.3. <i>Papéis temáticos múltiplos</i>	127
15. <i>Classes</i>	131
15.1. <i>Para que classificar?</i>	131
15.2. <i>Classificação por traços</i>	134
15.3. <i>Reaprendendo as classes</i>	135
16. <i>Classes e funções</i>	137
16.1. <i>Funções, classes e o contexto</i>	137
16.2. <i>Potencial funcional</i>	138
16.3. <i>Cada forma pertence a uma (e só a uma) classe</i>	140
17. <i>Construções</i>	141
17.1. <i>Montando estruturas</i>	141
17.2. <i>Definição das construções</i>	142
17.3. <i>As construções do português</i>	143
17.4. <i>Uma gramática de construções</i>	144
17.5. <i>As construções e a valência dos verbos</i>	145
17.6. <i>Revisitando o núcleo do predicado</i>	147
18. <i>O léxico</i>	151
18.1. <i>O tamanho da tarefa</i>	151
18.2. <i>Léxico e gramática</i>	154
18.3. <i>Conteúdo de um item léxico</i>	156
18.4. <i>A língua e o mundo</i>	157

3ª PARTE: GRAMÁTICA DO PORTUGUÊS

19. Classes: alguns exemplos	161
19.1. Advérbios	161
19.2. Preposições	164
19.3. Conjunções	165
19.4. Mais complexidades	167
19.5. Nominais	168
20. Sintagmas oracionais	169
20.1. Orações subordinadas	169
20.2. Estruturas oracionais especializadas	170
20.2.1. <i>Subjuntivo</i>	170
20.2.2. <i>Infinitivo</i>	172
20.2.3. <i>Gerúndio</i>	173
20.3. Infinitivo e gerúndio com verbos auxiliares	173
21. Período composto	175
21.1. Sumário: o período composto	175
21.2. Limites da oração subordinada	175
21.3. Limites da oração principal	176
21.4. Resgatando a teoria (mas não a prática) tradicional	177
22. Classes: <i>Todos, esses e cinco</i>	181
22.1. Três palavras	181
22.2. Classificando	182
22.2.1. <i>Todos</i>	182
22.2.2. <i>Esses</i>	183
22.2.3. <i>Cinco</i>	184
22.3. Definindo as classes	185
22.4. Revisitando a complexidade	187
23. Construções de tópico	189
23.1. Falhas da análise sintática	189
23.2. Tópico sentencial	189
23.2.1. <i>O que é o tópico sentencial</i>	189
23.2.2. <i>Tópico sentencial, agente e sujeito</i>	191
23.2.3. <i>Mecanismos de topicalização</i>	193
24. Tópico discursivo	195
24.1. O que é tópico discursivo	195
24.2. Função comunicativa do tópico discursivo	197
24.3. Tópicos, sujeitos e a classificação das línguas	200
Bibliografia	203
Índice remissivo	205

PRÓLOGO

Muito poucas vezes tive a oportunidade (e o privilégio) de conversar com o professor Mattoso Camara (1904-1970), introdutor dos estudos lingüísticos modernos no Brasil. Numa dessas vezes, comentamos¹ um livro recém-publicado, no qual o autor se propunha “trocar em miúdos” a lingüística, para benefício de alunos que achavam muito difícil o livro de Mattoso, *Princípios de lingüística geral*. Nessa ocasião, Mattoso se mostrou cético quanto à possibilidade de trocar em miúdos um assunto tão técnico, abstrato e controverso quanto era (e continua sendo) a lingüística.

Não tenho dúvidas, mesmo agora depois de toda uma carreira ensinando lingüística, de que Mattoso tinha razão. No entanto, é preciso continuar tentando — mas sem nunca perder de vista que simplificação tem limites e que somente dar informações sobre lingüística não é *ensinar* lingüística. Uma ciência se ensina a quem está disposto a investigar, não a quem quer ficar sabendo coisas sobre o seu objeto. Em outras palavras, uma introdução à lingüística não é uma obra de divulgação², mas o primeiro passo na formação de um profissional. Conseqüência: apresenta sempre certo grau de dificuldade, e é obra para ser estudada, não somente lida.

É claro que não há desculpa para se produzir um texto desnecessariamente difícil. Ao escrever, tento lembrar-me sempre da afirmação de Karl Popper de que o autor não deve nunca defender seu texto da acusação de ser obscuro; é responsabilidade dele produzir o texto mais claro possível — ainda mais quando se trata de um texto introdutório. Mas, por outro lado, é preciso levar em conta a natureza de cada texto, e de cada disciplina. Não creio que ninguém se deite na cama com um texto de matemática

¹ O prof. Mattoso como o maior lingüista do Brasil, e eu como um pretensioso recém-formado de 25 anos.

² Não tenho nada contra obras de divulgação, e eu mesmo já escrevi algumas. Mas seu objetivo é outro: informar o leitor interessado a respeito de aspectos selecionados de uma disciplina, não iniciar seu treinamento como profissional da área.

ou de lógica simbólica com a intenção de lê-lo antes de dormir. Há livros feitos para serem estudados, e estes não são adequados para leitura amena, não se destinam à fruição estética nem servem para quem procura informação fácil sobre alguma ciência. Uma introdução não troca em miúdos uma disciplina; apenas procura apresentá-la, com algo de sua riqueza e complexidade, de maneira racional e digerível.

Esse é o objetivo do presente livro. Seu âmbito é mais restrito do que o de Mattoso, e se limita aos aspectos sintático-semânticos da linguagem — deixando de lado outras perspectivas, como a histórica, a social, a psicológica, a textual, e igualmente omitindo a descrição fonética, fonológica e morfológica. Tive que fazer esse recorte para não sacrificar a profundidade e a extensão das discussões, que considero fundamentais para que o leitor tenha uma idéia adequada do que é “fazer lingüística”. As outras áreas merecem, e em parte têm recebido, tratamento especial em manuais separados.

E (desculpem a pretensão) escolhi um título que ecoa o do livro de Mattoso, um livro que ainda compensa ler sessenta e cinco anos após sua publicação.

APRESENTAÇÃO

OBJETIVOS

Este livro é uma introdução aos estudos lingüísticos, dirigida primariamente a estudantes de graduação. Não é uma introdução a uma teoria particular; procuro transmitir aqui um conjunto de conhecimentos e habilidades que o lingüista deve adquirir já de partida, antes mesmo de tomar contato com uma das modalidades de teoria: cognitivista, gerativista, funcionalista, ou seja lá qual for.

Por outro lado, não há aqui nenhuma insinuação de que as teorias não sejam necessárias. O que há é a convicção de que um lingüista deve ser consciente das escolhas teóricas que faz, e que isso depende de modo fundamental de sua visão dos fatos — o que por sua vez depende de treinamento em observação, coleta e sistematização dos dados da língua. Este livro pretende ser um subsídio para esse treinamento, sem o qual o iniciante fica à mercê de idéias alheias, arriscando-se a nunca se conscientizar devidamente do caráter empírico da lingüística e da fragilidade das teorias.

Não pretendo discutir as razões pelas quais as teorias são frágeis, muito menos sugerir soluções. Ofereço aqui uma introdução às técnicas de análise lingüística: técnicas de obtenção de dados, de sistematização e de interpretação superficial, sem as quais, acredito, nunca vai ser possível elaborar teorias mais sólidas do que as atuais. Essas técnicas permitem elaborar uma descrição que tem utilidade em duas áreas principais: primeiro, leva ao conhecimento sistemático dos fatos de uma língua, o que interessa aos estudiosos dessa língua e também aos que a ensinam ou estudam para chegar a seu uso na prática; e, depois, fornece ao lingüista teórico uma base de dados confiável para construir e testar eventuais teorias.

A descrição, evidentemente, não é isenta de teoria. Existe, por trás dos conceitos incluídos neste livro, um grande corpo de princípios teóricos. Isso é inevitável, pois a alternativa seria pura e simplesmente acumular dados sem nenhuma interpretação, algo que seria pouco útil, e que por

outro lado não se justifica, já que os pontos teóricos geralmente aceitos são numerosos. Ninguém questiona, por exemplo, a conveniência de se dividir as palavras em classes, segundo seu comportamento gramatical; ninguém questiona que uma dessas classes (chamada “verbo”) inclui as palavras (mais exatamente, os lexemas) que “variam” em pessoa e número (*falto, faltamos*), em oposição às que variam em gênero (*vermelho, vermelha*). Como essas, há um imenso número de coisas que todo lingüista aceita tacitamente ao desenvolver seu trabalho. Ignorar esse consenso seria dar um passo atrás inteiramente desnecessário. As propostas deste livro são, portanto, fundamentadas em muitos pontos teóricos de aceitação geral.

Por outro lado, esse conjunto de pontos não constitui uma teoria unificada universalmente aceita, que até aí ainda não chegamos em lingüística. Por isso, não foi possível evitar a introdução de algumas posições mais específicas, e estas derivam principalmente da influência da corrente denominada “gramática cognitiva”, tal como se apresenta, principalmente, nos trabalhos de Ronald W. Langacker, cuja influência é evidente neste livro. Mas este livro não é uma introdução à gramática cognitiva — é apenas uma introdução às técnicas de análise lingüística, em que, quando inevitável, princípios e conceitos cognitivistas ocupam o lugar que teria que ser ocupado de qualquer maneira. Espero que o livro seja de utilidade também para estudantes e pesquisadores que favorecem outras escolas de pensamento lingüístico. Na impossibilidade de elaborar um livro verdadeiramente “neutro”, achei que essa era a melhor solução.

Apesar dessa situação um tanto incômoda no que diz respeito às teorias, é necessário reconhecer que alguma teorização precisa estar presente por trás de toda afirmação gramatical. Isso pressupõe um alto grau de coerência, de modo que deve ser possível extrair da discussão um sistema teórico que, embora parcial e provisório, faça sentido como um todo — afinal de contas, precisamos, pelo menos, caminhar na direção de uma teoria unificada. Do contrário, seria impossível fazer discussões racionais sobre os fatos da língua. Minha convicção é que a área de consenso mencionada acima é suficientemente ampla para permitir esse tipo de coerência. Nas áreas em que tal consenso é frágil, optei por adotar uma posição relativamente tradicional, mesmo quando tenho restrições quanto a ela. Assim, defini por exemplo as funções de “sujeito” e “objeto direto” da maneira usual, para não trazer à baila, em um livro introdutório, questões ainda mal resolvidas.

FOCALIZANDO A GRAMÁTICA

O foco escolhido para este livro são os estudos gramaticais *stricto sensu*. Isso significa que tive que deixar de lado áreas como a análise do discurso, a pragmática e as relações entre o conhecimento da língua e a cognição em geral, para privilegiar a sintaxe, a semântica e as relações entre elas. Fatores relativos ao conhecimento do mundo são mencionados com freqüência, mas apenas quando são relevantes para o estudo da morfossintaxe e da semântica. Também deixei de lado a fonologia, a sociolingüística, a psicolingüística e a lingüística histórica. Essa decisão obedece à opinião de que para o iniciante é mais fácil desenvolver técnicas de análise em áreas nas quais já tem alguma competência — quando mais não seja, por seus estudos anteriores de gramática tradicional —, para depois passar a estudos que requerem uma formação técnica mais intensiva, como a fonologia, ou uma certa mudança de perspectiva frente ao objeto, como a análise do discurso. As habilidades a serem desenvolvidas com a ajuda deste livro, porém, são de utilidade geral, aonde quer que o estudante seja levado por seus interesses posteriores. Pelas mesmas razões, evitei, tanto quanto possível, o uso de formalizações que, mesmo quando úteis, costumam ter como efeito desviar a atenção do estudante dos problemas centrais. Uma vez adquirida uma compreensão básica dos fenômenos gramaticais, a formalização, onde se justifica, pode ser encarada em função de sua eventual utilidade.

O QUE INCLUIR?

De todos os problemas que tive que enfrentar ao escrever este livro, o mais difícil foi o de selecionar o que incluir e o que deixar de fora. À primeira vista tudo parece indispensável, mas é essencial manter o livro dentro de limites bastante estreitos, de modo que possa ser estudado em, no máximo, dois semestres, não só por causa da limitação do currículo, mas também por se levar em conta que é desejável que os alunos tenham alguma exposição a outras áreas da lingüística, além da gramática. Assim, tive que me contentar com uma lista muito restrita de tópicos, e certamente alguns professores a acharão inadequada.

Em diversas passagens, o professor achará que os fatos foram simplificados. Isso é inevitável em uma obra introdutória como esta; a preocupação de ser exato e completo a cada passo só leva, a meu ver, a textos de complexidade proibitiva para o iniciante. Assim, solicito alguma tolerân-

cia para com regras, classificações e princípios menos do que ótimos, em favor do treinamento e da informação básica que eles representam. Pelas mesmas razões, a ordenação da matéria segue mais um critério didático do que conceitual; por isso tive que dividir certos assuntos muito importantes (como a classificação das formas) em mais de um bloco, introduzindo o tema em nível elementar e voltando a ele posteriormente quando se tornou possível um tratamento mais aprofundado.

SOBRE O TREINAMENTO DO LINGÜISTA

As técnicas descritivas são (ou deveriam ser) um componente essencial do treinamento do lingüista, sem importar sua área de atuação ou sua orientação teórica. Escrevi este livro para cobrir uma lacuna na formação atual dos lingüistas, que freqüentemente são informados sobre aspectos teóricos, mas se mostram lamentavelmente deficientes no que se refere à análise das línguas propriamente dita. E acho importante que o estudante tenha condições de trabalhar com dados em função de seu conhecimento lingüístico, evitando a atitude tão freqüente de que os fatos são propriedade de alguns especialistas: gramáticos profissionais, professores ou, em casos patológicos, os "comandos paragramaticais" definidos por Bagno (2003). É essencial transmitir a idéia de que o falante de uma língua conhece, ainda que implicitamente, os fatos dessa língua, e portanto tem autoridade para fazer julgamentos de aceitabilidade.

Por isso, toda a exemplificação e discussão deste livro têm como centro a língua portuguesa; a referência a outras línguas é ocasional. A ênfase é toda na análise do português — mais especificamente, naquilo que distingue o português das outras línguas. Evidentemente, isso não significa que os princípios aqui expostos não tenham importância para a análise das línguas em geral; apenas, escolhi a análise do português como porta de entrada aos estudos lingüísticos, partindo da observação de que princípios gerais são mais bem assimilados quando o estudante consegue relacioná-los com exemplos concretos de análise elaborados com base em dados conhecidos.

Além de fornecer noções básicas de análise e exemplificação tirada da estrutura do português, o texto objetiva desenvolver no estudante uma certa sensibilidade aos fenômenos lingüísticos, algo que lhe permita saber o que observar ao se ver confrontado com dados. Essa é a finalidade principal da ampla exemplificação oferecida.

Ao organizar o texto, admiti certa dose de redundância, de maneira que o leitor encontrará por vezes o mesmo assunto mencionado e discutido em mais de um lugar. Isso facilita a leitura, evitando a irritação de ter que voltar repetidamente a capítulos anteriores — o que também prejudica a concentração e o esforço de compreensão global de cada capítulo. Também com vistas a facilitar o uso, procurei apresentar a matéria em capítulos tão breves quanto possível.

As notas de pé de página são dirigidas mais ao professor do que ao aluno, de maneira que este pode ler o livro e saltar as notas, sem prejuízo da compreensão.

SUGESTÕES: COMO USAR O LIVRO EM AULA

As duas primeiras partes (a introdução, mais os capítulos 1 a 18) foram planejadas para serem vistas em um semestre normal de 60 horas-aula. Esses capítulos devem fornecer ao aluno uma formação básica em gramática, o que constitui a meu ver a melhor porta de entrada aos estudos lingüísticos.

A terceira parte (capítulos 19 a 24) é composta principalmente de questões de gramática portuguesa, discutidas de maneira um pouco mais aprofundada, e às vezes sem chegar a uma solução satisfatória (é preciso lembrar sempre que a estrutura do português é um território inexplorado!). Com turmas mais preparadas, podem ser vistas no mesmo semestre, depois de terminadas as duas primeiras partes. Alternativamente, podem ser parte do material de leitura em um segundo semestre de gramática. Ou, ainda, podem servir ao professor como uma espécie de banco de exemplos, selecionados à medida que o curso avança, e dosados segundo a capacidade da turma.

AGRADECIMENTOS

Para terminar, gostaria de agradecer o auxílio prestado por alguns amigos que tiveram a gentileza de comentar partes deste livro durante sua elaboração. Sou grato a Lúcia Fulgêncio, que leu o texto e fez sugestões valiosas (literalmente a cada página), obrigando-me a rever a organização final e um grande número de detalhes; e também a Yara Bruno e George Manes Pereira, que detectaram diversas obscuridades que tentei sanar. Finalmente, agradeço à Parábola Editorial (Andréia Custódio e Marcos

Marcionilo) pela oportunidade de publicar este livro, apresentando-o aos alunos e professores dessa matéria tão mal-falada, mas no fundo tão fascinante, que é a gramática.

M.A.P.

*Pontifícia Universidade Católica de Minas Gerais
Maio de 2006*

INTRODUÇÃO: GRAMÁTICA E LINGÜÍSTICA

0.1. DE QUE TRATA ESTE LIVRO

Este livro é uma introdução à gramática — ou, mais exatamente, ao pensamento gramatical: trata das teorias e métodos utilizados para descrever a organização formal e semântica da língua. E a gramática é uma parte da lingüística; mas como as duas não se identificam, vou, antes de mais nada, delimitar nosso campo de interesse neste livro.

A disciplina que chamamos “lingüística” se ocupa da linguagem humana sob muitos aspectos. Assim, caem dentro do campo da lingüística temas como:

- a aquisição da linguagem durante os primeiros anos de vida e os mecanismos mentais que nos permitem falar e entender uma língua (psicolingüística);
- as variedades de uma língua segundo as regiões em que é falada (dialetologia);
- as diferentes variedades de uma língua utilizadas por pessoas de diferentes classes sociais, faixas etárias e grupos de atividade (sociolingüística);
- a evolução de uma língua ao longo de sua história (lingüística histórica, ou lingüística diacrônica);
- a maneira pela qual se organizam os textos falados ou escritos para atender a suas necessidades comunicativas (análise do discurso);
- as condições de uso da linguagem em relação com a identidade do emissor e do receptor, assim como com o conhecimento que cada um tem sobre o mundo e sobre a situação em que se dá a comunicação (pragmática);

Marcionilo) pela oportunidade de publicar este livro, apresentando-o aos alunos e professores dessa matéria tão mal-falada, mas no fundo tão fascinante, que é a gramática.

M.A.P.

Pontifícia Universidade Católica de Minas Gerais
Maio de 2006

INTRODUÇÃO: GRAMÁTICA E LINGÜÍSTICA

0.1. DE QUE TRATA ESTE LIVRO

Este livro é uma introdução à gramática — ou, mais exatamente, ao pensamento gramatical: trata das teorias e métodos utilizados para descrever a organização formal e semântica da língua. E a gramática é uma parte da lingüística; mas como as duas não se identificam, vou, antes de mais nada, delimitar nosso campo de interesse neste livro.

A disciplina que chamamos “lingüística” se ocupa da linguagem humana sob muitos aspectos. Assim, caem dentro do campo da lingüística temas como:

- a aquisição da linguagem durante os primeiros anos de vida e os mecanismos mentais que nos permitem falar e entender uma língua (psicolingüística);
- as variedades de uma língua segundo as regiões em que é falada (dialetologia);
- as diferentes variedades de uma língua utilizadas por pessoas de diferentes classes sociais, faixas etárias e grupos de atividade (sociolingüística);
- a evolução de uma língua ao longo de sua história (lingüística histórica, ou lingüística diacrônica);
- a maneira pela qual se organizam os textos falados ou escritos para atender a suas necessidades comunicativas (análise do discurso);
- as condições de uso da linguagem em relação com a identidade do emissor e do receptor, assim como com o conhecimento que cada um tem sobre o mundo e sobre a situação em que se dá a comunicação (pragmática);

- a aplicação dos resultados da lingüística à solução de problemas práticos, como o ensino de línguas, o aprendizado da leitura ou mesmo a elaboração de linguagens de computador (lingüística aplicada); e vários outros.

Neste livro não nos ocuparemos de nenhum dos temas listados acima; nosso assunto será

- o estudo da organização interna dos enunciados lingüísticos de cada língua, tanto no que diz respeito à sua forma (realizada através dos sons da fala ou de símbolos escritos) quanto no que diz respeito ao seu significado (gramática).

E mesmo esse assunto não vai ser abordado no total. Vamos precisar entrar em detalhes de como analisar os fatos da língua, com muitos exemplos, e por isso vamos deixar de lado dois aspectos importantes da gramática, a saber:

- o estudo dos sons da língua (fonética) e de sua organização na pronúncia (fonologia);
- o estudo da estrutura interna da palavra (morfologia).

O que é que sobra? Como se verá, muita coisa. Nosso estudo gramatical vai se concentrar em dois aspectos, a saber:

- o estudo da maneira como as palavras se organizam para formar frases e outras unidades menores (sintaxe); e
- o estudo do significado que os usuários da língua atribuem às estruturas sintáticas (semântica).

No decorrer do livro, não faltarão exemplos. Mas, para adiantar, vou dar um, para que fique claro o que é que se entende por "sintaxe" e "semântica". Vamos examinar a frase

[1] Helena incomodou Serginho à tarde.

Fazem parte da sintaxe afirmações como as seguintes:

(a) a posição da palavra *Helena* é relativamente fixa, porque ela não pode ser transportada para outra posição na sentença. Ou seja, as sentenças abaixo são rejeitadas pelos falantes:

[2] *Incomodou Helena Serginho à tarde.

[3] *Incomodou Serginho Helena à tarde.

[4] *Incomodou Serginho à tarde Helena.

(O sinal * é usado para marcar uma forma que não é aceita pelos falantes).

(b) A posição do elemento *à tarde* é mais livre, e ele pode ocorrer em mais de um lugar na frase:

[5] Helena incomodou Serginho à tarde.

[6] À tarde, Helena incomodou Serginho.

[7] Helena, à tarde, incomodou Serginho.

(c) A forma do verbo, *incomodou*, depende do que vem logo antes dele; com *Helena*, temos *incomodou*, mas com *eu* teríamos *incomodei*:

[8] Eu incomodei Serginho à tarde.

Todas essas afirmações, e outras mais, descrevem a forma da sentença, isto é, os aspectos diretamente observáveis, sem dizer nada sobre seu significado. Tradicionalmente, são chamadas de afirmações sintáticas, e o conjunto de todas as afirmações sintáticas constitui a sintaxe da língua.

Agora vamos mudar de ponto de vista e fazer afirmações sobre o significado da frase

[9] Helena incomodou Serginho à tarde.

Dela pode-se dizer coisas como:

(d) ela indica um evento que se deu no passado;

(e) o falante está falando de outra pessoa (em oposição a *eu incomodei Serginho à tarde*, onde o falante se refere a si mesmo);

(f) trata-se da ação de "incomodar" (em oposição a *Helena beliscou Serginho à tarde*);

(g) Helena praticou uma ação, e Serginho sofreu essa ação.

e assim por diante.

As afirmações (d), (e), (f) e (g) já não têm a ver com a forma da sentença, mas com seu significado. Não são coisas que se pode ver diretamente na seqüência de sinais gráficos, mas são coisas que uma

peessoa que sabe português percebe bem claramente. Só que aparecem na sua mente, não em sua vista ou ouvido. Essas afirmações se chamam semânticas.

Neste livro teremos que lidar tanto com fenômenos sintáticos quanto com fenômenos semânticos. E, além do mais, teremos que estudar a relação que existe entre uns e outros; essas relações se chamam simbólicas. Um exemplo de relação simbólica é a afirmação de que

(h) Com o verbo *incomodar*, o elemento que vem logo antes do verbo se refere a quem pratica a ação, e o elemento que vem logo depois se refere a quem sofre a ação¹.

Isso explica a diferença de significado entre

[10] Helena incomodou Serginho.

e

[11] Serginho incomodou Helena.

Em [10] Helena praticou a ação, mas em [11] foi Serginho.

Todos os pontos de vista mencionados são maneiras legítimas e necessárias de se estudar a linguagem; eles se completam em uma tentativa ampla de compreender o fenômeno. A divisão em áreas é, entretanto, inevitável, porque a linguagem é um fenômeno extremamente complexo. Um lingüista profissional, é claro, deve ter pelo menos alguma noção das principais áreas da lingüística, mas, dadas as limitações humanas, nenhum consegue abarcar todas elas dentro de suas atividades de pesquisa. Aqui, vamos estudar apenas a gramática, considerada como uma das áreas da lingüística; mas não devemos perder de vista a existência dos outros aspectos da linguagem, que interagem com a gramática de maneira complexa².

¹ Essa regra está bastante simplificada, é claro.

² Para algumas áreas já existem bons manuais no mercado brasileiro: a antologia de Mussalim e Bentes (2001) fornece um apanhado útil das diversas áreas da lingüística. Para a fonologia, há o livro de Silva (1999); Ilari e Galdi (1985) é uma boa introdução aos conceitos básicos da semântica, e na mesma área saiu recentemente Cançado (2005); Faraco (2005) é uma introdução à lingüística histórica; para a sociolingüística, o leitor pode consultar Tarallo (1985); e para a morfologia temos Basilio (1987 e 2004).

0.2. DESCRIÇÃO E PRESCRIÇÃO

Vimos acima que a lingüística se ocupa de muitos aspectos da linguagem e de seu uso; um aspecto do uso da linguagem de que a lingüística não se ocupa é a distinção entre o “certo” e o “errado” na língua.

O ensino do português muitas vezes difunde a crença de que existe uma maneira “certa” de usar a língua, e que essa é a única maneira aceitável; todas as outras são “erradas”, devem ser evitadas. Isso é reforçado por colunas em jornais, gramáticas escolares, livros de “não erre mais” e a pressão social de todo momento. Essa atitude, com suas perniciosas conseqüências, tem sido objeto de crítica por parte de lingüistas e professores, mas continua muito presente na escola e na vida³.

Não há a menor base lingüística para a distinção entre “certo” e “errado” — o lingüista se interessa pela língua como ela é, e não como ela deveria ser. Imagine-se um historiador que descobre que determinado povo antigo praticava sacrifícios humanos. Ele, pessoalmente, pode desaprovar esse costume, mas nem por isso tem o direito de afirmar que os sacrifícios não ocorriam — um fato é um fato, e precisa ser respeitado. No entanto, quantas vezes não nos dizem que a palavra *chipanzé* “não existe” (porque o “certo” seria *chimpanzé*)? Dizer isso é desrespeitar o fato de que milhões de pessoas dizem *chipanzé*. Um lingüista parte sempre de fatos, e a cada passo verifica suas teorias em confronto com eles: se muitos falantes dizem *chipanzé*, então ele precisa registrar esse fato, e levá-lo em conta em sua descrição e teorização. E se todo mundo diz *me dá ele aí*, essa é uma estrutura legítima da língua falada do Brasil, e precisa figurar na descrição. Vemos aqui então uma oposição entre gramáticas prescritivas (ou normativas), que tentam estabelecer como as pessoas devem falar ou escrever, e gramáticas descritivas, que procuram descrever como é que as pessoas realmente falam ou escrevem.

A oposição entre “certo” e “errado” muitas vezes corresponde, no fundo, à oposição — essa, sim, legítima — entre língua falada e língua escrita. É fato (e portanto temos que respeitar) que a gente não escreve como fala. E se é um fato, deve figurar em algum ponto de uma gramática completa da língua. Mas se é errado escrever *me dá ele aí* em uma carta formal de

³ Ver, por exemplo, Bagno (2000; 2003), e também os capítulos 1, 4 e 6 de meu livro *Sofrendo a gramática* (Perini, 1997) e o capítulo 4 de *A língua do Brasil amanhã e outros mistérios* (Perini, 2004).

pedido de emprego, é igualmente errado sentar na mesa do bar e dizer *dê-me esse copo*. Cada variedade da língua é apropriada em seu contexto próprio, e os falantes sabem isso muito bem, tanto é que empregam com toda a segurança a variedade adequada à situação do momento: ninguém fala como escreve, e ninguém escreve como fala.

Isso, já que é um fato, merece ser descrito e eventualmente ensinado. Mas note-se a diferença: não se trata de dizer que *me dá ele* é “errado”, mas que é uma forma coloquial, usada na fala. Diga-se, de passagem, que as formas faladas são usadas em uma variedade muito maior de situações, em ocasiões muito mais numerosas, por um número muito maior de falantes do que as formas escritas. Assim, elas são as representantes mais genuínas da língua do Brasil.

É preciso acrescentar que as formas “certas” defendidas por algumas pessoas nem sempre são as formas correntes na escrita, mas muitas vezes são simples ficções ou desejos de algumas pessoas. Assim, Sacconi (1990) afirma peremptoriamente que não existe a palavra *fleugmático* (o “certo” seria *fleumático* ou *flegmático*). Mas Neves (2003), autora que, em vez de sonhar com um mundo ideal, pesquisa o mundo real, mostra, a partir de dados, que na língua escrita ocorrem tanto *fleugmático* quanto *flegmático*⁴.

Diz-se às vezes que os lingüistas são permissivistas para quem “tudo vale, desde que haja comunicação”. Não é verdade. Por exemplo, praticamente ninguém questiona a conveniência de se ensinar o uso do português padrão escrito (desde que limitado aos contextos em que ele é socialmente aceito, ou seja, à língua escrita e a situações muito formais). O português padrão é, queiramos ou não, a nossa língua erudita, e, no que pese seu caráter exclusivamente escrito, está aí para ficar. O que se defende é o respeito aos fatos: a língua falada também existe e constitui um objeto de estudo interessante e importante.

Se neste livro o leitor encontrar exemplos como *minha sobrinha divorciou o ano passado* (em vez de *se divorciou*), não se trata de distração ou ignorância⁵, mas resultado da convicção de que a língua falada do Brasil é tão digna de atenção por parte dos pesquisadores quanto qualquer outra língua, falada ou escrita.

⁴ Curiosamente, Sacconi (1990, p. 39) afirma que “os nomes próprios estão sujeitos às mesmas regras dos nomes comuns”, em particular no que diz respeito a levarem acento gráfico; como exemplos, dá, entre outros, *Miron, Estêvão, Luís, Sônia*. No entanto, ele mesmo assina “Luiz Antonio”.

⁵ Não que eu seja isento de tais coisas, claro. Mas não nesse exemplo.

Um lingüista, portanto, não deve fazer julgamentos de valor a respeito de seu objeto de estudo — para ele, qualquer variedade da língua tem interesse, desde que realmente exista e seja usada (ou tenha sido usada) por uma comunidade. Uma pessoa que não consegue se libertar da sensação de que certas formas da língua são “feias”, “erradas” ou de alguma maneira desagradáveis deveria procurar outra profissão que não a de lingüista ou professor de línguas.

0.3. O QUE É UMA “GRAMÁTICA”?

Nossa disciplina, portanto, é a gramática. Mas essa palavra tem mais de um significado na linguagem cotidiana, de maneira que vale a pena esclarecer o que é que vai significar para nós.

Pode-se distinguir três sentidos principais da palavra “gramática”, a saber:

(a) gramática se entende muitas vezes como o nome de uma disciplina que procura estabelecer o que é “certo” e o que é “errado” na língua (esse tipo de gramática já foi criticado na seção 0.2). É esse o sentido da palavra quando alguém diz que “essa forma do verbo *ver* é contra a gramática”, ou “a gramática manda usar o pronome depois do verbo”. Essa, como vimos, é uma gramática prescritiva (ou normativa).

O sentido acima de gramática não é considerado neste livro. Mas há outros dois sentidos que são importantes para nós:

(b) chama-se gramática um sistema de regras, unidades e estruturas que o falante de uma língua tem programado em sua memória e que lhe permite usar sua língua. Nesse sentido, gramática é parte do nosso conhecimento do mundo. Por exemplo, os falantes do português do Brasil têm em sua memória uma regra que diz que a palavra *um* pode aparecer antes de palavras como *livro* ou *elefante*, mas não antes de palavras como *janela* ou *cabrita*. Por isso, rejeitamos as seqüências **um janela* e **um cabrita*, porque são formas que violam aquela regra. Note-se que não aprendemos isso na escola, nem estudando gramática [no sentido (a)] — isso é parte de nosso conhecimento do português, que adquirimos na infância juntamente com a localização do banheiro da nossa casa ou quais as comidas de que gostamos mais. A parte desse conhecimento que governa nosso uso da língua⁶ e que memorizamos sem nem perceber é denominada gramática internalizada.

⁶ Em certos aspectos; temos também, claro, o léxico, além das regras que controlam o uso adequado da língua em diversas situações etc.

E, finalmente,

(c) chama-se também gramática a descrição, feita por um lingüista, do sistema mencionado no item (b) acima. Nesse sentido, pode ser um livro, mas é bem diferente das gramáticas escolares a que estamos acostumados. Em vez de receitar o que se deve ou não se deve dizer, registra como se fala realmente, retratando e sistematizando os fatos da língua. A esse tipo de gramática se dá o nome de gramática descritiva.

Este livro se ocupa da gramática descritiva; e, inevitavelmente, também se ocupa da gramática internalizada, já que uma é a descrição da outra.

1ª PARTE:

NOÇÕES BÁSICAS

FATOS E HIPÓTESES

Vamos iniciar nossa excursão pela lingüística examinando uma distinção fundamental: a diferença entre fatos e hipóteses. Essa diferença não é, em si, particularmente difícil de entender; mas acontece que, aplicada à análise da língua, ela freqüentemente contraria idéias preconcebidas que recebemos quando estudamos gramática pela primeira vez. O difícil, então, não é tanto aprender quanto desaprender certas noções incorretas que temos a respeito da estrutura da língua — e às quais certas pessoas se apegam com enorme tenacidade.

1.1. DE AVES E HIPÓTESES

Para ilustrar essa diferença, vamos pegar primeiro um exemplo fora da lingüística. Digamos que um zoólogo vá estudar pela primeira vez as aves de uma ilha. Ele observa alguns fatos: algumas aves, de plumas verdes, sobrevivem comendo frutas e sementes; outras, de plumas pretas, devoram as aves verdes, assim como insetos e pequenos répteis. A partir da observação desses fatos, ele já pode construir uma hipótese, a saber: (a) naquela ilha, as aves se dividem em vegetarianas e carnívoras; (b) as vegetarianas são verdes, as carnívoras são pretas. Note-se a diferença entre o fato de que algumas aves verdes comem frutas e algumas aves pretas comem carne e a hipótese de que existem esses dois tipos de aves na ilha. A hipótese pode ser verdadeira, ou não. Por exemplo, se, continuando suas observações, o zoólogo descobre que há algumas aves verdes carnívoras, ele vai ter que reformular a hipótese; ou então, pode descobrir que algumas aves carnívoras também comem plantas, além de carne, e nesse caso a classificação precisa ser complementada com a informação de que há na verdade naquela ilha três tipos de aves: vegetarianas, carnívoras e onívoras, que comem tanto carne quanto plantas.

É fácil ver que os fatos observados pelo pesquisador são inquestionáveis (a menos que a gente queira dizer que ele não observou bem, por exemplo). Mas as hipóteses são necessariamente provisórias — mesmo a nova hipótese, que divide as aves em três grupos, pode ainda ter que ser reformulada com base em novos fatos. Como não há maneira de observar *todos* os fatos — todas as aves da ilha, em todos os seus atos de alimentação —, na verdade nunca se chega a uma hipótese definitiva. Mas isso não é realmente necessário; chega um ponto em que podemos confiar nas hipóteses, por terem sido freqüentemente testadas e confirmadas por novas observações.

Esse exemplo é fácil de compreender não apenas por ser simples. É fácil também porque não contraria expectativas nossas — como nunca estudamos a classificação das aves da ilha em questão, não temos que substituir conhecimentos anteriores pelos novos conhecimentos que nos são fornecidos.

1.2. CLASSIFICANDO PALAVRAS

Dei acima um exemplo típico (embora muito simplificado) do trabalho de um cientista ao lidar com dados. Vamos agora examinar um caso exatamente paralelo do ponto de vista lógico, mas que, em vez de tratar de aves e seus hábitos alimentícios, trata de palavras e sua relação com outras palavras na mesma frase. A ilha onde vivem essas palavras se chama “língua portuguesa”, e o pesquisador, nesse caso, é um lingüista. Vamos lidar, novamente, com fatos e hipóteses, da mesma maneira que fizemos no exemplo anterior.

O nosso lingüista observa que há algumas palavras que servem para designar coisas, como *livro* e *folha*; e há outras que servem para atribuir alguma qualidade às coisas, como *interessante* e *murcha*. Assim, quando se fala de um *livro interessante* ou uma *folha murcha*, estamos falando de livros e folhas, respectivamente, mas atribuindo a cada coisa uma qualidade. O lingüista, então, levanta uma hipótese: as palavras se dividem em duas classes, as que designam coisas (e que ele chama de “substantivos”) e as que atribuem qualidades (que ele chama de “adjetivos”). Isso está bastante de acordo com o que aprendemos na escola. Mas é preciso, antes de mais nada, notar uma coisa importante, e essa, em geral, não se aprende na escola: é um fato que algumas palavras designam coisas e outras atribuem qualidades; mas a classificação construída pelo lingüista com base nesse fato é uma hipótese — e, como toda hipótese, pode ser questionada, testada, e talvez refutada. Em geral, nos acostumamos a pensar que é um

fato que *livro* é um substantivo e *interessante* é um adjetivo — mas isso não é um fato, é uma hipótese.

Tanto é uma hipótese que pode ser — e de fato já foi — contestada. Por exemplo, dentro do grupo de palavras que tradicionalmente chamamos “substantivos” e “adjetivos”, já se mostrou que há mais do que dois tipos, mesmo levando-se em conta apenas as duas propriedades de “designar coisas” e “atribuir qualidades”. Existem palavras, como *velho*, que fazem as duas coisas: essa palavra atribui qualidade em *um livro velho*, mas em *aquele velho* designa uma coisa (no caso, uma pessoa).

Velho é um “substantivo” ou um “adjetivo”? Esta pergunta não tem resposta dentro da hipótese mencionada acima, que só nos fornece duas classes onde colocar as palavras. O problema não está na palavra *velho*, mas na hipótese, que é excessivamente simples. Lembremo-nos do zoólogo, quando descobriu que algumas aves comiam plantas e também carne. Ele não podia colocar essas aves em nenhuma das duas classes que já tinha criado; então, foi obrigado a reformular sua hipótese, complicando-a e criando uma terceira classe de aves (a das “onívoras”). Nossa situação perante palavras como *velho* é idêntica: temos duas classes, “substantivos” e “adjetivos”, e *velho* não cabe em nenhuma delas. A saída é criar uma terceira classe — o que vai implodir a dicotomia tradicional entre “substantivos” e “adjetivos”. Se quisermos manter as duas designações tradicionais, vamos ter que inventar uma nova — digamos, “ambivalentes”, pois podem funcionar dos dois modos¹.

Agora, ao classificarmos as palavras, temos que chamar *livro* de “substantivo”, *interessante* de “adjetivo” e *velho* de “ambivalente”. E é bom notar que em português os ambivalentes são bem mais numerosos do que os substantivos ou os adjetivos: alguns exemplos são *amigo*, *trabalhador*, *jovem*, *cego*, *brasileiro* e muitas outras palavras que se comportam gramaticalmente da mesma maneira que *velho*.

1.3. A GRAMÁTICA É FEITA DE HIPÓTESES

Mas, alguém vai dizer, espera aí! Nós sabemos que *velho* é um adjetivo! Como é que vamos agora rejeitar uma classificação que consta das

¹ O termo “ambivalente” não é parte da terminologia lingüística; uso-o aqui apenas para ilustrar o problema.

gramáticas e dos dicionários? De onde veio essa nova classe de “ambivalentes”? Se não está nas gramáticas, não existe.

A resposta é que, rigorosamente falando, nem os ambivalentes, nem os substantivos, nem os adjetivos realmente “existem” na língua: o que existe são os sons da fala (ou as letras no papel) e os significados que a eles se associam. Entidades puramente gramaticais como “substantivos” etc., são hipotéticas². E uma hipótese pode ser mudada sempre que for necessário. Se concordarmos que *velho* e *livro* não têm comportamento idêntico (porque *velho* pode exprimir uma qualidade, e *livro* não), temos que admitir que são gramaticalmente diferentes; e se concordarmos que *velho* e *interessante* não se comportam da mesma maneira (porque *velho* pode designar uma coisa, e *interessante* não), aqui também temos uma diferença gramatical. Ou seja, *velho* é diferente de *livro* e também de *interessante* — logo, temos três tipos de palavras, e não dois. Uma gramática que só distingue duas classes é deficiente, porque deixa de exprimir certas diferenças que são importantes para o conhecimento da língua. Imagine uma pessoa que não sabe que *velho* e *livro* são diferentes; ela pode tentar usar *livro* para atribuir uma qualidade, e o resultado vai ser uma construção errada em português: seguindo o modelo de *um estudo interessante*, ela vai dizer **um estudo livro* (em vez de, digamos, *um estudo livresco*).

A saída que consiste em dizer que *velho* é adjetivo, mas pode “ser usado” como substantivo em certos casos falha por duas razões: uma delas é que confirma que *velho* é diferente de *interessante*, que não pode “ser usado como substantivo” (e portanto não pode pertencer à mesma classe). A segunda razão vamos deixar de lado por ora; ela é devidamente estudada no capítulo 16. Mas a primeira já basta: só faz sentido estabelecer classes se for para colocar em cada uma as palavras que funcionam de maneira semelhante. É por isso que cada classe tem uma definição, como “as palavras que designam coisas” etc. Se *velho* funciona de maneira diferente de *interessante* e também de *livro*, precisamos de três classes onde colocar essas três palavras.

Isso, não custa repetir, é uma hipótese. Os conceitos gramaticais com os quais lidamos — “sujeito”, “verbo”, “substantivo”, “oração” etc. — são

² A maioria dos lingüistas acredita que categorias tais como “verbos”, “substantivos” etc. existem, de alguma maneira, como aspectos da nossa programação cerebral, e há muita evidência em favor dessa posição. Mas aqui não precisamos nos ocupar dessa questão, embora ela seja também fundamental para a compreensão do funcionamento da linguagem. Nosso ponto de vista neste livro é basicamente metodológico.

entidades hipotéticas, criadas pelos gramáticos em sua tentativa de descrever a língua. Os fatos da língua são coisas que se pode observar, e que podem ser reconhecidas sem se saber lingüística: por exemplo, podemos dizer *o livro* mas não **a livro*; o passado de *corro* é *corri*, e o de *faço* é *fiz*; *velho* pode ser tanto a designação de um ser quanto uma qualidade; e assim por diante. Os fatos são diretamente observáveis através do uso que os falantes fazem da língua; as hipóteses são explicitadas pelos lingüistas, e pretendem representar o conhecimento que os falantes têm (sem saber), e que controla o seu uso da língua³.

Mas as hipóteses precisam estar firmemente baseadas em fatos, e é isso que lhes dá importância. Uma hipótese sem base nos fatos não tem importância nenhuma. O trabalho do lingüista é, no fundo, o de formular hipóteses e fundamentá-las o melhor possível nos fatos da língua. Uma hipótese é a de que certas palavras do português são masculinas, outras femininas. Essa hipótese foi criada para explicar por que podemos dizer *o livro*, mas não **a livro*, assim como *a carta* e não **o carta*. Mas é necessário distinguir sempre o fato da hipótese: é fato que não se pode dizer **o carta*. Já a explicação que tradicionalmente se dá a esse fato é baseada numa hipótese: “**o carta* é inaceitável porque *carta* é feminino, e *o* é masculino; e em seqüências como essa temos que ter os dois elementos no mesmo gênero”.

Para resumir: a gramática (entendida como a descrição da estrutura de uma língua) é um conjunto de hipóteses. A função dessas hipóteses é fornecer uma imagem compacta da língua, de maneira que se possa, até certo ponto, prever o que os falantes aceitam e o que eles não aceitam. A partir da hipótese de que *carta* é feminino, prevemos que os falantes rejeitam **o carta certo*, mas aceitam *a carta certa*. A partir da hipótese de que “o verbo concorda com o sujeito”, prevemos que os falantes vão dizer *eu cheguei ontem*, e não **eu chegou ontem*. Todas essas hipóteses podem ser questionadas; os fatos não podem. Assim, alguém pode dizer que o que está errado em **eu chegou ontem* não é a concordância do verbo com o sujeito, mas outra coisa qualquer (aqui entra a imaginação, para elaborar outra hipótese). Mas não se pode dizer que **eu chegou ontem* é bom português, porque é um fato que os falantes rejeitam essa frase como mal formada.

³ Pode parecer contraditório dizer que alguém conhece alguma coisa sem saber, mas isso é comum em muitas áreas do conhecimento. Por exemplo, uma pessoa pode saber nadar, andar (ou dar o nó na gravata), sem conseguir explicar como é que o faz.

1.4. O QUE É QUE É TÃO DIFÍCIL?

Se você pretende ser um linguísta (ou apenas um cidadão devidamente informado sobre a sua língua), vai ter que se libertar de alguns preconceitos muito espalhados e fortemente enraizados. O principal desses preconceitos é o de que as afirmações encontradas em uma gramática são todas factuais, e portanto incontestáveis. Algumas afirmações realmente expressam fatos, como por exemplo a de que o passado de *corro* é *corri*. Mas outras são hipóteses, que podem ou não ser adequadas. Assim, se em uma gramática encontramos a afirmação de que “os pronomes possessivos *meu, seu, nosso* podem vir ou não precedidos de artigo definido”, isso pode ser testado. Se verificarmos que os falantes realmente aceitam seqüências como *o meu pai, a nossa casa*, e também *meu pai, nossa casa*, então a hipótese está confirmada; se os falantes rejeitarem algumas dessas seqüências, então a hipótese precisa ser rejeitada ou pelo menos modificada para dar conta dos dados.

Essa é uma das funções das hipóteses: dar conta dos dados. Isso quer dizer que elas substituem uma lista completa dos dados. Essa lista seria impossível de elaborar, claro, porque os dados são em número indeterminado — não é possível estabelecer quantas frases existem em português, muito menos fazer a lista delas. Por isso, construímos hipóteses, utilizando coisas como classes, funções e regras gramaticais, e essas hipóteses nos dizem antecipadamente o que vamos encontrar na língua. Em vez de dar uma lista de seqüências possíveis como *a aba, o abacate, ... a zebra, o zorro*, e seqüências impossíveis como **o aba, *a abacate* etc., construímos uma hipótese, dizendo que

os nominais se dividem em femininos e masculinos; os femininos aceitam o artigo *a*, e os masculinos o artigo *o*.

Essa hipótese tem a forma de uma regra. Ela se baseia em pelo menos uma outra hipótese, que é a da existência de uma classe dos nominais, que deve estar devidamente definida em alguma parte da gramática; para nós, basta saber que os nominais incluem os substantivos, adjetivos e “ambivalentes”. Mas, em algum momento, todas essas hipóteses precisam estar justificadas (em termos de sua utilidade da descrição) e cuidadosamente fundamentadas nos fatos, isto é, nas ocorrências de formas da língua tais como se observam na realidade.

1.5. AS HIPÓTESES E A ESTRUTURA DA LÍNGUA

Eu disse acima que uma das funções das hipóteses é substituir a listagem completa dos dados. Não se pense, contudo, que seja essa a única função

das hipóteses. Elas têm uma outra função, extremamente importante, que é a de nos revelar as grandes linhas da estrutura da língua. A idéia é que a língua está programada de alguma forma em nossa mente, como uma parte do nosso conhecimento do mundo. E a língua se organiza na mente através de classes, funções, regras etc. — conseqüentemente, quando formulamos a hipótese de que as palavras do português se colocam em certo número de classes, e definimos cada uma dessas classes, estamos ao mesmo tempo fazendo hipóteses sobre aspectos da estrutura do conhecimento em nossa memória.

Existe muita evidência em favor dessa presença das unidades gramaticais em nossa mente, o que faz com que a análise gramatical tenha importância não apenas na área da gramática, mas também na área dos estudos da estrutura e funcionamento da memória. Neste livro, entretanto, vamos estudar as unidades, regras, funções e princípios gramaticais sem relacioná-los sistematicamente com sua representação mental. Aqui nos limitamos aos aspectos descritivos, isto é, procuramos explicitar como a língua se manifesta exteriormente; o estudo, ainda que preliminar, dos aspectos mentais exigiria (no mínimo) outro livro.

DE ONDE VÊM OS DADOS?

2.1. DADOS E HIPÓTESES

Um lingüista, para seu estudo da língua, parte de dados. Assim como o zoólogo do capítulo 1 observa aves e seu comportamento, o lingüista observa palavras¹ e suas associações em frases e enunciados maiores, utilizando essas observações para construir sua análise da língua². Mas de onde vêm esses dados?

É bom observar desde logo que o objetivo do lingüista não se reduz à apresentação de dados. O que ele procura é uma descrição da estrutura da língua: o conjunto de regras, elementos, classes e princípios que governam as associações dos diversos elementos da língua e seu significado. Por exemplo, ao considerar a frase

[1] O gato sujou a poltrona.

ele precisa levar em conta que muitos aspectos da organização dessa frase decorrem de conhecimentos dos falantes — conhecimentos que só existem na mente desses falantes e que, portanto, não são diretamente observáveis. Ele vai dizer coisas como:

- (a) em *o gato*, o artigo *o* vem antes de *gato* em virtude de uma regra que estabelece que o artigo vem sempre antes do nominal que determina: *gato o é inaceitável;
- (b) nessa mesma seqüência, a forma do artigo (*o*, e não *a*, *as*, *os*) depende de traços morfológicos de *gato*; se tivéssemos *gatas*, o artigo seria *as*;

¹ Assim como fonemas, morfemas etc., que não precisamos considerar aqui.

² Não quero dizer que o lingüista “tire” a análise dos dados. Ele precisa igualmente utilizar uma teoria que justifique essa análise — o trabalho de pesquisa envolve uma interação constante de fatos e teorias, processamento *top-down* e *bottom-up*.

- (c) entende-se que se fala de um único gato porque a palavra gato não apresenta um -s final; se fosse gatos, estaríamos falando de mais de um gato;
- (d) entende-se que o evento relatado ocorreu no passado por causa da terminação da palavra sujou; se fosse suja, entenderíamos que se trata de um evento habitual.

e assim por diante.

As partes sublinhadas de (a) – (d) acima são hipóteses, tanto é assim que fazem previsões. As hipóteses, como já sabemos, são um objetivo essencial do trabalho de análise lingüística.

2.2. COLETANDO DADOS

Mas as hipóteses não são gratuitas, e dependem da observação de dados para sua confirmação. Como exemplo, vamos tomar a hipótese de (a), ou seja;

o artigo vem sempre antes do nominal que determina;

admitindo-se que já sabemos o que é um “artigo” e um “nominal”, além da relação de “determinar”.

Para verificar essa hipótese, temos que recorrer aos dados. Isso, no caso, pode ser feito de três maneiras:

(a) Exame de *corpus*

Podemos tomar um *corpus* de enunciados falados ou escritos, ou seja, textos tirados da fala (em geral transcritos) ou da escrita. Procuramos aí casos em que o artigo venha depois do nominal que determina: por exemplo, *gatos os* etc. Se não encontrarmos nenhum, podemos dizer que a hipótese foi confirmada.

Note-se que o exame de *corpus* não pode ser feito mecanicamente, porque pode ocorrer que um artigo venha depois de um nominal, como na frase

[2] Esquipéria deixou por um momento as tarefas domésticas.

No entanto, esse caso não conta, porque *as* não determina *momento*, mas *tarefas*³. Para saber isso, temos que utilizar nosso conhecimento de falan-

³ Ou, como se verá no capítulo 3, *momento as* não forma um constituinte.

tes da língua (nossa competência). Logo adiante, vamos ver que nunca é possível escapar totalmente do uso da competência que temos da língua ao coletar dados. Isso dito, entretanto, a ocorrência (ou não ocorrência) em *corpus* é um argumento muito forte para confirmar ou refutar uma hipótese.

(b) Testagem

Podemos também recorrer à competência de outros falantes, perguntando a eles o que acham. Por exemplo, podemos dar a alguém uma lista de construções que incluem o artigo depois do nominal e perguntar se são aceitáveis. Se todos, ou quase todos, as rejeitarem, teremos argumentos em favor da hipótese. Se muitos as aceitarem, teremos argumentos contra.

(c) Introspecção

Podemos também fazer a mesma coisa, mas com nós mesmos. Isso é o que se faz com maior frequência, e neste livro apelo constantemente para a introspecção dos leitores. Por exemplo, quando eu disse em um capítulo anterior que a frase

[3] *Incomodou Serginho Helena à tarde

é inaceitável, estava de certo modo convidando o leitor a verificar isso por si mesmo. Em casos como [3], em que a inaceitabilidade é muito clara, esse procedimento não apresenta problemas.

2.3. VANTAGENS E PERIGOS

Essas são as três principais fontes de dados para o investigador gramatical. Todas elas apresentam vantagens e também perigos, e é bom esclarecer isso, já que ninguém escapa de utilizar um ou outro desses recursos com frequência.

2.3.1. O *corpus*

O uso do *corpus* tem a vantagem de neutralizar os desejos do pesquisador. Quando este se vale da introspecção, ele pode (mesmo sem perceber) aceitar frases que cabem dentro da sua análise, ou rejeitar frases que a contrariam. Em certos casos, isso dificilmente acontece, porque os julgamentos são bem claros; assim, não há falante que rejeite [1], ou que aceite [4]:

[1] O gato sujou a poltrona.

[4] *Gato o sujou poltrona a.

Mas há casos muito mais difíceis de julgar. A aceitação varia de falante para falante; alguns ficam simplesmente na incerteza, e outros aceitam hoje a frase que rejeitaram a semana passada. Todo lingüista conhece essa situação, e nesses casos é um conforto encontrar a frase em questão ocorrendo com frequência em um *corpus*, porque é uma indicação de que ela faz parte da língua. Um exemplo típico de dificuldade de julgamento se observa em frases com infinitivo flexionado; o que é aceitável, [5] ou [6]?

[5] A professora viu os gatos comer.

[6] A professora viu os gatos comerem.

O exame de *corpus* vai nos dar uma resposta objetiva. No português escrito antigo, por exemplo, ocorrem ambas as estruturas, e a de [5] é a mais comum (isso provavelmente vai contra o palpite de muita gente)⁴.

Outra coisa que o *corpus* nos fornece é uma indicação da frequência das diversas formas; é a partir dele que se pode dizer, por exemplo, que a forma *cheguei* é muito mais frequente do que *chegaria*, e que *chegará* praticamente não aparece na língua falada.

A grande vantagem do *corpus* é, portanto, sua imparcialidade — ele reflete a realidade sem preconceitos teóricos. Mas, por outro lado, aparecem no *corpus* muitas coisas que não fazem, nem podem fazer, parte da gramática da língua. Por exemplo, encontramos em um *corpus* de língua falada⁵ a seguinte passagem:

[7] Eu já [...] um tempo desse mais — de — de tentar, sabe — de tentar, porque meu patrão tinha — ia viajar, ele ia de caminhão — ele é fazendeiro, ele ia levar uns gado lá pro exterior [...]

É evidente que nem tudo o que encontramos aqui é resultado da aplicação das regras da gramática portuguesa. Em uma transcrição desse tipo,

⁴ O levantamento da ocorrência dessa construção no português dos séculos XV-XVII está em Said Ali (1908). Para o português falado atual não há, que eu saiba, nenhum levantamento.

⁵ O de Hora e Pedrosa (2001). Regularizei um pouco a grafia, para facilitar a leitura.

encontramos traços de hesitação, de mudança de plano no meio da frase, gaguejamentos etc., que certamente não devem ser incluídos em uma descrição da língua. Quando o lingüista estuda esses textos, ele recorre, inevitavelmente, a seu conhecimento da língua (sua competência) para “filtrar” os traços irrelevantes. Assim, como se vê, na pesquisa gramatical não há maneira de escapar da introspecção, utilizada de uma maneira ou de outra.

Outro problema sério na pesquisa de *corpus* é que as pessoas não costumam proferir espontaneamente as frases de que o lingüista precisa em determinado momento. Principalmente quando se trabalha com sintaxe ou semântica, é impraticável examinar milhares de páginas de transcrições (para não dizer colhê-las e transcrevê-las) para cada estrutura que pesquisamos. Uma estrutura pode ser totalmente aceitável e assim reconhecida por todos os falantes, e ainda assim não ocorrer nem uma vez em um *corpus* de grande extensão.

2.3.2. Introspecção

A introspecção, como já apontei, pode nos levar a conclusões bastante fora da realidade, porque não se baseia em observações objetivas. Por exemplo, em trabalhos recentes, de lingüistas competentes, encontrei julgamentos com os quais não concordo em absoluto, como

[8] *O João gostou da comida que comeu.

[9] Esses professores parecem que eles gostam da Maria.

Para mim, julgando-as também introspectivamente, [8] é aceitável e [9] é pelo menos estranha, mas pelo menos algumas pessoas acham o oposto. Aqui seria necessário proceder a testes ou, se possível, recorrer a ocorrências em um *corpus*.

No entanto, a introspecção (seja direta, seja através de testagens) é inevitável, porque ao utilizá-la podemos selecionar as estruturas que estamos estudando, ao contrário do *corpus*, como apontei acima. Como sair desse dilema?

É preciso encarar uma realidade: estudar uma língua é, em última análise, estudar um fenômeno psicológico, portanto não observável diretamente. Só temos acesso a suas manifestações exteriores e aos julgamentos sobre elas, e daí temos que deduzir as características do fenômeno estudado: a língua, considerada como sistema de conhecimento armazenado na memória. Para isso, é preciso lançar mão de todos os recursos disponíveis, sem esquecer a devida cautela, sabendo que lidamos com dados bastante

inseguros. O *corpus*, as testagens e a introspecção são recursos a ser utilizados com consciência de suas vantagens e de suas fraquezas.

Neste livro, lanço mão sistematicamente da introspecção do leitor, evitando o quanto possível casos duvidosos. Essa é uma escolha inevitável, dado o caráter de um livro de introdução; mas fique aqui a advertência de que colher dados em lingüística não é um problema trivial e precisa ser abordado com toda a seriedade que merece.

2.4. AGARRE-SE AOS FATOS!

Uma palavra final: as recomendações prescritivas que compõem grande parte do ensino tradicional de português não servem como fonte de dados. O que nos interessa são os fatos da língua, não os ideais lingüísticos que algumas pessoas abrigam.

Se o que estamos estudando é o português brasileiro falado, uma frase como

[10] Os menino tão lá embaixo.

é aceitável, simplesmente porque ocorre — é normalmente usada pelos falantes (e não apenas pelos falantes “incultos”). O fato de que as gramáticas tradicionais condenam formas como *os menino* ou *tão* (no sentido de *estão*) não é relevante. O lingüista estuda a língua como ela é, e não idéias a respeito de como a língua deveria ser.

Isso não quer dizer que um lingüista seja contrário ao ensino do português padrão. Significa apenas que ele estuda realidades; e uma realidade é que a frase [10] é aceitável em determinadas situações de uso da língua (conversando com os amigos, por exemplo). É também verdade, e portanto deve ser reconhecido, que frases como [10] não se usam em textos formais, principalmente escritos (onde, por convenção, se usa outra variedade da língua, chamada “padrão”). Igualmente, não se usaria em uma conversa informal a frase

[11] A professora pede que discutamos a proposta.

Nenhum brasileiro usaria [11] em uma conversa, e isso é um fato.

Conseqüentemente, um estudioso da gramática portuguesa precisa definir, previamente, que variedade da língua vai estudar. Isso feito, ele deve trabalhar com os fatos. Se estuda o português padrão escrito, [11] faz parte de seus dados, mas não [10]; se estuda o português brasileiro falado, precisa incluir [10], mas não [11].

FORMA E SIGNIFICADO

3.1. DE ONDE TIRAMOS O SIGNIFICADO?

Quando examinamos uma frase, a primeira coisa que vemos é uma série de palavras. Mais tarde veremos que há muito mais em uma frase do que isso; mas por ora vamos nos deter no exame das palavras.

À primeira vista, podemos pensar que a relação entre as palavras e seus significados é simples: cada palavra serve para comunicar um significado. Assim, a palavra *mão* serve para evocar, na mente do ouvinte ou do leitor, a idéia de uma parte do corpo, *cachorro* designa um tipo de animal, e *ontem* se refere ao dia que passou por último.

Mas mesmo se aceitarmos essa relação simples entre palavra e significado, logo vamos encontrar problemas. Assim, qual é o significado da palavra *de*? Por exemplo, essa palavra aparece em

-
- [1] A mão de Vera
 - [2] O copo de plástico
 - [3] Eu gosto de Vera
 - [4] Vera veio de avião.
 - [5] Vera veio de São Paulo.
 - [6] Vera chegou ontem de noite.
-

Parece que em cada caso a palavra *de* significa uma coisa diferente. Mas como é que podemos identificar o significado de *de* em cada um desses exemplos?

O processo é mais ou menos o seguinte: em primeiro lugar, a gente parte da própria palavra — cada palavra tem um significado, ou, mais exatamente, um conjunto delimitado de significados (a que podemos chamar área semântica). A área semântica de algumas palavras é bastante

restrita; assim, *fotômetro* significa um tipo de aparelho, e muito pouco mais. Isso ocorre com muitas palavras de significado concreto, ou então com termos técnicos, como *carburador* ou *caule*.

Mas uma palavra como *de* tem uma área semântica tão ampla e difusa que à primeira vista somos tentados a afirmar que *de* pode significar qualquer coisa. Pensando melhor, vemos que não é isso: *de* não pode significar 'fotômetro', por exemplo. Acontece que o significado de *de* é de tipo muito abstrato, e pode ser expresso, inicialmente, como 'relação entre dois outros elementos'. Assim, nos exemplos acima, *de* expressa uma relação entre *mão* e *Vera* ("posse"), entre *chegou* e *avião* ("meio"), entre *chegou* e *São Paulo* ("proveniência") e assim por diante — em [3], inclusive, não parece significar nada. Mesmo entre as relações, nem todas podem ser expressas por *de*: nunca encontramos essa palavra com o significado de 'sem', ou 'contra', noções normalmente veiculadas por outras preposições.

Podemos então admitir que cada palavra tem uma área semântica mais ou menos definida, e que o receptor (ouvinte ou leitor) parte daí para realizar o processo que chamamos de interpretação das estruturas que ouve ou lê.

O fato de a área semântica ser mais ou menos restrita faz com que a imensa maioria das palavras tenha a capacidade de exprimir diversos significados, às vezes muito diferentes. Por exemplo, veja-se o significado de *mão* em

[7] A mão de Vera ficou machucada.

[8] Ele nunca dirige na mão.

[9] Me dá uma mão aí!

[10] Essa parede vai receber uma mão de tinta.

[11] Olha o problema que foi cair na minha mão!

Uma olhada no dicionário vai mostrar que esse é o caso com a maioria das palavras, principalmente com as mais comuns. Para chegar à interpretação correta de cada palavra, o receptor vai ter que lançar mão¹ de uma série de recursos, entre os quais:

(a) A estrutura formal (sintática) da frase

Assim, em

[3] Eu gosto de Vera.

1. Mais uma!

de vem logo depois de um verbo, o que significa que está introduzindo um complemento desse verbo, e não, por exemplo, exprimindo o fato de que alguma coisa pertence a Vera (como em *a casa de Vera*).

(b) As outras palavras que ocorrem na frase

No caso de [3], o verbo é *gosto*, e o receptor sabe (através de seu conhecimento da língua) que o complemento (*de Vera*) denota "a coisa de que se gosta"; e que *de*, nesse caso, não significa nada independentemente. E em

[10] Essa parede vai receber uma mão de tinta.

ele sabe que *mão de tinta* é uma expressão que significa "uma camada de tinta". Ele chega a essa conclusão por causa da presença de *tinta* em determinada posição estrutural em relação a *mão*. Igualmente ele pode distinguir *de São Paulo* como "proveniência", e *de avião* como "maneira de viajar", por saber que *São Paulo* é um lugar fixo, que não se move, e *avião* é um meio de transporte.

(c) Outros aspectos do conhecimento do mundo

Finalmente, o receptor de certo modo exclui interpretações que não fazem sentido. Por isso é que ao ouvirmos

[1] A mão de Vera

entendemos que Vera é a "possuidora" da mão, mas ao ouvirmos

[2] O copo de plástico

entendemos *de plástico* como o "material" de que o copo é feito, já que não faz sentido entender que o copo "pertence ao plástico".

Como se vê, existe um mundo de fatores envolvidos na interpretação do significado de cada palavra. Boa parte da lingüística se dedica justamente a explicitar esses fatores.

3.2. UMA TEORIA ERRADA SOBRE AS PALAVRAS E O CONTEXTO

Voltemos por um momento às palavras. Eu disse acima que "cada palavra tem uma área semântica mais ou menos definida" — ou seja, não

existe uma relação direta entre cada palavra e um significado. Essa observação é mais importante do que parece.

Alguns linguístas se impressionam muito com casos como os das frases [7] a [11], que vimos acima e repito aqui:

[7] A mão de Vera ficou machucada.

[8] Ele nunca dirige na mão.

[9] Me dá uma mão aí!

[10] Essa parede vai receber uma mão de tinta.

[11] Olha o problema que foi cair na minha mão!

a ponto de dizerem que uma palavra não tem significado nenhum fora de contexto. Essa afirmação não é correta, mas isso não se vê à primeira vista; por isso, vou discuti-la brevemente, nesta seção.

Primeiro, é preciso notar que as possibilidades de significado de uma palavra nunca são ilimitadas; *mão*, como vimos, significa muitas coisas, mas nunca é sinônimo de *pé*, nem de *chatear*, nem de *Brasil*. Isso quer dizer que quando um ouvinte vai interpretar alguma das frases acima, ele já parte de um conhecimento prévio do que é que *mão* pode e não pode significar.

Segundo, se todas as palavras só tirassem seu significado do contexto, as frases seriam incompreensíveis. Digamos que a frase seja:

[7] A mão de Vera ficou machucada.

De acordo com essa teoria², a primeira palavra, *a*, poderia significar qualquer coisa; o significado dessa palavra vai depender das outras palavras da frase. Mas a segunda palavra, *mão*, também pode significar qualquer coisa. Logo, nem *a* ajuda a compreensão de *mão*, nem vice-versa. E isso vai se reproduzindo para todas as outras palavras da frase — o ouvinte simplesmente não teria nenhum ponto inicial de referência para estabelecer o significado de nenhuma das palavras da frase, e por isso não conseguiria entendê-la. Mas isso não é verdade: é possível entender [7] muito bem, e mesmo a palavra *mão* tem um significado bem claro, e o mesmo para todos os falantes.

Finalmente, sabemos que muitas palavras, mesmo isoladas de contexto, nos dão uma idéia razoavelmente precisa do seu significado; assim,

Nunca vi escrito que as palavras só tiram seu significado do contexto, mas já ouvi isso dito por profissionais de certo gabarito; portanto, acho bom refutar a idéia aqui.

quando digo *machucada*, sabemos mais ou menos do que se trata. Pode ser um estado físico (*mão machucada*), e pode ser um estado psicológico (*ela ainda está machucada com o divórcio*), e muito pouco mais.

Todas essas indicações sugerem que cada palavra delimita um conjunto de significados possíveis³, sua área semântica, que a separa de outras palavras parecidas. Assim, *mão* pode significar uma parte do corpo, a direção do tráfego em uma rua, uma camada de pintura etc., mas nunca a parte do corpo que usamos para ver (*olho*), nem o lugar por onde passam os carros (*rua*), nem o objeto usado para pintar (*pincel*). O significado de cada palavra, portanto, está longe de ser indefinido. O papel do contexto é selecionar, das várias possibilidades semânticas que muitas palavras têm, a que é a mais plausível naquele contexto.

Em resumo, o significado de uma palavra nem é produto inteiramente do contexto em que aparece, nem é (geralmente) totalmente determinado de antemão.

3.3. O SIGNIFICADO DA ESTRUTURA

3.3.1. Ordem dos termos

Mas o significado das palavras não esgota o que se entende quando se interpreta um enunciado. A própria maneira como se organizam as palavras em frases e outras unidades tem um significado próprio. Por exemplo, as frases seguintes têm as mesmas palavras, mas não significam a mesma coisa:

[12] O periquito comeu meu gato.

[13] Meu gato comeu o periquito.

Neste caso, a ordem dos elementos em relação ao verbo tem seu significado próprio. Quando o verbo é *comer* (na voz ativa), o elemento que vem logo antes do verbo é o que pratica a ação (é o agente), e o que vem logo depois sofre a ação (é o paciente). Trocando a ordem desses termos, troca-se também o papel de cada um na interpretação da sentença.

Às vezes a mudança de significado é mais sutil, como nos exemplos:

[14] Ele é um grande jogador.

³ Fugindo do problema do que vem a ser exatamente um significado!

[15] Ele é um jogador grande.

que têm significados diferentes, como é fácil ver. Novamente, é a ordem dos termos que determina a diferença.

O que vimos acima são alguns exemplos de como a estrutura de uma expressão lingüística (como por exemplo uma frase) tem também um significado, que se soma ao das palavras individuais para compor o significado da expressão. Ou seja, o usuário da língua aprende não apenas o significado de cada palavra, mas também o das estruturas da língua (as construções, estudadas no capítulo 17).

3.3.2. Constituintes

Outro aspecto do significado da estrutura é sua organização em “partes naturais”, ou constituintes. Por exemplo, ao processarmos a frase

[16] A casa de Sandrinha é verde.

atribuímos significados parciais a certas seqüências de palavras. Assim, *a casa de Sandrinha* recebe um significado (uma interpretação), e o mesmo ocorre com *é verde*. Já as seqüências *casa de* ou *Sandrinha é* não recebem interpretação como unidades, mas dão antes a impressão de serem compostas de pedaços desconexos da frase. E em [16] a seqüência *Sandrinha é verde* não recebe interpretação nenhuma, muito embora esteja materialmente presente em [16]. Ou seja, ao ouvirmos [16] entendemos que a casa é de Sandrinha; que essa casa é verde; mas não entendemos que Sandrinha é verde.

Isso se deve ao fato de que o receptor, ao receber uma seqüência de palavras, procura organizá-la de acordo com seus conhecimentos da língua (e do mundo). Isso se faz em várias frentes: o receptor procura na seqüência estruturas sintáticas corretas da língua, e também procura evitar seqüências cuja interpretação seja absurda ou implausível.

Se o receptor considerar *Sandrinha é verde* uma unidade, ela será uma oração. Isto é, terá que ser analisada como sujeito + predicado, o que segundo as regras da língua é uma oração — as regras da língua estão presentes na mente do usuário, e incluem uma espécie de receita de como construir orações. Mas então o que fazer com *a casa de*? Esse pedaço fica sobrando, o que normalmente não é aceitável.

Poderíamos achar que é a estranheza semântica de *Sandrinha é verde* que está por trás do problema. Mas não é isso: aqui a motivação é puramente estrutural, porque em

[17] A casa de Sandrinha é alta.

não entendemos que *Sandrinha é alta*, embora essa seqüência tenha significado perfeitamente normal. O que impede o receptor de entender *Sandrinha é verde* em [16] como um constituinte é o fato de que se ele fizer isso não será possível “fechar” a frase inteira dentro de uma estrutura sintática existente na sua gramática internalizada (vão sobrar partes). O receptor, então, precisa procurar alguma outra solução, que é a de juntar *Sandrinha* ao que precede, produzindo *a casa de Sandrinha*, que agora pode ser juntada a *é verde*, formando assim uma estrutura completa.

Dizemos então que *a casa de Sandrinha* é um dos constituintes da frase [16], porque tem uma personalidade sintática (é o que chamamos um sintagma nominal, e é também o sujeito da oração), e porque recebe uma interpretação semântica coerente (refere-se a um objeto específico, a saber, uma casa). *Sandrinha é verde*, na frase [16], não é um constituinte, e por isso não tem análise sintática nem interpretação semântica — ao ouvir [16], ninguém fica pensando que Sandrinha é verde. Outras seqüências de palavras em [16] que não formam constituintes são *casa de* e também *Sandrinha é*. As seqüências que formam constituintes são chamadas sintagmas; assim, em [16] *a casa de Sandrinha* é um sintagma, assim como *é verde*, mas *Sandrinha é verde* ou *casa de* não são sintagmas.

Teremos que voltar aos constituintes mais tarde (capítulo 11), porque eles são básicos para a análise gramatical. Por ora, fique a observação de que segmentar as seqüências recebidas em constituintes é uma das operações necessárias para a compreensão dessas seqüências.

3.4. AMBIGÜIDADE

Um fenômeno importante ligado à relação forma/significado é o da ambigüidade — ou seja, o fato de que uma forma pode ter mais de um significado. Consideremos o seguinte exemplo:

[18] Rafael pediu bolo e sorvete de chocolate.

Esta é uma frase ambígua: pode-se entender que Rafael pediu bolo de chocolate, ou então que pediu qualquer tipo de bolo (mais sorvete de chocolate). Outro exemplo de ambigüidade é

[19] Tem um banco ali na esquina.

A ambigüidade pode decorrer de duas causas principais: ou a frase tem uma palavra que significa mais de uma coisa (ambigüidade léxica) ou a própria estrutura permite mais de uma interpretação (ambigüidade estrutural). No caso de [19], a ambigüidade é léxica, porque a palavra *banco* pode significar “objeto que serve para a gente sentar em cima” ou então “estabelecimento de crédito”. Já no caso de [18] a ambigüidade é estrutural: quando consideramos os dois sentidos, nenhuma palavra muda de significado, mas a própria construção muda: em um dos significados, a coisa pedida é “bolo de chocolate”, e no outro é apenas “bolo”.

O OBJETIVO DA DESCRIÇÃO

4.1. SELECIONANDO O QUE É RELEVANTE

O princípio básico da descrição gramatical é o de descrever as formas, os significados e as relações entre forma e significado (relações simbólicas). Por exemplo, nossa descrição explicita que o verbo *abrir* aparece em diversos ambientes gramaticais, ilustrados pelas frases abaixo:

-
- [1] Rodolfo abriu a janela.
 - [2] A janela, Rodolfo abriu.
 - [3] A janela foi aberta por Rodolfo.
 - [4] A janela abriu.
-

Essas frases diferem quanto às funções dos diversos sintagmas nominais; *a janela*, por exemplo, é objeto direto em [1], objeto direto topicalizado (isto é, colocado no início da sentença) em [2], e é sujeito em [3] e [4].

Por outro lado, cada uma dessas frases apresenta certos traços semânticos característicos, como: em [2] *a janela* se entende como o tópico da sentença (ou seja, entende-se que se está falando, principalmente, da janela); em [1], [2] e [3] *Rodolfo* denota o agente (isto é, o elemento que pratica a ação). *A janela* é o paciente (o elemento que sofre a ação) em todas as frases.

Esses são fatos da língua, e como fatos são inegáveis. Estão apresentados acima da maneira mais simples e despojada possível. É provável que essa apresentação neutra deixe de capturar muitas generalizações importantes, mas as generalizações, como afirmações gramaticais que são, são hipóteses, e portanto dependem de dados para sua validação. A descrição que usamos neste livro pretende apresentar os dados da maneira mais isenta possível; eles poderão ser a base para validar as hipóteses que forem elaboradas para explicá-los.

Nosso sistema de apresentação dos dados se baseia na descrição das formas da língua tais como são diretamente observáveis. Por exemplo, em

[3] A janela foi aberta por Rodolfo.

todos concordam que *a janela, por Rodolfo, foi aberta* são constituintes, e que *aberta por* não é um constituinte; todos definem *a janela* como "sujeito" por causa de sua posição e/ou sua relação de concordância com o verbo — e assim por diante. Igualmente, é de aceitação geral que o fato de que *a janela é paciente e Rodolfo é agente* é resultante da função sintática de cada sintagma. Note-se que algumas dessas afirmações são hipóteses (não é um fato que *a janela seja o sujeito de [3]*), mas são hipóteses amplamente aceitas. Não é possível trabalhar sem hipóteses — o pesquisador simplesmente não saberia o que procurar na massa dos dados —, mas, por sorte, o que já existe, e que é considerado como estabelecido pelos lingüistas em geral, não é pouco. A metodologia exposta neste livro se baseia em grande parte nessas hipóteses de aceitação geral.

Para ilustrar a importância de partir de hipóteses de trabalho, vamos ver um exemplo. Digamos que vamos começar com uma frase, pronunciada em determinado momento por determinada pessoa; ela diz:

[5] O leite ferveu.

Queremos fazer uma análise lingüística dessa frase, partindo de fatos bem estabelecidos, captados auditivamente. Mas a quantidade de traços observáveis é muito grande, e grande parte terá que ser desprezada. Por exemplo, são traços da frase [5], pronunciada por aquela pessoa naquele momento:

- (a) a qualidade da voz (baixa, rouca, um pouco anasalada...);
 - (b) o fato de que a frase durou exatamente 1 segundo;
 - (c) o fato de que a frase foi dita às 4 da tarde;
 - (d) a presença de *o* (pronunciado [u]) logo antes de *leite*;
 - (e) o fato de que *o leite* vem logo antes de *ferveu*;
 - (f) o fato de que o ouvinte entende que [5] é uma afirmação sobre o leite;
 - (g) o fato de que o ouvinte entende que se trata de um evento passado;
- etc. etc. etc.

Tudo isso está lá, ou no sinal acústico ou na representação mental que ele evoca na mente do ouvinte. Mas será que tudo isso tem que ser levado

em conta em uma análise gramatical de [5]? É claro que não. Já sabemos (ou seja, partimos dessa hipótese) que a qualidade da voz ou o horário em que uma frase é dita não têm a menor possibilidade de ser gramaticalmente relevantes; portanto, excluimos os traços (a) e (c). O mesmo se diga da duração do enunciado — o que nos libera de (b). Os traços restantes podem ser relevantes: (d) é importante para se descrever a concordância, por exemplo (*o leite* e não **a leite*); (e) nos dá uma indicação de que *o leite* deve ser o sujeito da frase; (f) caracteriza *o leite* como o tópico do enunciado; (g) situa o evento narrado no eixo temporal, o que sabemos que é codificado pela morfossintaxe do verbo (*ferveu*, e não *está fervendo* ou *vai ferver*).

Ou seja, não precisamos perder nosso tempo com três dos sete traços acima, embora todos eles estejam legitimamente presentes no sinal auditivo que recebemos ou na representação mental que construímos a partir dele. O que nos permite essa seleção de traços são hipóteses; mas hipóteses tão bem estabelecidas que nem pensamos em contestá-las.

Mas há hipóteses bem mais elaboradas, como por exemplo: esperamos que a seqüência ouvida possa ser dividida em palavras; e que cada uma dessas palavras pertença a uma classe, ou seja, tenha propriedades gramaticais semelhantes às de outras palavras (*leite* e *caldo*, por exemplo, ou então *ferveu* e *derramou*). Esperamos também que as diferentes palavras estejam em uma ordem mais ou menos fixa: podemos dizer *o leite ferveu* ou *ferveu o leite*, mas nunca **o ferveu leite* nem **leite ferveu o*. Estamos preparados para encontrar relações formais entre algumas palavras: por exemplo, *o* está relacionado com *leite* porque ambos estão no masculino singular, e *ferveu* tem essa forma por alguma razão ligada à presença de *o leite* logo antes: se fosse *todas as leiteiras*, teríamos que dizer *ferveram*.

Tudo isso são hipóteses de aceitação praticamente universal. Um dos objetivos deste livro é formular essas hipóteses, que servem de orientação inicial ao lingüista empenhado em descrever uma língua. São instrumentos de análise que todo lingüista precisa dominar, e sem os quais ele se arrisca a cometer os erros mais primários ao tentar descrever a gramática de uma língua.

4.2. O CARÁTER EXPLÍCITO DA ANÁLISE

A descrição dos fatos observados precisa ser o mais explícita possível. Uma descrição é explícita quando suas conseqüências são bem claras e podem ser avaliadas em confronto com os dados.

Para exemplificar esse ponto, digamos que estamos classificando as palavras, e queremos definir uma dessas classes. A definição pode ser algo como:

Os vocábulos gramaticais que servem para relacionar duas orações ou dois termos semelhantes da mesma oração chamam-se conjunções [Cunha, 1976: 533]¹.

Essa é, com pequenas variações de detalhe, a definição de “conjunção” encontrada nas gramáticas escolares. Vamos ver que essa definição, embora talvez não seja totalmente incorreta, não é suficientemente explícita.

Para uma definição ser realmente útil, ela tem que permitir a uma pessoa que não conhece a coisa definida (isto é, no caso, não sabe o que é uma conjunção) identificá-la com base na definição. Caso contrário, para que definir? É aqui que o caráter explícito se torna crucial. Uma coisa que precisamos saber para aplicar a definição de “conjunção” é o que vem a ser um “vocábulo gramatical”. A gramática em exame diz que são os vocábulos “constituídos de morfemas gramaticais” [p. 89]. Agora, naturalmente, surge o problema de saber o que é um morfema gramatical. Segundo a mesma gramática, trata-se de morfemas “de significação interna”, e se opõem a morfemas lexicais, definidos como

representantes ou símbolos básicos de tudo o que os falantes distinguem na realidade subjetiva ou objetiva [Cunha, 1976: 88].

Acabamos de bater na parede do caráter pouco explícito das definições. Como não sabemos o que é “significação interna” (nem a gramática a define), ficamos com a definição de morfema lexical; o morfema que não for lexical é gramatical, e pronto.

Mas a definição de morfema lexical é, ela própria, tão pouco explícita que não permite tomar decisões claras. Diante de um morfema, digamos, o morfema *nunca* (isto é, uma palavra formada de um só morfema), temos que saber se ele é “representante ou símbolo básico” de algo que “os falantes distinguem na realidade subjetiva ou objetiva”. Se for, será um morfema lexical, se não for será um morfema gramatical. Mas como decidir? O que significa ser “representante ou símbolo básico” de alguma coisa? E que idéia não faz parte do que distinguimos na realidade subjetiva ou objetiva? Acho que não preciso ir mais longe para deixar claro que alguma coisa não funciona na definição dada. E se a

¹ Vou citar e criticar frequentemente a gramática de Celso Cunha, não porque seja pior ou melhor do que as outras, mas porque, no que nos interessa, as diferenças entre as várias gramáticas escolares são muito pequenas, e a de Cunha é provavelmente a mais conhecida.

definição de “morfema gramatical” não funciona, a de “vocábulo gramatical” também não funciona. Como resultado, a definição de “conjunção” também se torna impossível de aplicar com clareza.

Esse exemplo ilustra uma qualidade que as afirmações gramaticais (uma regra, uma definição ou o que quer que seja) precisam ter para serem explícitas: é essencial que elas lancem mão apenas de termos claramente compreensíveis e bem definidos. Não é o caso das definições de “conjunção”, “vocábulo gramatical” e “morfema gramatical” vistas acima.

A gramática de Cunha dá mais uma pista sobre o que é um vocábulo gramatical, na seguinte passagem:

Na classe dos vocábulos gramaticais se incluirão o artigo, os pronomes, as preposições, as conjunções [...] [Cunha, 1976: 90].

Essa definição é menos confusa do que a anterior. Mas, se a examinarmos com cuidado, vamos encontrar um problema: ela diz que entre os vocábulos gramaticais estão as conjunções. E o que é uma conjunção? Ora, vimos acima que para saber o que é uma conjunção (segundo a mesma gramática) precisamos saber, antes, o que é um vocábulo gramatical. Acabamos, portanto, presos em um beco sem saída: para identificar os vocábulos gramaticais precisamos saber identificar uma conjunção, e para identificar uma conjunção é necessário saber identificar um vocábulo gramatical. Isso se chama “circularidade”, e é outra coisa que torna a definição não explícita.

Uma afirmação gramatical precisa ser também suficiente para que qualquer estudioso a aplique; não pode omitir informação importante. Por exemplo, imaginemos a situação seguinte: um gramático afirma que “o sujeito de um verbo de ação é sempre o agente”. Alguém contesta essa afirmação, mostrando que *comer* é um verbo de ação, mas em

[6] Os bolinhos foram comidos pelas formigas.

o sujeito (*os bolinhos*) não é agente. Aí o gramático responde: “Ah, mas é porque aí é voz passiva”. Mas na afirmação original não havia a restrição de que na voz passiva as coisas são diferentes; o gramático escapa das objeções usando fatores que não haviam sido fornecidos ao usuário. Isso se chama “argumento de carta na manga”, e é igualmente inaceitável na descrição gramatical. Em resumo, uma descrição gramatical precisa ser expressa em linguagem ao mesmo tempo clara, lógica e aplicável. Ou seja, explícita.

Dito isso, não se espere encontrar esse caráter explícito plenamente realizado em todos os trabalhos de lingüística, mesmo nos mais importantes. Trata-se de um ideal, e muitas vezes é impossível chegar a ele. Nos melhores casos, as descrições são pouco explícitas em virtude do nosso conhecimento insuficiente do próprio fenômeno estudado. Mas é importante tentar sempre ser o mais explícito possível, mesmo porque a alternativa, como vimos, é acabar não dizendo nada de efetivo. Diante de duas afirmações gramaticais, devemos sempre avaliá-las e escolher, em um primeiro momento, a mais explícita; essa é a mais facilmente testável, e portanto aquela que tem possibilidades de nos dizer alguma coisa sobre a língua. Afirmações pouco explícitas tendem a se reduzir a simples exercícios estilísticos, sem valor descritivo.

A virtude principal das afirmações gramaticais explícitas é que elas podem ser testadas, e por isso podem nos levar a um conhecimento melhor do problema estudado. A testagem se aplica principalmente a regras — as definições são mais instrumentos para a formulação posterior de afirmações descritivas. Considerando uma regra, se ela for verdadeira, e conseguirmos mostrar que é, é claro que aprenderemos alguma coisa; mas se ela for falsa, e conseguirmos mostrar que é falsa, também aprenderemos alguma coisa — ou seja, uma afirmação explícita sempre tem alguma coisa a dar (ainda que seja às vezes um pouco frustrante).

Já uma afirmação não explícita não tem nenhuma serventia. Uma afirmação desse tipo não pode, na verdade, ser plenamente compreendida, e portanto não pode ser testada. Não é possível mostrar que ela está certa, nem que está errada; em outras palavras, ela é inútil em termos de avançar nosso conhecimento da língua. Voltemos à definição de “conjunção” vista acima:

Os vocábulos gramaticais que servem para relacionar duas orações ou dois termos semelhantes da mesma oração chamam-se conjunções.

Se não pudermos saber o que é um vocábulo gramatical, não podemos aplicar essa definição em casos concretos. Portanto, não podemos nem sequer verificar se essa definição é ou não conveniente para ajudar na descrição da língua. Tudo porque a definição não é explícita o bastante para ser aplicada sem dificuldade à classificação das palavras.

COMPONENTES DO SIGNIFICADO

5.1. ESTRATÉGIAS DE INTERPRETAÇÃO

O significado que transmitimos através da fala é captado pelo ouvinte a partir de vários tipos de elementos que funcionam como pistas, entre as quais figura, como vimos no capítulo 3, a contribuição das palavras individuais e da estrutura das frases. Mas as pistas de que o receptor dispõe não param aí. Voltando aos exemplos

[1] A mão de Vera estava em cima da mesa.

[2] O copo de plástico estava em cima da mesa.

vimos que a relação expressa pela preposição *de* é de “posse” em [1] e de “material” em [2]. Isso não pode ser decorrente da palavra *de*, pois é a mesma. Nem pode ser inferido da estrutura, pois, novamente, é a mesma: artigo + substantivo + *de* + substantivo, tudo isso formando um constituinte (um sintagma nominal). De onde tiramos, então, que a relação expressa por *de* é diferente nos dois casos? A informação fornecida pelas palavras mais a estrutura gramatical não é suficiente.

Aqui entra o nosso conhecimento do mundo, que nos permite distinguir interpretações que fazem sentido das que não fazem, ou que são tão improváveis que acabam sendo rejeitadas. Sabemos que *Vera* é normalmente o nome de uma pessoa, e que as pessoas normalmente têm mãos; daí, inferimos que *de* em [1] designa “posse”. E sabemos que *copo* designa um objeto que pode ser feito de *plástico* (mas não pode ser posse do plástico, porque isso não faz sentido), e portanto entendemos que *de* em [2] designa “material”¹.

¹ Estou atribuindo o significado “posse” à preposição, mas há outras alternativas, como atribuí-lo à própria construção. Não é possível discutir isso aqui, e estou convencido de que

Ou seja, o significado que tiramos de uma expressão lingüística é resultado da aplicação de operações mentais que não são todas de caráter estritamente lingüístico — no sentido de que nem todas se baseiam na cadeia sonora ou gráfica (a seqüência de palavras) percebida pelo receptor. Parte dessas operações se baseia em nosso conhecimento geral do mundo.

Os detalhes desse processo não são bem conhecidos, de modo que temos que reconstruí-lo através de conjecturas. Vamos tomar o exemplo [3]:

[3] A mão de Vera

Podemos imaginar que o receptor, para processar esse sintagma, se vale de duas estratégias básicas:

Estratégia léxico-gramatical:

Com base nas indicações formais do sinal lingüístico e no seu conhecimento da língua, atribuir uma ou mais interpretações semânticas à expressão percebida.

Estratégia cognitiva²:

Com base no seu conhecimento geral do mundo, avaliar cada uma das interpretações semânticas resultantes da estratégia léxico-gramatical e rejeitar uma ou mais interpretações, se for o caso.

Não precisamos entender essas estratégias como se realizando sucessivamente, uma depois da outra, ou uma se aplicando ao resultado da outra. Antes, devemos concebê-las como dois tipos de recurso que o receptor aplica a cada momento, simultaneamente ou em mistura, durante seu processamento das expressões lingüísticas. Elas se distinguem apenas pelo tipo de conhecimento de que lançam mão: conhecimento da língua ou conhecimento do mundo extralingüístico³. Na exemplificação abaixo, terei

a preposição é, pelo menos, um componente indispensável para a interpretação do sintagma: *as mãos em Vera* será normalmente interpretado como referindo-se às mãos de outra pessoa, não às de Vera.

² A designação dessa operação como cognitiva não é muito adequada, porque nosso conhecimento léxico-gramatical é também, a rigor, cognitivo. Mas como falta uma palavra que designe o conhecimento não-léxico-gramatical, vou ter que usar cognitiva mesmo.

³ Estou certo de que muitos lingüistas e psicólogos vão contestar essa distinção que faço entre dois tipos de conhecimento. Mas é necessário assumir uma posição qualquer, e em uma área tão pouco conhecida não há nenhuma que escape a objeções.

que tratar das estratégias separadamente, cada uma por vez — mas isso é apenas uma conveniência de exposição, não um retrato fiel do que realmente se passa em nossa mente.

Voltando ao exemplo [3], a estratégia léxico-gramatical nos dará como resultado, entre outras coisas, que

- (a) essa seqüência é um sintagma nominal, e portanto denota uma “coisa” (e não uma ação, ou uma relação entre coisas);
- (b) existe uma relação entre a mão e Vera, porque essas palavras estão conectadas por uma preposição;
- (c) essa relação está dentro da gama de possibilidades semânticas da preposição *de* (por exemplo, não pode se tratar de uma mão *sem* Vera, nem de uma mão *contra* Vera, ou *ao lado de* Vera; mas pode ser uma relação de “posse”, e também de “material”);
- (d) dos itens léxicos usados, sabemos que *mão* pode ser uma parte do corpo (pode ser também outras coisas, como vimos!), e *Vera* é uma pessoa.

Trata-se ainda, é claro, de uma interpretação um tanto vaga; mas é fundamental para que se possa passar à estratégia cognitiva, que nos informará que

- (e) dentre as relações possíveis expressas por *de* (“posse”, “material” e outras), a de “posse” é a que dá resultado mais aceitável em termos de plausibilidade.

Acabamos, portanto, preferindo a interpretação “posse” para *de*, de preferência às outras que são em princípio possíveis. Assim, o significado “mão feita de Vera” é excluído por ser absurdo; mas se em vez de *Vera* tivéssemos *pedra*, a interpretação preferida no sintagma *a mão de pedra* seria “material”: *mão de pedra* significa “mão feita de pedra”.

5.2. OS DOIS COMPONENTES

Como se vê, o significado que conseguimos extrair de uma expressão lingüística provém de uma colaboração de vários fatores, nem todos eles propriamente lingüísticos (segundo a definição tradicional de “lingüístico”). Em situações reais de comunicação, é provável que nunca se usem exclusivamente recursos lingüísticos ou recursos não-lingüísticos, isto é, provavelmente submetemos todos os enunciados que recebemos ao escrutínio da estratégia léxico-gramatical e também ao da estratégia cognitiva.

Neste livro, estudamos principalmente os detalhes da estratégia léxico-gramatical; mas nem sempre será possível escapar de incluir em nossa análise aspectos mais amplamente cognitivos.

É tradicional estudar separadamente esses dois componentes do processo de interpretação, embora, a bem dizer, nem todos os lingüistas aceitem essa separação. Qualquer que seja a posição, entretanto, acredito que ela facilita a exposição, porque existem conhecimentos claramente gramaticais, e conhecimentos claramente não-gramaticais. Assim por exemplo, a forma dos verbos (*chego, cheguei, chegar...*) é um elemento claramente gramatical, ou seja, uma parte do nosso conhecimento da língua portuguesa; por outro lado, eu saber que meu amigo está na cama com uma perna quebrada não pode ser incluído no meu conhecimento da língua. Mas ambos esses tipos de informação colaboram para transmitir o significado; se meu amigo disser

[4] Você consegue fechar essa janela?

tenho que lançar mão dos dois tipos de conhecimento para saber o que, no fim das contas, ele está querendo me dizer. Por um lado, a forma do verbo *conseguir* (isto é, *consegue*) me informa que ele não está me perguntando sobre algum fato passado; se a pergunta fosse

[5] Você conseguiu fechar essa janela?

a resposta poderia ser *consegui, sim*, porque tenho que entender a pergunta como um pedido de informação. Mas como o verbo está no presente — e eu sei que meu amigo não pode levantar da cama e ir fechar a janela —, posso entender [4] como um pedido para que eu feche a janela.

Note-se como os dois tipos de fatores colaboram para a transmissão da mensagem, sem no entanto se confundirem. Temos que reconhecer, portanto, que o processo de compreensão não é puramente gramatical. Só que isso não impede que se possa separar com alguma clareza os fatores gramaticais dos fatores não-gramaticais que nos permitem pôr o processo em funcionamento.

ONDE ENTRA A GRAMÁTICA

6.1. REGRAS GRAMATICAIS

A estratégia léxico-gramatical depende, para funcionar, de um conjunto bastante complexo de operações menores, geralmente chamadas regras¹. Em resumo, a estratégia léxico-gramatical utiliza todo o nosso conhecimento da estrutura da língua, tanto no que diz respeito à forma quanto ao significado, além do conhecimento de todo o vocabulário da língua (o léxico)². A estratégia cognitiva é ainda mais complexa, e inclui todo o nosso conhecimento do mundo, mas por ora não precisamos nos ocupar dela; vamos nos concentrar no conteúdo da estratégia léxico-gramatical.

Uma parte significativa dos recursos dessa estratégia é, portanto, o que chamamos gramática da língua. E a gramática pode ser entendida como um conjunto de instruções para construir e interpretar unidades: sintagmas, frases, orações e períodos.

Essas instruções (as regras) são tradicionalmente distinguidas em quatro tipos principais, a saber:

regras fonológicas, que descrevem a pronúncia das formas da língua;
regras morfológicas, que descrevem a estrutura das palavras;
regras sintáticas, que descrevem a organização das sentenças;
regras semânticas (ou regras de interpretação), que relacionam as formas descritas pelas regras morfológicas e sintáticas com seus respectivos significados. O mais correto seria chamar essas regras de simbólicas, já

¹ Neste livro utilizo a noção de "regra" no sentido de "regra descritiva", seguindo o uso da tradição recente. Trata-se, a meu ver, de uma metáfora útil — que, aliás, não se opõe à posição cognitivista; ver, por exemplo, Langacker (1987: 532).

² Para nem falar dos clichês e das expressões idiomáticas!

que o que elas fazem é estabelecer a relação entre forma e significado (ou seja, colocá-los no que o linguísta Ronald Langacker chama de “espaço simbólico”). Mas o termo regra semântica parece estar consagrado.

Neste livro, não trato das regras fonológicas, e só de passagem das regras morfológicas, concentrando-me nas regras sintáticas e semânticas.

6.2. DEFINIÇÕES

Vamos considerar as regras sintáticas e semânticas, que são o assunto principal deste livro. Por exemplo, há uma regra sintática que determina o seguinte:

a forma *eu* é usada como sujeito; a forma *me* como objeto

Por exemplo:

[1] Eu chamei o advogado.

[2] O advogado me chamou.

Para aplicar essa regra, evidentemente, temos que saber com exatidão o que é um “sujeito” e um “objeto”. Como não podemos dar a lista de todos os sujeitos ou objetos da língua, é necessário lançar mão de definições.

Assim, definimos, por exemplo, o sujeito como o termo com o qual o verbo concorda, de modo que fica claro por que *eu*, e não *o advogado*, é o sujeito em [1], e por que *o advogado* é o sujeito em [2]. A noção de sujeito é importante porque várias regras gramaticais se referem a ela. Vimos acima uma regra sintática, que descreve o uso de *eu* e *me*; temos também na língua uma regra semântica que estipula que o sujeito do verbo *chamar* (na voz ativa) denota o ser que pratica a ação — por isso, entendemos que eu pratiquei a ação de chamar em [1], mas em [2] essa ação foi praticada pelo advogado.

As definições não são regras, mas servem para explicar o que se entende por determinado termo técnico, como por exemplo “sujeito”. As definições são indispensáveis para caracterizar funções (como a de sujeito ou a de objeto direto) e classes (como a dos substantivos ou a dos verbos); e as regras se referem constantemente às classes e funções (como se vê na regra dada logo acima, que faz referência à função de sujeito).

Entre as noções que precisam ser definidas em uma gramática temos: classes de palavras como verbo, preposição e pronome; e funções sintáticas

como sujeito, complemento e predicativo. As funções e as classes, juntamente com suas definições, são estudadas nos capítulos 12 a 16.

6.3. REGRAS PARA QUE?

É possível fazer uma lista de todas as palavras de uma língua (os grandes dicionários, como o *Aurélio* e o *Houaiss*, chegam perto disso), mas não é possível fazer uma lista de todas as frases (ou sintagmas) da língua. Estas são em número indefinido, e todos os dias ouvimos ou lemos frases que nunca vimos, e não obstante as entendemos perfeitamente. Por exemplo, quem já viu a frase seguinte?

[3] O gato preto coçou o meu camelo.

Mas não há problemas em interpretá-la e reconhecê-la como uma boa frase portuguesa.

Conseguimos fazer isso, claro, porque embora [3] seja uma novidade, suas partes não são novidade. E as partes podem ser distinguidas em dois tipos: palavras (*o*, *gato*, *coçou* etc.) e construções. Entre as construções que ocorrem em [3], temos:

sintagmas nominais: *o gato preto*, *o meu camelo*;

oração: *o gato preto coçou o meu camelo*.

Cada uma dessas construções forma um constituinte (ver 3.3.2) que pertence a determinado tipo (sintagma nominal, oração etc.) e se compõe de determinados elementos. Pois bem: um dos papéis das regras gramaticais é justamente descrever a maneira de montar palavras de modo a formar constituintes. Como vimos em 3.3.2, nenhuma regra descreve a montagem de *coçou o*, porque, embora essa seqüência esteja presente em [3], ela não forma um constituinte.

O mesmo vale para as regras semânticas: conseguimos atribuir significado à frase [3] e a suas partes (*o gato preto*, *o meu camelo* etc.) porque se compõem de palavras conhecidas, organizadas de maneiras conhecidas — ainda que o resultado final seja novo para nós.

Vamos examinar alguns aspectos do processamento da frase [3]. Esse processamento se faz, na parte que nos interessa, através do que chamamos estratégia léxico-gramatical — que, como diz o nome, inclui não apenas a

aplicação de regras, mas também o reconhecimento de itens léxicos. Os itens léxicos são palavras, morfemas e expressões maiores, mas aqui, para simplificar a exposição, vamos nos limitar às palavras.

Podemos começar com o seguinte: como é que sabemos que o gato é que era preto, e não o camelo?

A resposta não é difícil de imaginar: é porque a palavra *preto*, que indica a qualidade de “ser preto”, está junto a *gato*, e não a *camelo*. Se trocarmos a palavra de lugar, a qualidade se transfere do gato para o camelo:

[4] O gato coçou o meu camelo preto.

Isso ilustra o funcionamento de uma regra que diz mais ou menos o seguinte:

uma palavra que exprime qualidade acrescenta essa qualidade a coisa cujo nome estiver logo antes dela.

Essa regra³ é parte do nosso conhecimento do português. Nós a pomos em prática para descobrir, em uma frase como [3] ou [4], qual dos dois animais é preto. Essa é uma regra semântica, porque nos fornece a relação (simbólica) entre um fato de forma (“estar logo depois de uma palavra”) e um fato de significado (“acrescentar uma qualidade à coisa mencionada”).

Isso pode parecer insuportavelmente óbvio, mas é só porque descreve um conhecimento que dominamos muito bem. Mas imaginemos um estrangeiro que não conhecesse essa regra: essa pessoa, ao ler [3], não saberia qual dos animais é preto. Em algumas línguas, a posição adjacente não é condição para que uma palavra qualifique a outra. Por exemplo, em latim podemos ter uma frase como

[5] *Nigrum camelus fricauit felem.*

Nigrum quer dizer ‘preto’; *camelus* é ‘camelo’; *fricauit* é ‘coçou’; e *felem* é ‘gato’. Mas a frase significa ‘o camelo coçou o gato preto’ (o latim não tem artigos correspondentes ao português *o*) — apesar de *nigrum* estar junto de *camelus*, qualifica *felem*. Acontece que em latim a regra é diferente: o

³ Que deve ser entendida como uma primeira aproximação. A regra real é bem mais complicada.

que vale não é a posição, mas a forma do qualificativo. No caso, *nigrum* e *felem* terminam ambos em *-m*, e portanto devem ser semanticamente relacionados um ao outro⁴. Um brasileiro lendo [5] vai tender a achar que o camelo é preto; e um romano lendo a frase portuguesa [3] poderia ficar meio perdido, sem saber quem é preto, o camelo ou o gato. Observe-se que para chegar à decisão correta no caso de *preto* em [3] ou [4], é indispensável fazer uso da regra gramatical. É apenas a regra que nos permite interpretar essas frases.

Vamos voltar ao exemplo [3]:

[3] O gato preto coçou o meu camelo.

Outra coisa que essa frase nos diz é que o gato preto praticou a ação de coçar, e o camelo sofreu essa ação. Aqui novamente temos o direito de perguntar: como é que a gente fica sabendo disso?

A ordem dos elementos é o elemento decisivo também nesse caso, porque se dissermos

[6] O meu camelo coçou o gato preto.

os papéis ficam trocados, com o gato saindo coçado, em vez do camelo. Deve haver, então, uma regra que se refere à ordem dos elementos nessa frase, e que conforme a ordem nos informa quem praticou a ação.

Para formular essa regra, é preciso explicitar quais são os elementos cuja ordem é importante, e também deixar claro em relação a que ponto de referência essa ordem se define. Temos que chegar a alguma coisa como

com verbos do tipo de *coçar*, o sintagma nominal que vem logo antes do verbo se entende como o agente, e o sintagma nominal depois do verbo se entende como paciente;

onde agente é o elemento que pratica a ação, e paciente o que sofre a ação.

O problema é que temos que definir o que é um sintagma nominal e o que é um verbo. Aqui enfrentamos o problema já mencionado, de que

⁴ Espero que os latinistas me perdoem essa exposição tremendamente simplificada dos fatos; mas é só como ilustração de um sistema diferente do português.

⁵ Estou evitando usar as noções de “sujeito” e “objeto direto”, porque o que funciona nessa frase é realmente uma relação de ordem pura e simples.

as regras estão sempre se referindo a classes, e dependem de que essas classes sejam claramente definidas. Um falante sabe muito bem distinguir um sintagma nominal de outros tipos de sintagmas, tanto é que quando ouve

[7] Ontem cocei o gato preto.

não vai imaginar que *ontem* denote o elemento que pratica a ação; como não é um sintagma nominal, *ontem* não pode receber a função semântica de agente.

Observe-se como a regra semântica parte do aspecto formal (posição dos sintagmas em relação ao verbo) e relaciona esses fatores a elementos semânticos (agente e paciente). O usuário da língua utiliza seu conhecimento de regras como essa para construir ou entender enunciados. Em muitos casos, isso é suficiente para a compreensão; em outros, igualmente numerosos, o conhecimento das regras tem que ser complementado com aspectos do conhecimento do mundo — como vimos nos exemplos do capítulo 5.

Em resumo, precisamos conhecer as regras porque elas são parte do mecanismo mental que nos permite relacionar formas com significados e, portanto, compreender e transmitir mensagens.

6.4. SINTAGMAS NOMINAIS E VERBOS

Como o sintagma nominal (mais conhecido pela abreviatura SN) vai aparecer com frequência daqui por diante, vou defini-lo desde já. O sintagma nominal (SN) é uma seqüência de uma ou mais palavras que pode ser sujeito, objeto direto ou complemento de uma preposição. E ele se define, além disso, por determinado tipo de estrutura interna: por exemplo, um artigo mais um nominal, como em *o camelo*⁶. É claro que o ouvinte parte da estrutura interna como seu primeiro passo para identificar os SNS de uma frase.

Por exemplo, as seqüências entre colchetes nas frases abaixo são SNS:

[8] [O gato] passou entre [a mesa] e [aquele armário branco].

[9] [Café] faz mal para [o seu estômago].

[10] [Mamãe] serviu [uma torta] a[os convidados].

[atenção: o SN é *os convidados*; *a-* é a preposição]

⁶ O SN pode ter várias outras estruturas, exemplificadas por *Maria, minha mãe, o rapaz que morava com você* etc. E também pode ter outras funções, por exemplo predicativo.

[11] [O vizinho do 201] vive dando [presentes] para [a menina do 303].

Semanticamente, o SN tem a propriedade de se referir a uma coisa (o que inclui pessoas e seres vivos em geral). Por exemplo, *o gato, aquele armário branco e mamãe* designam “coisas”, nesse sentido. Mas *passou, para você, rapidamente e logo depois do meu aniversário* não designam coisas, não podem ser sujeito, objeto direto e complemento de preposição⁷, e por isso não são SNS.

Quanto ao verbo, ele pode ser identificado através de diversas características. A principal é que o verbo pertence a um lexema (um conjunto de palavras) cujos membros se diferenciam por terminações de pessoa (*serviu, servi, servimos, serviram*) e tempo (*serviu, serve, servirá*).

⁷ O sintagma *para você* contém uma preposição, mas não pode ocorrer depois de uma: **em para você, *de para você* etc.

PRINCÍPIOS DE ANÁLISE

7.1. O QUE É UMA LÍNGUA?

Em poucas palavras, a descrição de uma língua é a explicitação da relação que existe entre os significados e as formas dessa língua. As formas são percebidas pelos sentidos: pelo ouvido, quando faladas, e pela vista, quando escritas. Assim, a forma *gato* é manifestada pela pronúncia [‘gatu] ou pela grafia *gato*. E os significados, embora não sejam percebidos pelos sentidos, de certo modo aparecem claramente em nossa mente: o significado “gato” é o conceito de determinado tipo de animal, que associamos à forma percebida *gato*. Vamos usar a convenção “gato” (entre aspas) para representar um conceito, ou seja, a idéia de gato; e *gato* (em itálico) para representar uma forma falada ou escrita: uma seqüência de sons ou de letras¹. A pronúncia, quando é necessário referir-se especificamente a ela, vem entre colchetes, em transcrição fonética: [‘gatu].

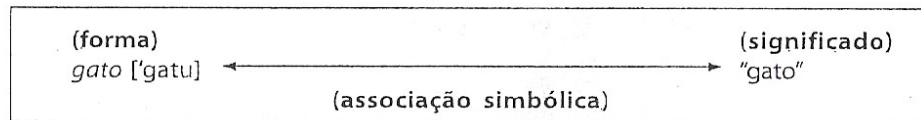
Uma língua é um sistema de relacionamentos entre formas e significados. Em português, o conceito de “gato” (seu significado) se relaciona com a forma *gato*; em francês, o mesmo conceito de “gato”² se relaciona com a forma *chat*; em inglês, esse conceito se relaciona com a forma *cat*; em latim com a forma *felis* e assim por diante. Ao usarmos ou ao estudarmos uma língua, estamos sempre, de uma maneira ou de outra, lidando com associações entre formas e significados³.

¹ Além disso, *gato* representa um lexema, ou seja, a unidade complexa que chamamos normalmente “a palavra *gato*” (ver o conceito de lexema no capítulo 11). Para não complicar excessivamente o texto, não vou fazer todas as distinções que seriam, em princípio, necessárias.

² Mais ou menos o mesmo conceito! Ver a nota seguinte.

³ Cada língua tem, evidentemente, suas formas próprias. O que muita gente não sabe é que cada língua tem também seus significados próprios. Por exemplo, o português *madeira* significa uma coisa que não existe em francês associado a uma forma única, porque *bois* quer dizer ‘madeira’, mas também ‘bosque’. Por outro lado, o francês distingue *espérer* e *attendre*, ambos traduzidos por *esperar* em português, mas que têm significados diferentes.

Uma conseqüência disso é que o lingüista precisa estar sempre consciente do caráter das afirmações gramaticais que faz: algumas se referem à forma, outras ao significado, e outras ainda se referem à relação entre forma e significado. Como já vimos, as afirmações que se referem à forma se chamam formais; as que se referem ao significado, semânticas; e as que se referem à associação entre forma e significado se chamam simbólicas. Um conjunto constituído de uma forma mais um significado relacionados um com o outro se chama um signo⁴. O signo correspondente à palavra *gato* pode ser representado graficamente assim:



O falante da língua conhece (através das regras gramaticais e das listas do léxico) quais são as formas, os significados e as associações simbólicas possíveis, e isso lhe permite usar a língua para se comunicar. Saber português, por exemplo, é saber que existe uma palavra que se escreve *gato*; saber que essa palavra designa um certo tipo de animal; saber que a palavra *gato* é masculina — ou seja, co-ocorre com o artigo *o* e com o qualificativo *branco*; saber que existe a palavra *gata*, que designa a fêmea do mesmo animal; saber que a palavra *gata* co-ocorre com *a*, *branca* etc. De posse de conhecimentos como esses, o falante consegue produzir e compreender enunciados em português. Descrever uma língua é explicitar esse conhecimento que existe na mente dos falantes.

7.2. AFIRMAÇÕES FORMAIS, SEMÂNTICAS E SIMBÓLICAS

Antes de passar adiante, vejamos alguns exemplos. A gramática se compõe de afirmações de caráter formal, semântico e simbólico. Exemplos de afirmações formais são:

- (a) O gerúndio é uma forma terminada em *-ando*, *-endo*, *-indo* ou (para alguns verbos) *-ondo*;
- (b) Verbo é uma palavra que aceita o sufixo *-mos*;
- (c) A palavra *gramática* é proparoxítota;
- (d) As letras dobradas (*ss*, *rr*) são separadas cada uma em uma sílaba: *car-ro*, *pas-sa*.

⁴ Essa é a designação mais comum. Langacker (1987, p. 81) prefere falar de símbolo.

Note-se como cada uma delas só se refere ao aspecto formal dos elementos tratados; não há nenhuma menção ao significado que os elementos ou suas combinações possuem. As afirmações (a) e (b) são definições (de gerúndio e de verbo, respectivamente); a afirmação (c) descreve um aspecto da pronúncia da palavra *gramática*; e a afirmação (d) dá uma regra de separação de sílabas. Para aplicar qualquer dessas definições ou regras, basta observar a forma dos enunciados da língua, sem necessidade de compreendê-los.

Afirmações semânticas, por outro lado, referem-se ao significado; por exemplo:

- (e) "Agente" é o elemento que pratica uma ação;
- (f) "Referencial" é um elemento que designa um ser;
- (g) "Futuro" é um evento ou estado entendido como válido para algum momento posterior ao momento da fala.

A observação da forma, por si só, não ajuda a identificar agentes, nem referenciais, nem o momento de um evento relatado. Para isso temos que compreender o significado do enunciado em questão. Por exemplo, para saber se temos um "futuro" na frase

[1] Eu vou te procurar logo que puder.

temos que compreendê-la, ou seja, precisamos ter acesso a seu significado. Igualmente, só sabemos que *livro* é referencial porque conhecemos o significado dessa palavra. Mas podemos verificar, apenas pelo exame da frase e sem compreendê-la, que em [1] não há nenhuma ocorrência de gerúndio, já que não há aí nenhuma forma terminada em *-ndo*.

A diferença entre afirmações formais e semânticas pode ser testada com alguém que não sabe a língua: qualquer estrangeiro pode verificar se um elemento vem ou não logo antes de outro; mas ele não vai conseguir nos dizer se *livro* é ou não referencial, ou se o sintagma *o gavião* expressa ou não o agente em

[2] O gavião arranhou meu gato.

porque ele não conhece a semântica de *livro*, nem a regra para encontrar o agente dentro de uma sentença.

Finalmente, temos as afirmações simbólicas — que, no fundo, são as mais importantes, porque tratam do objeto principal da lingüística, que são os signos. Exemplos de afirmações simbólicas são:

- (h) Em *os descobrimentos portugueses*, a palavra *portugueses* denota o agente;
- (i) O possessivo se refere ao nominal que lhe fica contíguo;
- (j) *Bode* é o nome de um tipo de animal;
- (k) O sufixo *-rei* dos verbos exprime a idéia de que alguma coisa ainda vai acontecer;
- (l) A palavra *livro* pode ser usada como referencial.

Agora temos afirmações que estabelecem relações entre formas e significados. Em (h) temos uma afirmação particular, referente a uma frase específica, mas em (i) já temos uma regra, que estabelece uma relação simbólica em termos gerais, válidos para muitas estruturas da língua. A regra (i) nos permite descobrir qual dos elementos da frase [2] é que é “meu”: é o gato, porque o nominal *gato* está contíguo ao possessivo *meu*. A afirmação (i) pode ser considerada uma regra da gramática portuguesa.

Para explicar melhor: a regra (i) parte de uma observação formal: a posição do possessivo (um possessivo é uma das palavras *meu, seu, nosso*) em relação ao nominal. E a mesma regra estabelece uma relação entre essa observação formal e um fato semântico: o que é “meu” (ou “seu”, ou “nosso”) é a coisa denotada pelo nominal. É por isso que em [2] o que é meu é o gato, mas em [3] é o gavião que é meu:

[2] O gavião arranhou meu gato.

[3] Meu gavião arranhou o gato.

Temos aqui mais um exemplo de como a própria ordem dos elementos tem seu significado, e participa de relações simbólicas.

Os exemplos (h) e (i) mostram que a relação entre forma e significado vai muito além das simples palavras. Um dos aspectos da forma é a maneira pela qual as palavras se associam, em determinada ordem e de determinada forma (por exemplo, depois de *eu* temos que colocar *cheguei*, e não *chegou*). Tudo isso pode ter conseqüências semânticas, e é outro aspecto do relacionamento simbólico.

REGRAS DESCRITIVAS

8.1. DESCREVENDO A REALIDADE DA LÍNGUA

Grande parte da descrição de uma língua consiste de regras. Mas o que se entende por “regra” em lingüística não é a mesma coisa que se aprende na escola, aquelas regras que dizem como se deve e como não se deve falar ou escrever. Um lingüista não se preocupa com a maneira como as pessoas deveriam usar a língua: o objetivo do lingüista é descrever a maneira como as pessoas realmente usam a língua, sem fazer julgamentos do que é “certo” ou “errado”.

Desse ponto de vista, errado é apenas aquilo que os falantes rejeitam, como dizer **uma livro*, ou então **gato gavião meu arranhou*. O lingüista tenta formular as regras de maneira a excluir da gramática essas seqüências rejeitadas, ao mesmo tempo incluindo as seqüências aceitáveis como *um livro* e *o gavião arranhou meu gato*. Em outras palavras, o lingüista estuda o que é, e não o que algumas pessoas acham que deveria ser. A lingüística é o estudo de um aspecto do mundo, e afirmar que é errado dizer *os livro* (como todos nós dizemos com freqüência)¹ faz tão pouco sentido quanto dizer que é errado respirar oxigênio. Respirar oxigênio, assim como dizer *os livro*, não é “certo” nem “errado” — é um fato, e como tal não pode ser questionado. Não é relevante para nós que algumas pessoas não gostem desse fato. Eu, por exemplo, não gosto que os cachorros babem, mas eles babam. Paciência!

8.2. AGENTES E PACIENTES ANTES DO VERBO

As regras que nos interessam aqui, portanto, são regras descritivas, e por meio delas vamos descrever alguns fatos da língua. Por exemplo, uma regra estipula que

¹ Falando, claro. Na escrita o que ocorre é sempre *os livros*.

Regra 1

O sintagma nominal que vem logo antes do verbo na forma ativa exprime o agente.

(Como já vimos, agente é o elemento que pratica a ação)

Isso quer dizer que um falante do português, ao ouvir a frase

[1] Meu gavião arranhou o gato.

fica sabendo que se trata de uma ação praticada pelo meu gavião, e não pelo gato; o gato, no caso, é o paciente (elemento que sofre a ação). Essa regra é confirmada não consultando as gramáticas, mas verificando o que os falantes do português entendem quando escutam essa frase: eles concordam que [1] comunica a idéia de que o gavião fez alguma coisa com o gato, e não vice-versa.

A regra vale não apenas para a frase [1], mas para muitas frases em que *arranhar* aparece na forma ativa:

[2] O gavião arranhou meu gato.

agente

[3] Eu arranhei o canário do papai.

agente

[4] Esse canário, o gato arranhou.

agente

Aí surge um problema inesperado: encontramos uma frase como

[5] O gato me arranhou.

e o elemento imediatamente antes do verbo é *me*, mas quem praticou a ação foi o gato. Será que a regra está errada?

Pode ser, claro, mas talvez seja apenas necessário fazer uma modificação na maneira como a regra foi formulada. Verifica-se que essa exceção só acontece com algumas palavras do português, a saber, *me*, *te*, *se*, *nos* (e, no português escrito, *o*, *a*, *os*, *as*, *lhe* e *lhes*)²:

² Não vou nem mencionar a forma *vos*, que não existe no português do Brasil.

[6] O gato nos arranhou.

[7] O gato te arranhou.

Assim, é possível adaptar a regra, acrescentando uma exceção:

Regra 1b (revisão)

O sintagma nominal que vem logo antes do verbo *arranhar* na forma ativa exprime o agente, exceto quando esse sintagma é *me*, *te*, *se* ou *nos*; nesse caso, o agente é o sintagma que vem antes de *me*, *te*, *se* ou *nos*.

Não precisamos nos afligir com essa introdução de exceções em uma regra que era tão simples no início. Nosso objetivo não é chegar às regras mais simples possíveis, mas às regras que descrevem os fatos da maneira mais realista. Se os fatos são complexos, as regras que os descrevem também têm que ser complexas.

Aliás, os fatos referentes ao verbo *arranhar* são ainda mais complicados. Em certos casos, o sintagma que precede o verbo não representa o agente, mas o paciente (isto é, o ser que sofre a ação). Nas frases que vimos até agora, o paciente (quando não representado por *me*, *te*, *se*, *nos*) é expresso pelo sintagma que vem depois do verbo:

[2] O gavião arranhou meu gato.

Aqui o paciente é *meu gato*.

Mas agora temos uma frase como:

[8] O menino arranhou todo.

Ao interpretar essa frase, entendemos *o menino* como o paciente — ou seja, o menino não fez nada com ninguém, mas saiu arranhado³. Isso vai ter que ser acrescentado à regra que interpreta o agente e o paciente do verbo *arranhar*:

Regra 1b (acrescimo)

quando *arranhar* for acompanhado por apenas um SN, esse SN é paciente; quando há dois SNs, o SN que vem logo antes do verbo é agente, e o outro é paciente.

³ Isso vale para o dialeto mineiro. Em outras partes do Brasil, a frase seria *o menino se arranhou todo*, mas o argumento continua valendo, em forma ligeiramente modificada: o sujeito é paciente.

Agora tenho uma boa notícia: as regras elaboradas para a interpretação do agente e do paciente de *arranhar* valem, sem modificação, para muitos outros verbos do português, como *quebrar*, *esquentar*, *desanimar*, *engordar* etc. Mas não valem para outros verbos, como *matar*: com esse verbo o SN pré-verbal é sempre agente, nunca paciente:

[9] A onça matou o cachorro.

[10] Onça mata.

Para o verbo *matar* o acréscimo à regra 1b não é necessário. Para *matar* e muitos outros verbos, a regra 1b (revisão) é a que funciona. Isto é, a regra 1b com o acréscimo só é válida para parte dos verbos da língua (*arranhar*, *esquentar*, *quebrar*...); outros (*matar*, *comer*, *escrever*...) se contentam com a regra sem o acréscimo.

SINTAXE E SEMÂNTICA

A questão do relacionamento entre forma e significado é provavelmente a mais importante da gramática; por isso, temos que voltar a ela com frequência. Vamos agora considerar a hipótese de que todo traço formal corresponde a algum conteúdo semântico, ou seja, nenhum traço formal na língua seria semanticamente inútil; como veremos, as indicações são de que essa hipótese é apenas uma aproximação, não uma verdade categórica.

9.1. FORMAS SEM SIGNIFICADO

Vimos que a gramática estabelece relacionamentos simbólicos (entre formas e significados). Esse relacionamento é, muitas vezes, bastante indireto. Por exemplo, existem elementos formais que não se relacionam com nenhum elemento semântico; é como se fossem uma espécie de “ruído” do sistema, coisas sem utilidade aparente para a relação simbólica, mas que não obstante existem, e precisam ser descritas.

Por exemplo, os verbos regulares em português se dividem em três conjugações, exemplificadas por *mandar*, *vender* e *abrir*. Cada conjugação tem um sistema completo de sufixos, de modo que, por exemplo, o pretérito perfeito pode terminar em *-ou* (*mandou*), em *-eu* (*vendeu*) ou então em *-iu* (*abriu*). Essa diferença formal não corresponde a coisa nenhuma no espaço semântico — isto é, o fato de um verbo pertencer à primeira, à segunda ou à terceira conjugação não nos diz nada a respeito de seu provável significado.

Outro exemplo são as preposições que acompanham obrigatoriamente certos verbos. O complemento do verbo *gostar* vem precedido da preposição *de*: *gosto de marmelada*; o do verbo *acreditar* requer a preposição *em*: *acreditei em você*; o do verbo *comer* vem sem preposição: *ela comeu a pizza*;

e contar, no sentido de “ter confiança”, requer *com*: *conto com todos vocês*. Essa exigência de uma ou outra preposição (ou nenhuma) não se correlaciona, que se saiba, com nenhum traço do significado desses verbos.

Esses casos mostram que formas sem significado existem na língua, e que, portanto, a hipótese mencionada acima não é uma verdade geral.

9.2. FORMAS COM SIGNIFICADO

Já outras marcas formais se relacionam com significados, tendo portanto valor simbólico. Um exemplo é a ocorrência do sufixo *-s* nos nominais, que se relaciona com a noção de “plural”: *cachorro* designa um animal, *cachorros* designa mais de um. Ou seja, a presença do sufixo *-s* é um sinal para que o receptor ative a noção de “plural” em sua mente. E, como a relação é mesmo complexa, em alguns casos a presença de *-s* não tem nada a ver com isso, como na palavra *férias*, que não denota um conjunto de objetos. E em outros casos o conceito de “plural” está presente mesmo sem que haja o sufixo *-s*, como nos coletivos: *turma*, *grupo*, *rebanho* etc.

Alguns verbos podem vir seguidos de preposição, mas em condições diferentes das que vimos acima para *gostar* e *acreditar*. Assim, pode-se dizer

-
- [1] Papai chegou a São Paulo.
 [2] Papai chegou de São Paulo.
 [3] Papai chegou a cadeira para perto da janela.
-

e em cada frase a preposição usada depois de *chegou* (ou a ausência de preposição no caso de *chegou a cadeira...*) veicula um significado diferente, o que mostra que as preposições, nesses casos, têm valor simbólico.

Em outras palavras, nem toda forma tem significado, mas muitas têm — e não se sabe qual é a proporção, na língua, entre formas com significado e formas sem significado. Essa é mais uma razão pela qual temos que estar sempre conscientes da diferença entre afirmações formais, semânticas e simbólicas. A cada passo podemos nos defrontar com um traço formal que se correlaciona, de diversas maneiras, com algum traço semântico; mas às vezes encontramos um traço formal que não se correlaciona com coisa alguma no plano semântico. Nosso conhecimento da estrutura da língua não está suficientemente avançado para nos permitir fazer afirmações seguras nesse terreno, de modo que o mais prudente é considerar que cada caso é um caso. Meu

palpite é que em geral as diferenças formais correspondem a diferenças de significado; mas só em geral, e as exceções são importantes e talvez numerosas.

O princípio prático a ser seguido é o seguinte: uma forma não tem significado até que se consiga descobrir um. Isso deixa bem claro como a conclusão é provisória — quem sabe um dia alguém vai descobrir que todos os verbos da segunda conjugação têm um traço de significado em comum? Eu duvido, mas quem sou eu para prever o futuro?

9.3. A SINTAXE RESIDUAL

A partir daí, é possível desenvolver um princípio para orientar o trabalho de descrição. A idéia central é a de que as relações simbólicas precisam ser formuladas inevitavelmente, qualquer que seja a teoria adotada; elas fazem parte dos fatos da língua. Conseqüentemente, formulamos o seguinte princípio:

Princípio da sintaxe residual (para a descrição das formas da língua)

Sempre que for possível descrever um traço de forma em termos de traços de significado a ele associados, essa descrição deve ser preferida a uma descrição que considere a forma independentemente do significado. Isso significa que só devem ser descritos exclusivamente em termos de forma os traços formais não associados a traços semânticos.

Vejamos alguns exemplos para ilustrar o funcionamento desse princípio. Um caso simples é o da ordenação dos nominais em casos como

-
- [4] Animal mamífero
 [5] *Mamífero animal
-

Temos acima sintagmas formados de dois nominais cada — relembro que os nominais são tradicionalmente chamados “substantivos” e “adjetivos” (e outras coisas). Mas aqui estou adotando uma nomenclatura mais correta.

À primeira vista, pode-se pensar que alguns nominais precisam ser marcados quanto à posição em que ocorrem no sintagma: *animal* seria do

¹ Venho utilizando esse princípio há vários anos, e durante algum tempo acreditei ter sido o primeiro a formulá-lo. Recentemente, contudo, encontrei exatamente o mesmo princípio formulado com muita clareza em Jackendoff (1983: 13), com o nome de *Grammatical Constraint*. Pereant qui ante nos nostra dixerunt!

grupo I, que precisa vir antes de *mamífero*, que é do grupo II. Se fizermos isso estaremos descrevendo os fatos — mas, relembro, toda descrição é uma hipótese, e portanto temos o direito de procurar outra, talvez melhor.

Essa divisão dos nominais nos grupos I e II é algo que teríamos que introduzir na gramática apenas para dar conta da ordenação observada em [4] e [5]; ela não tem nenhuma outra utilidade. Mas examinemos os dados com mais cuidado. Comparemos [4] e [5] com outros casos análogos:

-
- [6] Médico geriatra
 [7] *Geriatra médico
 [8] Queijo mussarela
 [9] *Mussarela queijo
-

Note-se que, a se manter nossa hipótese atual, teremos que marcar também essas palavras, colocando *médico* e *queijo* no grupo I, *geriatra* e *mussarela* no grupo II. Isso, porém, não vai resolver o problema, simplesmente porque palavras como *médico* e *animal* (grupo I) podem aparecer em segundo lugar no sintagma, como em

-
- [10] Amigo médico
 [11] Produto animal
-

e a palavra *geriatra* (Grupo II) pode aparecer em primeiro lugar:

-
- [12] Geriatra competente
-

Como sair desse dilema?

Se considerarmos a semântica desses sintagmas, verificaremos uma coisa interessante: nos casos aceitáveis [4], [6] e [8], o nominal que vem em segundo lugar denota um subconjunto próprio do conjunto denotado pelo primeiro; isto é, todo mamífero é animal; todo geriatra é médico; e toda mussarela é queijo. A recíproca, naturalmente, não é verdadeira: há animais que não são mamíferos, médicos que não são geriatras, queijos que não são mussarelas.

Isso sugere uma outra maneira de formular a restrição de ordenação observada: podemos dizer que nessas seqüências o elemento mais restritivo precisa vir sempre depois do menos restritivo. Essa análise funciona melhor

do que a anterior, e com uma vantagem extra: agora não precisamos inventar uma divisão (grupo I/grupo II), que só serve mesmo para descrever a ordem dos nominais. Estamos aproveitando uma diferença semântica que precisa, de qualquer maneira, ser explicitada em algum momento. Ou seja, todo falante do português sabe que *mussarela* exprime um subconjunto dos queijos; isso é parte de seu conhecimento da língua (e do mundo).

Nossa nova hipótese é que há uma regra de interpretação em português que faz com que os nominais mais restritivos sejam sempre colocados mais para o fim do sintagma. Essa regra é responsável pela má formação de sintagmas como [5], [7] e [9]: nessas frases ela nos obriga a atribuir a *médico* um significado mais restritivo do que o de *geriatra*, mas isso se choca com os conhecimentos que temos. É como se alguém dissesse: **os melhores geriatras são os médicos*; compare-se com *os melhores médicos são os geriatras*, que é aceitável. Em outras palavras, a inaceitabilidade de [5], [7] e [9] é atribuída não a uma regra sintática de colocação de elementos no sintagma nominal, mas à sua má formação semântica (ou, talvez melhor dizendo, à sua inadequação ao nosso conhecimento do mundo).

Há alguma evidência adicional em favor da regra mencionada acima. No sintagma

-
- [13] Vestido novo importado
-

não há nenhuma relação de subconjunto entre os dois elementos *novo* e *importado*: há vestidos importados que não são novos, e também há vestidos novos que não são importados. Por isso mesmo, podemos dizer igualmente:

-
- [14] Vestido importado novo
-

mas em certas situações discursivas a ordem pode se tornar fixa. Dos dois diálogos seguintes, só [15] é plenamente aceitável; [16] soa muito estranho:

-
- [15] — Por que você não quer este vestido novo?
 — Porque eu queria um vestido novo importado.
 [16] — Por que você não quer este vestido novo?
 — *Porque eu queria um vestido importado novo.
-

Acontece que a primeira frase estabelece que estamos tratando de vestidos novos; e em [15] a segunda, aceitando esse tópico, menciona um

subconjunto dos vestidos novos, ou seja, os vestidos novos importados. Já a resposta dada em [16] não funciona porque pressupõe que se está falando primariamente de vestidos importados, para especificar um subconjunto deles. Aqui se vê como o segundo elemento depois de *vestido* é interpretado como mais restritivo do que o primeiro².

Para concluir, voltamos ao caso de

[4] Animal mamífero

[5] *Mamífero animal

No caso de [5], a regra vai atribuir a *animal* um significado mais restritivo do que o de *mamífero* — ou seja, a regra atribui a *animal* o significado de “subconjunto dos mamíferos”. Para uma pessoa totalmente ignorante de zoologia, isso pode passar; mas se o ouvinte tem alguma noção do que é um mamífero, o significado atribuído pela regra se choca com seu conhecimento, e por isso o sintagma é rejeitado. Já [4] é aceito porque o significado atribuído a ele pela regra (mamíferos são um subconjunto dos animais) é correto, segundo nosso conhecimento do assunto.

Resumindo, essa explicação tem duas vantagens importantes sobre a anterior (que dividia os nominais em grupos I e II): primeiro, ela é mais econômica, porque evita a criação desses grupos. E, depois, ela se encaixa melhor com o que sabemos da interpretação dos nominais no sintagma e do nosso conhecimento do mundo. O princípio da sintaxe residual foi formulado a fim de favorecer esse tipo de análise sobre análises puramente formais, que só devem ser adotadas quando não houver outro remédio (é o caso da divisão dos verbos em conjugações, visto na seção anterior).

Implicitamente, sempre houve em gramática alguma coisa semelhante a esse princípio, embora sem explicitação ou aplicação sistemática. Por exemplo, sabemos que o objeto direto dos verbos *comer* e *beber* são bem diferentes: dizemos *comi a pipoca* e *bebi o guaraná*, mas não **comi o guaraná* e **bebi a pipoca*. Uma maneira de descrever esse fenômeno seria subclassificar os nominais do português segundo possam ser objeto de um ou outro desses verbos: em uma das classes teríamos *pipoca, bacalhau, pão* e *rapadura*, e na outra *guaraná, chope* e *água*. Essa é a solução formal para o problema.

² Essa regra é provavelmente um dos aspectos do princípio geral de que elementos com maior carga de informação tendem a ser colocados mais para o final da sentença.

Mas acontece que se pode ver claramente uma solução melhor: nominais que designam líquidos ocorrem com *beber*, e os que designam sólidos ocorrem com *comer*. Essa formulação nos libera de acrescentar à descrição do português uma subclassificação dos nominais só para dar conta das exigências dos verbos *comer* e *beber*. Além disso, ela lança mão de traços semânticos independentemente necessários, já que a especificação de *água* como o nome de um líquido, e *rapadura* como o nome de um sólido é inevitável: isso faz parte do significado básico dessas palavras, de todo modo. Como aqui a diferença de economia é muito grande, nunca ninguém pensou em aplicar a solução formal ao problema.

Para resumir: em muitos casos é possível descrever os fatos formais como sendo consequência, de alguma maneira, do significado; e outros casos, isso não é possível — por exemplo, as conjugações verbais, ou a exigência de preposição com verbos como *gostar*, *confiar* etc. Não é possível prever, antes de examinar um fenômeno, qual será a melhor solução; é necessário examinar caso por caso. Sempre que ambas as formulações forem possíveis, a formulação semântica (ou, mais precisamente, simbólica) deve ser preferida, porque esta se baseia em fatos que precisam, inevitavelmente, ser incluídos como parte de nosso conhecimento da língua.

RELAÇÕES SIMBÓLICAS NA MORFOLOGIA E NA SINTAXE

Vamos examinar brevemente algumas outras formas do português, procurando relacioná-las com significados, de maneira a poder formular regras de alcance simbólico. Devo avisar que nenhum dos problemas aqui apresentados vai chegar a receber uma solução completa — nosso objetivo é ilustrar o tipo de raciocínio que fundamenta a pesquisa em lingüística.

10.1. NOMEAÇÃO E FALAÇÃO

Nosso primeiro exemplo são as palavras *nomeação*, *falação*, *viajação* e *começão*. A primeira dessas palavras está em todos os dicionários, definida como “ato ou efeito de nomear”. Vamos deixá-la de lado por um momento, para examinar as outras três.

Falação, tal como a utilizamos coloquialmente, se refere à ação de falar, mas com a idéia de que se trata de algo muito intenso, freqüente ou mesmo excessivo:

[1] Não é possível trabalhar com essa falação.

Comparando *falação* com as outras palavras do grupo, podemos levantar a hipótese de que se compõe de duas partes: *fala-* mais *-ção*. A primeira parte ocorre no verbo *falar*, e deve ser a responsável pela identificação do tipo de ação referida (a ação de “falar”). Estabelecemos então a correspondência simbólica entre a forma *fala-* e o conceito de “falar”. Dessa maneira, a segunda parte da palavra, *-ção*, deve estar relacionada simbolicamente com o conceito de “ação freqüente ou excessiva”. Temos, pois, uma análise para a palavra *falação*, que podemos representar assim:

Formas:	fala-	-ção
Significados:	"falar"	"ação freqüente ou excessiva"

Mas não basta analisar uma palavra, evidentemente. Queremos chegar a uma regra de valor geral, que possa ser aplicada ao maior número possível de casos. Podemos usar a análise de *falação* para analisar *viajação* e *começão*: examinando a forma dessas palavras, levantamos a hipótese de que se relacionam com os verbos *viajar* e *comer*, respectivamente, da mesma maneira que *falação* se relaciona com *falar*. Formalmente isso faz sentido, pois em todos os casos basta retirar o *-r* final do verbo e acrescentar o sufixo *-ção*.

Vamos ver se a hipótese funciona. Se essas palavras são realmente paralelas a *falação*, elas devem designar ações freqüentes ou excessivas. Para verificar isso, podemos colocá-las em contexto, para ver que efeito produzem:

-
- [2] A Marina não suporta mais a *viajação* do marido.
 [3] Você engordou por causa dessa *começão* de doces.
-

Vemos que o efeito é o mesmo de *falação*: tanto *viajação* quanto *começão* denotam ação freqüente ou excessiva. Ou seja, as três palavras examinadas até agora têm uma composição formal semelhante (infinitivo do verbo, menos *-r*, mais *-ção*) e um significado semelhante (ação freqüente ou excessiva). Isso significa que, pelo menos como hipótese, temos o direito de formular uma regra (que vamos chamar, para efeito de referência, "Regra de *-ção*");

Regra de *-ção*: O sufixo *-ção*, acrescentado a um verbo no infinitivo (menos o *-r* final) significa "ação freqüente ou excessiva".¹

De posse dessa regra, podemos agora examinar o léxico do português à procura de outros casos. Esses são muito numerosos, como atestam as frases seguintes:

-
- [4] A menina vai endoidar com essa *fazeção* de prova.
 [5] Vamos parar com a *tocação* de bateria de noite?
 [6] É tanta *lavação*, que ele vai gastar o carro.
-

Em todos esses casos, temos o sufixo *-ção* acrescentado ao verbo sem *-r*, com o significado esperado. Ou seja, até agora nossa regra se confirmou brilhantemente. Agora precisamos procurar problemas.

¹ Ao apresentar o fenômeno aqui como uma regra, lembro que meu objetivo é descritivo. Naturalmente, poderá ser mais conveniente representar esse fenômeno como resultado da junção de dois itens léxicos, com conseqüente amalgamamento de seus significados.

Espera aí! Procurar problemas?

Exatamente. Essa é a melhor maneira de verificar a adequação de uma hipótese (e uma regra, lembro; é uma hipótese): colocá-la à prova em situações difíceis. Uma boa regra é a que resiste às provas. E uma regra que não resiste ainda assim pode ser útil de várias maneiras: pode servir de base para uma nova regra, nascida da reformulação da regra anterior, ou pode ser simplesmente a indicação de que estamos no caminho errado. Tudo isso é informação útil para o pesquisador; afinal, ele está procurando a solução mais adequada, não a mais fácil.

No caso da regra que estamos tentando elaborar, temos um problema com a palavra *nomeação*: em uma frase como

-
- [7] A sua *nomeação* só vai sair em julho.
-

embora *nomeação* realmente exprima o ato de nomear, não há nenhuma sugestão de que se trate de uma ação "freqüente ou excessiva" — trata-se de nomear alguém, e só isso.

Uma saída imediata seria marcar *nomeação* como uma exceção à regra de *-ção*. Mas, se pensarmos melhor, não deve ser só isso. Essa palavra pode ocorrer em outros contextos, como em

-
- [8] O partido tem que acabar com essa *nomeação* louca de parentes e amigos.
-

Aqui já aparece o ingrediente semântico "ação freqüente ou excessiva", o que mostra que o verbo *nomear* se comporta como *falar*, *fazer*, *tocar* e *gastar*, sofrendo normalmente o efeito da regra de *-ção*.

Mas como *nomeação* também aparece designando simplesmente o ato de nomear, sem mais especificações (como em [7]), temos que reconhecer que *nomear* é diferente de, por exemplo, *falar*, porque *falação* sempre tem o significado de "ação freqüente ou excessiva". Talvez os verbos se dividam em dois grupos, os que aceitam dois significados quando associados ao sufixo *-ção* (por exemplo, *nomear*), e os que só admitem um significado (como *falar*).

Mas quanto à regra de *-ção*, ela se sai bastante bem desse teste, porque afinal de contas até mesmo *nomear* sofre o seu efeito — a diferença é que *nomear* também sofre o efeito de algum outro fator, talvez outra regra,

quem sabe? Não vamos investigar isso mais aqui, porque o que nos interessa é submeter a regra de *-ção* a testes, e pelo primeiro ela passou.

Há pelo menos mais uma dificuldade com essa regra: é que com alguns verbos a associação com o sufixo *-ção* simplesmente dá mau resultado. Por exemplo, de *ser* não se pode formar **seção* (*seção* existe, mas não tem nada a ver com a “ação [?] de ser”); o mesmo com *estar*, porque *estação* certamente não é a “ação freqüente e excessiva de estar”. Portanto, parece que a regra de *-ção* tem, afinal de contas, exceções. Podemos simplesmente apontar essas exceções e dar o assunto por encerrado. Marcaríamos cada verbo da língua como podendo ou não sofrer a ação da regra de *-ção*. Isso equivaleria a considerar o comportamento excepcional de *ser* e *estar* como uma característica formal, sem correlato semântico: esses verbos escapariam à regra de *-ção*, sem que houvesse nenhuma razão semântica por trás desse fato. Uma exceção, em suma.

Mas agora eu gostaria de lembrar o princípio prático de análise mencionado no capítulo 9: uma forma não tem significado até que se consiga descobrir um. Ora, isso se aplica igualmente a regras gramaticais, assim como a suas exceções. Ou seja, antes de marcar alguns verbos como exceções puras e simples, devemos pelo menos tentar relacionar essa pretensa exceção com algum fator semântico (ou seja, aplicamos o princípio da sintaxe residual).

No caso, acho que é possível encontrar uma: o fato de que *ser* e *estar* não podem exprimir nenhum tipo de ação. Afinal de contas, *ser rico* ou *estar cansado* são estados, e não ações — ao contrário de *nomear*, *falar*, *fazer*, *lavar* e *tocar*. E como o sufixo *-ção* acrescenta exatamente o ingrediente semântico “ação freqüente ou excessiva”, deve ficar claro por que *ser* e *estar* não o aceitam: não existe a ação de “ser”, nem a de “estar”. Não precisamos nem sequer modificar a regra de *-ção*. Basta deixar que funcione normalmente: quando for se aplicar ao verbo *ser*, a regra vai dar como resultado a forma *seção*, que significa “ação freqüente ou excessiva de ser”. Mas aí vai surgir um problema: esse significado é impossível, porque a “ação de ser” não existe. Assim, a forma *seção* (nesse significado) acaba nem sendo formada, porque não significa nada de útil.

Descobrimos uma razão semântica para a não-aplicação da regra de *-ção* aos verbos *ser* e *estar* — é como se disséssemos que a regra não se aplica nesses casos, mas não é culpa dela. Ela pode ficar formulada do mesmo jeito que está, e em certos casos não vai dar resultado por razões de má-formação semântica. Casos assim são chamados casos de filtragem semântica, e há muitos deles na língua.

Vamos deixar, então, a regra de *-ção* como foi formulada acima; e vamos considerá-la como uma parte da gramática do português, até prova em contrário². A discussão feita nesta seção deve ser suficiente para mostrar como as formas e os significados interagem para definir mensagens possíveis e impossíveis em português.

10.2. BONS RELÓGIOS E RELÓGIOS JAPONESES

Agora vamos passar a outro fenômeno que tem a ver com a ordem das palavras. Sabemos que é possível dizer tanto [9] quanto [10]:

[9] Um relógio bom.

[10] Um bom relógio.

Deixando de lado o artigo *um*, que tem posição fixa, vemos que o termo que exprime o nome da coisa (*relógio*) e o que exprime a qualidade (*bom*) podem ocorrer em qualquer ordem. O termo que designa uma coisa é dito referencial; o que exprime a qualidade é o restritivo. A função mais importante do restritivo é restringir o tipo de ser de que se está falando. No caso acima, estamos sempre falando de um *relógio* (referencial), mas não de qualquer tipo de relógio: trata-se de um relógio *bom* (restritivo). Pelo que se observa em [9] e [10], podemos levantar a hipótese de que a ordem dos referenciais e restritivos é livre: podemos colocar qualquer deles em primeiro lugar.

Mas essa primeira hipótese é inadequada, por causa de sintagmas como

[11] Um relógio japonês

[12] * Um japonês relógio

Como se vê, nesses exemplos o referencial tem que vir obrigatoriamente antes do restritivo. Logo, temos que reformular nossa hipótese inicial de que a ordem é livre.

Uma possibilidade é que cada palavra que possa ser restritiva venha já de fábrica marcada como podendo ou não ocorrer antes do referencial.

² Pelo menos um detalhe vai ter que ser mudado: a referência a “ação” na formulação da regra. Isso porque há muitos verbos que não exprimem ação e que aceitam o sufixo; por exemplo, *morrer*, como em *começou uma morreção de gado na minha fazenda*. Talvez, em vez de “ação”, se deva falar de “ação ou evento”.

Ou seja, o falante, ao aprender a língua, precisa aprender caso por caso quais as possibilidades de ordenação de cada nominal: *bom* pode vir antes do referencial, mas *japonês* não pode. Em outras palavras, teríamos aqui uma marca formal, de modo que as palavras em questão seriam marcadas sem nenhuma razão semântica. Já vimos que há casos assim como, por exemplo, a divisão dos verbos em conjugações (capítulo 9). Se for isso, então não há muito mais o que fazer. Dizemos que as palavras do grupo de *bom* podem ocorrer nas duas posições, e as do grupo de *japonês* só podem ocorrer depois do referencial; seguem-se as listas.

Mas quem sabe essa diferença formal que observamos entre os nominais pode ser correlacionada com algum traço de significado? Seria mais econômico (pois o traço de significado já está lá), e também mais interessante, pois nos diria algo de novo sobre a língua, isto é, que existe uma correspondência entre determinado traço formal e determinado traço semântico de certas palavras.

Vamos pois explorar a semântica de *bom* em oposição a *japonês*. Uma diferença bastante evidente é que *bom* exprime uma qualidade, e pode ser modificado por um intensificador: *um relógio muito bom*, *um relógio não muito bom*, *um relógio melhor* (= *mais bom*) *do que o meu* etc. Por outro lado, *japonês* não exprime uma qualidade, mas antes a proveniência: [11] não diz quais são as qualidades do relógio, mas apenas onde ele foi fabricado. É também bem mais difícil intensificar esse nominal: ??*um relógio muito japonês*, ??*um relógio mais japonês do que o meu* etc. são bastante esquisitos. Quem sabe essa diferença semântica está na raiz da diferença formal observada?

Pelo menos uma coisa parece certa — nenhum restritivo com sentido de “proveniência” pode ocorrer antes do referencial:

-
- [13] Aquele vinho francês/*aquele francês vinho
 [14] Algumas armas americanas/*algumas americanas armas
 [15] O jabá pernambucano/*o pernambucano jabá
-

e assim por diante.

Isso quer dizer que já temos uma regra a formular, uma regra de caráter simbólico, porque relaciona forma com significado:

Regra de ordenação Q-R:

Um restritivo com significado de “proveniência” só pode ocorrer depois do referencial.

Pelo menos para esses restritivos, não precisamos mais elaborar uma lista; se o nominal exprime proveniência, só pode ser colocado depois.

A pesquisa mostrou que existem vários traços de significado que têm o mesmo efeito. Além do traço “proveniência”, temos os seguintes:

“Autoria”:

[16] As comédias shakespearianas/*as shakespearianas comédias.
 (o qualificativo *shakespeariano* expressa o autor das comédias).

“Agente”:

[17] A iniciativa chinesa/*a chinesa iniciativa
 (*chinesa* expressa o agente da iniciativa).

“Posse”:

[18] O palácio presidencial/*o presidencial palácio
 (*presidencial* expressa o possuidor, ou melhor, ocupante, do palácio).

Assim, podemos concluir que pelo menos boa parte do fenômeno é simbólico, ou seja, deve ser descrito em termos de relação forma/significado. A regra de ordenação Q-R é adequada, mas precisa ser completada, da seguinte maneira:

Regra de ordenação Q-R (revista):

Um restritivo com significado de “proveniência”, “autoria”, “agente” ou “posse” só pode ocorrer depois do referencial.

As condições de posicionamento do restritivo em relação ao referencial são bastante complexas, e foram formuladas em Perini *et al.* (1996); ali se mostra que, além dos fatores semânticos, há também fatores puramente formais em jogo. Sem entrar em detalhes, vou dar dois exemplos bem simples:

³ “Q” para qualificativo, “R” para referencial.

[19] Alta burguesia/*burguesia alta

Não existe, que se saiba, nenhuma razão semântica para que *alta* tenha que vir antes de *burguesia* nesse sintagma. Assim, marca-se a seqüência como um caso excepcional — ainda mais que *alto* pode vir posposto em outros casos: *telhado alto*, *mulher alta*.

Outro exemplo é:

[20] Um professor ruim

[21] *Um ruim professor

A palavra *ruim*, de significado qualificativo, só pode aparecer depois do referencial. Isso não tem nada a ver com seu significado, tanto é assim que *mau*, que é praticamente um sinônimo de *ruim*, pode aparecer nas duas posições (em geral, com alguma mudança de significado):

[22] Um professor mau

[23] Um mau professor

Aqui não temos saída: é preciso marcar *ruim* como restrito à posição posposta, e isso sem nenhuma correlação semântica.

A discussão acima tentou dar uma idéia do tipo de fator que está envolvido nessas condições de posicionamento, e de como os fatores semânticos interferem na construção formal das expressões da língua.

2ª PARTE:

INSTRUMENTOS DE ANÁLISE

NÍVEIS E UNIDADES

11.1. PALAVRA, LEXEMA, SINTAGMA

Quando um lingüista examina seus dados, ele de certo modo sabe o que procurar: a teoria lingüística lhe fornece muitos pontos de referência, de maneira que ele não precisa começar do zero. Por exemplo, ele sabe de antemão que as palavras são unidades relevantes para a análise; que as palavras se agrupam em unidades maiores (chamadas sintagmas); que os sintagmas se agrupam em orações etc. Neste capítulo, vou apresentar e definir as principais unidades normalmente utilizadas na análise gramatical, mostrando o modo como elas se articulam para formar unidades de nível cada vez mais alto.

A palavra nós todos conhecemos: ela é marcada na escrita através de espaços em branco antes e depois. Na fala não há espaços, e as palavras se pronunciam todas grudadas, o que às vezes apresenta dificuldades de análise que não podem ser discutidas neste livro. Para efeitos da presente exposição, vou partir da palavra (escrita), dando-a como conhecida.

Para efeitos de estudo da gramática, é importante distinguir palavra e lexema — duas noções bem diferentes, que são no entanto confundidas na linguagem usual. Uma palavra é uma forma individual, com uma representação fonológica ou gráfica única; um lexema é uma classe de palavras relacionadas de determinada maneira. Por exemplo, *flor* e *flores* são duas palavras, mas pertencem ao mesmo lexema.

Várias palavras pertencem a um lexema único quando elas se relacionam gramaticalmente de maneira muito sistemática e regular. *Flor* e *flores* são formas do mesmo lexema, porque é muito sistemático na língua um nominal ter singular e plural. Já *flor* e *florista* pertencem a lexemas diferentes porque essa relação (com o sufixo *-ista* significando “pessoa que trabalha com”) só ocorre em alguns casos, e de maneira pouco regular: temos

música e musicista, mas *sapato* e *sapateiro*, *mecânica* e *mecânico*, *medicina* e *médico*; e para *casa* não temos nada semelhante (*caseiro* não é uma “pessoa que trabalha com casas”, e **casista* não existe), nem para *parede*, *céu* ou *minhoca*. Mas (quase) todos os nominais têm plural, e todos os verbos têm infinitivo, presente do indicativo e primeira pessoa do plural — razão pela qual *beber*, *bebo* e *bebemos* são três palavras que pertencem ao mesmo lexema; mas *bebida* ou *beberrão* são lexemas independentes.

Uma palavra é formada de um ou mais morfemas: por exemplo, já sabemos que *falação* é formada de *fala-* mais *-ção*. *Fala-* é um morfema¹ que faz parte do verbo *falar*, e se encontra também em *falava*, *falasse* etc. O morfema *-ção* tem o significado de “ação freqüente ou excessiva”. Outro exemplo é *desmatamento*, onde *des-* dá uma idéia de “evento oposto”, *-mata-* é o mesmo morfema que se encontra em *mata* e *matagal*, e *-mento* significa “ação ou processo”. *Gata* pode ser dividido em *gat-*, que veicula a idéia básica de “gato”, e *-a*, que aí denota o sexo feminino. E a palavra *sempre* não pode ser dividida, sendo formada de um só morfema.

O estudo da formação das palavras se denomina morfologia, e não é abordado neste livro. Basta, por ora, ter uma idéia básica do que é um morfema, e de como os morfemas se agrupam para formar palavras. Aqui vamos partir do agrupamento das palavras em sintagmas. O estudo dos sintagmas, sua composição e sua estruturação em unidades maiores faz parte da sintaxe.

O sintagma é um constituinte (como definido no capítulo 3) menor do que uma oração, e composto de uma ou mais palavras. Assim, na frase

[1] Meus dois irmãos vão mudar para o Rio de Janeiro.

há algumas seqüências de palavras que nos dão a sensação de grupos naturais, e outros que não dão essa sensação. Assim, preferimos a segmentação [2] à segmentação [3]:

[2] [Meus dois irmãos] [vão mudar] [para o Rio de Janeiro].

[3] [Meus dois] [irmãos vão] [mudar para o] [Rio de] [Janeiro].

¹ Ou, mais exatamente, uma seqüência de dois morfemas, já que o *-a* final geralmente se analisa como “vogal temática”.

As partes de [2] parecem ter, cada uma, um significado claro, ao passo que a segmentação ilustrada em [3] nos dá a impressão de pedaços desconexos. Isso é porque em [2] a frase foi dividida em seus sintagmas (ou constituintes), ao passo que em [3] a divisão dos sintagmas não foi respeitada².

Pode-se testar a noção intuitiva de sintagma perguntando a qualquer falante do português se existe na língua a seqüência *trabalho o*. Ele provavelmente vai responder que não, que o que existe é *o trabalho*. No entanto, a seqüência *trabalho o* ocorre normalmente em português, como em

[4] Sem **trabalho o** dia custa muito a passar.

A seqüência *trabalho o* não é percebida como unidade porque não forma um sintagma; a segmentação de [4] é a seguinte:

[5] [Sem trabalho] [o dia] [custa muito a passar].

Como se vê, as duas palavras, *trabalho* e *o*, ficam em constituintes separados. Por isso sua seqüência não é sentida como uma unidade, e tendemos a dizer que ela não ocorre na língua. Esse caráter intuitivo da divisão em constituintes é muito importante para a análise, e se relaciona com o fato de que cada um deles tem um significado coeso.

Vamos examinar mais algumas propriedades dos sintagmas. Uma das mais importantes é que os sintagmas não são todos do mesmo tipo; outra é que eles se encaixam uns dentro de outros, como bonecas russas. Vamos examinar essas duas propriedades nas seções seguintes.

11.2. TIPOS DE SINTAGMAS

Os sintagmas, como as palavras, têm comportamento gramatical diferenciado. Alguns sintagmas se comportam (mais ou menos!) como os nominais, podendo ser semanticamente referenciais, e ocupando posições típicas de nominais na sentença. Por exemplo, o sintagma *meu vizinho da frente* pode ocorrer na oração nas mesmas posições que *Fred*:

[6] [Meu vizinho da frente] estuda medicina.

[7] [Fred] estuda medicina.

² Em geral, entende-se por sintagma um constituinte formado de pelo menos uma palavra, mas menor do que uma oração.

Para ser mais exato, não se trata de substituir o sintagma *meu vizinho da frente* pela palavra *Fred*, mas antes pelo sintagma composto apenas da palavra *Fred*. Esse tipo de sintagma se denomina sintagma nominal (tradicionalmente abreviado SN); em termos de gramática tradicional, um sintagma nominal é tudo aquilo que pode ser sujeito, objeto direto ou complemento de uma preposição³. Nas frases acima, *meu vizinho da frente* e *Fred* são sujeitos de suas orações; as outras duas funções são ilustradas nas frases seguintes:

-
- [8] Vou procurar [meu vizinho da frente] (**objeto direto**).
- [9] Vou procurar [Fred] (**objeto direto**).
- [10] Elza está saindo com [meu vizinho da frente] (**compl. da prep. com**).
- [11] Elza está saindo com [Fred] (**compl. da prep. com**).
- [12] O lote foi comprado por [meu vizinho da frente] (**compl. da prep. por**).
- [13] O lote foi comprado por [Fred] (**compl. da prep. por**).
-

Um sintagma nominal pode ter composições bastante variadas, como mostram os exemplos abaixo:

[Fred]
[meu vizinho da frente]
[o gato]
[o livro que você comprou]

Essas são apenas quatro das muitas possibilidades de estruturação de um sintagma nominal. É fácil verificar que todos esses sintagmas podem ocorrer nas três funções citadas, e todos eles podem ser usados para se referir a coisas⁴.

Já um sintagma adjetivo tem propriedades tradicionalmente atribuídas aos chamados “adjetivos”: semanticamente é restritivo, e sintaticamente é o modificador de um nominal⁵. Nas frases abaixo, as seqüências entre chaves são sintagmas adjetivos:

³ Como vimos, o sintagma nominal também pode ser predicativo, mas esta não é uma função exclusiva dos sintagmas nominais.

⁴ Relembro que “coisa”, no sentido aqui utilizado, inclui seres vivos, inclusive pessoas.

⁵ “Modificador” é uma das diversas funções englobadas pela gramática tradicional sob o rótulo de “adjunto adnominal” (ver 13.4). Há também sintagmas adjetivos não restritivos, mas é melhor deixá-los de lado por ora.

-
- [14] Guarde aqui o livro [azul].
- [15] Guarde aqui o livro [do Pedrinho].
- [16] Guarde aqui o livro [da capa azul].
- [17] Guarde aqui o livro [que você comprou].
-

Note-se que os sintagmas se encaixam uns dentro dos outros; por exemplo, em [17], o sintagma adjetivo [*que você comprou*] está inserido dentro do sintagma nominal {*o livro [que você comprou]*}. Essa possibilidade de inserir sintagmas dentro de sintagmas é examinada na próxima seção.

Paralelamente a esses dois tipos, há outros tipos, como o sintagma adverbial. Em princípio, seria fácil defini-lo: seria o sintagma que se comporta como um advérbio. Assim, se a palavra *sempre* é um advérbio, a seqüência *quase todo dia* seria um sintagma adverbial, pois ocupa as mesmas posições na oração e tem significado análogo (“tempo”):

-
- [18] Meu carro [sempre] dá problema.
- [19] Meu carro [quase todo dia] dá problema.
-

Isso é bastante claro. O problema com os sintagmas adverbiais é que ninguém sabe ao certo o que é um advérbio — tudo indica que essa classe tradicional de palavras encerra vários tipos muito diferentes de itens, sendo provavelmente um conjunto de classes, não apenas uma classe⁶.

Há ainda outros tipos, como o sintagma verbal, que corresponde ao “predicado” da gramática tradicional:

-
- [20] Carminha [viaja para a Espanha amanhã].
-

Finalmente, é bom observar que a oração é também um tipo de sintagma, pois é um dos constituintes da sentença (do período). O que se disser do sintagma em geral vale, portanto, também para as orações.

A noção de sintagma é básica em todas as teorias lingüísticas. Na gramática tradicional, ela é usada mas não explicitada, de maneira que soa

⁶ Muitos sintagmas adverbiais e também muitos sintagmas adjetivos são compostos de uma preposição mais um SN: *do Toninho* (sadj), *em Belém* (sadv). Esses sintagmas recebem muitas vezes a designação de sintagma preposicionado, que não é a rigor uma classe definida funcionalmente, mas apenas uma expressão conveniente que se usa quando a classe do sintagma não é clara, ou quando não importa para a discussão do momento.

como novidade para quem começa a estudar lingüística. Para ver como a divisão em sintagmas é essencial na análise tradicional, basta observar que, em uma frase como

[21] Meu carro quase todo dia enguiça no centro da cidade.

a análise sintática tradicional atribui funções a *meu carro* (sujeito), *quase todo dia* (adjunto adverbial), *quase todo dia enguiça no centro da cidade* (predicado) etc. Mas as seqüências *carro quase* e *dia enguiça* não têm função nenhuma na oração — justamente porque não constituem sintagmas. Nesse ponto, todas as teorias estão de acordo: somente os sintagmas (e as palavras) são unidades sintáticas, podem pertencer a classes e ter funções, e somente essas seqüências recebem interpretação semântica⁷. As outras seqüências, como *carro quase*, *dia enguiça*, *centro da* etc., não têm nenhum *status* na gramática e inclusive, como vimos, tendem a ser ignoradas pelos falantes. Ou, em outras palavras, as expressões lingüísticas não são (apenas) seqüências de palavras, e não podem ser descritas como se fossem; elas são seqüências de constituintes, estruturados da maneira que veremos na próxima seção.

Quando dizemos que há vários tipos de sintagmas, estamos dizendo que eles se distribuem em várias classes, exatamente como as palavras. As classes de palavras (verbos, preposições, pronomes etc.) são apenas um caso especial das classes de formas, que incluem os sintagmas. Assim, os sintagmas de uma classe (como as palavras de uma classe) têm traços importantes em comum — por exemplo, ocorrem em posições semelhantes na oração e seu significado tem alguma coisa em comum. No caso dos sintagmas nominais, todos eles podem ter as três funções mencionadas acima (sujeito, objeto direto, complemento de preposição), e todos podem ser usados como referenciais, designando um ser ou uma coisa qualquer. Como “ser referencial” é também uma função (semântica), podemos dizer, em geral, que uma forma se classifica de acordo com seu potencial funcional, isto é, o conjunto das funções que cada um pode desempenhar. Nos capítulos 15 e 16, examinamos com mais cuidado o porquê e o como da classificação das formas, com uma definição mais cuidadosa de “potencial funcional”.

⁷ Esta última afirmação sofre certas restrições, mas sua recíproca é sempre verdadeira: um constituinte sempre corresponde a um significado coeso.

11.3. ASSOCIAÇÃO DOS SINTAGMAS

Os sintagmas se associam para compor estruturas maiores. Assim, podemos juntar um sintagma nominal a um sintagma verbal para formar uma oração:

[22] [Meu carro] [enguiça todo domingo].

sint. nom.

sint. verbal

A composição típica (não única) de uma oração é justamente essa: um sintagma nominal mais um sintagma verbal.

Mas, além de ocorrerem lado a lado, os sintagmas também ocorrem uns dentro dos outros. Esse fato é particularmente importante, mas não é representado com clareza na gramática tradicional, e tende a ser esquecido. Por exemplo, a seqüência marcada em [23] é um sintagma adverbial:

[23] Sérgio está morando [com a minha prima].

No entanto, dentro do sintagma adverbial *com a minha prima* temos um sintagma nominal, *a minha prima* (verifique que essa seqüência tem todas as propriedades dos sintagmas nominais). Assim, a estruturação pode ser mostrada assim:

{ com [a minha prima] }.

Usei as chaves, { }, para marcar o sintagma adverbial, e os colchetes, [], para marcar o sintagma nominal que está dentro desse sintagma adverbial.

Esse processo de encaixamento pode chegar a um grau bastante alto de complexidade. Por exemplo, a seqüência marcada em [24]

[24] Estou trabalhando [com um amigo do pai de Bia].

se analisa como o resultado do encaixamento de seis sintagmas: *com um amigo do pai de Bia* é um sintagma adverbial; este contém o sintagma nominal *um amigo do pai de Bia*; este, por sua vez, contém o sintagma adjetivo *do pai de Bia*, que contém o sintagma nominal *o pai de Bia*; este contém o sintagma adjetivo *de Bia*, que contém o sintagma nominal *Bia*:

[25] { com [um amigo { d (o pai {de [Bia]}) }] } }.

Vemos, portanto, que a linguagem tem a propriedade de, inclusive, encaixar sintagmas dentro de outros sintagmas da mesma classe. Essa é uma propriedade fundamental da linguagem humana, e se denomina recursividade. Ela permite que a quantidade de enunciados possível em uma língua seja, em princípio, ilimitada (porque sempre se pode colocar mais um sintagma dentro dos que já temos).

Esse encaixamento de unidades dentro de outras tem conseqüências curiosas para a segmentação: por exemplo, dentro de que sintagma(s) está a palavra *Bia* em [24]? Sabemos que está, para começar, no próprio sintagma [*Bia*], que é composto de apenas uma palavra. Mas *Bia* também está dentro do sintagma [*de Bia*]; e do sintagma [*o pai de Bia*]; está igualmente dentro de [*do pai de Bia*] e de [*um amigo do pai de Bia*]; e, finalmente, está dentro de [*com um amigo do pai de Bia*]⁸. Se eliminarmos a palavra *Bia* de qualquer desses sintagmas, ele fica incompleto: a seqüência *o pai de* não é um sintagma (como aliás nos diz a nossa intuição). Ou seja, uma palavra ou qualquer outra forma pode perfeitamente fazer parte de diversos sintagmas — não há nada de incoerente nisso (pode-se comparar essa situação à da cidade de Brasília, que fica dentro do Distrito Federal, e também dentro do Brasil, dentro da América do Sul etc.). A lista dos sintagmas nominais de [25] é a seguinte:

[*um amigo do pai de Bia*]

[*o pai de Bia*]

[*Bia*]

11.4. SINTAGMAS: UNIDADES DE FORMA E DE SIGNIFICADO

Os sintagmas têm coesão semântica e formal. Semântica porque nos dão a impressão de alguma coisa que “faz sentido”, e essa impressão pode ser explicitada com certa clareza. E formal porque, em geral, podem ocorrer em determinadas posições sintáticas bem definidas, com função específica. Por exemplo, sabemos que *a minha prima* é um sintagma nominal. Essa mesma seqüência de palavras se encontra em diversas posições na frase, como:

⁸ Além disso tudo, está também dentro da oração completa.

[26] [A minha prima] telefonou ontem.

[27] Você vai encontrar [a minha prima]?

[28] A Nádia é que é [a minha prima].

[29] O carro d[a minha prima] é aquele ali.

É como se esse grupo de palavras fosse carregado daqui para ali, sem se desintegrar. Agora, se pegarmos em [29] a seqüência *carro da minha*, não veremos o mesmo tipo de coesão: fica difícil utilizar essa seqüência (sem *o* antes e *prima* depois, claro) em diversas posições em diversas sentenças. É que *carro da minha* não forma um sintagma. Paralelamente, *a minha prima* tem um perfil semântico bem definido. Essa seqüência se refere a um ser (uma pessoa, no caso), ao contrário de *carro da minha*, que parece antes a junção de dois pedaços que não se ligam.

Em resumo: uma frase se estrutura em unidades, isto é, palavras, sintagmas, orações, que podem ocorrer umas ao lado das outras e inclusive umas dentro de outras. Essas unidades têm valor formal (são peças utilizadas na montagem das seqüências) e valor semântico (são parte da informação passada pelo emissor ao receptor, e que permite a este último construir o significado final das sentenças).

11.5. ESTRUTURA SINTAGMÁTICA E A MONTAGEM DO SIGNIFICADO

Podemos agora voltar à frase vista no capítulo 3, a saber,

[30] A casa de Sandrinha é verde.

Ela se segmenta nos seguintes sintagmas:

[31] { [A casa (de Sandrinha)] [é verde] }.

Essa segmentação faz sentido em termos da descrição formal, porque todos os constituintes estabelecidos têm propriedades gramaticais coerentes. Assim, *a casa de Sandrinha* ocorre em posições onde tipicamente cabem outros sintagmas nominais como *a folha*, *esse livro*, *papagaio* ou *aquele carro que você queria comprar*. *De Sandrinha* pode ser substituído por sintagmas adjetivos como *nova*, *que fica em cima do morro* ou *com muitas janelas*. *É verde* corresponde em ocorrência a sintagmas verbais como *pegou fogo*, *fica na esquina* e *tem oito quartos*.

Mas não é só isso: a segmentação é fundamental para a interpretação semântica, isto é, para o processo pelo qual o receptor deriva o significado da estrutura. Assim, *a casa de Sandrinha* é a maneira usada nessa frase para trazer à mente do receptor uma entidade (uma casa determinada); *é verde* é usado para atribuir a essa entidade uma qualidade (“verde”); e *de Sandrinha* restringe a casa mencionada, de maneira que ficamos sabendo que não é toda e qualquer casa que é verde, nem alguma casa não especificada, mas uma casa identificável como sendo a de Sandrinha.

Por outro lado, a frase contém a seqüência *Sandrinha é verde*, mas, como essa seqüência não é um sintagma, não recebe interpretação semântica. Conseqüentemente, o significado “Sandrinha é verde” não faz parte do significado dessa frase: o receptor não entende que Sandrinha é verde, e na verdade não fica sabendo nada sobre as eventuais qualidades de Sandrinha. Isso mostra como a estruturação em sintagmas é essencial para a compreensão.

11.6. PROPRIEDADES DOS CONSTITUINTES

Os constituintes são básicos para a análise gramatical, e é importante saber identificá-los com segurança. Sem eles, nenhuma análise gramatical faz sentido. Por isso, é bom conhecer suas propriedades gramaticais (formais e semânticas), porque é em grande parte com elas que lidamos ao construir descrições gramaticais. A seguir dou algumas propriedades dos constituintes que os distinguem de seqüências que não formam constituintes.

11.6.1. Constituintes, uns dentro dos outros

A primeira, e a mais importante, dessas propriedades é a possibilidade de ocorrência de constituintes dentro de constituintes. Vimos alguns exemplos acima; agora vamos retomar o tópico.

Por exemplo, sabemos que *a casa de Sandrinha* é um SN, e que contém outro SN, a saber, *Sandrinha*:

[32] [A casa de (Sandrinha)].

E sabemos que esse SN *a casa de Sandrinha* pode ser, por exemplo, sujeito de uma oração:

[33] [A casa de Sandrinha] é muito pequena.

Mas por que é que o sujeito de *é muito pequena é a casa de Sandrinha*, e não apenas *Sandrinha* (que, como vimos, também é um SN)?⁹ E, tomando o lado semântico da pergunta, por que é que sabemos que [33] afirma que *a casa é que é pequena*, e não a própria Sandrinha?

A resposta é que, quando temos um constituinte de determinada classe (digamos, um SN) dentro de outro da mesma classe, o maior é o único que vale para qualquer relação gramatical, seja formal seja semântica, com outros elementos da oração. Assim, para efeitos de determinação da função sintática de [32] na oração, só se leva em conta o sintagma maior (*a casa de Sandrinha*). O sintagma menor (*Sandrinha*) só funciona para efeitos de relações internas ao SN maior — assim, podemos dizer que *Sandrinha* é complemento da preposição *de*.

Outro exemplo é o sintagma adjetivo *de janelas de peroba*, na frase

[34] Comprei uma casa [de janelas (de peroba)].

Sabemos que *de janelas de peroba* é um sadj, e que *de peroba* também é. No entanto, o modificador de *casa é de janelas de peroba*; e, no que diz respeito ao significado, entende-se que a casa tem janelas de peroba, mas não que a casa é de peroba. o sadj *de peroba* só vale como modificador de *janelas*, não de *casa*, ou seja, seu âmbito de validade se limita ao interior do sintagma a que pertence.

Esses exemplos ilustram o caráter hierárquico da estrutura de constituintes: eles se encaixam uns dentro dos outros, de tal maneira que o maior tem as propriedades de sua classe dentro do domínio maior, e o menor (ou os menores, porque pode haver vários) funciona apenas dentro de seu domínio mais restrito.

11.6.2. Constituintes e o contexto sintático

Uma observação que deve ser feita é que uma seqüência, tirada de seu contexto sintático, não forma, por si só, um constituinte. Assim, a seqüência

[35] Camilo e Fabrícia

⁹ Relembro que um SN pode ser composto de apenas uma palavra, no caso, *Sandrinha*.

pode ou não ser um constituinte, dependendo de onde estão inseridos. Essa seqüência é um constituinte na frase

[36] Camilo e Fabrícia têm um rancho de avestruzes.

mas não em

[37] Convidei Camilo e Fabrícia veio também.

O que acontece com a seqüência *Camilo e Fabrícia* em [36] é que a conjunção *e* coordena os dois SNS *Camilo* e *Fabrícia* para formar um SN maior, que é o sujeito de *têm um rancho de avestruzes*. Já em [37], *Camilo* pertence à primeira estrutura (a oração *convidei Camilo*), *Fabrícia* pertence à segunda (*Fabrícia veio também*), e a conjunção *e* está coordenando as duas orações, e não *Camilo* mais *Fabrícia*.

Como se vê, uma estrutura gramatical é muito mais do que apenas uma seqüência de elementos: é, entre outras coisas, uma hierarquia de constituintes. “Constituinte” é uma noção estrutural, e só faz sentido dentro de uma estrutura.

FUNÇÕES SINTÁTICAS

12.1. FUNÇÕES¹

Uma unidade gramatical — por exemplo, uma palavra — pode ocorrer em relações diferentes com as outras unidades da mesma frase. Essas relações podem se basear em diversos fatores: pode ser uma relação de ordem (uma unidade vem antes ou depois de outra); de presença de certos morfemas, como no caso da concordância, em que dois elementos têm os morfemas de feminino e plural: *casas brancas*; ou de significado, como quando uma palavra expressa uma qualidade atribuída a alguma coisa: *casa branca*, isto é, “casa que tem a qualidade de ser branca”, onde a coisa é designada por *casa* e a qualidade é expressa por *branca*. Em todos esses casos, diz-se que cada uma das palavras envolvidas tem uma função diferente. Neste capítulo e no próximo, vamos considerar algumas das funções mais importantes de caráter formal; no capítulo 14 trataremos das funções semânticas.

Devo advertir, antes, que nestes capítulos não apresento uma análise completa da estrutura da oração. Este livro não é uma gramática portuguesa, mas uma introdução ao pensamento gramatical, e por isso me concentro nos argumentos utilizados para justificar as diferentes funções, sem a preocupação de listá-las na íntegra.

12.2. RELAÇÕES DE ORDEM

12.2.1. Sujeito e objeto direto

Em uma frase como

[1] Os gatos arranharam Toninho.

¹ As funções sintáticas são apresentadas neste capítulo segundo uma concepção bastante tradicional. Há boas razões para contestar as definições de diversas funções, mas não podem ser incluídas em um texto introdutório.

podemos distinguir três constituintes principais: *os gatos*, *arranharam* e *Toninho*. *Os gatos* e *Toninho* são semelhantes porque são ambos sintagmas nominais. Eles aparecem em determinada ordem: *os gatos* logo antes do verbo, *Toninho* depois. Essa diferença de ordenação já pode ser considerada uma diferença de função — isto é, *os gatos* e *Toninho* têm funções diferentes porque aparecem em lugares diferentes na frase.

Mas a diferença não pára aí. Examinando agora o verbo, vemos que ele está em determinada forma (chamada “terceira pessoa do plural”), e que essa forma se relaciona com um dos SNs da frase: com *os gatos*, e não com *Toninho*. Tanto é assim que se mudarmos o primeiro SN, poderemos ter que mudar também a forma do verbo:

[2] O gato arranhou Toninho.

[3] Eu arranhei Toninho.

[4] Nós arranhamos Toninho.

Já se mudarmos o segundo SN, não observaremos nenhuma consequência na forma do verbo; compare-se [2] com

[5] O gato arranhou meus sobrinhos.

ou então compare-se [1]

[1] Os gatos arranharam Toninho.

com

[6] Toninho arranhou os gatos.

Temos aqui outro tipo de relação formal, que confirma a idéia anterior de que *os gatos* e *Toninho* em [1] têm funções sintáticas diferentes. O SN que (nesse caso) vem logo antes do verbo, e com o qual o verbo concorda, é chamado sujeito; o outro SN se chama objeto direto.

O leitor certamente já conhece essas funções. Mas devo chamar sua atenção para o seguinte: aqui estamos definindo “sujeito” e “objeto direto” (assim como as outras funções sintáticas) em termos puramente formais. O sujeito de uma frase não é nem “o ser que pratica a ação”, nem “o

elemento do qual se afirma alguma coisa”. É o sintagma que aparece em determinada posição na frase (nem sempre logo antes do verbo, como veremos) e com o qual o verbo concorda.

É verdade que nas frases vistas o sujeito exprime sempre o ser que pratica a ação (o que arranha, não o que é arranhado). Mas isso nem sempre acontece; assim, em

[7] Toninho apanhou dos vizinhos.

quem pratica a ação são os vizinhos, mas o sujeito é *Toninho*: sabemos isso por causa da posição de *Toninho* na frase, e também por causa da concordância do verbo. E em

[8] Toninho é magro.

não se exprime ação nenhuma; e no entanto temos um sujeito, que é novamente *Toninho*. Em outras palavras, a noção de “sujeito”, assim como as funções sintáticas em geral, não se identificam com funções semânticas como “agente”². Apesar de haver em geral alguma correlação entre o significado e a função sintática de um sintagma, é necessário separar sempre bem nitidamente esses dois tipos de função. “Sujeito” é uma função formal, e tem a ver com a ordem das palavras e com a concordância; “agente” é uma função semântica, e tem a ver com o papel que um ser desempenha dentro do evento descrito por uma sentença. “Sujeito” é uma das funções que podem ser desempenhadas por um sintagma nominal; “agente” é um papel desempenhado por uma pessoa, um animal etc.

Voltamos aqui a um ponto importante: é preciso separar bem nitidamente a descrição formal da descrição semântica, antes de estabelecer suas relações — que existem, e são importantes para a análise (volte ao capítulo 3 para lembrar por quê).

A definição de sujeito não é uma coisa tranqüila. O sujeito às vezes é definido apenas como o elemento com o qual o verbo concorda, e isso fica claro em frases como

[1] Os gatos arranharam Toninho.

² As funções semânticas serão estudadas no capítulo 14.

Em [1], é claro que há uma relação de concordância entre *os gatos e arranharam*, e que *arranharam* não concorda com *Toninho*. Seria essa, então, a base para que o ouvinte identifique, em [1], *os gatos* como o sujeito (e portanto, nesse caso, o agente) de *arranhar*.

Mas na frase

[2] O gato arranhou Toninho.

a relação de concordância não é clara, porque *arranhou* pode, por sua forma, estar concordando tanto com *o gato* quanto com *Toninho*. Apesar disso, analisa-se sempre *o gato* como sujeito em [2].

Seria de esperar, então, que houvesse alguma dúvida por parte dos usuários da língua quanto à identificação do sujeito de [2]. Por exemplo, já que com esse verbo o sujeito é o agente, poderíamos prever que as pessoas ficassem em dúvida quanto a quem praticou a ação. No entanto, essa frase não é ambígua, e *o gato* é entendido sempre como agente da ação de arranhar. Logo, a concordância não pode ser a única indicação da função de sujeito: temos que levar em conta a ordem dos sintagmas em relação ao verbo, pelo menos em alguns casos.

O objeto direto seria simplesmente o SN com o qual o verbo não concorda, e isso funciona para [1]. Mas, novamente, há problemas com a frase [2], onde tudo indica que é a ordem dos sintagmas que dá a indicação do papel de paciente de *Toninho* — ou seja, o objeto direto talvez também tenha que ser definido em termos de ordem.

Esse dilema ainda não recebeu solução satisfatória. A situação é curiosa: na maioria dos casos, todos concordam quanto à identificação do objeto direto, mas não se consegue fundamentar essa identificação em definições precisas. É necessário encontrar essas definições, é claro; por ora, contudo, vamos nos contentar com a observação de que as funções de sujeito e objeto têm algo a ver com a ordem dos SNs relativamente ao verbo, e talvez com a relação de concordância com o verbo³.

³ Estou convencido de que a concordância nunca é o fator determinante das funções de sujeito e objeto direto; mas não posso entrar nesse assunto aqui, porque nos levaria muito longe do tema principal do capítulo.

12.2.2. Complicações

A definição de sujeito comporta certas complicações. Uma delas é a seguinte: embora o SN imediatamente antes do verbo seja analisado como sujeito, existe uma exceção, representada pelos pronomes oblíquos. Por isso, em

[9] Os gatos me arranharam.

o sujeito é ainda *os gatos*, apesar da presença de *me* (um sintagma nominal) logo antes do verbo. Essa é uma característica especial dos pronomes oblíquos, que têm comportamento idiossincrático dentro da língua (já falamos do aspecto semântico dessa idiossincrasia em 8.2).

Outra complicação é que o sujeito (isto é, o elemento com o qual o verbo concorda) muitas vezes aparece depois do verbo. Isso só ocorre quando não há na frase outro SN que possa ser confundido com o sujeito; quando a oração comporta dois SNs, o sujeito precisa ficar em sua posição pré-verbal:

[10] Chegou o carteiro.

[11] O carteiro trouxe uns livros.

[12] *Trouxe o carteiro uns livros.

Em [10] *o carteiro* é o sujeito (e o verbo concorda com ele). Mas quando há outro SN (como em [11], onde aparece o objeto direto *uns livros*), o sujeito não pode aparecer depois do verbo, só antes, como mostram [11] e [12]. Esse fenômeno do sujeito posposto é bastante complexo, e não poderemos estudá-lo aqui⁴.

12.2.3. Orações sem sujeito

O sujeito, como vimos, é uma função formal, sintática, e se aplica a um sintagma nominal presente na oração. Assim, o sujeito de [1],

[1] Os gatos arranharam Toninho.

⁴ Frases como [12] ocorrem em certos tipos de texto, como em poesia tradicional. Na fala e na prosa normal, entretanto, não são possíveis.

é o SN *os gatos*, não o conceito de “gatos”, nem o conceito “elemento que pratica a ação”. O verbo concorda com esse SN, e não com o conceito de “gatos”.

Isso tem uma consequência importante: em uma frase como

[13] Comprei um aparelho de DVD.

não há sujeito, porque não há na oração nenhum SN com o qual o verbo concorde, nem nenhum SN posicionado logo antes do verbo. Essa frase tem um agente (“eu”, que pratiquei a ação de comprar); ela está falando de alguém (de mim, que sou um dos assuntos principais da frase). Mas, segundo nossa definição, o sujeito não é o elemento que pratica a ação (a que chamamos “agente”), nem o elemento do qual se afirma alguma coisa (que se chama “tópico”). Sujeito é o SN que se coloca logo antes do verbo (ou em outras posições determinadas, conforme o caso), e com o qual o verbo concorda. Ora, em [13] não se encontra nenhum SN com essas características, logo a oração não tem sujeito. Deixem-me repetir: [13] tem agente (“eu”, não a palavra *eu*, que não está ali, mas o conceito da pessoa que fala); tem tópico (novamente, “eu”). Mas não tem sujeito.

Já a frase

[14] Eu comprei um aparelho de DVD.

que é virtualmente sinônima de [13], tem sujeito: o SN *eu*, que está ali, logo antes do verbo, e com o qual o verbo concorda.

Essas duas sentenças, [13] e [14], têm o mesmo significado; e estão ambas falando de “mim”. Mas não são idênticas em tudo: existe entre elas uma diferença formal, representada no caso pela presença do sintagma nominal *eu*. Gramaticalmente, isso é expresso dizendo que [14] tem sujeito, e [13] não tem.

A gramática tradicional só admite a falta de sujeito em orações do tipo

[15] Choveu ontem.

e outras parecidas. Temos que concordar que não há sujeito aí, já que o verbo não concorda com nenhum outro sintagma da oração. Mas esse critério nos leva, necessariamente, a analisar como sem sujeito a oração

[13] Comprei um aparelho de DVD.

Por que, então, a gramática tradicional insiste em ver em [13] — mas não em [15] — um “sujeito oculto”? A resposta é que em [13] sabemos que houve alguém que praticou a ação de “comprar”, ao passo que em [15] se entende que ninguém praticou a ação de “chover” (que, aliás, não é ação nenhuma).

Mas essa argumentação tem uma falha: o que [13] tem, e [15] não tem, é agente, não sujeito; e, como já vimos, essas são noções nitidamente diferentes, e precisam ser mantidas separadas na análise. Temos que admitir (pois é um fato) que não há agente em [15], e há em [13]. A questão de haver sujeito é outra, e, se admitirmos que o sujeito é o SN que precede o verbo, ou com o qual o verbo concorda, temos que aceitar que nem [15] nem [13] têm sujeito.

12.2.4. Predicativo do sujeito

O predicativo do sujeito aparece em frases como

[16] A casa está meio desarrumada.

O sintagma *meio desarrumada* aparece depois do verbo, e nesse lugar às vezes pode aparecer um SN:

[17] A casa é um palácio.

No entanto, mesmo em [17] não se analisa esse termo como objeto direto, mas como predicativo do sujeito⁵. Há algumas diferenças que distinguem o predicativo do sujeito do objeto direto. No essencial, essas diferenças mostram que existe relação formal e semântica entre o predicativo do sujeito e o sujeito, ao passo que não há relação direta de nenhuma natureza entre objeto direto e sujeito.

Essa relação aparece, às vezes, sob a forma da concordância: em [16], *meio desarrumada* concorda com *a casa* (feminino, singular). Note-se que nunca há concordância entre sujeito e objeto direto.

⁵ Em um livro anterior (Perini, 1995), chamei essa função “complemento do predicado”; como muitas pessoas acharam incômoda a mudança de nomenclatura, volto aqui à designação tradicional.

Do ponto de vista semântico, observa-se novamente que existe uma relação. No caso de [16], *meio desarrumada* atribui uma qualidade ao sujeito *a casa*; em [17], exprime-se uma relação de semelhança entre *a casa* e *um palácio*; e em

[18] Beth é a rainha da Inglaterra.

o predicativo do sujeito exprime uma relação de identidade (correferência) entre *Beth* e *a rainha da Inglaterra* — essencialmente, diz-se que são a mesma pessoa.

Compare-se essa situação com a que aparece em

[19] Beth serviu o vinho.

O sintagma *o vinho*, que analisamos como objeto direto, não tem nenhuma relação formal com *Beth*, nem semântica: não há nada entre *Beth* e o vinho; só o verbo é que liga esses dois sintagmas.

Por outro lado, há uma relação clara entre o sujeito e o predicativo do sujeito. Essa relação se manifesta nas condições de boa formação semântica. Exige-se certo tipo de congruência entre o conceito correspondente ao sujeito e o conceito correspondente ao predicativo do sujeito. Por isso é que não podemos dizer

[20] *Beth é / são dois cachorrinhos brancos,

e é estranho dizer

[21] ?*Minha camisa é jurídica⁶,

mas não há problemas em dizer

[22] Esse problema é jurídico.

Não existe nenhuma restrição desse tipo entre sujeito e objeto direto. O objeto direto depende, sintática e semanticamente, do verbo, e não do sujeito.

⁶ Exceto, é claro, em contextos metafóricos ou poéticos.

O predicativo do sujeito, como se vê, pode ser representado por um SN (como em [17] e [18]), e também por um sintagma adjetivo (como em [22] acima). Já o objeto direto é sempre um SN.

12.3. RELAÇÕES DE CONCORDÂNCIA

Vamos voltar à concordância, isto é, à exigência de que uma palavra assumira determinada forma por causa de sua relação com outra palavra da mesma frase. Assim, em *casas brancas*, as duas palavras estão no que chamamos feminino plural, e em *carro branco* temos ambas no masculino singular. Não é possível dizer **casas branco*, por exemplo, nem **carro brancas*. Fala-se, nesse caso, de concordância nominal⁷. Já a concordância verbal se dá entre o verbo e o sujeito, como *eu cheguei*, *ela chegou*, *nós chegamos*: mudando-se o sujeito, é necessário mudar a forma do verbo.

Nos casos de concordância, tradicionalmente se definem duas funções, a de termo regente e a de termo regido. A idéia é que um dos termos “obriga” o outro a concordar: *casas* “rege” *brancas*, e *eu* “rege” *cheguei*. A concordância é a base para a definição de algumas funções sintáticas: o sujeito (que já vimos acima como o elemento que ocorre logo antes do verbo) pode também ser definido como o elemento com o qual o verbo concorda em número e pessoa. O sujeito seria, portanto, o termo regente e o verbo, o termo regido.

O mesmo se diz dos casos de concordância nominal: em um SN como *o gato branco*, diz-se que *o* e *branco* concordam com *gato*. *Gato* seria o núcleo do sintagma nominal, o termo regente que controla o gênero e o número dos outros dois: *o*, que é determinante, e *branco*, que é modificador. Assim, as relações de concordância estão na base de algumas das funções sintáticas tradicionais.

Resta observar que em *o gato branco* a gramática tradicional chamaria tanto *branco* quanto *o* de “adjunto adnominal”. No entanto, esses dois termos têm comportamento gramatical muito diferente (basta lembrar suas possibilidades de ordenação dentro do sintagma), e portanto devem ser analisados como tendo funções distintas (conforme será mostrado no capítulo 13).

⁷ Aqui estou me referindo ao português padrão. A concordância é diferente na língua falada, onde *as casa* é aceitável. Mas mesmo nessa variedade há concordância, porque nunca se diz **casa branco*, nem **casa brancas*, nem **a casas*.

FUNÇÕES SINTÁTICAS: MAIS EXEMPLOS

Neste capítulo, vou apenas ampliar a exemplificação dada no capítulo anterior, definindo mais algumas funções sintáticas. Repito que a lista não é completa, e deve ser entendida apenas como ilustração dos princípios de análise envolvidos.

13.1. O NÚCLEO DO PREDICADO

Toda oração contém um verbo, e o verbo desempenha a função de núcleo do predicado. Seguindo a tendência geral, não vou propriamente definir o núcleo do predicado; apenas estabelecer que essa é a única função possível de um verbo na oração. Assim, cada oração tem um e só um núcleo do predicado. Por exemplo, o núcleo do predicado é o termo sublinhado nas orações abaixo:

[1] Meu tio viajou para a Argentina.

[2] Meu tio é rico.

[3] Meu tio tinha uma casa em Búzios.

No entanto, há casos que precisam ser explicados; o que fazer com frases como as seguintes?

[4] Meu tio vai viajar para a Argentina.

[5] Meu tio tem viajado para a Argentina.

[6] Meu tio está viajando para a Argentina.

Nesses casos, admite-se que o núcleo é o segundo verbo da seqüência (*viajar, viajado, viajando*), ou seja, o que está em forma nominal. O outro verbo (*vai, tem, está*) se chama auxiliar, e não constitui um núcleo do predicado.

Essa é a solução tradicional, e é correta, mas naturalmente tem que ser justificada. A justificação envolve algumas noções que ainda não estudamos; por isso, vou pedir paciência ao leitor e adia-la até o capítulo 17. Ali veremos por que frases como [4] – [6] contêm apenas uma oração, e portanto apenas um núcleo do predicado; e veremos por que esse núcleo é a forma nominal, não a forma conjugada (*vai, tem, está*).

Mas é preciso fazer mais uma observação acerca do núcleo do predicado. Em frases como

[7] Meu tio é argentino.

o núcleo do predicado tem que ser *é*, que é o único verbo presente. Essa é a análise que vamos adotar; mas ela contraria a gramática tradicional, que analisa *argentino* como o núcleo do predicado. Isso provém do fato de que a gramática tradicional concebe o núcleo do predicado como uma função semântica, e não uma função sintática, como fazemos aqui. Não precisamos nos preocupar com essa diferença; basta saber que se uma oração só tem um verbo, esse verbo deve ser automaticamente analisado como o núcleo do predicado.

13.2. PREDICATIVO DO OBJETO

Vimos que em orações com dois SNS, como

[8] O gato arranhou Toninho.

um desses SNS é sujeito e o outro objeto direto. As orações em geral comportam um máximo de dois SNS, e portanto essas duas funções costumam ser suficientes. Mas há algumas orações que contêm três SNS; por exemplo,

[9] Marília considera Ronaldo um gênio.

Aqui temos três SNS, *Marília, Ronaldo* e *um gênio*. A ordenação é fixa: *um gênio* precisa aparecer sempre depois do objeto direto:

[10] *Um gênio, Marília considera Ronaldo.

[11] *Marília considera um gênio Ronaldo.

[12] *Marília um gênio considera Ronaldo.

Esse terceiro SN pode ser chamado de predicativo do objeto¹. Ele se caracteriza por sua posição, obrigatoriamente após o objeto direto. Semanticamente, o predicativo é sempre qualificativo: não se refere a um novo participante, mas acrescenta uma qualidade ao objeto. Por isso mesmo, pode ser representado por um nominal de valor qualificativo (coisa que nunca acontece com o sujeito ou o objeto direto):

[13] Marília considera Ronaldo inteligente.

13.3. ADVERBIAIS

Além dos termos representados por SNS e dos representados por verbos, a saber, o auxiliar e o núcleo do predicado, a oração contém mais alguns termos. Um grupo importante é o representado pelos adjuntos (e complementos) adverbiais; trata-se de um grupo de funções bastante diferenciadas, que têm em comum a possibilidade de serem expressas por sintagmas compostos de preposição + SN, ou então por itens tradicionalmente chamados “advérbios”. Veremos no capítulo 19 que a classe dos “advérbios” não tem coerência; por ora, vamos aceitá-la provisoriamente, para efeitos da exemplificação, dando-lhes a designação não comprometida de “elementos adverbiais”.

Exemplos de elementos adverbiais na oração são:

[14] Meu pai tocava violão no banheiro.

[15] Amanhã vamos ter um almoço na casa do Noca.

[16] Ela vive chorando por causa do namorado.

[17] O Eduardo escreve muito bem.

[18] Vovó vai ao jogo com o Dr. Azevedo.

Como se vê, esses elementos se compõem de preposição + SN ou de advérbio, e exprimem muitas relações semânticas diferentes: “lugar” (em [14], ou o segundo adverbial de [15]); “tempo” (o primeiro adverbial de [15]); “causa” (em [16]); “modo” (em [17]); ou “companhia” (em [18]).

¹ Aqui também estou voltando à nomenclatura tradicional; em Perini (1995), chamei essa função simplesmente de “predicativo”.

Fala-se também de elementos adverbiais que exprimem “finalidade”, “intensidade”, “dúvida”, “instrumento”, terminando a lista em geral com um “etc.”. Essa multiplicidade semântica, um tanto fora de controle, é típica dos elementos adverbiais da oração — não porque eles sejam mesmo semanticamente indefinidos, mas porque não foram ainda devidamente estudados. A variedade semântica é grande, mas sintaticamente muitos deles têm comportamento bastante semelhante, o que nos autoriza atribuir a esses uma função sintática única — até melhor juízo, claro, já que faltam estudos pormenorizados.

13.4. TERMOS DO SINTAGMA NOMINAL

Naturalmente, as funções sintáticas não se definem apenas no nível da oração. Outras classes de sintagmas, como o sintagma nominal, têm estrutura interna, sendo formadas de termos relacionados por funções sintáticas. Vamos examinar agora alguns dos traços da organização interna do SN.

Alguns dos elementos internos do sintagma nominal se definem exclusivamente em termos de ordem. Assim, existem algumas palavras como *o*, *um*, *esse*, *aquele* (e seus femininos e plurais) que se caracterizam por aparecerem em primeiro lugar no SN, como em

[19] A última casa da rua

[20] Um professor gordo

[21] Aquelas camisas horrorosas que você usa

Note-se que não é possível colocar essas palavras em nenhum outro lugar no SN sem causar inaceitabilidade:

[22] * Última a casa da rua / *professor um gordo / *camisas horrorosas aquelas²

Dizemos então que *o*, *um*, *esse*, *aquele* e mais algumas palavras ocorrem sempre na função de determinantes.

² No Sul, algumas pessoas usam *aquele* posposto: *o professor aquele...* Por outro lado, *camisas horrorosas, aquelas* (com vírgula!) pode ser usado, mas nunca como SN: é antes uma frase independente.

Para dizer a verdade, existe um caso em que o determinante ocorre em segundo lugar no SN — é quando aparece a palavra *todo*, que é a única da língua que pode ocupar a função de predeterminante³:

[23] Todas essas camisas horrorosas

[24] Todos os professores

Temos então aqui duas funções dentro do SN, que se definem puramente em termos de posição: o predeterminante é sempre o primeiro elemento do SN, e o determinante é o primeiro se não houver predeterminante, caso contrário é o segundo.

Vamos admitir que cada SN inclui um núcleo (núcleo do SN, ou nSN), que é o elemento com o qual vários outros termos do SN concordam⁴. Em [19], o nSN é *casa*, em [20] é *professor*, em [21] é *camisas* e assim por diante⁵.

Ocorrendo em geral depois do núcleo encontramos o modificador; por exemplo, *gordo* em [20] é modificador. Ao contrário do nSN, do predeterminante e do determinante, o modificador pode ocorrer repetido dentro do mesmo SN:

[25] Um professor [de matemática] [gordo]

Aqui temos dois modificadores, *de matemática* e *gordo*. Como se vê, o modificador pode ser representado por um nominal (*gordo*) ou por uma seqüência de preposição + SN (*de matemática*) — ou seja, o modificador é representado por um sintagma adjetivo (ver 11.2).

Um SN pode ter uma estrutura bastante complexa, incluindo diversos termos de funções diferentes. Por exemplo, o SN seguinte

[26] Todos aqueles professores de matemática gordos

se analisa assim:

³ Na verdade existe outra, *ambos*; mas é uma palavra pouco usada hoje em dia, pelo menos na fala.

⁴ Vários, mas não todos: *de matemática* em [26] não concorda com nada.

⁵ A definição dada aqui de nSN é provisória. A definição do nSN envolve fatores ligados ao significado, que não podem ser incluídos em um texto introdutório. O leitor interessado encontrará uma exposição detalhada do problema em Perini *et al.* (1996) pp. 75ss.

[27] Todos aqueles professores de matemática gordos

PDet Det nSN Modif Modif

Essas funções não esgotam a estrutura do SN, que ainda pode conter outros termos, com funções diferentes. Como não é nosso objetivo oferecer uma análise completa do português, entretanto, vamos ficar por aqui em nosso exame da estrutura interna do SN.

13.5. PREPOSIÇÕES E SEUS COMPLEMENTOS

Mais um exemplo: em sintagmas precedidos de preposição, como os marcados em

[28] Elaine trabalhava para uma firma estrangeira.

[29] A roupa do menino ficou toda rasgada.

podemos distinguir dois termos principais: a preposição e o complemento (que é sempre um SN). E a ordem é obrigatoriamente essa: preposição primeiro, e complemento depois; não se admite inversão nenhuma:

[30] *Elaine trabalhava uma firma estrangeira para.

13.6. RESUMO: FUNÇÕES SINTÁTICAS

Esses exemplos mostram como as funções sintáticas são importantes na língua, determinando o significado, estabelecendo uma ordem obrigatória para certos itens léxicos ou definindo a concordância. Chamo desde já a atenção do leitor para uma característica importante das funções (que será retomada e expandida no capítulo 16), a saber, elas relacionam elementos presentes em uma estrutura. Assim, a função de sujeito relaciona um SN com um verbo (ou um predicado) presente na mesma oração. Trata-se de uma relação entre elementos presentes na seqüência formal — é portanto uma relação que só pode ser estabelecida dentro de um contexto (uma frase). Como veremos no capítulo 16, nem todas as relações gramaticalmente relevantes são dessa natureza; as relações de classe, ao contrário das de função, associam elementos que não estão presentes na mesma oração: compare-se a relação de função sujeito/verbo (*Denise* tem a função de sujeito de *fugiu* em *Denise fugiu*) com a relação de classe que existe entre *fugiu*, *almoça*, *dormia*, *chegou* (todos pertencentes à classe dos verbos, mas que não precisam estar presentes na mesma sentença).

FUNÇÕES SEMÂNTICAS

14.1. AGENTE, PACIENTE¹

Quando ouvimos a frase

[1] O gato arranhou Toninho.

entendemos que *o gato* é o ser que pratica a ação e *Toninho* designa a pessoa que sofreu a ação. Essas relações são também funções — funções semânticas, desta vez. Esse tipo de função se chama papel temático². Podemos definir o papel temático como a relação de significado que liga uma palavra que exprime ação, estado ou evento (muitas vezes um verbo, mas nem sempre) com as unidades que exprimem os participantes dessa ação, estado ou evento. Por exemplo, em [1] *o gato* tem o papel temático de agente, que é o elemento que pratica a ação, e *Toninho* tem o papel de paciente, que é o elemento que sofre a ação.

Não custa repetir que é necessário distinguir claramente entre as funções sintáticas (formais) e as funções semânticas. Em [1], *o gato* é sujeito, e é também o agente, mas sujeito e agente não são sinônimos. Primeiro, porque eles nem sempre coincidem: na frase

[2] Toninho foi arranhado pelo gato.

o agente é o gato, mas *o gato* não é o sujeito da oração (o sujeito é *Toninho*).

¹ Este capítulo está muito esquemático — talvez mesmo esquelético. Mas, dado o atual estado dos estudos no que diz respeito às funções semânticas (em especial os papéis temáticos), não parece haver outra alternativa.

² Também se encontram as designações “papel semântico”, “relação temática” e “papel-theta”.

Mas a razão principal para se distinguir funções sintáticas de funções semânticas é outra. Vimos no capítulo 3 que as construções têm significado; e as construções consistem de unidades conectadas através de funções. Ora, o ouvinte precisa partir de formas para chegar aos significados — isso porque o significado não pode ser transmitido diretamente. Assim, em uma frase como [1], o que o ouvinte percebe imediatamente é que *o gato* vem primeiro, antes do verbo, e *Toninho* vem depois; além disso, ele sabe que o verbo *arranhar* pertence a determinada classe³. A partir dessa informação é que ele vai deduzir (com base em seu conhecimento da língua) que *o gato* designa o agente, e *Toninho* o paciente. Ou seja, “sujeito”, “objeto direto” e as demais funções sintáticas são aspectos da organização formal das sentenças, perceptíveis na cadeia sonora ou gráfica que passa do emissor ao receptor. E “agente”, “paciente” e os demais papéis temáticos são aspectos da paisagem mental que existe originalmente na mente do emissor, e que o receptor reconstrói com base nas informações formais (entre outras coisas).

Voltamos aqui, como sempre, ao fato central da gramática: a relação que se estabelece entre um objeto físico (som, marcas no papel) e uma estrutura de conceitos. Para estudar essa relação, é essencial distinguir cuidadosamente os dois planos (ou espaços): o formal e o semântico. A noção de “sujeito” pertence ao espaço formal, a de “agente” pertence ao espaço semântico; a relação que existe entre esses elementos é uma relação simbólica. São as relações simbólicas que nos permitem interpretar seqüências de sons em termos de conceitos, ou seja, compreender as mensagens que percebemos pelo ouvido ou pela vista.

14.2. LISTA DOS PAPÉIS TEMÁTICOS

Existem muitos papéis temáticos além de “agente” e “paciente”. Dou abaixo uma lista que certamente não é completa, mas que inclui os papéis mais frequentemente encontrados na literatura⁴:

³ Isto é, tem determinada valência; ver o cap. 17 a respeito.

⁴ Seguindo Jackendoff (1972), com algumas alterações. Ao contrário de Jackendoff, não uso a designação tema para um dos papéis temáticos (que inclui, entre outras coisas, o paciente), preferindo distinguir vários papéis, em particular o paciente (o elemento que sofre a mudança de estado associada com o evento expresso pelo verbo, ou seja, o que “sofre a ação”) e o localizando (o elemento cuja localização é expressa, como *a cerveja* na frase *a cerveja está na geladeira*).

▼ Agente, como já vimos, é o elemento que pratica uma ação, como *o governador* nas frases

[3] O governador lavou os pratos.

[4] O governador resolveu o problema.

Discute-se se o agente deve ser sempre dotado de vontade e poder de decisão próprios. Assim, há quem duvide que haja um agente em

[5] O peso das caixas amassou o capô.

Para nossos objetivos, entretanto, vamos considerar todos esses casos como incluindo o papel temático de “agente” (*o governador* em [3] e [4] e *o peso das caixas* em [5]).

▼ Paciente se define como a entidade que sofre o efeito de uma ação ou evento, mudando de estado em consequência dela. Temos pacientes típicos em

[5] O peso das caixas amassou o capô.

[6] O capô amassou.

[7] O mecânico amassou o capô.

O capô é o que sofre o efeito do amassamento, mudando de estado (de não-amassado para amassado). Como se vê, não há necessariamente um agente para que haja paciente: em [7] *o mecânico* é agente, mas em [6] não há agente expresso. Observe-se também que o paciente é objeto direto em [5] e [7], mas sujeito em [6].

▼ Localizando é o elemento cuja posição é expressa, ou seja, do qual se indica a localização. Assim, *os caixotes* é localizando em

[8] Os caixotes estão no porão.

Em geral, os papéis temáticos são interpretados de maneira muito ampla, incluindo, por exemplo, movimento ou posição virtual ou metafórica. Assim, *a solução* se analisa como localizando na frase

[9] Eu estava com a solução na cabeça.

porque aí se exprime o “lugar” onde está a solução (a cabeça).

▼ Local especifica o lugar em que se dá o evento ou estado expresso pelo verbo; na maioria dos casos, esse papel é representado por um advérbio de lugar ou por um sintagma precedido de preposição:

[10] Elas moram em Uberaba.

[11] Elas moram aqui.

Pode-se dizer que palavras como *aqui* e *lá* são especializadas em exprimir esse papel temático.

Na frase [9],

[9] Eu estava com a solução na cabeça.

o local é *na cabeça*. Como se vê, geralmente um local se associa a um localizando.

Um problema a discutir a respeito do papel temático local é se ele serve também para a expressão do tempo. Há bastante evidência de que usamos expressões locativas para situar alguma coisa no tempo, como por exemplo em

[12] Eva nasceu em Belo Horizonte / Eva nasceu em 1976.

Isso acontece principalmente com preposições. Por outro lado, há também expressões específicas de tempo, diferentes das de lugar, como *hoje*, *amanhã*, *agora*, além das expressões específicas de lugar já mencionadas (*aqui*, *lá*). Mas não existem preposições específicas de tempo — nem mesmo *depois de* e *antes de*⁵, porque estas podem ser usadas para exprimir lugar, como em

[13] Bicas fica logo depois de Juiz de Fora.

▼ Fonte / meta: os papéis de fonte e meta aparecem na frase

[14] Papai viajou a pé de Ouro Preto a Mariana.

A fonte é a origem de um movimento (*de Ouro Preto* na frase acima), e a meta é o destino final desse movimento (*a Mariana*). Seguindo o

⁵ Chamadas tradicionalmente “locuções prepositivas”.

paralelismo já observado entre as expressões de espaço e as de tempo, temos fonte e meta também em

[15] Papai trabalha das 8 da manhã até as 4 da tarde.

E como o movimento no caso pode ser real ou metafórico, há também fonte e meta nas frases

[16] O professor ensinou a matéria aos alunos.

fonte meta

[17] Os alunos passaram do entusiasmo ao tédio.

fonte meta

Em [16], a matéria se “movimenta” do professor aos alunos, e em [17] os alunos se “movimentam” de um para outro estado de espírito. Essa aplicação metafórica procura refletir o comportamento de frases como [16] e [17], que funcionam, gramaticalmente falando, da mesma maneira que as frases que expressam movimento literal.

▼ Experienciador / causador de experiência

Há mais dois papéis temáticos importantes, os de experienciador e causador de experiência, que se vinculam a verbos de significado sensorial ou emocional. Assim, em

[18] Ana ama Luís.

[19] Ana viu Luís.

às vezes se fala de *Ana* como agente e de *Luís* como paciente. Mas a relação semântica presente nessas frases é nitidamente diferente: nem *Ana* pratica uma ação, nem *Luís* a sofre, mudando de estado; na verdade, nem “amar” nem “ver” são ações, e é mais correto dizer que *Luís* afetou (involuntariamente) *Ana* do que o oposto.

Para descrever esses casos, vamos dizer que nas duas frases *Ana* tem o papel de experienciador, e *Luís* o de causador de experiência. A experiência, no caso de [18], é emocional, e em [19] é sensorial. Também em frases que exprimem conhecimento, como

[20] Ana acreditou em Luís.

geralmente se analisa o sujeito como experienciador e o complemento como causador de experiência.

▼ Instrumento

Finalmente, pode-se mencionar o papel temático instrumento, que relaciona o verbo com um objeto (concreto ou não) utilizado para desempenhar uma ação:

[21] O homem abriu a porta com um pé-de-cabra.

[22] O orador convenceu o público com seus argumentos.

Para resumir, dou a lista dos nove papéis temáticos definidos acima:

▼ localizando

▼ agente

▼ paciente

▼ local

▼ fonte

▼ meta

▼ experienciador

▼ causador de experiência

▼ instrumento

Além desses, podemos acrescentar mais um:

▼ tema (o elemento que sofre a movimentação).

O tema é o papel temático de *o embrulho* em *ela levou o embrulho para a mesa*.

Os papéis temáticos são geralmente entendidos como relações entre o verbo (ou melhor, o predicado — o conceito expresso pelo verbo) e seus complementos. Mas há também relações de papel temático, por exemplo, entre nominais, como em

[23] A construção da ponte pelos engenheiros.

onde o nominal *construção* expressa uma ação, com os participantes *da ponte* (paciente) e *pelos engenheiros* (agente). A relação semântica entre *construção* e *pelos engenheiros* em [23] é exatamente a mesma que se encontra em

[24] Os engenheiros construíram a ponte.

Isso não quer dizer que [23] e [24] sejam sinônimos; há diferenças importantes, mas a relação de papéis temáticos entre os elementos é a mesma.

14.3. PAPÉIS TEMÁTICOS MÚLTIPLOS

Cada sintagma que é constituinte de uma oração (exceto o próprio verbo) tem um papel temático dentro da pequena peça teatral que a oração descreve. Assim, em

[25] Alda lavou os pratos rapidamente depois do almoço.

Alda é agente, *os pratos* é paciente, *rapidamente* é modo (um papel temático que não estudamos), *depois do almoço* é local⁶.

Mas, além disso, em muitos casos pode-se distinguir mais de um papel temático para o mesmo sintagma. Por exemplo, na frase

[26] Alda pagou oito mil reais a Renato por um carro.

temos que *Alda* é não apenas agente (pois foi ela que tomou a iniciativa de fazer o pagamento), mas também a fonte (do dinheiro) e a meta (do carro); igualmente, *Renato* é meta (do dinheiro) e fonte (do carro). Todas essas relações semânticas são necessárias para a plena compreensão da frase: se alguém não entender que Alda praticou uma ação (a de pagar), e que o dinheiro passou de Alda a Renato e o carro de Renato a Alda, simplesmente não entendeu a frase.

Em muitos desses casos (inclusive em [26]), os papéis temáticos múltiplos se relacionam a um evento múltiplo. Uma frase como [26] exprime na verdade dois eventos de transferência: a transferência do carro de Renato para Alda e a transferência dos oito mil reais de Alda para Renato.

⁶ Estou tomando "papel temático" em sentido amplo, como sinônimo de "relação semântica com o evento descrito". Segundo algumas teorias, nem todos esses papéis são realmente atribuídos aos sintagmas em questão. Aliás, relembro que as relações de tempo são, ao que tudo indica, analisáveis em termos do papel temático lugar.

As entidades envolvidas em ambos os eventos (Alda e Renáto) têm um papel temático em cada um deles (Alda é fonte do dinheiro e meta do carro, e Renato o oposto).

Outro exemplo de papel temático múltiplo é

[27] O cachorro arrastou para debaixo da cama.

O cachorro é agente, porque foi para debaixo da cama por seus próprios meios; mas é também paciente, porque é o elemento que sofre movimento. Ao contrário de [26], aqui temos apenas um evento, e não obstante *o cachorro* tem dois papéis temáticos.

Vimos na seção precedente que na frase

[16] O professor ensinou a matéria aos alunos.

o professor é fonte e *aos alunos* é meta. Mas também é claro que *o professor* também é agente, pois praticou a ação de ensinar a matéria; ou seja, o sintagma *o professor* tem dois papéis temáticos superpostos.

Mais um exemplo claro de papel temático múltiplo é o do verbo *arrumar* (no sentido de “vestir”). Com esse verbo, o sujeito pode ser o agente:

[28] A mãe arrumou a menina para a festa.

Aqui, claramente, quem praticou a ação foi a mãe (que é, portanto, o agente), e quem a sofreu foi a menina (que é o paciente).

Mas esse mesmo verbo pode ser usado em outra construção, exemplificada por

[29] A menina arrumou em menos de cinco minutos.

E agora a menina praticou e sofreu a ação: foi ela que arrumou, e ela que ficou arrumada. E agora? Como explicitar o papel temático do sintagma *a menina*? A única maneira é aceitar que esse sintagma tem dois papéis temáticos: é agente e paciente, simultaneamente.

Por outro lado, em [29] é bastante claro que *a menina* é sujeito, e não objeto direto, do verbo *arrumar*. Assim, o dilema se coloca apenas no nível

semântico: o sujeito de *arrumar*, em certos casos, recebe o duplo papel temático de agente e paciente.

Talvez agora alguém diga: mas a frase [29] está errada! Não é assim que se diz. O certo é *a menina se arrumou...*, com *se*. A esses eu respondo que a frase [29] não está errada, porque é assim que se fala, pelo menos em Belo Horizonte. Aqui (e, estou certo, em muitos outros lugares do Brasil) uma frase como [29] é usada por gente de todas as classes sociais, todos os níveis de escolaridade, em todas as situações informais. Nessas áreas, o *se* só é usado (por alguns!) em momentos solenes como discursos de formatura (que, aliás, ninguém escuta mesmo). Em outras partes do Brasil, o *se* é usado normalmente em tais frases; mas o fato permanece que para milhões de brasileiros [29] é a maneira normal de falar⁷.

Algumas teorias negam a necessidade dos papéis temáticos múltiplos, mas na minha opinião eles são indispensáveis para a representação adequada do significado de muitas sentenças. E, além disso, esses papéis múltiplos são representados na sintaxe de maneira sistemática — por exemplo, com o verbo *comprar*, se a construção for ativa, o sujeito é sempre ao mesmo tempo agente, meta (do objeto comprado) e fonte (do dinheiro). Isso significa que esses papéis múltiplos não são apenas algo que a gente entende a partir do contexto geral, mas são codificados através de estruturas lingüísticas específicas — no caso, através dos complementos de um verbo específico (*comprar*). São, portanto, parte legítima da estrutura gramatical da língua.

⁷ Para uma discussão mais detalhada desse fenômeno de rejeição por parte de certas pessoas das formas que elas mesmas usam, ver a Introdução e, principalmente, o livro de Marcos Bagno (2003).

CLASSES

15.1. PARA QUE CLASSIFICAR?

Todas as gramáticas incluem uma classificação de palavras. Isso sugere que a classificação das palavras (e das formas lingüísticas em geral) deve ser uma parte muito importante da descrição, e qualquer lingüista concordaria que realmente é. No entanto, quase nunca se consideram algumas perguntas básicas: para que classificar as formas lingüísticas? De onde vem a importância dessa classificação? E que princípios a norteiam? A partir deste capítulo, vamos enfrentar a questão fundamental (mas às vezes ignorada) do como e para quê da classificação das formas lingüísticas.

Para começar bem do início, vamos considerar o seguinte fato: boa parte do nosso conhecimento do mundo se compõe de classificações. Estamos a todo momento utilizando classificações implícitas para lidar com os conceitos que nos permitem compreender o mundo. Assim, ao entrar em uma sala e ver um objeto, posso deduzir que é uma cadeira e que serve para sentar. Mesmo que esse objeto específico seja diferente em vários detalhes de todas as outras cadeiras que já vi, em geral não tenho dificuldade de classificá-lo como um exemplo da classe das cadeiras. Para isso, utilizo certas características que considero importantes: o objeto tem quatro pernas, assento e encosto, está em determinado local dentro de uma sala (não pendurado no teto, por exemplo), é feito de madeira (não de papel ou de gelo, por exemplo), e assim por diante. Essa análise me permite colocá-lo em uma classe que já tenho previamente armazenada na memória, e que chamo de classe das cadeiras.

O mesmo acontece com as formas lingüísticas: uma das coisas que temos que fazer para compreender uma frase é classificar suas partes. Ao ouvir ou ler *o gato dormiu debaixo da cama*, precisamos classificar *o gato* como um sintagma que pode se referir a um ser, e *debaixo da cama* como

um sintagma que exprime um local; para isso, classificamos esses sintagmas formalmente (como “sintagma nominal” e “sintagma adverbial”, respectivamente). Temos que saber que a palavra *dormiu* não pode designar um ser, nem um local, mas sim um evento — e, por essa razão e por outras, acabamos classificando-a como “verbo”. A classificação das formas é importante em gramática porque ela é um dos principais instrumentos que utilizamos para construir, processar e compreender as expressões lingüísticas.

Mas a classificação de um objeto (lingüístico ou não) não é uma coisa absoluta, invariável. Nossa capacidade de classificação é extremamente flexível, porque precisa se adaptar às necessidades do momento. Se eu chamo o animal que tenho em casa de “cachorro”, e o do vizinho também de “cachorro” (apesar de o meu ser um pinscher e o do vizinho um labrador), devo ter alguma razão para isso. De algum ponto de vista, esses dois animais tão diferentes ficam juntos em uma classe. Mas é exatamente isso: de algum ponto de vista. Para levar no veterinário, digo pelo telefone apenas que é um cachorro. Mas se o ponto de vista for servir como animal de guarda, aí tenho que dizer que o meu é um pinscher (e portanto não tem grande serventia para isso). Ou seja: o ponto de vista condiciona a classificação.

Digamos que vamos classificar palavras. Não podemos simplesmente perguntar quais são as classes de palavras do português. Vai depender do ponto de vista — digamos que o objetivo da classificação seja colocar as palavras em determinada ordem no dicionário. Aí, a classificação vai ser: palavras começadas por *A*, palavras começadas por *B* etc. A ordem alfabética é a classificação “correta” nesse caso. Mas se o objetivo for escrever uma gramática da língua, sabemos muito bem que a primeira letra da palavra não tem a menor relevância: as palavras começadas por *A* não se comportam gramaticalmente de maneira semelhante. Já para os aficionados de palavras cruzadas, o número de letras de cada palavra tem importância: eles distinguem palavras de uma letra, de duas letras, de dez letras etc. E para um poeta que quer metrificar seus versos é o número de sílabas que vai ter importância.

A moral dessa história é que, antes de atacar o problema da classificação das palavras e formas do português, é necessário estabelecer os objetivos dessa classificação. Mesmo dentro do objetivo geral de escrever a gramática da língua, os objetivos podem variar segundo o tipo de fenômeno que está sendo descrito no momento.

Para dar um exemplo: digamos que vamos classificar os verbos da língua. Se no momento estamos descrevendo a morfologia (a flexão das palavras), uma classificação fundamental é a de verbos da 1ª, 2ª e 3ª conjugações, já que a morfologia do verbo depende da conjugação a que ele pertence: *andamos*,

comemos, *abrimos*. Para a descrição da morfologia, também é importante classificar os verbos em regulares e irregulares. Ou seja, em português os verbos se classificam em regulares e irregulares; e em verbos de 1ª, 2ª e 3ª conjugação. Essa classificação é importante na gramática, e não podemos passar sem ela.

Mas a importância dessa classificação fica na morfologia. Passando à descrição dos complementos que cada verbo pode ter, temos que considerar objetivos diferentes, e conseqüentemente uma classificação diferente. Alguns verbos podem aparecer com objeto direto, como *comer* (*eu comi a pizza*), e outros não podem, como *desaparecer*. Para descrever esse fenômeno, a conjugação a que o verbo pertence não tem a menor importância. Aqui vamos distinguir as classes dos verbos “transitivos” e “intransitivos”, e nunca “transitivos da 1ª conjugação”, “transitivos da 2ª conjugação” etc., porque a conjugação a que pertence um verbo não afeta o seu comportamento segundo o aspecto que estamos descrevendo no momento.

Por conseguinte, não existe uma maneira única, correta e exclusiva, de classificar os verbos (ou outras formas da língua): existem necessariamente várias, porque há vários pontos de vista possíveis no estudo dos verbos. Na memória do falante, a entidade que chamamos “o verbo *comer*” é um objeto complexo, e está classificado segundo várias propriedades, entre as quais:

- (a) é um verbo;
- (b) é regular;
- (c) é da 2ª conjugação;
- (d) aceita objeto direto;
- (e) não exige objeto direto;
- (f) expressa um tipo de ação;
- etc., etc., etc.

Segundo o objetivo descritivo do momento, cada uma dessas características é ou não relevante. Para descrever a fonologia, nenhuma das características listadas acima para *comer* tem relevância; para descrever a morfologia, só valem as características (a), (b) e (c); para a sintaxe, valem (d), (e) e muito provavelmente (f)².

Isso vale para todas as formas lingüísticas: na verdade, cada uma delas é um feixe de características (chamadas traços) como (a) – (f) acima. Desse feixe, o lingüista seleciona alguns traços que considera relevantes para a tarefa descritiva que tem em mãos no momento. Como resultado, o núme-

¹ Essas são as classes tradicionais. Veremos, a seu tempo (cap. 17), que são insuficientes.

² Isso porque os verbos que exprimem ação têm comportamento sintático típico.

ro e a identidade das classes de formas da língua varia segundo o objetivo descritivo estabelecido; e os objetivos variam segundo a parte da língua que está sendo examinada. É por isso que não faz sentido perguntar simplesmente quais são as classes de palavras do português; uma classificação só faz sentido em função de um objetivo específico. É como perguntar como se classificam os animais: mamíferos e répteis? domésticos e selvagens? úteis e nocivos? terrestres e aquáticos? perigosos e inofensivos? grandes e pequenos? Depende do que queremos dizer deles, claro.

15.2. CLASSIFICAÇÃO POR TRAÇOS

Para classificar uma forma lingüística, portanto, lidamos com traços, e cada traço é uma propriedade dessa forma — propriedade fonológica, morfológica, sintática ou semântica, conforme o caso. Todos os traços são considerados como presentes na representação mental da forma na memória do falante; ele seleciona os traços de que vai precisar a cada momento.

Por exemplo, se um falante vai construir uma frase, e pretende incluir nela o verbo *comer*, ele sabe que o pretérito perfeito é *comi*, porque o verbo é da 2ª conjugação. Sabe também que pode ou não colocar na frase um objeto direto; se o verbo for *desaparecer*, sabe que não pode colocar objeto direto. Esses são, por assim dizer, dois “momentos” no processo de construção da frase; a cada momento, o falante utiliza as informações relevantes para aquela tarefa, e não outras³. Como se vê, as diferentes classificações não são incompatíveis; antes, são todas necessárias, cada uma em seu momento. Mas isso nos obriga a classificar por objetivos, e como os objetivos mudam segundo a parte da gramática que estamos considerando, não faz sentido perguntar, fora da situação de análise, quais são as classes de palavras do português.

Na situação de análise, porém, não só faz sentido, como é essencial. Nunca conseguiremos construir uma descrição coerente e econômica da conjugação dos verbos sem distingui-los em conjugações; e nunca conseguiremos descrever a estrutura da frase sem distinguir sintagmas nominais (como *o gato*) de sintagmas adverbiais (como *debaixo da cama*). A razão principal é que a gramática precisa fazer afirmações gerais, não a respeito de palavras ou sintagmas individuais, mas a respeito de classes inteiras. Por exemplo, sabemos que *gato* faz o plural *gatos*; *xícara* faz o plural *xícaras*; *verde* faz o plural *verdes* e assim por diante. Essa regra de acréscimo de um *-s* se aplica

³ “Momentos” abstratos, entenda-se. Tudo indica que essas tarefas são levadas a efeito paralelamente, mas separamos os momentos para facilidade de descrição.

a dezenas de milhares de palavras da língua, e é evidente que a gramática não pode simplesmente listar todas elas; isso seria não apenas muito trabalhoso, mas principalmente deixaria escapar uma generalização, ou seja, uma regra: a de que os nominais fazem o plural acrescentando um *-s*.

Igualmente, não diremos que o sintagma *o gato* pode ser sujeito de uma oração (acrescentando que os sintagmas *os coelhos*, *as xícaras* e *a gata que arranhou meu pé* também podem). Em vez disso, formulamos uma regra que estipula que qualquer sintagma nominal pode ser sujeito; e definiremos o sintagma nominal em algum lugar da gramática.

Mas para formular essas regras temos que lançar mão de classes, como a classe dos nominais, a dos verbos, a dos sintagmas nominais e muitas outras. Sem a definição das classes, portanto, uma gramática seria proibitivamente complexa e, ao mesmo tempo, deixaria de expressar muitos traços importantes da estrutura da língua, pois se reduziria a uma listagem de casos individuais⁴.

Em princípio, nada disso é exclusivo da gramática. Utilizamos nossa capacidade de classificar a todo momento para fazer afirmações gerais: não dizemos que Arnaldo, Beatriz, Carlos, ... Xavier, Yara e Zezinho precisam pagar suas mensalidades até o dia 5, mas que os alunos da escola precisam fazer esse pagamento — utilizamos uma classe para fazer uma afirmação compacta e econômica, e também mais informativa, pois ficamos sabendo não apenas que aquelas pessoas precisam fazer um pagamento, mas também que é porque são alunas da escola. Nosso uso da língua é um dos aspectos de nossa atividade cognitiva, baseada, entre outras coisas, na organização de conceitos em classes.

15.3. REAPRENDENDO AS CLASSES

Para quem teve alguma instrução em gramática tradicional — ou mesmo em lingüística moderna —, as considerações acima podem parecer bastante estranhas. Acontece que as classes estão no âmago da descrição gramatical, e estamos muito habituados a partir delas para pensar os fenômenos da língua. A própria comunicação entre os lingüistas depende de um vocabulário que inclui os nomes tradicionais das classes: temos que falar de “substantivos”, “advérbios” e “pronomes” para nos entender entre nós. Agora alguém vem dizer que essas classes não existem?

Aqui estamos novamente enfrentando o problema já mencionado no capítulo 1, onde ficou avisado que para se tornar lingüista uma pessoa tem

⁴ Não há nada de errado em listar casos individuais. Mas isso é tarefa do léxico (ver o cap. 18).

que “se libertar de alguns preconceitos muito espalhados e fortemente enraizados”. Assim como um estudante de medicina tem que se acostumar com a idéia de que sair na chuva não causa gripe (a gripe é causada por um vírus) e um historiador tem que saber que não foi Colombo quem provou que a terra é redonda (isso já era conhecido muitos séculos antes dele), um lingüista deve, desde o princípio, se acostumar com a idéia de que as palavras de uma língua não se distribuem simplesmente, para todos os efeitos, em mais ou menos dez classes.

A classificação tradicional, como tentei explicar acima, não se sustenta. A classificação é necessariamente condicionada aos objetivos descritivos do momento, e não há limite previamente estabelecido ao número de maneiras que uma forma pode ser classificada. É claro que existe uma lista finita de traços importantes, de modo que a classificação não é aberta; mas a lista de traços ainda não é conhecida, e uma das tarefas da pesquisa lingüística é justamente estabelecê-la. E, ao desempenhar essa tarefa, temos que levar em conta que não se trata de uma classificação de valor autônomo, mas uma representação de algo que existe na mente dos falantes, e que lhes permite (junto com outras coisas) utilizar corretamente a língua. Ora, sabemos que nossa mente funciona de maneira extremamente flexível: somos capazes de utilizar um sistema de traços de maneira que ele funcione em cada caso em particular. Nossa mente é oportunista, e sabe aproveitar os dados disponíveis para chegar a decisões da maneira mais econômica: por isso é que podemos identificar um amigo apenas pela voz, ou pelo vulto, ou pelo barulho que ele faz ao subir a escada. É que os traços que nos permitem identificá-lo são variados e estão armazenados* de tal forma que podem ser utilizados conforme a situação.

É esse sistema que temos que aplicar também ao uso da língua — já que a língua é uma parte do nosso sistema de conhecimento. Assim, a palavra *pálida* é um proparoxítono, um nominal, um feminino, um singular, designa uma qualidade, mas não é o nome de uma coisa, nem tem pretérito perfeito. Cada um desses traços a coloca junto a certas palavras, e a separa de outras, e essas classificações se cruzam de várias maneiras. O fato de designar uma qualidade coloca *pálida* junto com *jovem*, mas essas mesmas duas palavras se distinguem porque *jovem* pode designar uma coisa, e *pálida* não (nunca dizemos **tem uma pálida na sala*, mas podemos dizer *tem uma jovem na sala*).

Em suma, temos em mãos um objeto, a língua, de grande complexidade; e nossa descrição não pode deixar de reconhecer e incorporar essa complexidade, sob pena de ser seriamente inadequada. Aqui não se trata apenas de aprender algo novo, mas de desaprender algo que estudamos desde o curso primário. É preciso reconhecer que isso não é fácil; mas é necessário.

CLASSES E FUNÇÕES

16.1. FUNÇÕES, CLASSES E O CONTEXTO

Os tipos de funções gramaticais que vimos são bastante diferentes uns dos outros: por exemplo, as funções baseadas em ordem são puramente formais, ao passo que os papéis temáticos são funções semânticas. Mas todas as funções têm uma importante característica em comum: são relações entre dois ou mais elementos presentes em uma mesma sentença. As classes, por outro lado, não são relações entre elementos presentes na mesma sentença.

Explicando melhor: as funções são relações (de ordem, de concórdância, de papel temático etc.) que existem entre as unidades dentro de uma sentença específica. Uma função é definida em termos de um contexto. As classes, ao contrário, se definem em relação à estrutura da língua, independentemente de contextos particulares. Isso pode se verificar com um teste simples: perguntamos a uma pessoa que tenha um mínimo de instrução gramatical a que classe pertence a palavra *João*; ela vai responder que é um “substantivo”. Note-se que essa pergunta faz sentido, e tem uma resposta. Mas digamos que a pergunta seja: que função sintática tem a palavra *João* (digamos, é sujeito ou objeto direto)? Essa pergunta, do jeito que foi formulada, não faz sentido, porque *João*, fora de contexto, não tem função. Para sabermos que função essa palavra tem, precisamos do contexto: que função tem a palavra *João* em *o gato arranhou João*? Aí já é possível responder: nessa oração, *João* é o objeto direto. O mesmo acontece se perguntarmos se *João* é agente ou paciente: isso não pode ser respondido sem um contexto. Mas é possível dizer, mesmo sem contexto, se *João* denota um ser ou uma qualidade, ou se pode denotar ambas as coisas. Ou seja, as classes são verificáveis fora de contexto, mas as funções só se verificam em contexto.

16.2. POTENCIAL FUNCIONAL

Isso tem a ver com a maneira como a informação fica armazenada em nosso cérebro e como a utilizamos para construir enunciados. Vamos considerar um falante diante da tarefa de construir uma sentença. Para isso, ele tem que utilizar determinados conhecimentos que guarda na memória. Entre esses conhecimentos, estão as palavras e demais unidades da língua.

Conhecer uma palavra (digamos, *João*) significa conhecer muita coisa. É saber que *João* é o nome de uma pessoa, mas é também saber onde e como essa palavra cabe em uma sentença, ou seja, quais são as funções que *João* pode desempenhar — não a função que está desempenhando no momento, mas a função ou funções que pode desempenhar. De posse desse conhecimento, o falante pode começar a construir a frase. Por exemplo, ele sabe que pode dizer *João me telefonou*, mas não **João noite eu quase não dormi*. Esta última frase seria boa com *essa* no lugar de *João*: *essa noite eu quase não dormi* — e um falante sabe disso porque também conhece as funções possíveis da palavra *essa*, que não pode ser livremente permutada com *João*.

Antes de começar a produzir seu enunciado, o usuário da língua precisa conhecer as palavras e as funções que cada uma pode desempenhar — ou seja, as palavras devidamente classificadas. É por isso que se diz que as unidades da língua se classificam por seu potencial funcional: uma classe é o conjunto das formas da língua que têm potencial funcional semelhante. As diferentes formas da língua, juntamente com as classes a que cada uma pertence, são parte de nosso conhecimento permanente.

Um exemplo é o nosso velho conhecido sintagma nominal (SN). O SN se define como o conjunto das formas que podem ser (note bem: podem ser) sujeito, objeto direto ou complemento de preposição. Outra propriedade importante do SN é que ele pode ser usado referencialmente, ou seja, para designar um ser (esta propriedade é semântica). As funções que o SN pode desempenhar constituem o seu potencial funcional. Assim, os SNS *Virgínia*, *o tampo da mesa branca*, *o computador que você comprou*, todos podem se referir a seres (animados ou inanimados)¹, e todos podem ser sujeito, objeto direto ou complemento de uma preposição.

¹ Há muitas complicações aqui, e a definição do que vem a ser, exatamente, o uso referencial não tem nada de simples; basta lembrar SNS como *a tristeza de Helena* ou *o fato de que choveu ontem*, que não se referem claramente a “seres”, ou “coisas”; mas podemos deixar esse problema de lado, por ora. A dificuldade central que enfrentamos para caracterizar

Já o sintagma adjetivo (sadj) pode ocorrer como modificador ou como predicativo, como mostram os exemplos abaixo:

-
- [1] Compramos um carro amarelo (modificador).
 [2] Esse carro está sujo (predicativo).
 [3] Compramos um carro de quatro portas (modificador).
 [4] Esse carro é de quatro portas (predicativo).
 [5] Ela mora na casa que fica logo depois da sapataria (modificador).
-

As formas *amarelo*, *sujo*, *de quatro portas* e *que fica logo depois da sapataria* se classificam como sintagmas adjetivos (sadj). O leitor atento notará que o modificador de [5] não pode ser predicativo. Ou seja, a afirmação de que os sadjs podem ser modificadores ou predicativos não é correta, ou então o modificador de [5] não é um sadj. Acontece que rótulos como “SN”, “sadj”, assim como “nominal” ou “pronome”, são abreviaturas de feixes de traços, cada traço exprimindo uma propriedade gramatical. As abreviaturas são aproximadas, e apenas a lista dos traços é que dá o comportamento gramatical exato de cada unidade lingüística.

Semanticamente, o sadj restringe o significado do SN de que faz parte ou a que se refere, facilitando sua identificação. Por exemplo, em [1] ficamos sabendo não só que se trata de um carro, mas de um carro amarelo (referência mais restrita); e em [3] ficamos sabendo que se trata de um carro de quatro portas.

Voltando ao potencial funcional como critério principal para a classificação: ele exprime o que uma forma pode ser — as funções que ela pode ocupar na sentença. Uma forma que esteja, em determinado contexto, ocupando uma dessas funções continua podendo ocupar outras (o que pode acontecer em outros contextos). Por exemplo, a palavra *João* pode ser sujeito ou objeto direto². Na frase

-
- [6] João ronca muito.
-

essa palavra é sujeito. Mas ela continua podendo ser objeto, em outra frase. Assim, mudando o contexto, pode mudar a função de *João*; mas sua classe

semanticamente as entidades gramaticais (classes, palavras, funções...) vem de que elas se relacionam não com os objetos, relações e eventos do mundo, mas com nossa visão desses objetos, relações e eventos.

² Mais rigorosamente: *João* pode ser núcleo de um SN, e o SN pode ser sujeito ou objeto.

é sempre a mesma, porque a classe se define por potencialidades, não por realidades presentes do contexto.

16.3. CADA FORMA PERTENCE A UMA (E SÓ A UMA) CLASSE

Pela mesma razão, uma forma não pode pertencer a mais de uma classe. Muitas vezes um gramático, percebendo a inadequação da doutrina tradicional, procura uma saída vinculando a classificação ao contexto. Um exemplo típico é o de palavras como *amigo* em

[7] Meu amigo vai telefonar às oito horas.

[8] Eu sempre prefiro consultar um médico amigo.

A solução tradicional é dizer que *amigo* é substantivo em [7] e adjetivo em [8]. Essa solução é incorreta, porque esconde o fato de que a palavra *amigo* tem o potencial funcional de palavras como *mesa*, mais o potencial funcional de palavras como *estomacal*. *Mesa* pode ser núcleo de um SN, mas não modificador; *estomacal* pode ser modificador, mas não núcleo; e *amigo* pode ser as duas coisas. A palavra *amigo* pode desempenhar mais de uma função, mas isso não significa que possa pertencer a mais de uma classe. As palavras que podem ser núcleos ou modificadores constituem uma classe, distinta das que só podem ser núcleos e das que podem ser só modificadores. Conclui-se que essas três palavras pertencem a três grupos de potencial funcional distinto — portanto, a três classes distintas. Por isso mesmo, *amigo* pode aparecer em contextos onde *estomacal* não pode (isto é, como núcleo de um SN) e também em contextos onde *mesa* não ocorre (como modificador). Temos aqui três classes de palavras, e não apenas duas; e cada uma dessas palavras pertence a uma dessas classes.

Isso vale em geral para todas as classes: cada palavra pertence a uma delas, e uma palavra não pode ter mais de um potencial funcional. Uma palavra pode, claro, ocupar mais de uma função, dados diferentes contextos³.

³ Ou seja, “classe” é uma relação paradigmática, e “função” é uma relação sintagmática, conforme a distinção estabelecida por Saussure (1916).

CONSTRUÇÕES

17.1. MONTANDO ESTRUTURAS

As funções sintáticas e semânticas, associadas às formas da língua, podem ser entendidas como parte das instruções de como construir estruturas da língua. Assim, pegando-se dois SNs e um verbo, pode-se montar uma oração colocando um dos SNs como sujeito e o outro como objeto direto desse verbo:

[1] Essas faixas amarelas enfeiam a pintura.

A oração [1] representa uma estrutura que pode ser representada assim:

[2] [SN V SN]_{oração}
sujeito obj. direto

Como a estrutura interna do SN é muito variada, podemos desenvolver uma oração da forma de [2] de muitas maneiras, como por exemplo:

[3] [Essas faixas amarelas]_{SN} enfeiam [a pintura]_{SN}

[4] [Ana]_{SN} beijou [Chico]_{SN}

[5] [Aquela aluna implicante]_{SN} procurou [o professor de Química]_{SN}

[6] [Élcio]_{SN} come [essas balas que grudam no dente da gente]_{SN}

[7] [O velhinho]_{SN} disse [que o barulho da festa deu dor de ouvido nele]_{SN}

Todas essas frases (e muitas mais) são exemplos de uma única construção, que é composta de um SN sujeito, seguido de um verbo, seguido de um SN objeto direto.

17.2. DEFINIÇÃO DAS CONSTRUÇÕES

Cada construção tem uma estrutura típica, definida em termos das classes de seus elementos e sua ordem (ver logo abaixo algumas dessas definições). Mas cada construção tem também um significado — ou seja, também no nível da construção, como no nível da palavra, cada forma se associa a um significado próprio. Assim, tomando a construção exemplificada em [3] a [6] (deixando de lado [7], na qual o SN depois do verbo contém uma oração; veremos esses casos no capítulo 21), podemos dizer que o primeiro SN — o sujeito — sempre exprime o elemento que “faz” alguma coisa; ou seja, o causador principal do evento: as faixas amarelas são a causa de a pintura ficar feia; Ana é a causa do evento de beijar (Chico apenas recebeu o beijo, talvez até sem querer); a aluna implicante praticou a ação de procurar o professor; Elcio pratica a ação de comer as balas. Como já sabemos, a esse elemento que causa o evento damos a designação de agente. “Agente” é uma função semântica, que pode ser desempenhada por um SN (e só por um SN)¹. Assim, podemos acrescentar ao potencial funcional do SN a propriedade de poder exprimir o agente de um evento. Já vimos a relação entre o agente e as funções sintáticas no capítulo 3; aqui vamos reconsiderá-la em função de uma entidade maior, a construção.

É necessário considerar o nível da construção, porque não se pode dizer pura e simplesmente que o agente é expresso pelo sujeito. Isso aparece com clareza em casos de orações com sujeito e sem objeto direto, como

[8] Marilene gritou.

Aqui, tal como nos exemplos anteriores, o SN sujeito expressa o agente: Marilene praticou a ação de gritar, ou seja, causou o evento descrito pela oração. Mas em

[9] Marilene engordou.

não se pode dizer que Marilene “fez” alguma coisa, nem que tenha causado o evento. Ela engordou por causa de alguma outra coisa, e Marilene é o ser que sofreu o processo descrito pelo verbo *engordar*, não o ser que o causou. Em outras palavras, em [9] o sujeito não é agente. A explicação é

¹ Em certos casos, por um SN precedido de preposição: *ele apanhou de Maria*, *o carro foi lavado pelos meninos*; mas em todos os casos o SN é que designa o ser que causa o evento.

que, apesar de [8] e [9] serem formalmente idênticas (sujeito + verbo), elas representam construções diferentes: em [8] temos a construção intransitiva, e em [9] a construção ergativa.

Como essas duas construções são formalmente idênticas, é claro que é preciso acrescentar alguma coisa que as distinga, a saber, algum traço de significado. Assim, podemos definir essas duas construções da seguinte maneira:

[10] **Construção intransitiva**

Definição:

SN V

Agente

[11] **Construção ergativa**

Definição:

SN V

Paciente

Como se vê, as definições incluem informação formal (SN + V) e informação semântica (o SN é agente em [10], paciente em [11]) —, ou seja, as definições são simbólicas, pois exprimem relações entre a forma e o significado.

Definições como as duas dadas acima nos dizem algo como: em português é possível construir uma oração juntando um SN mais um verbo; e nesse caso, dependendo do verbo, o SN pode ser agente ou paciente.

17.3. AS CONSTRUÇÕES DO PORTUGUÊS

Assim como existe a construção causativa sem objeto e a ergativa, o português dispõe de grande número de outras construções — ao que se sabe atualmente, e considerando apenas as que formam orações, há pelo menos várias dezenas, e deve haver muitas mais. Por ora, vamos nos limitar às duas vistas na seção precedente, mais uma que é particularmente importante, a saber:

[12] **Construção transitiva**

Definição:

SN V SN

Agente Paciente

Como se vê, é parecida com a intransitiva descrita em [10], só que agora com objeto. Um exemplo é

[13] Élcio comeu as balas.

Vamos utilizar essas três construções para ilustrar a importância da identificação e definição de construções para a gramática do português.

17.4. UMA GRAMÁTICA DE CONSTRUÇÕES

Quando dispusermos de uma lista completa das construções do português — suas estruturas formais e traços semânticos —, teremos em mãos uma espécie de núcleo da própria gramática da língua. Essa lista nos dará uma resposta à pergunta:

Que estruturas sintáticas existem na língua, e como são interpretadas?

Vai ser uma resposta parcial, porque há detalhes que não aparecem na definição de uma construção (por isso falei de “núcleo”, e não da gramática completa). Por exemplo, os fatos ligados à concordância não são incluídos na definição das construções, de modo que as duas frases abaixo representam a mesma construção:

[13] Élcio comeu as balas.

[14] Eu comi as balas.

A concordância, como vimos no capítulo 12, é descrita de modo mais conveniente através de regras.

Há também muitos detalhes de significado que não cabem na definição das construções; por exemplo, tomando

[15] A menina acariciou o cachorrinho.

[16] O caminhão matou o cachorrinho.

dizemos que o segundo *SN* (o objeto direto) é paciente nos dois casos, e que ambas as frases são exemplos da construção transitiva. Mas a relação semântica entre o verbo e o objeto não é exatamente a mesma: no caso de [16], temos uma mudança de estado (trágica), em que o cachorrinho passa de vivo a morto; mas em [15] o que se entende é apenas um contato superficial, sem que o cachorrinho mude de estado.

Isso é mesmo de se esperar, dadas as condições sob as quais interpretamos as expressões linguísticas — relembro os exemplos vistos no capítulo

3, que ilustram como nosso conhecimento do mundo colabora com o conhecimento da língua para produzir o sentido dos enunciados. Os exemplos [15] e [16] mostram que relações bem diferentes podem ser codificadas de maneira idêntica pela língua, de maneira que, lingüisticamente falando, podemos dizer que *o cachorrinho* é paciente nos dois casos, apesar da diferença de sentido. E podemos dizer que [15] e [16] são exemplos da mesma construção (transitiva).

17.5. AS CONSTRUÇÕES E A VALÊNCIA DOS VERBOS

Agora vamos considerar as construções do ponto de vista do receptor. O fato de algumas construções serem formalmente idênticas a outras levanta um problema: como é que o receptor sabe que em [8] o sujeito é agente, mas em [9] é paciente, se a estrutura é a mesma (*SN + V*)?

[8] Marilene gritou.

[9] Marilene engordou.

A resposta é que o verbo não é o mesmo nos dois casos, e o receptor conhece as propriedades diferentes de *gritar* e *engordar*. O verbo *gritar* tem sempre sujeito agente, ao passo que *engordar* pode ter sujeito agente ou paciente, conforme a construção. O sujeito de *engordar* é agente na frase

[17] A cerveja engordou Marilene.

porque aqui temos o objeto direto *Marilene*, mas o sujeito é paciente em [9], porque aí não há objeto direto.

O usuário da língua tem em sua memória não apenas a lista dos verbos (com seus significados etc.), mas ainda sua classificação em termos das construções em que cada um pode aparecer. Assim, ele precisa saber que *gritar* aparece na construção transitiva, e que *engordar* pode aparecer na transitiva ou na ergativa — e ainda na intransitiva, porque ao lado de [9], onde Marilene é paciente, temos

[18] Cerveja engorda?²

² Frases como [9] e [18] são ambíguas em princípio, mas as condições pragmáticas em geral as desambiguam.

onde o sujeito (*cerveja*) é agente.

Em outras palavras, os verbos se classificam em função das construções em que podem ocorrer. Assim, dizemos que

[19] *gritar* ocorre nas construções: transitiva³, intransitiva;

[20] *engordar* ocorre nas construções: transitiva, ergativa, intransitiva.

Os verbos apresentam uma grande variedade quanto às construções em que podem aparecer: além dos dois tipos ilustrados por *gritar* e *engordar*, temos verbos que só aparecem na transitiva (como *fazer*, que sempre tem objeto), e inclusive verbos que só aparecem na ergativa (como *morrer*). Tudo isso precisa estar na memória do falante quando ele organiza as sentenças que vai proferir.

Essa classificação dos verbos se chama sua valência: em [19] temos a valência de *gritar*, em [20] a de *engordar*. Como se vê, a valência de um verbo é o conjunto de construções em que ele pode ocorrer, e os verbos se classificam segundo a valência de cada um. E cada construção utilizada para exprimir a valência de um verbo se chama diátese; assim, a valência de um verbo é o conjunto de suas diátesses.

O falante da língua precisa conhecer a valência de cada verbo para poder utilizá-lo e compreendê-lo corretamente. Isso pode parecer uma tarefa imensa (dado que os verbos da língua são milhares), e é, mas é uma tarefa que as crianças enfrentam e realizam sem dificuldade; por exemplo, qualquer criança de 7 anos sabe que em [8] Marilene “fez” alguma coisa, mas em [9] “aconteceu” alguma coisa com ela. Ora, ela sabe disso porque conhece (implicitamente, claro) as valências dos verbos *gritar* e *engordar* — isto é, conhece as estruturas formais em que esses verbos podem aparecer e o papel temático de cada um dos complementos que acompanham o verbo.

Recapitulando a importância da formulação das construções, elas são um componente essencial da descrição da língua porque:

- (a) exprimem as possibilidades de ordenação, agrupamento e interpretação dos itens léxicos da língua;

³ *Gritar* ocorre na transitiva na frase *eles gritaram as instruções de cima da torre*.

- (b) exprimem parte das propriedades sintáticas e semânticas dos verbos da língua⁴.

Ou seja, ao sabermos que existe uma construção como a ergativa, exemplificada em

[9] Marilene engordou.

ficamos sabendo que:

- (a') uma oração pode ser formada de um sintagma nominal seguido de um verbo; nesses casos, o sintagma nominal pode receber o papel temático de paciente etc.
 (b') o verbo *engordar* pode ocorrer sem objeto e com sujeito paciente (ou seja, na construção ergativa, que é uma de suas diátesses).

As informações de (a') podem ser chamadas gramaticais, porque se referem à forma e ao significado de estruturas da língua; e as de (b') léxicas, porque descrevem propriedades particulares de uma palavra. Mas, como se vê, os dois tipos de informações estão intimamente unidos, de modo que não é possível, por exemplo, expressar as propriedades de um item léxico (*engordar*) sem se referir a classes e funções gramaticais (sujeito, objeto, paciente, verbo). Essa é a razão principal pela qual muitos linguistas negam a distinção entre gramática e léxico. Para nós, a distinção continua tendo utilidade, porque facilita a organização das afirmações descritivas sobre a língua: assim, em (a') temos basicamente regras gerais, e em (b') temos as propriedades particulares de uma palavra.

17.6. REVISITANDO O NÚCLEO DO PREDICADO

Agora podemos voltar a um assunto que ficou pendente no capítulo 13: a identificação do núcleo do predicado em frases como

[21] Meu tio vai vender pneus.

[22] Meu tio tem vendido pneus.

[23] Meu tio está vendendo pneus.

⁴ Outras palavras, que não verbos, também têm valências: nominais e certos advérbios.

Relembrando: defini o núcleo do predicado como “o verbo da oração”; mas em [21] — [23] temos dois verbos. Vimos que o núcleo do predicado, nesses casos, é o segundo verbo da seqüência (o que aparece em forma nominal: gerúndio, infinitivo ou particípio). Agora podemos ver o porquê dessa decisão.

Tomando [21] como exemplo: acontece que nessa frase se realiza a valência do verbo *vender*, e nenhuma outra; a presença da forma verbal *vai*, embora tenha sua importância, não afeta em nada a valência observada⁵. Em outras palavras, a valência de *vai vender* é idêntica à de uma forma simples do verbo, como *vende* por exemplo. Assim, temos:

-
- [21] a. Meu tio vai vender pneus.
 b. Meu tio vende pneus.
- [24] a. *Meu tio vai vender de pneus.
 b. *Meu tio vende de pneus.
-

etc. A substituição de *vai vender* por *vende* não acarreta nenhuma mudança de aceitabilidade. Conclui-se que frases como [21] realizam apenas uma valência; e como queremos atribuir a valência ao predicado, diremos que aí só há um predicado, e portanto só um núcleo do predicado.

Confronte-se essa situação com aquela em que há realmente dois predicados. Seja a frase

[25] Meu tio quer chegar antes das oito.

Aqui temos *quer* e *chegar*, duas formas verbais. Mas para que uma frase com essas duas formas seja aceitável, é necessário satisfazer as exigências tanto do verbo *querer* quanto do verbo *chegar*. Assim, podemos dizer:

[26] A encomenda chegou antes das oito.

o que mostra que *a encomenda* é compatível com *chegar* (já que *chegar* não exige que seu sujeito seja “animado”). Mas não se pode dizer

⁵ É bom observar que a argumentação aqui se refere apenas ao verbo *ir* quando usado em conexão com um infinitivo: *vou comprar*, *vamos trabalhar*. Quando usado como verbo de movimento (*vou para São Paulo*), o verbo *ir* tem valência.

[27] *A encomenda quer chegar antes das oito.

A inaceitabilidade de [27] precisa ser atribuída ao verbo *querer* (que exige sujeito “animado”). Isso mostra que, em frases com *querer* e *chegar*, ambos os verbos preservam seus predicados separados, e portanto seus conjuntos de exigências de valência. Analisamos, portanto, [25] como tendo dois predicados, e dois núcleos do predicado: *quer* e *chegar*. Ou ainda, dizemos que há aí duas orações, ao contrário de [21], onde só há uma.

Essa é a razão de se analisar *vender* em [21] como núcleo do predicado, e *vai* como auxiliar. Isso ocorre também em [22] e [23]; mas, como vimos, não se pode aplicar mecanicamente essa análise a qualquer seqüência de verbo + forma nominal (isto é, particípio, infinitivo ou gerúndio), porque em [25] temos *quer chegar* (um verbo mais um infinitivo) mas, como vimos, há dois núcleos do predicado e (dois predicados)⁶. O leitor poderá testar os casos de [22] e [23], para confirmar que também nesses casos o predicado é apenas um por oração.

⁶ Essa argumentação se encontra de forma mais desenvolvida em Pontes (1973).

O LÉXICO

18.1. O TAMANHO DA TAREFA

Vamos agora considerar brevemente o léxico, aquele grande repositório de informação sobre as palavras e outras unidades, que temos que memorizar para poder usar a língua com fluência.

Todos sabemos que conhecer uma língua envolve o aprendizado de um grande número de lexemas (ver definição de “lexema” no capítulo 11), que tradicionalmente são colocados em um léxico (isto é, em um dicionário mental). Aprendê-lo é, evidentemente, uma tarefa enorme: um falante instruído pode conhecer 20.000 lexemas ou mais, e provavelmente nenhum falante adulto sabe menos de 5.000¹.

A primeira mensagem deste capítulo é que a tarefa é muito, mas muito mais complexa do que parece, porque cada item léxico envolve diversos tipos de informação. E, apesar de tudo, conseguimos adquirir competência no léxico da língua quando somos ainda crianças, em alguns anos, independentemente de estudos formais, ou seja, apenas usando a língua no dia-a-dia. Vou começar dando uma idéia muito resumida do que significa conhecer o léxico de uma língua.

Por exemplo, todos conhecemos o lexema *tentar*, e sabemos usá-lo sem problemas, construindo frases corretas e compreensíveis. Mas o que significa isso, exatamente? Que conhecimentos nos permitem esse uso tão fácil e espontâneo do lexema a que chamamos “o verbo *tentar*”?²

¹ Um dicionário como o Aurélio tem cerca de 120.000 palavras, mas a maior parte são termos técnicos, regionais e arcaicos que pouca gente conhece. A estimativa de quantos lexemas um falante conhece é citada em Nation (2001).

² Estou pegando já o lexema, isto é, o conjunto de palavras que identificamos como formas do verbo *tentar*; nosso conhecimento da relação entre essas palavras (*tentar*, *tentei*, *tentássemos*, *tentando* etc.) é assunto da gramática, e não do léxico.

Primeiro, temos que saber que é um verbo; desse modo, não vamos tentar usá-lo do mesmo modo que *em* (preposição) ou *verde* (nominal). É esse conhecimento da classe a que pertence *tentar* que nos permite saber que [1] é uma boa frase do português, mas [2] e [3] não são:

[1] Tentei consertar o motor.

[2] *Ricardo comprou um vestido tentou sua esposa (*em vez de para* sua esposa).

[3] *Essa é uma oferta tentou (*em vez de oferta tentadora*).

Além disso, sabemos que *tentar* significa mais de uma coisa, como se vê nas frases

[1] Tentei consertar o motor. ("procurar")

[4] O ladrão tentou a janela da cozinha. ("experimental")

[5] Esse quindim está me tentando. ("provocar desejo")

Essas diferentes acepções precisam estar todas incluídas em nosso conhecimento do lexema *tentar*. Mas não basta: cada um dos diferentes significados do verbo corresponde a uma ou mais diáteses distintas. No sentido de "procurar", o complemento deve ser uma oração no infinitivo (*consertar o motor*), ou então um nominal abstrato (*tentei uma solução nova*); no sentido de "experimental", deve ser um SN, concreto (*a janela da cozinha*) ou abstrato (*suicídio*); no sentido de "submeter a tentação", precisa ser um SN referente a ser animado (*me*). Temos que saber tudo isso também.

Tentar se relaciona com nominais que significam "o ato de tentar", mas novamente cada acepção tem sua forma especializada. No sentido de "procurar", o nominal é *tentativa*; no sentido de "provocar desejo", é *tentação*.

Outra coisa que precisamos saber é que *tentar* exige um agente animado, e por isso se pode dizer [6], mas não [7]:

[6] A menina tentou ficar na cama.

[7] *A poeira tentou ficar na cama.

Além disso, *tentar* pode ocorrer com um complemento no infinitivo, como em [6] acima, e nisso se distingue de *explicar*, cujo complemento pode ser formado de *que* mais uma oração com verbo no indicativo:

[8] A menina explicou que estava cansada (*cf.* *a menina tentou que ficava na cama).

Em frases como [6], entendemos que a menina não apenas tentou, mas que ela é que ia ficar na cama. Nisso *tentar* se diferencia de *mandar*, porque com este último verbo os agentes são necessariamente distintos; em

[9] A menina mandou ficar na cama,

entendemos que uma pessoa mandou e outra ficou (ou devia ficar) na cama.

Outra coisa referente a esse verbo é que ele, quando significa "procurar", inclui a insinuação de que a coisa tentada não foi conseguida; assim, entende-se de [6] que a menina não conseguiu ficar na cama. Isso pode ser neutralizado, mas nesse caso é preciso acrescentar alguma informação explícita, por exemplo, completar [6] com *...e tanto tentou que acabou ficando*.

Tentar não apenas ocorre com complemento no infinitivo (como em [6]), mas também com objeto direto não oracional, como em

[10] Vou tentar a escalada do Everest.

E *tentar* aparece sempre com objeto direto (oracional ou não), ao contrário de *comer*, que pode ter ou não objeto:

[11] *Amanhã vou tentar sozinho³.

[12] Amanhã vou comer uma pizza sozinho.

[13] Amanhã vou comer sozinho.

Além disso, temos que saber também que *tentar* pertence à primeira conjugação, e que é regular. Tudo isso é necessário para conhecer o lexema *tentar*, que é apenas um dos lexemas da língua. Para cada um dos milhares de verbos que utilizamos, é preciso ter pelo menos a quantidade de informações mostrada acima; e para muitos a quantidade é muito maior. E olhe que *tentar* é um verbo bastante simples. Se o verbo fosse *dar*, aí sim teríamos um caso verdadeiramente complexo em mãos. Só para avaliar a riqueza semântica de *dar*, explicito o significado dessa palavra nas frases abaixo:

[14] Guto deu um presente para a irmã.

[15] Guto deu o bolo na namorada.

[16] Guto deu um passeio na beira do rio.

³ Uma frase como [11] só pode ocorrer em contexto anafórico, isto é, quando o SN pode ser recuperado relacionando-o com um outro SN em posição paralela que ocorre antes no texto: *ontem tentei a escalada com Chico, mas amanhã vou tentar sozinho*; mas [13] pode ocorrer independentemente.

- [17] Deu a louca no Guto.
 [18] Guto deu andamento ao projeto.
 [19] Guto nem deu as caras.
 [20] Guto deu a entender que pretendia viajar.
 [21] Deu na vista que o Guto desanimou.
 [22] Guto deu no cachorro até quase matar.
 [23] Essa idéia do Guto não deu certo.
 [24] Não deu pra ligar pro Guto hoje

etc., etc., etc. Acho que qualquer pessoa pode aumentar bastante essa lista. E em cada caso, o significado e o comportamento sintático de *dar* é diferente.

Cada palavra da língua apresenta sua complexidade própria: temos que saber em que construções ela se encaixa, o que ela significa em cada caso, como ela varia (no caso de lexemas complexos, como verbos e nomes). Além disso, precisamos saber se ela participa de expressões idiomáticas; por exemplo, *vida*, além de suas obrigações regulares, ainda aparece, com significado bem diferente, em expressões como

- [25] Você segue essa rua toda vida, até a pracinha.
 [26] Era uma questão de vida ou morte.
 [27] A bolsa ou a vida!
 [28] Estou feliz da vida / danado da vida / putado da vida (*mas não *preocupado da vida, nem *alegre da vida*).
 [29] Estudar lá é um atraso de vida.
 [30] Tem uma semana que ela não dá sinal de vida.

etc.

Precisamos adquirir toda essa informação para cada um dos itens que compõem o nosso léxico mental. Tudo isso é codificado em nossa memória, e em pouco tempo: uma criança de 5 anos já tem um léxico considerável, e consegue utilizar a língua falada com tanta facilidade quanto um adulto. Naturalmente, continuamos a aprender — e a esquecer — novos itens a vida toda.

18.2. LÉXICO E GRAMÁTICA

Nosso conhecimento da língua inclui, portanto, o conhecimento do léxico e da gramática; mas qual é exatamente a diferença entre léxico e gramática? Para muitos linguistas, nenhuma. Mas tradicionalmente se faz uma distinção, e para nossos objetivos podemos adotá-la, deixando as controvérsias teóricas para outro dia.

Vamos distinguir léxico e gramática seguindo a linha entre conhecimento generalizado (que pode ser expresso por regras) e conhecimento individualizado (que se vincula a unidades específicas). Voltando ao verbo *tentar*: vimos acima que esse item se liga semanticamente à forma nominal *tentativa*, que denota a ação de tentar (no sentido de “procurar fazer algo”). Ora, isso não pode ser reduzido a uma regra, porque os outros verbos se vinculam a outras formas nominais, com sufixos diferentes; e ao que tudo indica não se pode prever que sufixo vai aparecer. Por exemplo,

<i>tentar</i>	–	<i>tentativa</i>
<i>cortar</i>	–	<i>corte</i>
<i>pensar</i>	–	<i>pensamento</i>
<i>pintar</i>	–	<i>pintura</i>
<i>contar</i>	–	<i>contagem</i>
<i>realizar</i>	–	<i>realização</i>
<i>chegar</i>	–	<i>chegada</i>

Formas como **tentatura*, **pintagem*, **chegamento* etc. não são aceitáveis; por isso, esses itens precisam ser individualmente marcados para que se possa saber que sufixo nominalizador cada um deles aceita. Esse tipo de informação se coloca no léxico, que seria então um repositório de informações individualizadas.

O lexema *tentar*, como sabemos, inclui formas de tempo e de pessoa, formadas através de sufixos: no presente temos *tento* (2ª pessoa do singular), *tentam* (3ª pessoa do plural), no perfeito temos *tentei* (2ª sing.), *tentaram* (3ª sing.) e assim por diante. Mas isso não vale apenas para *tentar* — os mesmos sufixos, com os mesmos significados, valem para todos os verbos da primeira conjugação: *penso*, *pensam*, *pensei*, *pensaram*; *lavo*, *lavam*, *lavei*, *lavaram* etc. Aqui é possível formular uma regra, e por isso se coloca essa informação na gramática: o repositório das informações generalizadas.

Léxico e gramática são muito intimamente ligados, porque as características dos itens léxicos são em grande parte expressas em termos de categorias gramaticais. Por exemplo, uma palavra como *eu* é marcada (no léxico) como podendo ser sujeito de uma oração, mas não objeto⁴ — e sabemos que oração é uma classe gramatical, e sujeito e objeto são funções gramaticais. A palavra *ruim* pode ocorrer depois do núcleo de um SN (*um livro ruim*), mas não antes do núcleo (**um ruim livro*) — aqui vemos o uso das relações gramaticais de ordem e de núcleo do SN para caracterizar um item léxico.

⁴ Exceto em variedades em que frases como *ajuda eu aqui* são aceitáveis.

Assim, a base para a distinção entre léxico e gramática é a distinção entre informação individual e informação geral; mas nem sempre essa distinção é clara. Como vamos classificar um traço que vale para alguns verbos, mas não todos? Por exemplo, a conjugação de *pôr* é irregular, e poderia ser considerada uma marca léxica desse verbo. Mas as mesmas irregularidades valem para certo número de outros verbos: *dispor, compor, impor, repor, antepor, pospor, propor, contrapor, depor, expor, interpor, justapor, opor, supor, superpor, transpor*. São menos de 20 verbos, o que é pouco para constituir uma “conjugação”, mas muito para que sejam considerados pura e simplesmente exceções⁵. Estritamente falando, não há solução ideal, a não ser que se negue a distinção entre léxico e gramática — que é o que muitos lingüistas atuais fazem. Como é necessário tomar uma decisão, optei por manter a posição tradicional de distinguir léxico e gramática, o que é adequado para a maioria dos casos.

18.3. CONTEÚDO DE UM ITEM LÉXICO

Um item léxico pode ser visto como um pacote bastante heterogêneo de informações, explicitando as propriedades formais, a pronúncia, o significado e tudo o mais que sabemos a respeito de um lexema. Por exemplo, o lexema *cachorro* inclui informações como:

- ▼ a pronúncia é *kaxôrru*;
- ▼ é um nominal;
- ▼ é (gramaticalmente) masculino;
- ▼ tem potencial referencial (isto é pode se referir a um ser);
- ▼ tem potencial qualificativo (isto é, pode denotar uma qualidade);
- ▼ quando qualificativo, é pejorativo⁶;
- ▼ refere-se a um ser animado;
- ▼ relaciona-se com o feminino *cachorra* (e também com *cadela*);
- ▼ é um sinônimo mais usual de *cão*;
- ▼ evoca expectativas várias, tais como “fidelidade”, “ferocidade”, “proteção”...;
- ▼ o animal denotado pode pertencer a diversos tipos: vira-lata, beagle, pastor alemão, fila...

Algumas dessas informações são gramaticalmente relevantes, outras não; mas todas desempenham um papel na compreensão de enunciados que incluam a palavra *cachorro*. Por exemplo, o fato de denotar um ser animado é que permite que *cachorro* apareça como sujeito de *tentar*:

⁵ Sintomaticamente, a gramática tradicional oscila entre considerar *pôr* e seus compostos como irregularidades da segunda conjugação ou como uma conjugação menor separada.

⁶ Nesse uso, significa o que um cachorro não é: desonesto, indigno de confiança, inamistoso. Esse uso da palavra, naturalmente, diz mais dos seres humanos do que dos cachorros.

[31] O cachorro tentou passar por baixo do portão.

E nessa mesma frase *cachorro* é núcleo do SN *o cachorro* (sujeito de *tentou*) porque tem potencial referencial. Como também tem potencial qualificativo, pode ocorrer como modificador como em

[32] Aquele gerente cachorro demitiu o Sérgio.

18.4. A LÍNGUA E O MUNDO

Das características listadas acima, algumas se ligam a nosso conceito de “cachorro” enquanto ser do mundo real, enquanto outras se ligam especialmente a nosso conhecimento do lexema *cachorro*, e por isso são mais importantes do ponto de vista gramatical. Mas essa separação é mais de conveniência do que de princípio, e só tem valor aproximado. Por exemplo, o fato de *cachorro* ser “animado” tem importância gramatical (é o que permite a essa palavra aparecer como núcleo do sujeito de verbos como *tentar*), mas sem dúvida está também relacionado com fatos do mundo extralingüístico — afinal de contas, um cachorro é um ser animado.

O problema é que não se sabe como separar nitidamente esses dois tipos de conhecimento; não há dúvida de que existem coisas puramente gramaticais — por exemplo, o gênero feminino de *mesa* ou a irregularidade de *ir* frente à regularidade de *partir*. E há conhecimentos que não têm muito a ver com a gramática — por exemplo, o de que o cachorro é um mamífero, ou de que tem um tamanho mais ou menos delimitado (sempre maior que um besouro, mas menor que um cavalo). E há uma grande quantidade de conhecimentos que são enraizados na realidade, mas que têm importância para a maneira como construímos nossas frases: é o caso de que a palavra *cachorro* indica uma coisa (e não, por exemplo, uma relação entre coisas)⁷, e de que um cachorro é um ser animado. Por isso, nem sempre é possível separar no item léxico a informação gramatical dos conhecimentos que temos do mundo.

Isso é apenas um aspecto de um fenômeno mais geral: nosso conhecimento da língua é parte de nosso conhecimento do mundo. A separação total é impossível, mesmo porque é função da língua representar o mundo de alguma forma. Mas isso não quer dizer que a separação seja impossível em todos os casos: muitas vezes a língua segue seu caminho próprio, sem consideração de

⁷ Um exemplo de palavra que exprime uma relação entre coisas é *em*: no sintagma *uma verruga no nariz* exprimimos a relação que existe entre a verruga e o nariz.

aspectos do conhecimento do mundo extralingüístico. Ou seja, a língua, sua estrutura e seu vocabulário, é em grande parte arbitrária, sem motivação visível. Os exemplos de traços arbitrários da estrutura lingüística são muitos, como:

- ▼ o agente de certos verbos é colocado antes e não depois do verbo;
- ▼ nesses casos, o agente é indicado redundantemente, por um SN pré-verbal e por um sufixo do verbo;
- ▼ os substantivos se dividem em masculinos e femininos;
- ▼ o modificador precedido de preposição se coloca obrigatoriamente depois do núcleo de um SN;
- ▼ o nome do objeto que usamos para beber água é *copo*;
- ▼ os verbos se dividem em três conjugações, e muitos deles são irregulares;
- ▼ certos qualificativos (como *ruim*) só podem aparecer antes do núcleo do SN;
- ▼ algumas palavras têm o acento na última sílaba;

etc.

Todos esses traços (e muitíssimos outros) são puramente gramaticais; conseqüentemente, não se encontram em outras línguas, que organizam a informação de modo diferente.

Mesmo nos casos em que o traço gramatical é previsível a partir do significado, muitas vezes percebe-se um componente arbitrário que funciona juntamente com o componente motivado. Assim, sabemos que *cachorro* tem potencial qualificativo (no sentido de “desonesto, inamistoso”), e isso nos permite prever alguns aspectos do seu comportamento gramatical. Mas que aspectos? Não só que essa palavra vai aparecer qualificando outra, como em

[32] Aquele gerente cachorro demitiu o Sérgio.

mas que vai aparecer depois da palavra qualificada, e concordando em gênero e número com ela (*uma gerente cachorra, uns amigos cachorros*). Ora, isso já é arbitrário, pois em outra língua poderia ser diferente (colocado antes do núcleo, sem concordância, ou concordando em caso, por exemplo).

Como se vê, os dois aspectos são entrelaçados, e a separação se faz com objetivos de descrição. Não obstante, ela é muito conveniente, e em muitos casos a separação pode ser vista com clareza. Por isso, aqui e na maioria dos trabalhos de lingüística, o conhecimento da língua e o conhecimento do mundo são tratados (pelo menos em um primeiro momento) como coisas separadas. Esse estudo separado proporciona, inclusive, a possibilidade de melhor explicitar as relações entre esses dois componentes do conhecimento. Qualquer afirmação mais detalhada sobre o assunto terá que esperar até que se obtenha uma compreensão melhor do fenômeno.

3ª PARTE:

GRAMÁTICA DO PORTUGUÊS

CLASSES: ALGUNS EXEMPLOS

19.1. ADVÉRBIOS

As palavras que tradicionalmente se classificam como “advérbios” pertencem na verdade a diversas classes nitidamente diferentes quanto ao tipo de significado e quanto ao potencial funcional. Quanto à estrutura interna, freqüentemente não têm nenhuma, sendo formadas de um único morfema (*sempre, nunca, depois, aqui*) ou então se formam por meio do sufixo característico *-mente*. De qualquer forma, é impossível colocar todos os “advérbios” tradicionais em uma única classe. Vamos ver alguns exemplos.

Os dicionários e gramáticas citam como advérbios, entre muitas outras, as seguintes palavras: *depressa, sim, não, só, francamente*. Seria de esperar, portanto, que todas elas tivessem comportamento gramatical semelhante, de pontos de vista relevantes. Mas não é o caso, como é fácil mostrar.

Vamos pegar o critério do potencial funcional. Se essas palavras tivessem potencial funcional semelhante, poderiam se substituir mutuamente (com diferença de significado, claro) em muitos ambientes. Por exemplo, sabemos que *carro* e *camelo* são ambos “substantivos”; e realmente podemos substituir um pelo outro na maioria dos ambientes em que ocorrem:

[1] Meu carro está dando problema. / Meu camelo está dando problema.

[2] Ahmed tem dois carros. / Ahmed tem dois camelos.

[3] Fomos a Bagdá de carro. / Fomos a Bagdá de camelo.

É também de esperar que em certos ambientes essa substituição não seja possível, seja por razões de conhecimento do mundo, como em

[4] Vou pôr gasolina no carro. / ?? Vou pôr gasolina no camelo.

seja porque, afinal de contas, as classificações são apenas aproximadas. Mas uma classificação só faz sentido se houver um alto grau de coincidência no comportamento dos itens — no caso, potenciais funcionais muito parecidos.

Agora consideremos a lista dos “advérbios” tradicionais. Veremos que as diferenças de comportamento são tão grandes que a inconveniência de colocá-los todos em uma só classe é evidente. Por exemplo, *sim* e *não* ocorrem em ambientes profundamente diferentes:

[5] Essa loja não existe. / *Essa loja *sim* existe.

[6] Os não iniciados eram excluídos. / * Os *sim* iniciados eram excluídos.

[7] Não, eu não vou lá. / * *Sim*, eu *sim* vou lá. / * *Sim*, eu não vou lá.

Vemos que o item *não* se liga sintática e semanticamente a um verbo ou a um adjetivo. Também funciona como uma espécie de interjeição, e aí (e só aí) *sim* e *não* se parecem:

[8] Não, minha casa é aqui. / *Sim*, minha casa é aqui.

Depressa pode modificar um verbo, mas ao contrário de *não*, ocorre em geral depois do verbo:

[9] Do Carmo não dirige. / Do Carmo dirige *depressa*. / *Do Carmo *depressa* dirige.

Depressa pode ocorrer modificando um adjetivo, tal como *não*, mas também nesse caso ocorre depois:

[10] Um edifício não terminado / Um edifício terminado *depressa*.

Só pode modificar um adjetivo, um verbo e também um substantivo:

[11] Ele não faz nada, só fala.

[12] D. Marlene não é má, é só maluca.

[13] Só D. Marlene faz uma coisa dessas.

Depressa e *sim* nunca ocorrem modificando um substantivo.

Ao contrário de *depressa* e *só*, *francamente* pode ser usado com grande liberdade de posição na sentença, sem modificar nenhum de seus termos, mas expressando um ponto de vista do falante (nesse aspecto, *francamente* se assemelha a *sim*):

[14] *Francamente*, D. Marlene deve ser maluca.

[15] D. Marlene, *francamente*, deve ser maluca.

[16] D. Marlene deve ser, *francamente*, maluca.

[17] D. Marlene deve ser maluca, *francamente*.

Compare-se com

[18] *Sim*, D. Marlene deve ser maluca.

[19] D. Marlene, *sim*, deve ser maluca.

etc.

A palavra *não* pode expressar a opinião do falante, mas não tem essa liberdade de ocorrência:

[20] Não, D. Marlene deve ser maluca.

mas não

[21] * D. Marlene deve ser, não, maluca.

Não é preciso ir mais longe — esses poucos exemplos mostram a grande variedade de comportamentos gramaticais encontrada no grupo de palavras tradicionalmente chamadas “advérbios”. É claro que o que temos aqui não é uma classe, mas diversas classes. É bastante difícil descobrir traços gramaticais importantes que unam todas elas (ninguém até hoje conseguiu isso), tanto mais que o valor semântico desses itens também varia: alguns modificam qualidades (*bastante feliz*), outros ações (*trabalha cuidadosamente*), outros referentes (*só D. Marlene*); alguns expressam atitudes do falante (*francamente*); alguns negam ou afirmam (*sim, não*). Alguns são especializados: *redondamente* só ocorre na expressão *redondamente enganado*; outros ocorrem em grande variedade de ambientes. Chamar

a todos de “advérbios” é ignorar um mundo de fenômenos gramaticais importantes.

19.2. PREPOSIÇÕES

As preposições costumam ser definidas de forma bastante vaga, tanto é assim que, a se seguir a definição, fica difícil distingui-las das conjunções. Mas não é tão difícil assim definir as preposições. Conforme veremos, a função das preposições é, falando geralmente, a de promover mudanças de classe; nesses casos, vamos dizer que a preposição é um reclassificador.

Vamos começar com alguns exemplos. Sabemos que *minha fazenda* é um sintagma nominal, e que tem portanto um potencial funcional típico: pode ser sujeito, objeto direto ou complemento de uma preposição. Esta última propriedade pode ser exemplificada em

[22] Deixei três cachorros em minha fazenda.

Agora, *em minha fazenda* não é um sintagma nominal (*minha fazenda* é, mas *em minha fazenda* não). Não pode ser sujeito, nem objeto direto, nem complemento de outra preposição:

[23] *de em minha fazenda / *para em minha fazenda etc.

O potencial funcional do sintagma *em minha fazenda* é semelhante ao de algumas das palavras que se chamam tradicionalmente “advérbios”:

[24] Deixei três cachorros em minha fazenda.

[25] Deixei três cachorros aqui.

Ou seja, quando colocamos a preposição *em* na frente do sintagma nominal *minha fazenda*, o conjunto (preposição + SN) é algo que podemos chamar um “advérbio de lugar”. A função da preposição *em*, portanto, é a de fazer de um SN um advérbio de lugar (a preposição *em* tem outras funções, que vou ignorar para não complicar a exposição). Outras preposições que formam advérbios de lugar a partir de SNS são *de* (*cheguei de Teresina*), *para* (*vou mudar para o seu bairro*), *debaixo de* (*o gato está debaixo da mesa*), *entre* (*tem um parafuso entre as moedas*) e mais algumas.

Há também preposições que formam sintagmas do tipo adjetivo (ou seja, que podem ser modificadores). A mais importante é *de*. Assim, podemos dizer:

[26] A porteira de minha fazenda

O sintagma *de minha fazenda* não é um SN, e tem função semelhante à dos adjetivos:

[27] A porteira nova

[28] A porteira branca

Essas são as duas funções principais das preposições: transformar SNS em advérbios¹ ou em adjetivos. Em alguns outros casos, uma preposição pode aparecer sem que haja essa transformação, notadamente quando introduz o complemento de um verbo:

[29] Gosto de você.

[30] Conto com você.

[31] Acredito em você.

Na verdade, ninguém sabe por que essa preposição tem que aparecer aí.

19.3. CONJUNÇÕES

Agora passemos às conjunções. A situação aqui é menos confortável, porque a classificação tradicional é falha, e a chamada classe das conjunções inclui itens de comportamento bem diverso.

Algumas conjunções funcionam de maneira algo semelhante às preposições. Por exemplo, a conjunção *que* tem o poder de fazer de uma oração um SN. Assim, temos uma oração em

[32] Vai chover hoje.

mas se acrescentarmos *que*, a seqüência resultante tem o potencial funcional típico de um SN:

¹ De certo tipo, claro. Já vimos que a classe dos “advérbios” é um saco de gatos.

- [33] Neli disse que vai chover hoje. (objeto direto)
 [34] Que vai chover hoje é evidente. (sujeito)
 [35] A notícia de que vai chover hoje é falsa. (complemento de preposição)

Essas construções formadas através de *que* não são totalmente idênticas aos demais SNS, mas a semelhança de potencial funcional é suficiente para que as analisemos, em termos aproximados, como SNS — ou seja, possuem uma boa parte dos traços que caracterizam um SN “prototípico” como *a casa amarela*. Um item que funciona mais ou menos como *que é se*²:

- [36] Ninguém sabe se vai chover hoje.

Outras palavras tradicionalmente chamadas “conjunções” têm a propriedade de reclassificar orações, mas de outras formas. Assim, *quando* forma seqüências de valor “adverbial”, como em

- [37] Hélio saiu de casa quando Antônia começou a gritar.

A seqüência de *quando* mais a oração *Antônia começou a gritar* tem o potencial funcional de uma palavra como *ontem*:

- [38] Hélio saiu de casa ontem.

Vemos que *quando* mais uma oração forma uma construção equivalente a um advérbio de tempo. Assim, *quando*, *que* e *se* podem ser chamadas de reclassificadores, tal como em geral as preposições.

Todas as conjunções vistas acima são da subclasse tradicional das “subordinativas”. Há ainda as “coordenativas”, das quais a mais típica é *e*:

- [39] Hélio saiu de casa e Antônia correu atrás dele.

Essa palavra tem como função unir duas ou mais construções de mesma classe (oração + oração, SN + SN etc.), de tal modo que a seqüência resultante pertence à mesma classe dos componentes. Assim, a seqüência SN e SN é, ela própria, um SN, como se pode ver em

² Em uma de suas funções; *se* também pode introduzir um termo de valor “adverbial”, como em *só vou se você for*.

- [40] [Hélio]_{SN} foi ao cinema.

- [41] [Hélio e Antônia]_{SN} foram ao cinema.

A função de *e* é, portanto, nitidamente diferente da das chamadas “conjunções”; por isso, talvez seja melhor chamar palavras como *e* de coordenadores, e não conjunções. Uma palavra que se comporta de maneira semelhante à de *e* é *ou*.

19.4. MAIS COMPLEXIDADES

A descrição acima está longe de esgotar o que se pode dizer sobre a classificação dessas palavras. Voltando às preposições, temos que distinguir as que podem ocorrer em seqüência, como em

- [42] Ela sempre foi caridosa para com os doentes.

Aqui a primeira impressão é de uma preposição (*para*) cujo complemento igualmente se inicia por uma preposição (*com*). Isso refutaria a afirmação de que uma preposição sempre toma um SN como complemento. A solução tradicional, que evita essa objeção, é considerar *para com* uma preposição única, mas composta; haveria algumas dessas em português, chamadas “locuções prepositivas”. Mas os outros exemplos citados nunca se compõem de duas preposições, sendo em geral formadas de uma palavra de outra classe, combinada com uma ou mais preposições: *além de*, *em cima de*, *antes de*, *com que* etc. Evidentemente, há necessidade de pesquisa neste setor.

Há ainda uma observação interessante a fazer sobre as preposições. Algumas delas podem ocorrer sozinhas (isto é, sem um SN seguinte) em contexto anafórico, isto é, quando a referência do SN pode ser recuperada relacionando-o com outro SN em posição paralela que ocorre antes no texto. Assim, podemos dizer

- [43] Ela veio com o material e eu vim sem.

Entende-se, evidentemente, “sem o material”. Algumas preposições aceitam essa ocorrência sem SN:

- [44] Justino é a favor do governo, mas Faustina é contra.

Mas a maioria não aceita:

[45] *Ela veio sem o material e eu vim com.

[46] * Um dos aparelhos estava embaixo da mesa, e o outro estava sobre.

Como se vê, há aqui mais um critério de subclassificação do grupo a que chamamos “preposições”. Falta ainda fazer um mapa das possibilidades de ocorrência de preposições sozinhas.

19.5. NOMINAIS

Tenho usado mais ou menos livremente o termo nominal para designar uma classe de palavras — evitando os termos tradicionais adjetivo e substantivo. Já vimos (cap. 1) que essas duas classes tradicionais não conseguem exprimir o comportamento dos itens do ponto de vista de seu potencial funcional. Em particular, temos itens que podem ser núcleo do SN, itens que podem ser modificadores, e itens que podem ser as duas coisas; nossos exemplos, no capítulo 1, foram, respectivamente, *livro*, *interessante* e *velho*. Todos esses itens pertencem à classe dos nominais, e esta se define através da possibilidade de desempenhar uma das funções de núcleo do SN, modificador, ou as duas.

A vantagem de vincular assim essas duas funções vem do fato de que as palavras em português que têm uma dessas propriedades mostram grande facilidade de adquirir a outra. Assim, nomes próprios como *Kubitschek* parecem ser apenas a designação de um ser (ou seja, seriam exclusivamente referenciais); mas é possível utilizá-los como modificadores em expressões como *o governo Kubitschek*.

Historicamente, pode-se citar muitos exemplos de nominais qualificativos que adquiriram referencialidade: *celular* e *maternal* eram puramente qualificativos, mas acabaram sendo usados para designar coisas. O fenômeno oposto é menos comum, mas também se verifica: é o caso de *cabeça*, que designava apenas uma coisa (tipicamente, uma parte do corpo), mas que passou a ser usado como qualificativo: *um filme cabeça*. Essa flexibilidade é muito típica da sintaxe do português brasileiro, e afeta exatamente o grupo de palavras a que chamo nominais.

A classe dos nominais tem alguns traços gramaticais importantes em comum: por exemplo, só nominais podem fazer o plural em *-s*, e só nominais podem variar em gênero. Esses traços, entretanto, não valem para todos os nominais: o nominal *baita*, como em *um baita problema*, não varia em gênero. Encontramos aqui, como de costume, a complexidade que os estudos gramaticais ainda não enfrentaram devidamente.

SINTAGMAS ORACIONAIS

20.1. ORAÇÕES SUBORDINADAS

Vamos considerar o exemplo seguinte:

[1] Neli disse que vai chover hoje.

Vimos no capítulo 19 que a seqüência sublinhada (composta de uma conjunção mais uma oração) constitui um SN, que no caso é objeto direto de *disse*. As razões para analisar essa seqüência como um SN são:

- ▼ ela complementa o verbo *dizer*, o verbo *dizer* nunca aparece sem objeto direto;
- ▼ por isso, em [1] também deve haver um objeto direto de *dizer*;
- ▼ essa mesma seqüência, em outras frases, pode ser sujeito ou complemento de preposição:

[2] Que vai chover hoje é só a sua opinião. (sujeito)

[3] Acabou de dar no rádio a notícia de que vai chover hoje. (complemento da prep. *de*)

- ▼ ou seja, a seqüência *que vai chover hoje* é um SN, porque um SN é a unidade que pode ser sujeito, objeto direto ou complemento de preposição.

Vemos então que pode haver orações dentro de SNS; e como o SN se encaixa dentro de outra oração, pode haver orações dentro de orações. Essa propriedade de colocar unidades dentro de outras unidades da mesma classe é fundamental nas línguas humanas, e é o que nos permite construir um número ilimitado de frases distintas.

Outros exemplos de sintagmas com conteúdo oracional são:

- ▼ oração dentro de um sintagma adjetivo:

Temos isso na frase

[4] A notícia de que vai chover hoje preocupou titia.

A seqüência *de que vai chover hoje* funciona como um delimitador referencial para *notícia*, deixando claro que não se trata de qualquer notícia que preocupou titia, mas a notícia *de que vai chover hoje* (como poderia ser a notícia *divulgada pelo rádio*, ou a notícia *trágica*). Assim tem a função semântica de um adjetivo; e sintaticamente também se coloca em uma das posições do adjetivo, logo depois do núcleo do SN, tendo a função de modificador.

▼ Oração dentro de um sintagma adverbial:

[5] Recolhi a roupa quando a vizinha me mostrou as nuvens.

Aqui o sintagma adverbial é *quando a vizinha me mostrou as nuvens*, que contém a conjunção *quando* e a oração *a vizinha me mostrou as nuvens*. O *sadv* ocorre onde poderia ocorrer *no fim da tarde*, ou *ontem*, todos eles com o significado de “momento em que ocorreu o evento descrito pelo verbo”, e todos eles na mesma posição sintática.

Como se vê, colocamos na mesma classe unidades de estrutura interna muito diferente; assim, as seqüências *Paula*, *o gato*, *aquele gato branco*, *que vai chover amanhã* se classificam todas como *sns* — isso porque têm o mesmo potencial funcional (ver o capítulo 16 a respeito da importância do potencial funcional na classificação das unidades gramaticais). No caso, todas podem ser sujeito, objeto direto e complemento de preposição.

As orações que acabamos de ver, colocadas dentro de outra unidade (dentro de *sns*, *sadjs* ou *sadv*s) se denominam orações subordinadas. Ou seja, uma oração subordinada é aquela que faz parte de outra estrutura — por exemplo, de um *sn* que, por sua vez, é parte de outra oração maior. Uma oração subordinada é, portanto, parte de outra oração.

20.2. ESTRUTURAS ORACIONAIS ESPECIALIZADAS

20.2.1. Subjuntivo

Nos exemplos vistos acima, encontramos orações de forma usual inseridas dentro de outras estruturas; assim, a oração contida no sintagma

nominal *que vai chover amanhã* é “normal”, e poderia ocorrer sozinha como um enunciado independente: *vai chover amanhã*¹. Mas isso nem sempre acontece; há casos em que a oração inserida no sintagma tem uma forma especializada, só podendo ocorrer como subordinada, nunca como enunciado independente.

Uma dessas adaptações é o que se chama tradicionalmente “modo subjuntivo”. Na frase

[6] O chefe queria que vocês lavassem o pátio.

a seqüência sublinhada é objeto direto, logo um *sn*. Este é composto da nossa conhecida conjunção *que*, mais algo que se assemelha muito a uma oração: *vocês lavassem o pátio*. Essa seqüência parece uma oração porque tem um *sn* sujeito (*vocês*), seguido de um verbo, mais outro *sn*, que é o objeto direto.

Mas há um problema: *vocês lavassem o pátio* não poderia aparecer sozinha, como um enunciado independente; teria que ser algo como *vocês lavaram o pátio*, ou *vocês lavam o pátio*, com outra forma do verbo. Nesses casos, a solução tradicional é analisar *vocês lavassem o pátio* como oração, apesar de ser atípica porque não poderia ocorrer sozinha. A razão é que há traços que aproximam essa seqüência das orações: ela ocorre com a conjunção *que*, formando um *sn*; e apresenta vários traços da estrutura interna da oração como, no caso, um verbo, um sujeito e um objeto direto. O problema é apenas a forma do verbo, que aparece no subjuntivo. O subjuntivo, aliás, só aparece em orações subordinadas², e pode portanto ser considerado uma forma especializada, ao contrário do indicativo, que pode ocorrer tanto em subordinadas quanto em não-subordinadas. Outros exemplos de uso do subjuntivo em subordinadas são

[7] É uma pena que ela seja tão mal-humorada.

[8] Só vou telefonar quando você estiver em casa.

¹ Veremos no cap. 21 por que a oração é apenas *vai chover amanhã*, e não *que vai chover amanhã*.

² Há uma exceção, que são fórmulas de desejo como *Deus te ajude*. Um outro caso de forma especializada é o imperativo, que ao contrário do subjuntivo nunca aparece em orações subordinadas.

20.2.2. Infinitivo

Um segundo tipo de oração de tipo especializado, que só ocorre como subordinada, é a de infinitivo³. Temos exemplos nas frases

[9] O deputado declarou ser inocente.

[10] É uma pena esse restaurante ter fechado.

As seqüências sublinhadas apresentam muitas das características das orações em geral, embora nem todas; por exemplo, em certos casos não pode haver concordância verbal:

[11] *Os deputados resolveram entregarem os documentos.

No entanto, a maioria dos traços sugere que vale a pena chamar essas seqüências de orações, embora “reduzidas” porque não apresentam o potencial funcional completo das outras orações. Primeiro, em alguns casos a concordância se verifica:

[12] É uma pena esses restaurantes terem fechado.

Depois, o infinitivo mantém todos os complementos do verbo correspondentes (suas diáteses). Por exemplo, sabemos que o verbo *gostar* tem um complemento regido da preposição *de*:

[13] Ariadne gosta de banana.

Pois bem, quando *gostar* aparece no infinitivo é acompanhado pelo mesmo complemento:

[14] É uma pena Ariadne gostar de banana.

Outro exemplo: *engordar* tem sujeito agente quando há objeto direto na oração, caso contrário o sujeito é paciente:

³ Também aqui há uma exceção, quando o infinitivo se usa com valor imperativo: *olhar à direita!*

[15] A cerveja engordou Ariadne. (*a cerveja* é agente)

[16] Ariadne engordou. (*Ariadne* é paciente)

O mesmo se verifica se *engordar* aparecer no infinitivo:

[17] O perigo é a cerveja engordar Ariadne. (*a cerveja* é agente de *engordar*)

[18] O perigo é Ariadne engordar. (*Ariadne* é paciente de *engordar*)

Assim, temos no infinitivo a maior parte da estrutura oracional, o que nos autoriza a falar de “orações de infinitivo”. Como vimos (e com a pequena exceção do uso imperativo), essas orações só aparecem como subordinadas, ou seja, como parte de outra oração.

20.2.3. Gerúndio

A situação do gerúndio é algo semelhante à do infinitivo: pelas mesmas razões, geralmente se fala de “orações de gerúndio” (embora o gerúndio, ao contrário do infinitivo, nunca concorde com seu sujeito). Assim, as diáteses dos verbos *gostar* e *engordar*, vistas acima, se mantêm quando esses verbos estão no gerúndio:

[19] Minha noiva gostando do vestido, a gente pode comprar.

[20] Não vou ficar tomando cerveja, e ela me engordando o tempo todo.

[21] Você engordando um pouco, o vestido vai cair melhor.

A diferença é que, enquanto a oração de infinitivo é um SN, a oração de gerúndio é um sintagma adverbial. Nesses casos, não há conjunção; a própria forma do verbo sinaliza o caráter nominal ou adverbial da estrutura — uma oração de infinitivo é sempre um SN, uma de gerúndio é sempre um *sadv*⁴.

20.3. INFINITIVO E GERÚNDIO COM VERBOS AUXILIARES

Vimos acima que o infinitivo e o gerúndio podem ser núcleos de orações subordinadas. Mas, como quase nada é sintples na língua, essas

⁴ Alguns linguistas preferem dizer que o *-r* do infinitivo e o *-ndo* do gerúndio é que são os reclassificadores, exercendo função paralela à das conjunções.

mesmas formas podem aparecer em uma função bastante diferente, ao lado de verbos chamados "auxiliares". Por exemplo, nas frases

[22] Eduardo vai chegar às quatro horas.

[23] Eduardo estava dormindo.

há boas razões para se acreditar que há apenas uma oração, e não duas. Em primeiro lugar (e repetindo o que vimos no capítulo 17), observa-se que nesses casos só um dos verbos é que tem valência; o verbo auxiliar (*vai* em [22], *estava* em [23]) não faz exigências quanto aos complementos que podem ou não aparecer na oração. Ou seja, não é possível encontrar um complemento que seja aceito pelo verbo principal (*chegar*, *dormir*), mas que seja excluído pelo auxiliar (*vai*, *estava*).

A situação é oposta quando há duas orações; assim, na frase

[24] Eduardo me atendeu comendo.

tanto as exigências de *comer* quanto as de *atender* precisam ser satisfeitas. Assim, é estranho dizer

[25] ?? As galinhas me atenderam comendo.

porque, embora galinhas possam comer, em geral não atendem as pessoas. E igualmente é estranho

[26] ?? A secretária eletrônica atendeu comendo.

porque secretárias eletrônicas atendem, mas não comem. Conclui-se que em [24] temos duas orações, e portanto dois conjuntos de exigências a serem satisfeitas (as de *comer* e as de *atender*); já em [23] só existem as exigências de *dormir*, porque *estar* (como auxiliar) não faz exigências. Dizemos então que em [23] há uma oração apenas.

PERÍODO COMPOSTO¹

21.1. SUMÁRIO: O PERÍODO COMPOSTO

Neste capítulo vou resumir e comentar algumas idéias introduzidas nos últimos capítulos, e que muitas vezes são apresentadas de maneira confusa nas gramáticas tradicionais.

Conforme vimos, a língua permite colocar orações dentro de orações, de modo que na frase

[1] A menina disse que o cachorro está doente.

a oração *o cachorro está doente* está dentro da oração maior *a menina disse que o cachorro está doente*. Esse fenômeno é chamado subordinação. A oração que fica dentro da outra (ou seja, *o cachorro está doente*) é a oração subordinada, e a oração maior é a oração principal.

Até aí, tudo é razoavelmente claro. Mas o ensino tradicional introduziu certas práticas de análise que vão se eternizando, e precisam ser revistas, porque são fundamentalmente incorretas. O problema principal tem a ver com a delimitação das orações, e vou examiná-lo nas seções seguintes.

21.2. LIMITES DA ORAÇÃO SUBORDINADA

Primeiro, muitas vezes se afirma que a oração subordinada em [1] é *que o cachorro está doente*. No entanto, essa seqüência não tem a estrutura

¹ Este capítulo é do interesse de leitores que tenham alguma formação em análise sintática tradicional, porque o objetivo aqui é o de apontar e corrigir certas inadequações dessa análise.

de uma oração. Uma oração típica é formada de um sintagma nominal mais um sintagma verbal (um verbo, mais eventuais complementos). Se examinarmos [1], encontraremos, depois do *que*, esses elementos na ordem esperada: o SN é *o cachorro*, o verbo é *está*, o complemento (predicativo do sujeito) é *doente* — poderíamos inclusive, nesse caso, utilizar essa seqüência como uma oração independente:

[2] O cachorro está doente.

Tudo isso mostra que essa seqüência é uma oração. Mas a análise tradicional diz que a oração subordinada é *que o cachorro está doente*, e aí temos um *que* que fica sobrando: pelo que se sabe, as orações não começam com conjunções, mas freqüentemente com o sujeito.

Esse elemento *que*, portanto, não deve fazer parte da oração (não é parte nem do sujeito nem do predicado, para usar os termos tradicionais). *Que* é, como vimos no capítulo 19, uma partícula cuja função é a de formar um SN a partir de uma oração. Ou seja, em [1] a seqüência *que o cachorro está doente* não é uma oração, mas um SN (objeto direto de *disse*). A análise correta é que a oração subordinada é apenas *o cachorro está doente*, que tem a estrutura típica de uma oração — o que nos permite inclusive analisar *que o cachorro está doente* como um SN, dado que ocorre em posição e função típicas de SNS: poderíamos, por exemplo, substituí-la por um SN não oracional, como em

[3] A menina disse uma verdade.

21.3. LIMITES DA ORAÇÃO PRINCIPAL

Mas não fica aí a inadequação da prática tradicional de análise. Segundo a tradição, ao se analisar [1],

[1] A menina disse que o cachorro está doente.

distinguem-se duas orações: a subordinada, que, como vimos, é incorretamente segmentada como *que o cachorro está doente*, e a principal, que é dada como sendo *a menina disse*. Aqui também há uma incorreção, que passamos a examinar.

Podemos começar perguntando: o objeto direto de uma oração faz parte dessa oração? A resposta é, obviamente, “sim”. Assim, em

[3] A menina disse uma verdade.

uma verdade é objeto direto e faz parte da oração, que é *a menina disse uma verdade*. Isso é tão óbvio que fica até estranho fazer a pergunta; é como se perguntássemos: a cabeça da Flávia faz parte da Flávia? ou: a capital de Sergipe faz parte de Sergipe? É claro que sim: tanto a Flávia sem sua cabeça quanto Sergipe sem Aracaju estariam incompletos. Essa é a relação elementar parte-todo, que aplicamos a todo tipo de objeto minimamente complexo.

Tendo isso em mente, vamos voltar à afirmação (endossada pela gramática tradicional) de que, em [1], a seqüência *que o cachorro está doente* é o objeto direto do verbo *disse*. Ora, já sabemos que o objeto direto de uma oração faz parte dessa oração. Mas (segundo a análise tradicional) a oração principal — aquela cujo verbo é *disse* — é apenas *a menina disse*; e acontece que *que o cachorro está doente* não é parte de *a menina disse* (as partes de *a menina disse* são: *a*, *menina* e *disse*). Mas aqui teríamos um caso em que o objeto direto de uma oração não é parte dessa oração? Ou, talvez, um caso em que a cabeça de uma pessoa não é parte dessa pessoa? Alguma coisa deve estar errada.

Ainda mais quando consideramos a chamada “oração principal”, que seria (segundo essa análise) *a menina disse*. Essa seqüência tem todas as marcas de ser uma estrutura incompleta: temos aí sujeito e verbo, mas não objeto direto. E sabemos que o verbo *dizer* nunca aparece sem objeto direto:

[4] *A menina disse.

A solução é aplicar à frase [1] os mesmos princípios que aplicamos à frase [3]: a oração cujo verbo é *disse* é simplesmente a seqüência completa: em [3], a oração (a única, nesse caso) é *a menina disse uma verdade*; e em [1] a oração principal é *a menina disse que o cachorro está doente*, e a subordinada é *o cachorro está doente*.

21.4. RESGATANDO A TEORIA (MAS NÃO A PRÁTICA) TRADICIONAL

Vamos examinar as conseqüências da divisão em orações exposta acima.

Primeiro, podemos continuar dizendo que o objeto direto é (em todos os casos) uma parte da oração, já que

que o cachorro está doente

é uma parte de

a menina disse que o cachorro está doente

(exatamente como *uma verdade* é parte de *a menina disse uma verdade*). A oração principal de [1] é

a menina disse que o cachorro está doente

e a subordinada é

o cachorro está doente.

e a segunda é parte da primeira.

Quanto a *a menina disse*, não é a oração principal. Na verdade, não é uma oração. Antes, é uma oração amputada de seu objeto direto; um pedaço de oração, que nem sequer é um constituinte. Isso significa que *a menina disse* em [1] não tem função sintática nenhuma, porque só os constituintes têm função sintática (ver em 3.3.2 a definição e as propriedades dos constituintes).

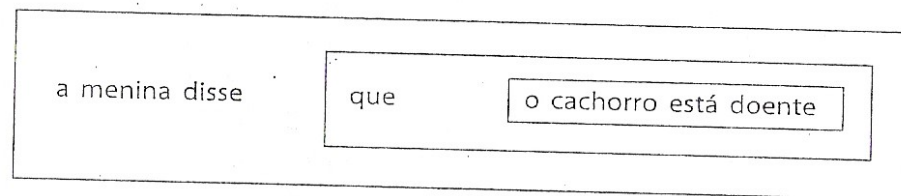
Depois, podemos continuar afirmando que uma oração subordinada é aquela que é parte de outra oração: *o cachorro está doente* é realmente parte da oração principal, que é *a menina disse que o cachorro está doente*.

Finalmente, podemos até mesmo salvar a afirmação tradicional de que a oração principal é aquela que “tem o sentido completo”: é o caso de *a menina disse que o cachorro está doente*, mas certamente não o de *a menina disse*.

O curioso é que as gramáticas tradicionais em geral fazem as mesmas afirmações (corretas): uma oração subordinada é aquela que é parte de outra oração; uma oração principal é aquela que tem outra oração dentro de si; uma oração à qual se retira o objeto direto se torna uma oração incompleta; o objeto direto é parte da oração a que pertence; e assim por diante. No entanto, em geral se analisa a estrutura [1] (incorretamente) como contendo a oração principal (??) *a menina disse*. Há aqui uma incoerência, que fica sanada se aceitarmos a análise exposta acima.

Em resumo, a análise de [1] nos revela a presença de duas orações: uma delas é subordinada, porque está dentro de outra: esta é *o cachorro*

está doente; a outra é principal, porque contém outra oração: trata-se de *a menina disse que o cachorro está doente*. A seqüência *que o cachorro está doente* é formada de *que* mais uma oração, e constitui um SN. Esse SN é objeto direto da oração principal (relembro que os objetos diretos são sempre SNS). Graficamente, fica assim:



O quadrado maior corresponde à oração principal; dentro dele temos um SN, representado pelo quadro imediatamente inserido no maior; e dentro do SN temos a oração subordinada, representada pelo quadrado menor. Como se vê, temos aqui um exemplo de “estruturas dentro de estruturas”, tal como exposto no capítulo 11.

CLASSES: TODOS, ESSES E CINCO

22.1. TRÊS PALAVRAS

Vamos agora considerar alguns itens que aparecem como componentes do sintagma nominal, tipicamente antes do núcleo. Por exemplo, pode-se dizer

[1] Todos esses cinco jogadores estão suspensos.

O SN é *todos esses cinco jogadores*, e o núcleo é *jogadores*. Aqui vamos nos ocupar dos elementos pré-nucleares, a saber, *todos*, *esses* e *cinco*; veremos que eles fazem parte de classes bem delimitadas.

Mas primeiro temos que considerar suas funções, que serão necessárias para estabelecer seus potenciais funcionais. Segundo a análise tradicional, esses três elementos teriam todos a mesma função, a de “adjunto adnominal”. Mas essa análise é inadequada, pois ignora o fato de que eles se comportam cada um à sua maneira, e bem diferentemente uns dos outros. Vou mostrar que, ao contrário do que diz a gramática tradicional, temos aqui três funções diferentes. Ou seja, a estrutura interna do SN é bem mais rica do que dá a entender a análise tradicional, que só inclui um núcleo e um número variável de adjuntos adnominais¹.

Para analisar esses itens, vamos observar suas possibilidades de ordenação (para *todos* e *esses*, isso já foi feito no capítulo 13, mas vou repetir parte da argumentação aqui). Limitando-nos à área antes do núcleo *jogadores*, verifica-se que *todos* só pode aparecer na primeira posição:

[2] *Esses todos cinco jogadores / *esses cinco todos jogadores

¹ Mais, ocasionalmente, um aposto, caso este faça parte do SN (a gramática não esclarece isso).

Já *esses* pode aparecer em segundo lugar (depois de *todos*) ou então em primeiro lugar, caso não ocorra *todos* em primeiro:

[3] Todos esses cinco jogadores / esses cinco jogadores / *cinco esses jogadores

Finalmente, *cinco* aparece depois de *todos* e *esses*, caso estes ocorram, e só pode aparecer em primeiro lugar se nenhum dos outros dois elementos ocorrer:

[4] Cinco jogadores

Essas relações posicionais se resumem assim:

Antes do núcleo, a ordem é fixa: *todos* + *esses* + *cinco*².

Como se vê, cada um desses itens tem seu comportamento gramatical próprio dentro do SN, tal como se pode julgar pelas relações de posição. Eles não podem ser colocados um na posição de outro, sob pena de inaceitabilidade. Em outras palavras, eles desempenham funções diferentes, e têm conseqüentemente potenciais funcionais diferentes. E por isso, nos termos vistos no capítulo 15, essas três palavras pertencem a classes diferentes.

22.2. CLASSIFICANDO

22.2.1. Todos

Mas a que classes, exatamente? Isso vai depender de outros itens que encontrarmos, que tenham os mesmos potenciais funcionais. Assim, vamos começar por *todos*, cujo potencial se resume³, como vimos, à ocorrência como primeiro termo do SN: que outras palavras da língua têm a mesma propriedade?

É interessante notar que, considerando a língua falada, não existe nenhuma outra palavra com o potencial funcional de *todos*; ou seja, pelo menos nessa variedade da língua, *todos* está sozinho em sua classe. Na

² *Todos* também pode, pelo menos para a maioria dos falantes, ocorrer depois do núcleo: *esses jogadores todos estão suspensos*. Vamos deixar de lado essa possibilidade, para manter a discussão simples; mas é claro que ela tem que ser considerada em uma análise completa.

³ Sempre deixando de lado sua ocorrência após o núcleo.

língua escrita, *todos* tem um companheiro, que é a palavra *ambos*, e apenas ela. Ou seja, de qualquer maneira trata-se de uma classe muito pequena, com um ou dois membros apenas. Logo abaixo veremos que talvez haja mais uma palavra na classe, mas não podemos ter certeza.

Pode parecer um desperdício definir uma classe para englobar apenas uma palavra; mas essas coisas acontecem. Quando colocamos *todos* (na língua falada) sozinho em sua classe, estamos com efeito afirmando que os falantes precisam aprender uma coisa a respeito dessa palavra que não vale para nenhuma outra, a saber, que ela tem a preferência total para ocorrer em primeiro lugar no SN. Uma pessoa que não souber isso não sabe construir SNs corretos na língua — vai tentar, por exemplo, dizer **esses todos rapazes*. Para evitar tais construções malformadas, é indispensável utilizar informação que se refere à palavra *todos*, e a nenhuma outra.

22.2.2. Esses

Agora vamos passar à palavra *esse*. Do ponto de vista que nos interessa, podemos dizer que ela se caracteriza por ocorrer em segundo lugar no SN quando há *todos*, e em primeiro lugar em caso contrário. Agora temos que procurar outras palavras que tenham o mesmo comportamento.

Existem algumas palavras assim, de modo que *esse* não é um solitário como *todos*. As palavras análogas a *esses* são *o*, *um*, *aquêle*, *este*⁴ e talvez *o* que interrogativo e exclamativo. Essas palavras ocorrem exatamente no lugar que *esse* pode ocupar no sintagma:

[5] Todos estes cinco jogadores estão suspensos.

[6] Todos os cinco jogadores estão suspensos.

[7] Todos aqueles cinco jogadores estão suspensos.

etc.

Essas frases nos dão a distribuição de *o*, *este* e *aquêle*. O item *um* não pode ocorrer no mesmo sintagma com *todos*, mas pode ocorrer em certos casos com o singular *todo*, e nesse caso aparece no mesmo lugar que *esse*:

[8] Isso se deve a [todo um complexo de fatores]_{SN}.

⁴ E seus femininos e plurais, evidentemente.

Se não houver *todo(s)*, essas palavras ocorrem em primeiro lugar, tal como *esse*. Já o *que* exclamativo e interrogativo⁵ é mais difícil de testar, porque não pode co-ocorrer com nenhuma forma de *todo*. O fato é que quando ocorre fica em primeiro lugar:

[9] Que jogadores! / Que jogadores?

Essa distribuição muito limitada dificulta a determinação de sua função no sintagma; a rigor, poderia ser idêntica à de *todos*. Em outras palavras, examinando [9] não podemos realmente saber se *que* está ali porque não há ocorrência de *todos* ou se é porque pertence à mesma classe de *todos*, e por isso está em primeiro lugar.

Como decidir? Simplesmente não há maneira de decidir, porque falta evidência. A se considerar a ordenação dos termos no SN, *que* pode ficar na classe de *todos* ou na de *esse*, e de qualquer maneira as coisas vão funcionar. Em inglês, o correspondente de *que*, a saber *what*, se comporta como *todos* (ou melhor, como *all*), porque pode aparecer antes de *a*, que corresponde ao nosso *um*:

[10] What a player! 'que jogador!' [literalmente: que um jogador!]

Mas os fatos de uma língua não são relevantes para outra, de modo que em português a questão tem que ficar em suspenso. Menciono a situação em inglês apenas para ilustrar como o problema poderia ser resolvido.

22.2.3. Cinco

Finalmente, vamos examinar as propriedades de *cinco*. Como vimos, esse item aparece em terceiro lugar (se houver *todos*); em segundo (se houver alguma palavra da classe de *esse*); ou em primeiro, se nenhuma dessas alternativas se verificar. Vamos procurar outras palavras com o mesmo comportamento.

Aqui, claro, não vai haver dificuldade: já devemos desconfiar que *quatro*, *seis*, *oito* e *vinte* devem funcionar como *cinco*, e isso é verdade:

⁵ Estou considerando apenas essa acepção do *que*, o que é uma simplificação a rigor inadequada; mas se formos examinar todos os usos da palavra *que*, como se deve fazer, acabaremos indo muito longe, e perdendo de vista o tema do capítulo. Assim, peço paciência para o pequeno truque pedagógico de "fingir" que o *que* exclamativo/interrogativo é uma forma separada de pleno direito.

[11] Todos esses quatro / seis / oito / vinte jogadores estão suspensos.

Parece então que *cinco* representa um grande número de elementos, a saber, os numerais cardinais.

Mas há também uma surpresa: a palavra *um* não funciona como *cinco*, porque pode ocorrer ao lado dos outros cardinais, e sempre antes deles:

[12] Vou convocar uns vinte jogadores.

Uma saída, claro, é dizer que há duas palavras *um*: uma é numeral (como *cinco*), a outra é artigo (como *o*); essa é a saída tradicional, e de certo modo faz sentido.

22.3. DEFININDO AS CLASSES

Como ficamos, então, no que diz respeito à classificação dos itens em exame: *todos*, *esse*, *cinco*?

Já sabemos que cada um deles pertence a uma classe diferente. Agora falta definir essas classes em termos do potencial funcional que caracteriza cada uma delas. Para expressar o potencial funcional, vou adotar algumas abreviações, a saber:

- [+1] significa "pode ocorrer em primeiro lugar no SN";
- [+2] significa "pode ocorrer em segundo lugar no SN";
- [+3] significa "pode ocorrer em terceiro lugar no SN".

Quando um item não tem a propriedade em questão, será marcado negativamente; assim, [-2] significa "não pode ocorrer em segundo lugar no SN" etc. Uma propriedade, positiva ou negativa, como [+2] ou [-1], se chama traço. O potencial funcional de uma palavra se expressa por um certo número de traços.

Agora podemos dar os potenciais funcionais dos itens discutidos; começando pelos casos claros, temos:

- todos* (e, no português escrito, *ambos*) tem potencial funcional [+1, -2, -3];
- o*, *esse*, *este*, *aquele* têm potencial funcional [+1, +2, -3];
- dois*, *três*, *quatro*, *cinco*... têm potencial funcional [+1, +2, +3].

Aí estão as três classes, com alguns de seus membros. No caso das duas primeiras classes, a listagem é quase completa (só faltam, possivelmente, os

casos duvidosos mencionados acima). Mas a terceira dessas classes tem um número maior de membros, porque são, pelo menos, todos os cardinais que se expressam por uma única palavra⁶ (com a possível exceção de *um*).

Os itens *que* e *um*, como vimos, são problemáticos; o exame do comportamento gramatical de *que* só permite saber que essa palavra é [+1], mas não nos fornece dados para estabelecer os outros traços — não temos meios de saber se *que* é [+2] ou [-2], porque ele nunca ocorre nem com *todos* nem com *esse* e seus companheiros de classe. Assim, o potencial funcional dessa palavra fica incompleto, e sua classificação é desconhecida.

A incerteza quanto à classificação de *que* e *um* pode parecer muito incômoda para as pessoas acostumadas a recorrer à gramática (ou aos gramáticos) em busca de respostas categóricas e definitivas a suas dúvidas. O que estou dizendo aqui (e que pode desagradar a alguns) é que nem tudo se sabe a respeito da estrutura da língua — pior, nem tudo pode ser descoberto, porque em certos casos falta evidência e não vemos onde obtê-la.

Mas essa é a situação que ocorre em todas as ciências empíricas. Há um ou dois anos foi descoberto um planeta em outro sistema solar. No momento⁷, entretanto, os astrônomos ignoram quase tudo a respeito desse planeta: seu tamanho exato, sua temperatura, os detalhes de sua superfície, a existência ou não de satélites etc. Essa ignorância não depõe contra os astrônomos, nem ninguém sonharia em negar à astronomia o caráter de ciência porque não resolveu esse problema. Uma ciência faz perguntas, e nem sempre pode respondê-las, porque as respostas dependem de evidência, e a evidência pode faltar. No caso de *que* e *um*, pelo menos temos as perguntas bem claras, mas não temos resposta, talvez nunca tenhamos, e isso tem que ser reconhecido. Relembro o que disse no capítulo 1: em gramática lidamos com hipóteses, e as hipóteses precisam ser validadas em termos de dados da língua. Quando faltam dados, qualquer hipótese é vazia de sentido, e é necessário deixar a questão em aberto. A alternativa seria dar uma resposta a todo custo — mas em lingüística não se faz afirmações sem fundamento; e para um lingüista não é vergonha dizer “não sei”.

⁶ Isso para excluir os compostos como *vinte e três*, *cento e dezessete* etc., que não nos interessam aqui porque estamos falando só da classificação de palavras.

⁷ Janeiro de 2006.

22.4. REVISITANDO A COMPLEXIDADE

Definimos três classes de palavras; mas que nomes vamos dar a elas? As três palavras que nos serviram de ponto de partida se classificam tradicionalmente como “pronome indefinido” (*todos*), “pronome demonstrativo” (*esses*) e “numeral” (*cinco*). Mas essas denominações não vão nos ajudar, porque *outros* e *vários* também se chamam “pronomes indefinidos”, e no entanto não funcionam como *todos*; *o* funciona como *esse*, mas não se chama “pronome demonstrativo”, sendo classificado como “artigo”. E ainda que *cinco*, *dois*, *três*, *vinte* sejam todos “numerais”, já vimos que *um*, também um “numeral”, não tem comportamento idêntico ao deles.

Esse é um dos problemas com a nomenclatura tradicional: como não pode corresponder à que se estabelece em termos de potencial funcional, usá-la pode levar a confusões. Mas há um outro problema, e esse é mais grave: simplesmente não há nomes tradicionais suficientes para todas as muitas classes que vamos precisar estabelecer. Como já vimos no capítulo 15, a classificação das formas varia segundo o objetivo descritivo, e não se pode prever de quantas classes vamos precisar — só sabemos que vão ser muitas. Por exemplo, vimos que no português escrito *todos* e *ambos* seriam semelhantes em comportamento gramatical. Mas a semelhança só é perfeita se nos limitarmos à possibilidade de posicionamento antes do núcleo: depois do núcleo *todos* soa muito melhor do que *ambos*:

[13] Os jacarés todos fugiram do zoológico.

[14] ?? Os jacarés ambos fugiram do zoológico.

Além disso, claro, *ambos* só tem plural, ao passo que *todos* tem singular:

[15] Toda a turma estava presente.

E então? *Todos* e *ambos* pertencem ou não pertencem à mesma classe?

Finalmente, para apontar um terceiro problema, é preciso observar que classificamos essas três palavras, cada uma em sua classe, através da observação de suas possibilidades de posicionamento. Mas há outras propriedades, igualmente importantes, que nos fornecem uma classificação diferente. Assim, *esse* e *todos* variam em gênero e número (*essa*, *essas*, *esses*; *toda*, *todas*, *todo*), ao passo que *cinco* é invariável — desse ponto de vista, *todos* e *esse* ficam juntos, em oposição a *cinco*. Rever, a esse respeito, o capítulo 15, onde se tratou da classificação por objetivos.

CONSTRUÇÕES DE TÓPICO

23.1. FALHAS DA ANÁLISE SINTÁTICA

Todos nós estudamos alguma coisa da estrutura da oração (a famigerada análise sintática), e temos alguma noção do que vem a ser um sujeito, um objeto direto, um adjunto adverbial etc. Em geral, temos a idéia de que essas funções definem as relações entre os constituintes, e que são mais ou menos suficientes para esgotar o assunto: todas, ou quase todas, as frases do português seriam analisáveis em termos dessas funções sintáticas. As exceções são chamadas “anacolutos” e relegadas ao gueto das estruturas malformadas.

Mas a maior parte da atividade lingüística é oral, não escrita — e é justamente na análise da língua oral que as funções tradicionais mostram suas limitações. Os problemas são muitos, mas aqui vou considerar apenas um, representado pelas construções de tópico. Para ser exato, não se trata de um problema, mas de dois, porque as construções de tópico podem ser distinguidas em dois tipos nitidamente diferenciados: os tópicos sentenciais e os tópicos discursivos. Vamos examinar cada um deles.

23.2. TÓPICO SENTENCIAL

23.2.1. O que é o tópico sentencial

Em uma frase como

[1] Café eu só tomo de manhã cedo.

há razões para analisar o sintagma *café* como objeto direto de *tomar*:

- a) o verbo *tomar* exige a presença de objeto direto, e portanto é de esperar que essa função esteja representada em [1];

b) não é possível acrescentar um objeto direto à frase:

[2] *Café eu só tomo leite de manhã cedo.

c) *café* tem o mesmo papel temático (isto é, paciente) que tem o objeto direto de *tomar* em casos claros:

[1] Café eu só tomo de manhã cedo.

pac.

[3] Eu só tomo café de manhã cedo.

pac.

Assim, tanto a frase [1] quanto a [3] são compostas dos mesmos termos sintáticos¹, além de serem semanticamente semelhantes (não propriamente sinônimas, como veremos). No entanto, não se pode dizer que [1] e [3] sejam idênticas: formalmente, o termo *café* está em posições diferentes, e semanticamente há uma diferença sensível: [1] é claramente uma declaração a respeito do café, ao passo que [3] é mais neutra, e pode ser entendida como uma declaração a respeito do próprio falante.

Assim, *café* em [1] é objeto direto, mas é qualquer coisa mais; a essa coisa mais chamamos tópico sentencial (e o restante do enunciado vai se chamar comentário)². O tópico sentencial não é uma função sintática como o objeto direto, e portanto não precisaremos dizer que *café* em [1] tem dupla função sintática. Trata-se, antes, de uma função comunicativa, cujo conteúdo aproximado é “aquilo sobre o qual se fala”; essa função comunicativa se expressa formalmente através da colocação do elemento no início do período, como está em [1].

A maioria dos termos da oração podem ser colocados no início (“topicalizados”), assumindo assim a função de tópico:

[4] De manhã cedo, eu só tomo café.

[5] Bonita ela certamente não é.

¹ Isso não é estritamente verdadeiro, porque a ordem também define a função sintática: e se a ordem difere, a função precisa ser diferente. Em termos tradicionais, contudo, [1] e [3] têm a mesma análise.

² Muitos autores denominam o tópico tema, e o comentário rema. Vou evitar essa terminologia, porque tema é freqüentemente utilizado em outro sentido, como denominação de um dos papéis temáticos.

[6] Do Chaves eu só conheço o nome.

Alguns termos, como o núcleo do predicado, não podem ser topicalizados:

[7] *Tomo, eu café só de manhã cedo.

[8] *Comeu, Solange os biscoitos.

Quanto ao sujeito, que em geral já está no início da sentença, em muitos casos fica difícil verificar se ele é ou não o tópico:

[9] O Chaves sempre me telefona para pedir dinheiro.

Certamente se pode entender essa frase como sendo “sobre o Chaves”, de modo que o mais provável é que se trate de um tópico sentencial; mas a marca formal (colocação no início do período) não é clara, porque o sujeito precisa estar ali de qualquer modo.

23.2.2. Tópico sentencial, agente e sujeito

A gramática tradicional define o sujeito como “o termo do qual se afirma alguma coisa” — definição que, como vimos, se aplica não ao sujeito, mas ao tópico sentencial. Mas a própria gramática tradicional não respeita essa definição, e em muitos momentos fala como se o sujeito fosse o elemento que pratica uma ação — o que, como sabemos, corresponde ao papel temático de agente. E, em outros momentos, parece considerar que o sujeito é o termo com o qual o verbo concorda. As seguintes passagens, tiradas de uma mesma gramática, ilustram essa confusão:

O sujeito é o termo sobre o qual se faz uma declaração [Cunha, 1981:81].

Aqui o sujeito é aquilo que chamamos tópico sentencial. Algumas páginas adiante encontramos:

Algumas vezes o verbo não se refere a uma pessoa determinada, ou por se desconhecer quem executa a ação, ou por não haver interesse no seu conhecimento. Dizemos, então, que o sujeito é indeterminado [Cunha, 1981: 84].

Agora o sujeito passa a ser o que chamamos agente: o elemento que designa o praticante da ação. E em outra página da mesma obra temos:

O verbo concorda em número e pessoa com o seu sujeito [...] [Cunha, 1981: 275].

Agora trata-se do sujeito enquanto elemento que controla a concordância verbal.

Como se vê, a gramática desliza de uma definição para outra, sem aviso. Se seguirmos estritamente todas essas passagens, torna-se impossível apurar qual é o sujeito em uma frase como

[10] O Danilo, os próprios irmãos não agüentam.

Se o sujeito é o termo sobre o qual se faz uma declaração, o sujeito de [10] será *o Danilo*; se for o termo com o qual o verbo concorda, será *os próprios irmãos*; e se for o termo que pratica a ação, não há sujeito em [10], porque essa frase não expressa uma ação. A saída, evidentemente, é distinguir essas três funções: em [10] *o Danilo* é tópico, *os próprios irmãos* é sujeito e não há agente.

Essa distinção é necessária também porque as três funções pertencem a níveis diferentes de análise. O sujeito é uma função formal (sintática), definida em termos de sua posição na seqüência e, secundariamente, pela concordância do verbo com ele. O agente é um dos papéis temáticos vistos no capítulo 14. E o tópico é uma função discursiva, que, embora seja representada formalmente (colocação no início do período), tem a ver com o papel do elemento na situação de comunicação³.

Isso pode ser visto com clareza se inserirmos essas frases em um contexto maior, digamos em um diálogo. Uma conversa bem natural poderia ser:

[11] Antônio: O que aconteceu com o Daniel?

Beatriz: O Daniel bateu o carro.

Note-se como ficaria estranho se a conversa fosse assim:

[12] Antônio: O que aconteceu com o Daniel?

Beatriz: O carro o Daniel bateu.

³ É interessante notar que os gramáticos já haviam percebido isso pelo menos desde o século XIX, e falavam de "sujeito gramatical" (= sujeito), "sujeito lógico" (= agente) e "sujeito psicológico" (= tópico). Essa distinção, no que pese sua importância, foi simplesmente esquecida, não figurando nas gramáticas atualmente adotadas nas escolas.

Alguma coisa deu errado em [12], mas o quê?

A resposta é que em [12] Antônio fez uma pergunta a respeito de Daniel, e a resposta de Beatriz se refere ao carro. Em outros termos, Antônio colocou Daniel como tópico do diálogo, e Beatriz respondeu como se o tópico fosse o carro.

Para nossos objetivos, vamos definir informalmente o tópico sentencial como

tópico sentencial é o termo da frase do qual se afirma (ou pergunta) alguma coisa.

23.2.3. Mecanismos de topicalização

Há diversos mecanismos gramaticais que, de uma maneira ou de outra, marcam determinado termo como tópico de uma sentença. Esses mecanismos não são propriamente intercambiáveis, tendo em geral valor discursivo um tanto diferente. Por exemplo, vimos que o posicionamento no início do período, como em

[1] Café eu tomo de manhã cedo.

marca o elemento inicial como sendo aquele "sobre o qual" a declaração se faz. É como se disséssemos: "no que diz respeito a café, tenho a declarar que tomo [café] de manhã cedo".

Outro recurso de topicalização é colocar o elemento realçado no início, mas precedido de uma forma do verbo *ser* e seguido de *que* (frases com essa estrutura se chamam clivadas):

[13] É café que eu tomo de manhã cedo.

Aqui a mensagem é um pouco diferente da de [10]. Não só o café é do que se fala, mas há um entendimento de que se falava de algo mais; por exemplo, [11] poderia ser a resposta a

[14] Disseram que você toma uísque de manhã cedo.

Ou seja, *café* está contrastado com *uísque* — diz-se então que *café* é o foco de contraste nessa frase. Essa contrastividade, típica das frases clivadas, não fica tão evidente em [1].

Outra maneira de topicalizar um elemento é colocá-lo no final da sentença. Esse recurso é bem menos freqüente do que os dois vistos acima, mas ocorre às vezes, como em

[15] Eu só tomo de manhã cedo, café.

Na fala, esse termo posposto é marcado por uma entonação muito baixa, soando como se fosse alguma coisa acrescentada à frase.

Essas estruturas são importantes por sua ocorrência freqüente na língua falada. Evidentemente, será preciso incluir na análise sintática esses recursos de topicalização, porque, se *café* é objeto direto em [1], em [13] e em [15], também é verdade que há diferenças sintáticas relevantes (a posição e/ou a presença dos elementos de realce *ser ... que*), sem falar nas diferenças de valor discursivo. Ou seja, não basta falar de "objeto direto": é também necessário especificar se se trata de um objeto não topicalizado, ou topicalizado como em [1], ou clivado como em [13], ou deslocado à direita como em [15]. Isso é importante para a descrição sintática e também para certos aspectos da descrição do significado: embora a topicalização nunca afete o papel temático do termo em questão, ela afeta seu valor no discurso, distinguindo o termo tomado como "assunto" principal da frase dos outros termos que são secundários desse ponto de vista.

TÓPICO DISCURSIVO

24.1. O QUE É TÓPICO DISCURSIVO

N o capítulo anterior, estudamos o tópico sentencial, ou seja, um termo da oração que é marcado, de uma maneira ou de outra, como tópico. Mas há uma outra maneira de marcar o tópico, e essa simplesmente não cabe na análise sintática, porque o termo em questão não encontra lugar na estrutura da sentença tal como tradicionalmente se entende — isto é, não há maneira de atribuir a esse elemento topicalizado uma função sintática como sujeito, objeto direto ou adjunto adverbial. Trata-se do tópico discursivo, muito freqüente na linguagem falada, embora raro na escrita. O estudo do tópico discursivo, e das estruturas em que ocorre, é algo que está por ser feito; mas já é possível mostrar alguns exemplos interessantes¹.

Para ir direto a um deles:

[16] Esse carro o motor tá precisando trocar².

Essa frase foi realmente usada na fala, e não houve problemas em compreendê-la. No entanto, ela não cabe na análise sintática tradicional, porque não há função disponível para *esse carro*. A única possibilidade, a julgar pela semântica, seria a de modificador de *motor*, e nesse caso teria que haver ali uma preposição: *o motor desse carro...* Mas não há preposição

¹ Os trabalhos mais importantes a respeito do tópico discursivo no português são os de Eunice Pontes (1986; 1987). Acho que valeria a pena dar designações diferentes aos dois tipos de tópico, pois são noções muito distintas; mas a tradição chama a ambos de "tópico", e temos que segui-la aqui.

² Alguns dos exemplos de tópico discursivo dados nesta seção provêm dos trabalhos de Pontes (1986; 1987); outros foram recolhidos por mim ou por meus alunos.

antes de *esse carro*, e nem poderia haver, porque [17] não soa natural na linguagem falada:

[17] *Desse carro o motor tá precisando trocar.

Temos que concluir que as funções tradicionais não bastam para analisar [16], que, no entanto, é uma frase legítima da linguagem falada. Vamos então chamar *esse carro* em [16] de tópico discursivo. Observa-se que o tópico discursivo, em muitos casos, é associado a uma oração completa:

[18] O Ricardo [eu cheguei pra ele e fiz uma pergunta] _{Oração}

[19] Essa escola [eu trabalhei lá mais de um ano] _{Oração}

[20] A Patrícia [ela não faz nada o dia todo] _{Oração}

Em termos de significado, muitas vezes o tópico discursivo funciona como um dos termos da oração — mas o termo está também presente na própria estrutura oracional, e como sabemos que uma oração não pode ter dois sujeitos nem dois objetos diretos, não podemos simplesmente encaixar o tópico discursivo na oração. É o caso de [20], onde o sujeito da oração é *ela*, que se refere à mesma pessoa que *a Patrícia*.

Em outros casos, o tópico discursivo, embora de certa forma repetido por um dos termos, não é sintaticamente marcado para a função desse termo: em [19], *essa escola* é repetido por *lá*, mas não apresenta a preposição *em*, que seria esperada para um adjunto de lugar; o mesmo em [18], onde *pra ele* repete *o Ricardo*, que entretanto não tem a preposição *pra*³.

Mesmo quando o tópico discursivo se associa, de uma maneira ou de outra, a um termo da oração subsequente, há características significativas que o distinguem do tópico sentencial, dentre as quais:

- primeiro; como vimos, o termo ao qual o tópico se associa está presente na oração, ainda que não possa ser normalmente repetido — por exemplo, não há dois sujeitos em [20];
- quando o termo associado é regido de preposição, na maioria dos casos o tópico não tem essa preposição, como se vê em [16];
- ao tópico discursivo não se aplicam as restrições sintáticas que limitam as possibilidades de topicalização. Por exemplo, vimos

³ *Pra* é a pronúncia normal da preposição que escrevemos *para*.

nos exemplos [7] e [8] que o núcleo do predicado (o verbo) não pode ser tópico sentencial; mas o verbo (sempre no infinitivo) pode ser tópico discursivo:

[21] Lavar, eu queria que você lavasse as xícaras pra mim.

[22] Vender, eu não vendo; mas posso emprestar.

- o caráter de foco de contraste que muitas vezes marca o tópico sentencial é menos freqüente com o tópico discursivo, que costuma ocorrer em enunciados onde não há nenhuma intenção de contrastar o elemento topicalizado com outros elementos possíveis.

Os casos de tópico discursivo vistos acima estão entre os mais favoráveis, e se associam a estruturas analisáveis, pelo menos em parte. Se simplesmente criássemos uma nova função sintática (“tópico discursivo”), talvez fosse possível encaixar frases como [16], [18], [19], [20], [21] e [22] na análise sintática. Mas há muitos casos em que isso não seria possível, e que mostram a inadequação dessa alternativa, e sugerem que nossa compreensão do fenômeno é ainda muito precária. Por exemplo,

[23] Essa minha barriga, só jejum.

[24] Dinheiro, só Jesus nessa casa.

[25] Diferente, esse homem do jeito que ele tá aí.

[26] A minha mãe, tudo beleza.

Aqui temos estruturas típicas da linguagem falada, com tópico discursivo associado a seqüências de análise ainda desconhecida⁴.

24.2. FUNÇÃO COMUNICATIVA DO TÓPICO DISCURSIVO

O tópico discursivo tem uma função comunicativa no enunciado, que pode ser descrita como

um quadro de referência espacial, temporal ou individual dentro do qual a predicação principal é válida [Chafe, 1976: 50].

Ou seja, em uma frase como

⁴ Se alguém tiver objeções contra essas frases — que aliás foram observadas na realidade —, releia a seção 0.2 (com atenção, desta vez).

[24] Dinheiro, só Jesus nessa casa.

que, à primeira vista, parece muito difícil de interpretar, ganha sentido em um contexto específico, lançando mão do quadro de referência fornecido pelo tópico discursivo. Ela vai ser entendida mais ou menos assim:

[no que se refere a dinheiro], só Jesus nessa casa.

De posse do quadro de referência, o receptor tentará inserir o resto da frase em uma "paisagem mental" coerente. Acredito que ele vai chegar a uma representação mental como "no que se refere a dinheiro nessa casa, só mesmo recorrendo à ajuda divina".

Outro exemplo é:

[27] Esse texto aí, repetição é pouco.

Aqui o comentário significa muito pouca coisa: *repetição é pouco*. Mas o tópico nos fornece um campo onde procurar a relevância do comentário: *esse texto aí*, especialmente em contexto escolar, deve se referir à necessidade de ler e entender um texto. Aí, o comentário adquire significado, levando a algo como "quanto a esse texto aí, não basta ler muitas vezes para entendê-lo".

Como se vê, a interpretação nesses casos é menos dependente da estrutura lingüística do que nos casos "clássicos" que examinamos nos outros capítulos. Depende em grande medida dos conhecimentos, expectativas e crenças que o emissor e o receptor têm em comum, somados a inferências lógicas e simples bom senso. Para isso, o papel do tópico discursivo é crucial, porque limita as possibilidades, facilitando a interpretação do restante do comentário. Ao estudarmos o tópico discursivo, estamos na verdade transpondo as fronteiras⁵ entre o estudo da estrutura da língua e o estudo das condições de uso da língua, chamado pragmática.

A linguagem falada recorre muito a elementos do contexto, justamente porque se realiza com o receptor e o emissor presentes, em geral sabendo algo um do outro. Isso facilita a expressão, pois já não é necessário explicitar um mundo de coisas que na escrita precisam ser expressas. Mas, por outro lado, esse caráter um tanto fragmentário do texto falado levanta proble-

⁵ Caso existam!

mas de compreensão: o texto é fragmentário, mas a compreensão precisa ser completa. A solução é lançar mão de conhecimentos extralingüísticos, tais como elementos do contexto físico (*esse negócio aí* — o falante aponta com o dedo), do conhecimento pessoal entre emissor e receptor (*a Leda* — que o receptor sabe que é a esposa do emissor), do contexto do discurso (*o papel ficou impossível* — dito por um fotógrafo comentando o preço do material usado em sua atividade), do contexto cultural (*isso só em dezembro* — dito por uma pessoa para explicar que só vai ter folga financeira quando receber o 13º salário).

A amplitude dos conhecimentos relevantes para interpretar um enunciado pode causar grandes dificuldades de compreensão. Nesse contexto, o tópico discursivo tem o papel de limitar as escolhas. Por exemplo, o comentário *o papel ficou impossível* será interpretado de maneiras distintas nos enunciados seguintes:

[28] Fotografia, o papel ficou impossível.

[29] Essa umidade, o papel ficou impossível.

Em [28], trata-se do preço do papel fotográfico. E em [29] talvez o falante esteja se referindo ao problema de usar o papel em uma impressora, em época de muita umidade, quando o papel tende a enrugar e agarrar.

Como se vê, o tópico discursivo pode mudar significativamente a interpretação de um comentário. A compreensão final é resultado da interação entre, de um lado, a interpretação semântica fornecida pelos elementos léxicos e gramaticais e, de outro lado, elementos do contexto, grandemente auxiliados pela delimitação feita pelo tópico discursivo. Tomando o exemplo [29]: o tópico nos informa que o que vem a seguir deve ter uma relação qualquer com a umidade; e nosso conhecimento do mundo, assim como nossa capacidade lógica, já exclui uma série de interpretações que, em princípio, seriam possíveis para o comentário. Por exemplo, entender que o papel ficou impossível de comprar não faria sentido dentro do contexto estabelecido pelo tópico discursivo *essa umidade*; mas entender que ele ficou impossível de usar por causa das conseqüências da umidade faz sentido — desde que se saiba o que a umidade faz com o papel (conhecimento do mundo) e se faça a ligação lógica desse conhecimento com as informações fornecidas pelo enunciado: se a umidade enrugou o papel ou gruda as folhas e se o falante é uma pessoa que usa papel que não

pode ser enrugado ou que precisa ficar bem solto (em uma impressora, por exemplo), essa “impossibilidade” do papel deve ter a ver com a impossibilidade de usá-lo para os fins relevantes, e portanto se conclui que ele deve ter ficado impossível de usar em uma impressora porque está enrugado ou grudado⁷.

Termino esta seção com mais alguns exemplos, especialmente saborosos, colhidos por meus alunos em conversas informais:

[30] Droga, ele tá afundado de novo.

[31] Namoro, menina nova só faz besteira.

[32] Esses cento e cinquenta reais que eu tou negativo, é sem contar com o cheque.

24.3. TÓPICOS, SUJEITOS E A CLASSIFICAÇÃO DAS LÍNGUAS

Ao estudarmos o tópico discursivo, já estamos de certo modo passando além dos limites estritos do estudo da gramática portuguesa. Não que o tópico discursivo não seja parte legítima da estrutura da língua — ao contrário, ele é indispensável a uma descrição que pretenda dar conta da estrutura da língua. Mas a relação entre tópico discursivo e comentário se manifesta de maneira bastante semelhante em muitas outras línguas, de certo modo concorrendo com a relação sujeito/predicado. Partindo dessa observação, alguns lingüistas (notadamente Li & Thompson, 1976) sugeriram que as línguas se distinguem em dois grandes grupos: aquelas em que a relação tópico/comentário é mais importante (línguas com predominância de tópico), e aquelas que dão maior proeminência à relação sujeito/predicado (línguas com predominância de sujeito).

Nas línguas com predominância de tópico, as frases seriam concebidas basicamente como uma seqüência de tópico discursivo + comentário, mais ou menos no modelo das frases portuguesas citadas neste capítulo. Um caso citado por Li & Thompson é o do chinês; eles dão o seguinte exemplo⁸:

⁷ Aqui não é possível ir longe no exame do processo de compreensão de enunciados maiores do que uma oração. Ao leitor interessado por essa questão fascinante e fundamental, recomendo começar pela leitura dos livros de Fulgêncio & Liberato (1992; 1996), que tratam do assunto em nível elementar.

⁸ A transcrição, que dou em forma simplificada, é de Li & Thompson; deve estar correta, já que Li é chinês. Para nós, é claro, a pronúncia exata não é importante.

[33] he chi tê pé' ô dà' jâ.

este campo arroz muito bom

A frase seria traduzida como

[34] O arroz é muito bom neste campo.

mas as relações gramaticais são diferentes: em chinês o tópico aparece primeiro (“este campo”), sem marca de locativo, embora acabe sendo entendido como o lugar onde o arroz é bom.

Mas [34] não é a única maneira de traduzir [33] em português; na língua falada, seria igualmente adequado dizer

[35] Esse campo o arroz é muito bom.

Ou, situando o enunciado em um ambiente mais comum,

[36] Esse restaurante a comida é boa.

Aqui *esse campo* e *esse restaurante* funcionam como tópicos discursivos. A construção portuguesa falada é paralela à chinesa, e não à portuguesa padrão exemplificada em [34].

Devemos deduzir daí que o português brasileiro falado é uma língua com predominância de tópico? Acho que ainda não. Embora não tenham sido feitos levantamentos amplos, uma olhada em algumas transcrições de fala brasileira mostra que as construções de sujeito são muito numerosas, mais do que as de tópico. Mas as construções de tópico — principalmente as de tópico discursivo — são frequentes na fala, e quase não ocorrem na escrita formal. Talvez isso denuncie uma tendência histórica do português do Brasil, que estaria no processo de se tornar uma língua com predominância cada vez maior das construções de tópico em detrimento das construções de sujeito: uma hipótese interessante a verificar.

BIBLIOGRAFIA

- BAGNO, Marcos (2000). *Dramática da língua portuguesa: tradição gramatical, mídia e exclusão social*. São Paulo: Loyola.
- (2003). *A norma oculta*. São Paulo: Parábola Editorial.
[Nesse volume, Bagno trata do uso da língua como instrumento de discriminação, acentuando o papel da escola, da imprensa e da sociedade no reforço de atitudes preconceituosas. Leitura recomendada para todos (principalmente para os que não concordam com ele).]
- BASILIO, Margarida (1987). *Teoria lexical*. São Paulo: Ática.
- (2004). *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto.
- CAMARA JR., J. (1969). *Princípios de lingüística geral*. Rio de Janeiro: Acadêmica.
[A primeira, e até hoje a melhor, introdução à lingüística publicada em português. A primeira edição é de 1941; e o autor morreu em 1970, mas o texto ainda tem partes notavelmente úteis para o leitor atual.]
- CANÇADO, Márcia (2005). *Manual de semântica — noções básicas e exercícios*. Belo Horizonte: Ed. UFMG.
- CHAFE, Wallace L. (1976). "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View", in Li (1976).
- COLE, Peter e MORGAN, Jerry (orgs.) (1975). *Syntax and Semantics 3: Speech Acts*. New York: Academic Press.
- CUNHA, Celso (1976). *Gramática da língua portuguesa*. MEC/FENAME.
- (1981). *Gramática de base*. MEC/FENAME.
- FARACO, Carlos Alberto (2005). *Lingüística histórica*. São Paulo: Parábola Editorial.
- FULGÊNCIO, Lúcia e LIBERATO, Yara G. (1992). *Como facilitar a leitura*. São Paulo: Contexto.
- (1996). *A leitura na escola*. São Paulo: Contexto.
[Os dois volumes de Fulgêncio e Liberato são uma excelente introdução à questão da compreensão de textos e leitura fluente.]
- HORA, Dermeval da e PEDROSA, Juliene L.R. (orgs.) (2001). *Projeto Variação Lingüística no Estado da Paraíba*. João Pessoa: Idéia.
- ILARI, Rodolfo e GERALDI, João Wanderley (1985). *Semântica*. São Paulo: Ática.
- JACKENDOFF, Ray S. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- (1983). *Semantics and Cognition*. Cambridge, MA: MIT Press.
- LANGACKER, Ronald W. (1987). *Foundations of Cognitive Grammar, Vol. I - Theoretical Prerequisites*. Stanford: Stanford University Press.
- LI, Charles N., org. (1976). *Subject and Topic*. New York: Academic Press.
- e THOMPSON, Sandra A. (1976). "Subject and Topic: a New Typology of Language", in Li (1976).
- MUSSALIM, Fernanda e BENTES, Anna Christina (2001). *Introdução à lingüística: domínios e fronteiras*. São Paulo: Cortez.
- NATION, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

- NEVES, Maria Helena M. (2003). *Guia de uso do português*. São Paulo: Ed. UNESP.
- NURC – *Documentação Eletrônica Organizada pelo Projeto NURC/SP*. [CD]
- PERINI, Mário A. (1995). *Gramática descritiva do português*. São Paulo: Ática.
- ____ et al. (1996). *O sintagma nominal: estrutura, significado e função*. Número especial da *Revista de estudos da linguagem*. Belo Horizonte: UFMG.
- ____ (1997). *Sofrendo a gramática*. São Paulo: Ática.
- ____ (2004). *A língua do Brasil amanhã e outros mistérios*. São Paulo: Parábola.
- PONTES, Eunice (1973). *Verbos auxiliares em português*. Petrópolis: Vozes.
- ____ (1986). *Sujeito: da sintaxe ao discurso*. São Paulo: Ática.
- ____ (1987). *O tópico no português do Brasil*. Campinas: Pontes.
[Estes dois livros de Pontes (1986, 1987) são a obra básica de referência para o estudo das construções de tópico discursivo no português do Brasil.]
- SAID ALL, Manuel (1908). *Dificuldades da língua portuguesa*. Rio de Janeiro: Acadêmica.
[Depois de quase um século, o livro de Said Ali ainda vale uma leitura atenta. Embora naturalmente só se ocupe do português escrito formal, é um modelo de discussão lúcida e bem informada, pelo maior lingüista brasileiro de sua época.]
- SACCONI, Luiz Antonio (1990). *1000 erros de português*. Ribeirão Preto: Nossa Editora.
- SAUSSURE, Ferdinand de (1916). *Cours de linguistique générale*. Paris: Payot.
- SILVA, Thais Cristóforo (1999). *Fonética e fonologia do português*. São Paulo: Contexto.
- TARALLO, Fernando (1985). *A pesquisa sociolingüística*. São Paulo: Ática.

ÍNDICE REMISSIVO

(A referência é a capítulos e seções de capítulos. Números em **negrito** remetem a capítulos inteiros que tratam principalmente do item em questão)

- A|
adverbiais 13.3
advérbio *ver classes*
ambigüidade 3.4
- C|
classes 15, 22
 advérbio 19.1
 conjunção 19.3
 de palavras 1.2
 e funções 16
 e traços 15.2, 22.3
 nominal
 finalidade das 15.1
 preposição 19.2
concordância 12.3
conjunção 19.3
constituintes 3.3.2
 e o significado 11.5
 propriedade dos 11.6
construções 17
 transitiva 17.3
 intransitiva 17.2
 e valência verbal 17.5
 ergativa 17.2
coordenador 19.3
corpus 2.2, 2.3
- D|
dialetologia 0.1
diátese 17.5
discurso, análise do 0.1
- E|
espaço formal 14.1
espaço semântico 14.1
estratégias 5.1
 cognitiva 5.1
- de interpretação 5.1
léxico-gramatical 5.1
- F|
fonologia 0.1
forma
 e significado 0.1, 3, 9
funções semânticas 14
funções sintáticas 12, 13
 determinante 13.4, 22.3
 e classes 16
 modificador 13.4
 núcleo do predicado 13.1, 17.6
 núcleo do SN 13.4
 objeto direto 12.2.1
 predeterminante 13.4, 22.3
 predicativo do objeto 13.2
 predicativo do sujeito 12.2.4
 quantificador 22.3
 sujeito 12.2.1
- G|
gerúndio 20.2.3
gramática 6
 descritiva 0.1
 e léxico 18.2
 internalizada 0.3
 objeto da 0.1
 prescritiva 0.2
- H|
hipótese 1.1, 1.3
- I|
infinitivo 20.2.2
introspecção 2.2, 2.3
item léxico 18.3
- L|
lexema 11.1, 18.1
léxico 6.1, 18

- e gramática 18.2
- lingüística
 - aplicada 0.1
 - campo da 0.1
 - histórica (ou diacrônica) 0.1
- M|
 - morfema 11.1
- N|
 - nominal 1.4, 19.5
- O|
 - objeto, *ver* funções sintáticas
 - oração
 - principal 21.1, 21.3
 - sem sujeito 12.2.3
 - subordinada 20.1, 21.1, 21.2
 - ordem
 - das palavras no SN 10.2
 - e significado 3.3.1
- P|
 - palavra 11.1
 - papéis temáticos 14.2
 - agente 3.3.1, 14.1
 - causador de experiência 14.2
 - experienciador 14.2
 - fonte 14.2
 - local 14.2
 - localizando 14.2
 - múltiplos 14.3
 - meta 14.2
 - paciente 3.3.1, 14.2
 - tema 14.2
 - período composto 21
 - potencial funcional 16.2
 - pragmática 0.1, 24.2
 - preposição 19.2
 - complemento da 13.5
 - psicolingüística 0.1
- Q|
 - qualificativo 10.2
- R|
 - reclassificador 19.2
 - recursividade 11.3, 11.6.1
- referencial 10.2
- regras
 - descritivas 8
 - finalidade das 6.3
 - fonológicas 6.1
 - gramaticais 6.1, 8
 - morfológicas 6.1, 10.1
 - semânticas (ou de interpretação) 6.1
 - sintáticas 6.1
- restritivos 10.2
- S|
 - semântica
 - área 3.1
 - e sintaxe 9
 - objeto da 0.1
 - significado
 - depreensão do 3.1
 - e forma 0.1, 7.1
 - signo 7.1
 - simbólica, relação 0.1, 10, 14.1
 - sintagma 11
 - oracional 20
 - sintagma adjetivo (s_{adj}) 11.2
 - sintagma adverbial (s_{adv}) 11.2
 - sintagma nominal (s_N) 6.4, 11.2, 13.4
 - sintagma verbal (s_v) 11.2
 - sintaxe
 - e semântica 9
 - objeto da 0.1
 - sintaxe residual, princípio da 9.3
 - sociolingüística 0.1
 - subjuntivo 20.2.1
 - subordinação 21.1
 - sujeito, *ver* funções sintáticas
- T|
 - testagem 2.2
 - tópico 23
 - discursivo 24
 - e a classificação das línguas 24.3
 - sentencial 23.2
 - topicalização 23.2.3
- V|
 - valência 17.5
 - verbo 18.1
 - verbo auxiliar 20.3

mas) em mais de um bloco, introduzindo o tema em nível elementar e voltando a ele posteriormente quando se tornou possível um tratamento mais aprofundado.

Além de fornecer noções básicas de análise e exemplificação tirada da estrutura do português, *Princípios de Linguística Descritiva — Introdução ao pensamento gramatical* objetiva desenvolver no estudante a sensibilidade aos fenômenos linguísticos, algo que lhe permita saber o que observar ao se ver confrontado com dados.

As duas primeiras partes foram planejadas para serem vistas em um semestre normal de 60 horas-aula. Esses capítulos devem fornecer ao aluno uma formação básica em gramática, o que constitui a melhor porta de entrada aos estudos linguísticos.

A terceira parte é composta principalmente de questões de gramática portuguesa, discutidas de maneira um pouco mais aprofundada, e as vezes sem chegar a uma solução satisfatória (é preciso lembrar sempre que a estrutura do português é um território inexplorado). Com turmas mais preparadas, podem ser vistas no mesmo semestre, depois de terminadas as duas primeiras partes. Alternativamente, podem ser parte do material de leitura em um segundo semestre de gramática.