

Mário A. Perini

**ESTUDOS DE GRAMÁTICA
DESCRITIVA**

as valências verbais

2007

Para
William Curtis Blaylock

SUMÁRIO

Apresentação

Proposta

Agradecimentos

PARTE I: PRELIMINARES TEÓRICOS

1. Por uma metodologia da descrição gramatical

- 1.1. Forma e conceito
- 1.2. Conhecimento e uso do contexto
- 1.3. O que levar em conta na análise?
- 1.4. Algumas conseqüências
 - 1.4.1. Dados e hipóteses
 - 1.4.2. Palavra e lexema
- 1.5. Descrição e teorização
 - 1.5.1. Relatórios descritivos
 - 1.5.2. Colaboração ou rivalidade?
- 1.6. Nota sobre o momento em lingüística
- 1.7. Em que basear uma metodologia?
- 1.8. Guias de descrição
 - 1.8.1. Observações preliminares
 - 1.8.2. Partindo de exemplos
 - 1.8.2.1. Sujeito “oculto”
 - 1.8.2.2. Topicalização

2. Posicionamento teórico

- 2.1. Teoria e descrição
- 2.2. Por uma gramática descritiva
 - 2.2.1. Necessidade de descrições
 - 2.2.2. Descrição e teoria(s) da linguagem
- 2.3. Regularidade e anomalia
 - 2.3.1. Anomalias
 - 2.3.2. Léxico e gramática
- 2.4. Postulados da análise
 - 2.4.1. Descrição superficial
 - 2.4.2. Fatos de forma e fatos de significado
 - 2.4.3. Sintaxe simples (“simpler syntax”)
- 2.5. Tipologia das regras

3. Categorização

- 3.1. Importância da taxonomia de formas
- 3.2. Correlação entre forma e significado
- 3.3. Insuficiência da classificação tradicional
 - 3.3.1. O que está faltando
 - 3.3.2. Gerativistas e estruturalistas
 - 3.3.3. Posição dos cognitivistas
- 3.4. Algumas questões básicas
 - 3.4.1. O que classificar?
 - 3.4.2. As classes na descrição gramatical
 - 3.4.3. Classificando por objetivos
 - 3.4.4. Classes e funções
 - 3.4.5. Exclusão do contexto anafórico
- 3.5. Calibrando regras

3.6. Taxonomia e descrição gramatical

3.7. Protótipos

4. Funções sintáticas: sujeito e objeto direto

4.1. Funções sintáticas

4.1.1. A Hipótese da Sintaxe Simples revisitada

4.1.2. Repensando as funções sintáticas

4.1.3. Para que funções sintáticas?

4.1.4. Redundância nos papéis temáticos

4.1.5. Identificação do sujeito

4.1.6. Identificação do objeto

4.1.7. Papel das valências

4.1.8. SNs e não-SNs

4.2. Fenômenos gramaticais ligados ao sujeito

4.2.1. Concordância

4.2.2. Clíticos

4.2.3. Sujeito posposto

4.2.4. O que é o sujeito

4.2.5. O que é a concordância

4.3. Complexidade: fatores semânticos

4.3.1. Nós = eu e ele

4.3.2. O núcleo do SN como controlador da concordância

4.4. Complexidade: fatores sintáticos

4.4.1. Coletivos

4.4.1.1. Coletivos e plurais

4.4.1.2. Silepse

4.4.2. Outros fatores não-semânticos

4.4.2.1. Marcados e não-marcados

4.4.2.2. Impessoais

4.4.2.3. Nós = a gente

4.4.2.4. Casos de “atração”

4.5. Como definir as construções: a notação ‘X’

5. Outras funções sintáticas

5.1. Objeto direto

5.1.1. Problemas de definição

5.1.2. Semântica do objeto direto: papéis temáticos

5.1.3. Comportamento sintático

5.1.4. O objeto direto como relação simbólica

5.1.5. O que é um objeto direto?

5.1.5.1. Regras de atribuição do papel Paciente

5.1.5.2. A representação sintática do Paciente

5.1.6. Complementos de medida

5.1.7. Conclusão

5.2. Notas sobre a construção ergativa

5.2.1. Sujeito ou tópico?

5.2.2. Ambigüidades

5.3. Outras funções sintáticas

5.4. Adjuntos e complementos adverbiais

5.5. Objeto indireto

5.6. Predicativos

5.7. Conclusão: funções sintáticas

5.8. Funcionamento e interação das regras

PARTE II: VALÊNCIAS VERBAIS

6. Resenha da literatura

- 6.1. A tradição gramatical
 - 6.1.1. Gramáticas
 - 6.1.2. Dicionários
- 6.2. Tesnière (1959)
- 6.3. Gross (1975) e Levy (1983)
- 6.4. Allerton (1982)
- 6.5. Levin (1993)
- 6.6. Goldberg (1995)
- 6.7. Vilela (1992)
- 6.8. Concluindo

7. Papéis temáticos

- 7.1. Preliminares
- 7.2. Por que e para que papéis temáticos?
- 7.3. Conceito e codificação
 - 7.3.1. Como a sintaxe vê o significado
 - 7.3.2. Diferenciando os papéis temáticos
 - 7.3.3. Relações conceptuais como elaboração dos papéis temáticos
 - 7.3.4. Caráter esquemático dos papéis temáticos
- 7.4. Contra os protótipos-como-contínuos
 - 7.4.1. Protótipos
 - 7.4.2. Protótipos e a descrição gramatical
 - 7.4.3. Perspectiva descritiva
- 7.5. Previsibilidade
 - 7.5.1. O que se pode prever?
 - 7.5.2. Arbitrariedade das funções sintáticas
- 7.6. Critérios de delimitação dos PTs
 - 7.6.1. Critério de necessidade
 - 7.6.2. Aplicando o critério de necessidade
 - 7.6.2.1. Agente e causador
 - 7.6.2.2. Lugar e tempo
 - 7.6.2.3. Lugar e estado
 - 7.6.2.4. Agente e paciente
 - 7.6.2.5. Meta e beneficiário
 - 7.6.2.6. Sumário parcial
 - 7.6.3. Critério de semelhança semântica
 - 7.6.4. Conteúdo semântico dos sintagmas
 - 7.6.5. Perigos da circularidade
- 7.7. PTs múltiplos
 - 7.7.1. Mais de um PT para o mesmo sintagma
 - 7.7.2. Os PTs nas construções complexas
 - 7.7.3. Sobre a representação formal dos papéis temáticos
- 7.8. Fazendo pesquisa em tempos de incerteza

Cap. 8: Construções. diáteses, valências

- 8.1. A apresentação dos dados
- 8.2. Construções: papel descritivo
- 8.3. Limitações do estudo
- 8.4. Contando verbos
- 8.5. Expressões idiomáticas
- 8.6. Valência do verbo ou do predicado?
- 8.7. Estrutura sintática

- 8.8. Definindo construções
 - 8.8.1. Definição formal
 - 8.8.2. Papel da semântica do verbo
- 8.9. O que incluir?
- 8.10. Definindo diáteses
 - 8.10.1. Diáteses como fator de subcategorização
 - 8.10.2. Graus de esquematicidade
 - 8.10.3. Previsibilidade
 - 8.10.4. Sintagmas preposicionados e sintagmas nominais
 - 8.10.5. Procedimento
 - 8.10.5.1. Preposições unívocas
 - 8.10.5.2. Preposições ambíguas
 - 8.10.5.3. Marcados e não marcados
 - 8.10.5.4. Ocorrência obrigatória
 - 8.10.6. Verbos que recusam sujeito
 - 8.10.7. Complementos especificados semanticamente?
 - 8.10.8. Inversões
- 8.11. Problemas
 - 8.11.1. Distinções relevantes
 - 8.11.1.1. Diáteses e não-diáteses
 - 8.11.1.2. Restrições seletivas
 - 8.11.1.3. Extensão de sentido
 - 8.11.1.4. Traços léxicos inerentes: os complementos
 - 8.11.1.5. Traços léxicos inerentes: o verbo
 - 8.11.1.6. O que fica de fora
- 8.12. Quando um verbo não tem valência
 - 8.12.1. Orações com dois verbos
 - 8.12.2. Os dois participípios
 - 8.12.3. A passiva é uma diátese verbal?
- 8.13. Concluindo
- 8.14. Notas para a pesquisa futura

9. Classes de verbos: problemas e planos

- 9.1. Classificando verbos
- 9.2. Verbos “transitivos”, “ergativos” e “transitivo-ergativos”
- 9.3. Casos de complementaridade
- 9.4. Expansão da classe dos transitivo-ergativos na língua falada
 - 9.4.1. Ergativos em português padrão
 - 9.4.2. Ergativos se tornam transitivo-ergativos
 - 9.4.3. Transitivos se tornam transitivo-ergativos
- 9.5. Efeitos sintáticos da ergatividade
- 9.6. Correlatos semânticos
 - 9.6.1. Relação complexa entre valência e significado
 - 9.6.2. Definindo semanticamente os transitivo-ergativos
 - 9.6.3. Relações entre diáteses
 - 9.6.4. Polissemia induzida
 - 9.6.5. Concluindo
- 9.7. Construção pseudo-ergativa
- 9.8. A chamada “construção média”
 - 9.8.1. Colocando o problema
 - 9.8.2. Testando uma hipótese
 - 9.8.2.1. Casos em que a ergativa é possível, mas não a média
 - 9.8.2.2. Casos em que a média é possível, mas não a ergativa
 - 9.8.2.3. Discussão dos exemplos
 - 9.8.3. A proposta de Camacho (2003)
 - 9.8.4. Sumário e conclusão

9.9. Terminologia

10. A valência na interface entre semântica e estrutura informacional

- 10.1. Introdução
- 10.2. Casos problemáticos
- 10.3. O ponto de partida da interpretação
- 10.4. O papel temático do tópico
- 10.5. Identificando o tópico
- 10.6. O mecanismo de atribuição do papel temático
- 10.7. Um sistema complexo
- 10.8. Exemplos: papéis temáticos atribuídos pragmaticamente
- 10.9. Tópicos sem papel temático?
- 10.10. Examinando (e rejeitando) uma alternativa
- 10.11. Breve resenha da literatura
- 10.12. Conclusão

11. Um catálogo de construções

- 11.1. O Catálogo das Estruturas do Português
- 11.2. Listas e regras
 - 11.2.1. Regras de correlação
 - 11.2.2. Regras de estrutura sintagmática
- 11.3. Pesquisando as regras por indução
- 11.4. Relevância da lista para a descrição
 - 11.4.1. Definição dos papéis temáticos
 - 11.4.2. Outras contribuições

12. Notação e lista de diáteses

- 12.1. Convenções de notação
- 12.2. A lista das diáteses do português
 - Construções X V SN
 - Construções X V
 - Construções X V SN SPrep
 - Construções X V SPrep
 - Outras construções
 - Construções complexas

Índice remissivo

Apresentação

Proposta

Os capítulos que se seguem se concentram principalmente em dois temas: a discussão de problemas metodológicos e um programa de descrição de parte da estrutura do português do Brasil.

Entenda-se “descrição” como a apresentação sistemática dos fatos da língua – não a elaboração ou validação de alguma teoria específica da linguagem. E não há aqui preocupação de caráter comparativo, tipológico ou universal; o trabalho visa basicamente a caracterizar uma língua (o português falado do Brasil) naquilo que a distingue das demais línguas. No entanto, como não se pode realizar descrições sem assumir algum ponto de vista teórico, adoto aqui, sem dogmatismo, as posições de Culicover e Jackendoff (2005).¹ Como esse trabalho será mencionado com frequência no presente texto, vou usar a abreviatura **C&J**. No entanto, como já aponte, minha perspectiva é descritiva, o que exclui, pelo menos na prática, certos aspectos da proposta de C&J que se referem à construção de uma teoria geral da linguagem, com pretensões universalistas. Não que eu rejeite esses aspectos em princípio, mas acredito, como se verá, que tais propostas são prematuras no estágio atual da pesquisa.

Outros trabalhos que foram importantes como base teórica deste livro, e que merecem menção, são Goldberg, 1995 (que desenvolve os trabalhos de Fillmore e seus colaboradores); Langacker (1987, 1991, 1999) e, principalmente na parte relativa aos papéis temáticos, Jackendoff (1972, 1990). No que diz respeito à atitude assumida frente aos fatos da língua, em especial a ênfase no levantamento de dados como única maneira de fundamentar as teorias, sigo a orientação geral de Gross (1975; 1979). Não adoto, porém, seus princípios descritivos, que me parecem excessivamente livres de restrições formais, levando a uma descrição detalhista ao extremo.

A posição descritivista que subjaz ao presente trabalho inclui pontos como os seguintes:

- A preocupação com a formalização e a elaboração de teorias altamente especificadas deve ceder lugar a uma necessidade mais urgente, a de definir e esclarecer questões fundamentais de análise. Dentre essas questões, posso mencionar: os princípios de taxonomia das formas lingüísticas; a definição operacional dos papéis temáticos; a questão da diferenciação entre adjuntos e complementos; a explicitação dos fatores que governam a ordem das palavras; a relação da ordem das palavras com as chamadas funções sintáticas.

- Os estudos lingüísticos sofrem de falta de evidência para fundamentar as teorias. Ou seja, faltam bases de dados cuidadosamente colhidos, sistematizados e descritos, cobrindo áreas significativamente grandes de línguas naturais particulares.

¹ O livro de Culicover e Jackendoff é um trabalho da mais alta relevância, e deve ter um grande impacto na pesquisa lingüística dos próximos anos.

- Para os objetivos acima discriminados, grande parte da estrutura da linguagem pode ser descrita em termos de estrutura formal aparente, representações semânticas acessíveis à introspecção e relações entre esses dois níveis; não há necessidade (no momento, pelo menos) de níveis intermediários de análise tais como os utilizados pela teoria gerativa clássica.

- A lingüística se encontra atualmente em um estágio que pode ser chamado de “história natural” – ou seja, estamos mais ou menos na situação da física antes de Newton, quando ainda não havia um paradigma que norteasse a interpretação dos fatos observados. Assim como os físicos daquela época, os lingüistas de hoje têm como tarefa principal levantar dados, sistematizá-los e encaixá-los, quando possível, em teorias parciais – em uma palavra, elaborar descrições (nossa tarefa é mais próxima da tarefa de Tycho Brahe e de Lineu do que da de Einstein e Darwin). A lingüística atual é dominada pela ilusão de que é possível queimar etapas e desenvolver teorias válidas sem um longo e laborioso trabalho prévio de levantamento e sistematização de dados. Este livro representa uma reação a essa atitude. Não sou, nesse particular, nem pioneiro nem original: muitos lingüistas estão e estiveram conscientes do problema, mas a meu ver não tiveram a repercussão merecida.

- É urgente elaborar estudos amplos de grandes fatias da estrutura das línguas. Esses estudos devem se concentrar em línguas particulares, sem ceder à tentação de aplicar a uma língua a análise de outra. A comparação entre as línguas, assim como a procura de universais, é um empreendimento válido, mas só pode ser realizado em um segundo momento. Não devemos comparar pequenas áreas da estrutura em diversas línguas, mas fazer confrontos abrangentes de grandes porções das estruturas. E isso, a meu ver, não pode ser ainda realizado por falta de descrições confiáveis de um número suficiente de línguas.

A língua estudada é o português do Brasil. Em consonância com os pontos expostos acima, não tento nenhum exame sistemático de outras línguas, e em princípio as análises aqui propostas só valem para uma língua. Por outro lado, tenho certeza de que muitos dos problemas discutidos são de validade geral, e podem ser transferidos, *mutatis mutandis*, para a análise de outras línguas. Mas é convicção minha que qualquer investigação dos eventuais traços universais das línguas depende crucialmente da existência de estudos descritivos de línguas particulares, suficientemente amplos e adequados; neste trabalho, tento fazer a minha parte.

As questões metodológicas e teóricas são tratadas na primeira parte (capítulos 1 a 5). Na segunda parte, aplico as idéias expostas ao exame detalhado de um problema importante: as valências verbais, com suas muitas ramificações na estrutura da língua. Nessa área, acredito que já se chegou ao ponto em que há alguma coisa de concreto a apresentar, no que diz respeito à análise em si e também no que diz respeito à maneira de realizar o trabalho de descrição. Aqui também, portanto, uma das motivações do texto é metodológica. A ênfase na metodologia vem de uma preocupação constante que me persegue, que é a impressão de que a lingüística pretende se tornar uma ciência amadurecida sem dar a devida atenção ao desenvolvimento de uma metodologia de pesquisa: seleção, coleta e categorização preliminar dos dados. Neste livro me concentro

de preferência na terceira das tarefas mencionadas, e procuro elaborar um sistema para expressar a categorização preliminar dos dados. Talvez minhas observações possam contribuir um pouco para minorar essa situação.

Mas acho importante fazer uma observação: embora o que se esteja propondo aqui seja essencialmente um sistema de notação, não é possível separá-lo de todo de uma concepção particular da linguagem. Caso contrário, nossa tarefa consistiria apenas de criar critérios facilmente aplicáveis, sem compromissos ulteriores. Mas dessa maneira o resultado seria também em grande parte irrelevante para efeitos de elaboração de teorias mais gerais do português e das línguas em geral.

Na medida em que a presente proposta se insere em alguma concepção da linguagem humana, ela pode ser rotulada de “cognitivista”, por aceitar a idéia de que o conhecimento da linguagem é parte integrada do conhecimento geral do mundo. Isso não quer dizer que não haja maneira de distinguir um do outro – essa é uma questão empírica, que só poderá ser respondida com o progresso da investigação em ambos os lados. Mas quer dizer que se deve levar em conta os resultados da pesquisa em áreas como a psicolinguística, a ciência cognitiva e a neurolinguística, seguindo o princípio de que não pode haver contradição entre o que se sabe sobre o conhecimento do mundo e o que se sabe sobre a estrutura da língua, considerada esta como parte do conhecimento do mundo.

É claro que isso só pode ser conseguido, nas atuais circunstâncias, muito parcialmente. Se a descrição gramatical é, em grande medida, uma atividade autônoma, não é assim em princípio,² mas antes em virtude do estágio atual da ciência da linguagem.

A valência verbal é um tema bastante estudado, que remonta, para o português, aos dicionários de Fernandes (1940 e 1950), contando com uma tradição de estudos teóricos: Tesnière (1959), Allerton (1982), Borba (1990), Vilela (1992). No entanto, muitos problemas básicos estão ainda à espera de solução, ou mesmo de formulação clara. Tento mostrar isso na segunda parte deste trabalho.

Deixei que o texto incluísse certo grau de redundância, com ocasional repetição de exemplos e argumentos. Fiz isso para dar aos capítulos alguma autonomia, facilitando sua leitura independente, sem necessidade de ficar voltando continuamente aos capítulos anteriores. Venho trabalhando na formulação das idéias aqui apresentadas há vários anos, e o texto, em sua forma presente, reflete os princípios que atualmente norteiam meu trabalho de pesquisa. Nenhum dos princípios aqui seguidos é definitivo, evidentemente; e como alguns deles são controversos, espero que este livro venha a servir como ponto de partida para alguma discussão.

Agradecimentos

Para terminar, gostaria de fazer alguns agradecimentos. Durante sua longa elaboração, este livro foi objeto de muitas discussões e conversas; não seria possível dar a lista de todas as pessoas que, sabendo ou não, contribuíram para o amadurecimento destas idéias. Mesmo assim, e correndo o risco de cometer omissões injustas, vou

² Como defendiam alguns dos estruturalistas americanos; cf. Bloomfield (1933: 17; 32).

mencionar alguns colegas que me ajudaram a desenvolver estas idéias e colocá-las em forma (espero) compreensível. Entenda-se, claro, que a responsabilidade pelo texto é só minha; muitas vezes recebi sugestões e não as acatei, e posso me arrependar um dia.

Em primeiro lugar agradeço a Lúcia Fulgêncio, minha colaboradora constante, sempre exigindo clareza e objetividade no texto, coerência e adequação nas idéias. Para ela, uma porção dupla do quindim da gratidão.

Apreendi muito em conversas com Yara Liberato, então na UFMG, principalmente no que diz respeito aos temas tratados no capítulo 6. Agradeço ainda às colegas Beth Saraiva e Márcia Cançado (UFMG) por suas sugestões e críticas. Meus alunos da PUC-Minas, em especial Yara Bruno, George Manes Pereira, Luciana Mazur e Bruno Lima, também trouxeram sua contribuição ao ler e comentar partes deste texto durante meus seminários de sintaxe portuguesa.

Este livro é dedicado a meu amigo e colega Curt Blaylock, de Champaign, Illinois, com quem aprendi muitas coisas sobre lingüística e muitas mais sobre a difícil arte de viver.

Belo Horizonte, 2007
M.A.P.

PARTE I: PRELIMINARES TEÓRICOS

Capítulo 1: Por uma metodologia da descrição gramatical

1.1. Forma e conceito

O estudo da estrutura de uma língua tem como objetivo explicitar o relacionamento realizado pelo falante entre um sistema de conceitos e certos estímulos sensoriais (fonéticos, gráficos...), que podemos chamar de “formas”. Essa posição, claro, não é novidade: repete apenas o que já se encontra em Saussure (1916) acerca da associação de imagens acústicas com conceitos. Mas seria preciso levar isso a extremos metodológicos, o que me parece que em geral não se faz.

Por exemplo, *casa* e *casas* são **formas** distintas. Percebe-se um relacionamento gramatical e semântico entre elas, mas isso é resultado de um processo ativo de análise realizado pelo falante, com base em seu conhecimento da língua; e esse processo não é consequência automática e simples da semelhança formal, porque também temos *trabalha* e *trabalhas*, onde a relação formal é a mesma (presença ou ausência de um *-s* final) mas a relação conceitual é outra (3ª pessoa x 2ª pessoa), e *Dante* e *dantes*, onde não há nenhuma relação conceitual sistemática.

Assim, uma pergunta fundamental é:

“De que é que o receptor dispõe, em um primeiro momento, para decodificar uma seqüência formal?”¹

Uma resposta a essa pergunta (muito parcial; ver seção 1.2 abaixo) é:

“Ele dispõe da seqüência formal, acessível aos sentidos, mais seu conhecimento da gramática e do léxico.”

Isso, evidentemente, não é pouco. Mas exclui muita coisa; por exemplo, exclui a estrutura de constituintes – o receptor pode até precisar dela, mas vai ter que trabalhar para obtê-la. Exclui também a separação das palavras: o receptor vai ter que produzi-la, a partir de seu conhecimento léxico e gramatical, mais certas pistas fonológicas – fonologicamente falando, em português *espeto* poderia ser uma palavra só, ou duas: *es+peto*, mas não **esp+eto* nem **espet+o*, porque nenhuma palavra termina em *p* ou em *t*.

Para dar um exemplo, vejamos a frase:

[1] O fazendeiro matou um patinho.

Há boas evidências em favor da hipótese de que a interpretação dessa frase requer uma estruturação de constituintes do tipo

¹ Aqui me coloco do ponto de vista do receptor. O mesmo vale, *mutatis mutandis*, para o emissor.

[2] [o fazendeiro]_{SN} { [matou]_v [um patinho]_{SN} }_{SV}

Isso é uma das razões pelas quais *o fazendeiro* é interpretado nessa frase como uma unidade semântica, ao passo que *matou um* não é. Mas essa segmentação não faz parte do sinal sensorialmente perceptível, ou seja, não é foneticamente representada de modo direto. O receptor vai ter que trabalhar para chegar a ela, mais ou menos assim:

- a) as palavras são *o, fazendeiro, matou...* etc. (aqui o receptor se louva em seu conhecimento do léxico, com algum adjutório das regras fonológicas);
 - b) *o* é tipicamente início de um SN;
 - c) *fazendeiro* é aceitável como continuação de um SN iniciado por *o*; e o SN pode ficar só nisso;
 - d) *matou* só pode ser forma do verbo *matar*; e essa forma é aceitável como um verbo seguindo-se ao SN *o fazendeiro* (tanto a estrutura sentencial quanto a concordância estão corretas);
 - e) a seqüência processada até o momento (*o fazendeiro matou*) é aceitável como início de uma oração, cujo sujeito é *o fazendeiro* e o núcleo do predicado é *matou*;
 - f) essa seqüência, assim estruturada, permite esperar outro SN, que será o objeto direto (a julgar pela valência do verbo *matou*);
- etc.

O que quero dizer é que, ainda que a estruturação em constituintes (a “árvore”) possa ser necessária para a compreensão, ela é resultado de um processo de análise realizado pelo receptor.²

Note-se que se trata da formulação de hipóteses, não da aplicação de regras categóricas – já que, por exemplo, *o fazendeiro* pode ser um SN, mas pode também ser apenas parte de um (*o fazendeiro magro de Piracicaba*). As hipóteses vão sendo confirmadas ou eventualmente reformuladas; um exemplo em que a hipótese inicial precisa ser reformulada é a frase *os fazendeiros entram a aplicação da nova lei*, onde o verbo pode ser interpretado inicialmente como forma de *entrar* (com base na maior ocorrência de *entrar* frente a *entravar*), mas tem que ser reinterpretado depois por causa da continuação da frase, que não é compatível com a presença do verbo *entrar*.

1.2. Conhecimento e uso do contexto

O esboço acima é muito simplificado; e, em pelo menos um aspecto, excessivamente simplificado. Quando se disse que o receptor “dispõe da seqüência formal, mais seu conhecimento da gramática e do léxico”, um fator da mais alta importância foi desprezado: o conhecimento extra-lingüístico, tais como o de que um fazendeiro é um ser animado, e portanto capaz de matar; de que é muito mais de se esperar que um fazendeiro mate um patinho do que que um patinho mate um fazendeiro;

² Esse processo corresponde à elaboração das “fatias tipo 2” propostas em Perini, Fulgêncio e Rehfeld (1984).

e mesmo o conhecimento de que a frase em questão não foi enunciada em uma conversação espontânea, mas como exemplo, em um livro de lingüística.

No caso em pauta, tais informações têm valor secundário. Mas em outros casos elas são decisivas. Vejamos por exemplo a frase

[3] Esse cobertor vai esquentar demais.

Se a frase for enunciada por alguém que está escolhendo o cobertor com que vai para a cama dentro de alguns minutos, entenderemos que o cobertor vai esquentar demasiadamente o seu corpo. Mas se a situação for uma em que alguém deixa o cobertor em algum lugar ao sol, será mais fácil entender que a pessoa está preocupada com a possibilidade de o cobertor ficar excessivamente quente. Nesse caso, como se vê, o contexto é decisivo para que o receptor possa decodificar corretamente a frase (isto é, compreender adequadamente o que seu interlocutor queria dizer: o cobertor é o agente ou o paciente?).³ Assim, convém reformular a resposta dada à pergunta

“De que é que o receptor dispõe, em um primeiro momento, para decodificar uma seqüência formal?”

Agora diremos que ele dispõe:

- da seqüência formal (acessível aos sentidos);
- de seu conhecimento da gramática e do léxico;
- de seu conhecimento geral do mundo (armazenado no que os psicólogos chamam “memória semântica”);
- de sua percepção do contexto natural e/ou social em que a seqüência é enunciada (armazenada na “memória episódica”).

Se antes o receptor já não dispunha de pouco, agora ele parece estar bem mais armado para enfrentar a tarefa de compreender as seqüências que lhe chegam aos sentidos. E é bom que seja assim, porque mesmo com todo esse aparato a compreensão – muito em especial a compreensão da fala espontânea – opera sob condições bastante desfavoráveis de tempo (precisa ser muito rápida), de qualidade (nem sempre o sinal é suficientemente claro) e de limitações do próprio aparelho perceptivo (interferência de fatores tais como o ruído ambiente, o próprio cansaço do receptor etc.).

1.3. O que levar em conta na análise?

Diante dessa grande riqueza de fatores envolvidos na decodificação de seqüências, é um problema recortar o que é e o que não é imediatamente relevante para o estudo gramatical. E, de fato, boa parte das controvérsias da lingüística moderna gira em torno da questão de o que levar em conta na descrição da estrutura gramatical das línguas.

³ As noções de “agente” e “paciente” serão discutidas mais adiante. Por ora, aceite-se “agente”, é o causador do evento descrito, e o “paciente” a entidade que sofre os efeitos desse evento.

Por razões puramente práticas, o pesquisador pode fazer um corte preliminar e desprezar certos tipos de fatores – correndo o risco de deixar escapar informações fundamentais. Esse risco é, a meu ver, inevitável, dadas as limitações do cérebro humano, que é incapaz de enfrentar como um todo problemas verdadeiramente complexos. Se levarmos em conta um número excessivo de fatores, os dados se apresentarão de forma tão complexa que não nos será possível apreendê-los a ponto de podermos formular hipóteses úteis.

Exemplificando, digamos que alguém deseja apurar quais são os elementos que permitem a interpretação da frase

[4] Você pode fechar essa janela?

como um pedido ou como uma pergunta. Ele pode, com alguma segurança, desconsiderar elementos como: (a) a pronúncia reduzida da vogal final de *janela*; (b) a marca de pessoa-número do verbo *pode* como “terceira pessoa do singular”; (c) a pronúncia normal *cê* do pronome que se escreve *você*; (d) o fato de que *janela* é feminino; e assim por diante. O lingüista faz isso não por alguma razão de princípio, mas porque a investigação lingüística até hoje não encontrou relevância nesses elementos para a determinação da força ilocucionária de um enunciado. Ele está arriscando, mas não muito. Por outro lado, não seria sábio desprezar certos traços do contexto extra-lingüístico em que a frase foi enunciada, porque esses, sim, podem (segundo o que se sabe) determinar se a frase é uma pergunta ou um pedido, por exemplo. Outra coisa que não deve ser desprezada é a entonação, que freqüentemente distingue perguntas de pedidos; e assim por diante.

Os exemplos acima sugerem que a questão de que fatores ou tipos de fatores devem ser levados em conta quando se investiga determinado aspecto de uma língua não está solucionada através de princípios básicos de validade geral. E, com efeito, acredito que o lingüista precisa estar preparado para eventualmente tomar decisões a esse respeito.

Um exemplo instrutivo é a concordância verbal. Tradicionalmente, acredita-se que é um fenômeno tipicamente formal, isto é, morfossintático, e que considerações fonológicas, semânticas e/ou discursivas não teriam lugar em sua análise. No entanto, pesquisas recentes têm abalado essa posição tradicional. Já se argumentou que a concordância verbal é em parte condicionada por elementos do discurso; e eu suspeito que o fenômeno é complexo, e que inclui, pelo menos, fatores morfossintáticos e semânticos (ver, por exemplo, a proposta de LaPointe, 1980, e o capítulo 4 abaixo).

Um exemplo de interferência semântica em um fenômeno tradicionalmente rotulado de morfológico é o caso do acusativo de certos substantivos em russo. Os substantivos masculinos terminados em consoante têm o acusativo idêntico ao nominativo; assim, *vek* ‘século’ tem o acusativo *vek*. Mas se o substantivo denota um ser animado, o acusativo é idêntico ao genitivo, e não ao nominativo: *čelovek* ‘homem’ tem o acusativo *čeloveka* (idêntico ao genitivo *čeloveka*; no mais, *vek* e *čelovek* se declinam da mesma maneira). Como isso só acontece com nomes que denotam seres animados, a diferença **morfológica** entre *čelovek* e *vek* é previsível a partir da **semântica** desses itens. Na sintaxe, pode-se citar o exemplo do uso da preposição *a* em espanhol com objetos diretos referentes a seres humanos: *vi a mi hermana*, mas *vi la casa*. A existência de tais fenômenos de interferência (semântica + morfologia ou sintaxe) não está em dúvida. O problema é averiguar sua extensão.

Por outro lado, eu não defenderia a idéia de que qualquer fator, de qualquer natureza, seja em princípio relevante para a análise de qualquer aspecto de uma língua (como quando se afirma que o estudo de qualquer aspecto da língua só faz sentido se se levar em conta fatores discursivos). Acho essa posição metodologicamente inconveniente, primeiro porque desconsidera resultados importantes da investigação lingüística – a tradição não é uma vaca sagrada, mas tampouco pode ser simplesmente atirada pela janela. E, depois, como se viu acima, porque leva a uma situação de acúmulo tal de informações a serem levadas em conta que acaba paralisando a pesquisa.

Como superar então esse dilema? Será que uma ciência amadurecida não tem a responsabilidade de fornecer a seus investigadores princípios de seleção dentre o mundo de elementos que compõem seu objeto de pesquisa? Um biólogo, quando parte para a tarefa de classificar organismos ainda desconhecidos, está de posse de um sistema de hipóteses bastante rico e estruturado – ele não espera encontrar seres vivos que ao mesmo tempo elaborem clorofila e amamentem seus filhotes, nem animais capazes de respirar CO₂ puro.

E o lingüista, como fica? O que é que sua ciência lhe fornece como ponto de partida em sua investigação? A resposta é que temos muita coisa em que nos basearmos: um corpo de conhecimentos talvez menos estruturado, mas não necessariamente menor do que aquilo de que o biólogo dispõe. Mas, como não tem havido muito interesse em listar e organizar o que se tem, cada pesquisador precisa construir sua própria metodologia: um evidente desperdício de esforço, para dizer o mínimo. Roman Jakobson⁴ falava dos “invariantes da lingüística”, ou seja, aquilo que, no frígir dos ovos, todos os lingüistas dão (explícita ou implicitamente) como indiscutível. Jakobson sustentava que há muito mais em comum entre os lingüistas das diferentes vertentes teóricas do que dá a entender a eterna controvérsia que recheia as revistas e os congressos. Creio que isso é verdadeiro, e levo minha crença ao ponto de afirmar que em algumas áreas é possível elaborar um estudo descritivo bastante amplo e elaborado sem entrar no terreno minado das discussões teóricas (gerativistas x cognitivistas x funcionalistas x estruturalistas...). Não que essas discussões sejam, em princípio, indesejáveis; mas há muito mais a fazer em lingüística do que construir, destruir, defender e atacar teorias da linguagem. Voltarei a esse ponto, que é muito importante, na seção 1.7.

Saliento que essa ênfase nos dados não implica na crença de que estes são o nosso objeto central de estudo. O objeto de estudo da lingüística é o conhecimento da língua, e este é, inevitavelmente, de natureza psíquica. Mas tais fenômenos psíquicos não são acessíveis à observação direta, e é por isso que temos que lidar com dados da produção e da recepção para abordá-los. Esses dados consistem de enunciados realizados de maneira concreta (em última análise, fonética), mas também intuições,⁵ como quando o falante reconhece que *Alda vai casar com Breno* e *Alda e Breno vão casar* veiculam exatamente a mesma informação.

1.4. Algumas conseqüências

1.4.1. Dados e hipóteses

⁴ Em uma conferência que ouvi nos anos 70.

⁵ Ou melhor, percepções semânticas. O termo “intuição” é bastante usado, entretanto, nesse sentido.

Vamos examinar algumas conseqüências da posição metodológica que estou assumindo frente à pesquisa lingüística. Como se viu, advogo uma análise que parte de dois tipos muito “concretos” de informação, independentes em grande parte de teorização prévia, e identificados com os dois extremos em que os enunciados são diretamente acessíveis ao receptor: o som e o significado. Neste ponto, é bom observar, estou seguindo uma tendência adotada hoje por muitos lingüistas, como Langacker (1987; 1991) e C&J; ver. 2.4 adiante.

Digamos, então, que a forma *casa*, enquanto sinal fonético, é sensorialmente perceptível, e que o conceito ou conceitos a ela associados são imediatamente acessíveis à mente do receptor. As complexidades são muitas, evidentemente; não quero parecer ingênuo quanto à imediatidade desses acessos, mas vamos deixar isso de lado por ora. O que quero dizer é que, qualquer que seja a posição teórica do pesquisador, ele não pode negar que: (1) *casa* se inicia com uma consoante velar surda, seguida de uma vogal oral, baixa etc. e (2) *casa* se relaciona com um conceito geral referente a certo tipo de construção. Informações desse tipo são pré-teóricas (gramaticalmente falando) e precisam ser levadas em conta por qualquer análise completa da língua.

Note-se que isso pressupõe que existe alguma coisa como o significado “básico” ou “literal” das formas lingüísticas, em oposição a significados metafóricos ou derivados. É isso mesmo; em geral é possível identificar um significado não-marcado, que pode ser depreendido pelos usuários da língua com um grau muito alto de consenso. Ao se ouvir *casa*, pensa-se imediatamente em uma construção, não em outra coisa qualquer. Uma confirmação interessante desse fato é a seguinte observação:

[...] em Ciência Cognitiva 101, a primeira coisa que faço é dar um exercício para meus alunos. O exercício é: Não pense em um elefante ! O que quer que você faça, não pense em um elefante. Nunca encontrei um aluno que conseguisse fazer isso. Cada palavra, como *elefante*, evoca um esquema, que pode ser uma imagem ou outro tipo de conhecimento.

[Lakoff, 2004: 3]⁶

Estou certo de que o elefante em que os alunos de Lakoff pensam não é metafórico, nem derivativo: trata-se do animal grandalhão que todos conhecemos. Isso, a meu ver, ilustra bem a existência de um significado não-marcado das formas lingüísticas (ou, pelo menos, da forma *elefante*).

Essa posição tem corolários metodológicos bastante importantes. Por exemplo, não podemos **partir** da noção de que *casa* e *casas* são a mesma palavra, ou variantes da mesma palavra. São fonologicamente distintas, e isso é suficiente para que sejam consideradas duas **palavras**, e não apenas uma.⁷ Eventualmente, poderão ser agrupadas em um único **lexema** – mas é fundamental entender que o lexema não é uma unidade básica, primitiva, da análise (aliás, nem a palavra é, embora a gente acabe admitindo que é, para evitar conversas muito compridas sobre segmentação, etc. Isso é apenas uma hipótese de trabalho, útil para facilitar a discussão). O lexema é resultado de uma opção teórica – ou seja, um construto hipotético.

⁶ Tradução minha (aqui e em todas as citações de obras em língua estrangeira).

⁷ Uso **palavra** no sentido em que alguns autores falam de *word form* (cf. Trask, 1992).

O fato, pois, é que *casa* e *casas* têm pronúncia diferente, e são portanto duas unidades diferentes em um primeiro estágio do processamento. Paralelamente, *morro*, que pode ser uma forma verbal (*morro de medo de avião*) ou um nome (*o morro do Livramento*), também é uma unidade nesse primeiro estágio, porque não há diferença sonora entre a forma verbal e o nome. Em outras palavras, *morro* é uma palavra só, embora se relacione com dois conceitos distintos (o evento de “morrer” e o objeto “morro”). E, finalmente, *oculista* e *oftalmologista* são duas palavras (pronunciam-se diferentemente) e se referem a um único conceito (são sinônimos).

Em resumo, dois tipos de unidades são “dados” da análise lingüística: formas fonológicas (mais precisamente, fonéticas) e conceitos. Tudo o mais são resultados da própria análise, e por conseguinte unidades hipotéticas, sujeitas a justificação e questionamento. Não se discute se *casa* começa com vogal ou consoante; começa com consoante, e isso é um dado. Não se discute se *casa* designa um objeto material. Mas pode-se discutir a classe gramatical a que pertence, porque isso depende de definições e posicionamento teórico. Parece bem estabelecido, por exemplo, que *casa* e *janela* devem pertencer à mesma classe gramatical; mas isso não é indiscutível. Em princípio, *janela* e *casa* podem pertencer a classes diferentes – desde que alguém mostre que isso leva a uma análise mais econômica da estrutura do português e a uma melhor inserção dessa análise no que se sabe das condições psicológicas, históricas e sociais do uso da língua.

1.4.2. Palavra e lexema

Para ilustrar melhor a distinção feita acima entre palavra e lexema, vou acrescentar alguns exemplos. Vimos que as formas *casa* e *casas* pertencem ao mesmo lexema, mas são palavras diferentes. Isso é gramaticalmente importante, tanto é assim que *casa* e *casas* não têm distribuições idênticas: no contexto *as ... pegaram fogo* só cabe *casas*, não *casa*.

Um exemplo bastante impressionante da diferença entre palavra e lexema é dado pelas formas casuais de línguas como o latim, o russo etc. Por exemplo, ‘cidade’ em latim é *urbs*; mas essa palavra não apenas se relaciona com seu plural *urbes*, mas também com o acusativo *urbem*, o genitivo *urbis*, o ablativo *urbe*, etc. Essas diferentes formas (“palavras”) precisam sem dúvida ser reunidas em uma só entidade lingüística de determinado nível, que se caracteriza pelas propriedades comuns a todas elas: por exemplo, seu significado de “cidade”. Esse é o lexema *urbs* (mais exatamente, *urbs* + *urbem* + *urbis* + ...), que aparece no dicionário.

Mas as propriedades gramaticais das diferentes palavras que compõem o lexema são distintas. Tanto *urbs* quanto *urbem* podem ser classificadas, em termos tradicionais, como substantivos; mas sua distribuição não é idêntica, porque *urbs* ocorre como sujeito ou predicativo:

haec urbs magna est ‘esta cidade é grande’
Neapolis urbs est ‘Nápoles é uma cidade’

ao passo que *urbem* ocorre como objeto direto ou após certas preposições:

urbem uideo ‘estou vendo a cidade’
imus in urbem ‘vamos para a cidade’

Por outro lado, a distribuição do genitivo *urbis* se assemelha mais à de um adjetivo do que à de um substantivo (em termos tradicionais, tanto os adjetivos quanto *urbis* podem ocorrer como adjunto adnominal). Como se vê, é necessário distinguir palavras de lexemas, porque em certo nível não é possível expressar o comportamento gramatical de um lexema. Se dissermos simplesmente que o **lexema** *urbs* pode ser sujeito, objeto direto, predicativo, adjunto adnominal e adjunto adverbial, estaremos ocultando a parte mais importante do fenômeno, porque essa afirmação se aplica não ao lexema *urbs*, mas às palavras *urbs*, *urbem*, *urbis*, *urbe*, etc.

O mesmo se pode dizer, em linhas gerais, de grupos de palavras como, em português, *eu + me + mim*; em francês, *qui + que*; em inglês, *John + John's*.

1.5. Descrição e teorização

1.5.1. Relatórios descritivos

A descrição apresentada na segunda parte deste livro pode ser considerada o plano de um relatório descritivo – do tipo que se encontra em geral nas publicações da maioria das ciências naturais (paleontologia, zoologia, botânica, geografia...) e que não eram raros nas revistas de lingüística até os anos 60. Esses relatórios apresentam o resultado de um trabalho de levantamento e sistematização de dados, e têm a preocupação de ser tão exaustivos quanto possível.

A exigência de exaustividade (na medida do possível) é importante para evitar a seleção de evidência, voluntária ou involuntária, que constitui uma das pragas da metodologia lingüística atual. O pesquisador, na preocupação de encontrar evidência que corrobore sua teoria, seleciona dados favoráveis com muito mais energia do que a que utiliza na procura de dados desfavoráveis, apresentando dessa maneira uma imagem deformada da realidade lingüística em estudo. O resultado é muitas vezes o paradoxo de análises que, ao serem lidas, parecem solidamente fundamentadas, mas que ao enfrentar oposição real mostram uma fragilidade inesperada. Os contra-exemplos pipocam de todo lado, em alguns casos levando ao embaraço de se ter que responder que “no meu idioleto é assim”. A recente ênfase no uso de *cópus* deve sem dúvida ser vista como um importante avanço metodológico, no que pesem as restrições que se pode levantar quanto ao uso exclusivo de dados de *cópus* na investigação lingüística, especialmente gramatical. O *cópus*, se usado com critério, pode funcionar como elemento regulador, ajudando a evitar as deformações apontadas acima.⁸

Minha opinião é que a pesquisa lingüística precisa ser muito mais baseada em dados do que tem sido nos últimos tempos; que faltam dados, sistematicamente descritos, que dêem apoio à maioria das análises e teorias; e que o trabalho de levantamento de tais

⁸ Por outro lado, fique claro que o *cópus* não pode bastar. O *cópus* é resultado de muito mais do que o conhecimento da língua, e é impossível tirar dados dele sem lançar mão de julgamentos de aceitabilidade. Por exemplo, ao trabalhar com *cópus* para o estudo da sintaxe, temos que eliminar repetições, hesitações, mudanças de plano etc., que não podem ser descritas em termos gramaticais.

dados é algo que se deve encorajar pelo menos tanto quanto o de elaboração de novos modelos e teorias. Em outras palavras: sem desprezar a explicação, falta descrição.

O trabalho descritivo não é, evidentemente, neutro do ponto de vista teórico. Coisas como a escolha do tema, o recorte dos dados e os aspectos considerados relevantes para a classificação são inevitavelmente dirigidos por uma posição teórica. A única maneira realmente não-teórica de descrever a língua seria listar os dados, pura e simplesmente, o que todos concordamos que não é possível. Além do mais, uma lista sem análise estaria desprezando os progressos verdadeiros da lingüística, o “terreno conquistado”, que não é nada pequeno.

No entanto, acredito que a teorização pode ser mantida em um nível relativamente geral, de modo a maximizar a utilidade dos dados apresentados. O resultado, espera-se, é um conjunto de imagens que refletem a estrutura da língua de modo mais direto do que estudos essencialmente movidos a teoria – que, como se sabe, muitas vezes deixam de ter utilidade assim que a teoria correspondente passa de moda. O trabalho teórico, repito, é necessário; mas, para que tenha algum significado, deve basear-se em uma base de dados suficientemente ampla, elaborada com um mínimo de direcionamento prévio específico. Além disso, cada teoria precisa assumir a responsabilidade de, pelo menos, dar conta de mais dados do que suas concorrentes. Só assim se conseguirá progresso real na compreensão das línguas naturais – em oposição ao que se vê com freqüência, a simples troca de uma teoria mal fundamentada por outra teoria igualmente mal fundamentada. O objetivo imediato de relatórios como os que proponho é oferecer descrições que possam pelo menos “salvar as aparências”, ou seja, ser compatíveis com uma porção significativa dos dados tais como podem ser observados.⁹ Esse não pode ser o objetivo último de um estudo lingüístico, mas precisa ser atingido em um primeiro momento, sob pena de comprometer a relevância do estudo para a compreensão do fenômeno.

As considerações feitas acima configuram uma concepção da lingüística como “história natural”: uma disciplina voltada de preferência para o levantamento e a sistematização de dados, reconhecendo a inexistência de um paradigma¹⁰ no estudo sincrônico da linguagem. É com efeito minha opinião que a lingüística se encontra em um estágio pré-científico se comparada com outras disciplinas tais como a física e a biologia. Não vejo por que isso nos deva preocupar: a situação não se deve a incompetência ou indolência dos lingüistas, nem à alegada juventude da disciplina (que, aliás, não tem nada de recente – remonta, pelo menos, ao período helenístico, isto é, ao terceiro século a.C.). O problema está, acredito, na natureza dos fatos estudados, na sua complexidade e na dificuldade de acesso à evidência. Assim como Bacon ou Aristóteles não precisavam envergonhar-se de sua física, nós hoje em dia podemos encarar o estado da lingüística com tranqüilidade.

Por outro lado, para se beneficiar de trabalhos que apresentam dados devidamente organizados não é necessário aceitar a posição assumida no parágrafo precedente. Tais levantamentos são de utilidade óbvia para qualquer pesquisador, pois fornecem um campo de testagem para análises e eventuais teorias gerais.

Um aspecto fundamental do trabalho descritivo é a taxonomia das formas. A teorização se baseia às vezes em taxonomias indefensáveis – como por exemplo um trabalho que pretenda examinar algum aspecto do uso dos “advérbios” sem previamente

⁹ A noção de “salvar as aparências” em ciência foi estudada por Duhem (1908).

¹⁰ No sentido de Kuhn (1962).

criticar a coerência e o significado gramatical dessa classe tradicional. Para pegar um exemplo concreto, encontramos no manual de Haegeman (1991) afirmações do tipo

“Substantivos e adjetivos não atribuem caso em inglês” [p. 162]

“[...] predicados adjetivais também induzem o alçamento.” [p. 295]

Afirmações como essas se encontram por toda parte; no entanto, os mesmos trabalhos não oferecem uma conceituação operacionalizável do que vem a ser um “adjetivo” (em especial, em que se distingue um adjetivo de um substantivo – questão nada trivial, como se tentou mostrar em Perini *et al.*, 1996). A situação é ainda mais grave com a pretensa “classe” dos pronomes, extremamente heterogênea, mas muitas vezes tratada como um grupo, sem maiores discussões, na literatura atual.

A classe dos verbos é certamente mais homogênea, muito nitidamente distinta das demais classes de palavras. Mas as subclasses de verbos são pouco conhecidas, embora a descrição da distribuição e da semântica dos verbos dependa crucialmente dessas subclasses. Na segunda parte deste livro estudo um aspecto da taxonomia dos verbos, a saber, suas valências.

Ao proceder à classificação dos verbos, é necessário estabelecer critérios para essa classificação; ou seja, propor uma análise. Trata-se de uma análise descritiva, baseada em traços, que, acredito, são reconhecidos como importantes pela maioria dos lingüistas de todas as tendências. Procurarei mostrar que esses traços precisam fazer parte do conhecimento prévio do falante, possibilitando a codificação e a interpretação de mensagens. Se a proposta for correta, a análise desenvolvida nos capítulos 7 a 13 constituirá uma base adequada para a cartografia de uma parte importante do léxico português, a saber, a classe dos verbos.

1.5.2. Colaboração ou rivalidade?

É evidente que a tarefa da lingüística não se esgota na descrição das línguas naturais. Além disso, coloca-se a questão de investigar o que as línguas naturais têm em comum, aquilo que toda e qualquer língua humana precisa ter; ou seja, é preciso construir uma teoria da linguagem humana.

À primeira vista se poderia esperar que essas duas tarefas fossem levadas a efeito em colaboração – alguns lingüistas se concentrariam na descrição das línguas particulares, outros partiriam dos dados coligidos e sistematizados pelos primeiros para elaborar hipóteses universais sobre a estrutura da linguagem humana. Isso foi expresso por Chomsky da seguinte maneira:

[...] o problema mais crucial para a teoria lingüística parece ser o de abstrair informações e generalizações de gramáticas particulares descritivamente adequadas e, sempre que possível, atribuí-las à teoria geral da estrutura lingüística [...]

[Chomsky, 1965: 46]

No entanto, se examinarmos o campo dos estudos lingüísticos tal como se apresenta hoje em dia, veremos que o panorama é muito diferente. Temos, por um lado, lingüistas dedicados à descrição das línguas particulares, alguns dos quais dão muito pouca atenção às teorias gerais da linguagem. Alguns mesmo, ao tentarem se inteirar do que acontece no campo dos lingüistas teóricos, se sentem desencorajados, ou até amedrontados pela diversidade de posições entre eles. Por exemplo, Orrin Robinson, em um estudo descritivo das línguas germânicas antigas, explica que vai se concentrar na morfologia e fonologia das línguas estudadas. Comenta a dificuldade de se estudar a sintaxe de uma língua antiga e pouco atestada, e acrescenta a seguinte passagem:

[...] com todas as mudanças que ocorreram, e continuam ocorrendo, na teoria sintática nas últimas décadas, não temos segurança da maneira correta de apresentar o que achamos que sabemos. Na falta de um quadro de referência estável para a descrição das línguas, poucos pesquisadores desejam se ocupar da descrição mesmo de uma língua viva, quanto mais de uma língua morta.

[Robinson, 1992: 163]

A situação é de preocupar: o estado dos estudos teóricos é tal que acaba afastando os lingüistas da própria descrição das línguas particulares. Evidentemente, não é dessa maneira que vamos conseguir desenvolver a teoria lingüística em bases sólidas.

Do outro lado do muro, os construtores da teoria da linguagem freqüentemente pecam, a meu ver, por não darem a devida atenção a trabalhos descritivos abrangentes. Um número excessivo de trabalhos se concentram em analisar e reanalisar (segundo a moda do dia) os mesmos setores, às vezes muito restritos, das mesmas línguas: basta aqui lembrar a insistência dos estudos que repetidamente se concentram na análise dos clíticos, das mini-oracões, das condições de movimento de “ α ” e outros problemas favoritos da teoria lingüística atual. Às vezes os pesquisadores procuram aplicar os resultados teóricos obtidos dessa maneira, e que em si já são bastante inseguros, a outras línguas, com o objetivo de mostrar que os construtos teóricos elaborados para uma língua são válidos para outras.

Este último tipo de investigação, aliás, exigiria muita cautela, porque um dos objetivos da teoria lingüística é o de mostrar o que é que as línguas têm em comum (os “universais da linguagem”). A meu ver, esse objetivo requer que as línguas sejam analisadas, em um primeiro momento, cada qual em seus próprios termos, com um mínimo de pressupostos de análise e com atenção escrupulosa a seus modos próprios de relacionar a face formal e a face semântica. Do contrário, estaremos afirmando previamente o que pretendemos investigar, e os eventuais resultados ficarão irremediavelmente viciados. Mas nem sempre se enxerga essa cautela. Alguns poderão dizer que essa exigência dificulta muito a tarefa do tipologista das línguas – uma objeção que é verdadeira sem ser relevante. Se a tarefa é complexa, simplificá-la a todo custo significa pôr em perigo os resultados.

1.6. Nota sobre o momento em lingüística

Tem-se discutido muito os “paradigmas” da lingüística, partindo da famosa proposta de Kuhn (1962). Frequentemente se fala do “paradigma gerativista” como mais ou menos vigente no momento (ou, para alguns, até recentemente). Na minha opinião, não há atualmente paradigmas em lingüística sincrônica. Se houve algo parecido com um paradigma, deve ter sido o modelo (diacrônico) da gramática comparativa do século XIX – mas mesmo esse só se sustenta mediante uma definição bastante restrita de “lingüística”, excluindo o trabalho de descrição e análise sincrônica das línguas. A atividade de descrição continuou durante todo o período, embora alguns (como Paul, 1884) lhe negassem caráter científico.¹¹

Segundo Kuhn, um paradigma é uma teoria unificada, que dá conta de uma parte considerável dos dados, e que é aceita como pressuposto básico pela quase totalidade dos pesquisadores. O primeiro paradigma, segundo Kuhn, foi criado por Newton para a física, e vigorou até que Einstein propôs uma nova teoria que dava conta de tudo o que Newton explicava, mais alguma coisa.

Nada disso ocorre em lingüística sincrônica. O gerativismo (como Chomsky não cansa de repetir) é, e sempre foi, minoritário. Nenhuma posição teórica abrangente é aceita pela quase totalidade dos profissionais da área. E quanto a dar conta dos dados, o triste fato é que muitos pesquisadores dão pouca atenção aos dados. Há pouco diálogo entre as diversas vertentes, e quando há é inconclusivo porque cada grupo está entrincheirado não apenas em seus pressupostos, mas também em seu conjunto selecionado de dados favoritos.

Há quem diga que não há falta de dados, o que há é falta de teorização. Discordo. Acho que faltam dados, e principalmente faltam dados sistematizados, classificados segundo algum critério amplamente aceitável. Obter e sistematizar esse corpo de dados de modo a possibilitar a comparação interlingüística é certamente difícil, mas é possível.

Mas sistematizar dados ainda não é construir teorias explicativas. O que foi que se fez em física antes de Newton? Ou em biologia antes de Darwin? Basicamente, me parece, o que se fez foi coletar e sistematizar dados, assim como construir teorias parciais, que cobriam uma parte restrita dos dados disponíveis; é o que chamei o estágio de “história natural” de uma ciência. Newton só pôde elaborar sua teoria porque dispunha de dados confiáveis, numerosos e sistematizados (bem como teorias parciais) legados pelos seus antecessores: Ptolomeu, Copérnico, Galileu, Brahe, Kepler e outros que, se trabalhavam sem um paradigma, eram não obstante cientistas, e grandes cientistas. Não há portanto nenhum demérito em trabalhar no recolhimento de dados e em elaborar teorias parciais.

Por conseguinte, e sem desmerecer o esforço, que deve ser contínuo, de elaborar teorias – afinal, não se sabe quando é que vamos estar preparados para o nosso primeiro paradigma – eu sugeriria que a maior parte do trabalho dos lingüistas deveria se concentrar na obtenção e sistematização de dados das milhares de línguas faladas hoje (e em outros tempos) no mundo. Isso foi sustentado com bastante veemência já nos anos 70 por Maurice Gross, que começou tentando elaborar uma gramática gerativa

¹¹ A gramática tradicional da época talvez fosse um paradigma, apresentando pelo menos o aspecto sociológico de ser aceita pela maioria dos profissionais; essa variedade de gramática, no entanto, não é mais adotada em trabalho de pesquisa há mais de 50 anos, embora continue firme no ensino médio.

(transformacional) do francês, e acabou chegando à conclusão de que faltava um trabalho descritivo preliminar. Gross observou que

Acumular dados não é, obviamente, um fim em si mesmo. Mas em todas as ciências naturais é uma atividade fundamental, uma condição necessária para avaliar a generalidade dos fenômenos.

[Gross, 1979: 862]

A crítica de Gross é por vezes muito severa, mas creio que era justa na época, e ainda hoje se aplica a boa parte da pesquisa lingüística, gerativa ou não. Diz ele:

Há [...] uma inconsistência comportamental fundamental entre (a) a idéia de que a gramática gerativa fornece uma base para descobrir a teoria da linguagem e (b) a total ausência de um programa de construção de gramáticas de línguas particulares. A abordagem gerativa, inicialmente justificada em razão de suas contribuições para tornar precisas uma grande variedade de procedimentos gramaticais, chegou a um ponto em que a pesquisa lingüística baseada no trabalho empírico sistemático é descartada como irrelevante. A obtenção de uma imagem tão completa e detalhada quanto possível de qualquer língua não é mais uma tarefa para essa “lingüística”.

[Gross, 1979: 871; aspas no original]¹²

Acredito que a situação é um pouco melhor hoje em dia, mas ainda me parece que a ênfase em teorização sobre dados insuficientes é ainda uma das marcas do trabalho em lingüística (e não apenas gerativa). É urgente, portanto, enfatizar muito mais a elaboração de descrições particulares tão amplas e abrangentes quanto possível.

Ora, isso pressupõe, entre outras coisas, uma metodologia que oriente a coleta de dados. E é sintomático que os lingüistas atuais dêem tão pouca atenção a esse particular. Um geólogo começa sua formação aprendendo a distinguir tipos de rochas, pelo seu aspecto, pela sua composição química etc. Já um lingüista habitualmente começa aprendendo teorias, e em geral nunca recebe treinamento sério na obtenção e sistematização de dados.¹³ No atual momento as teorias em lingüística não são muito confiáveis. Valem como hipóteses de trabalho, mas precisam estar constantemente em questão. Precisam ser testadas com seriedade, fazendo um esforço consciente no sentido de colocá-las em dificuldades, procurando dados que as refutem. A prática usual é a oposta: o pesquisador procura dados que confirmem sua teoria preferida. E, como resultado desse tipo de metodologia, os dados quase sempre acabam confirmando a teoria.

¹² Gross já nesse momento vinha pondo em prática suas idéias em uma série de trabalhos seus e de seus colaboradores; um bom exemplo é Gross (1975), resenhado no capítulo 7.

¹³ É preciso notar, no entanto, que a chamada “lingüística de córpus” representa um passo importante na direção correta. Mas o uso de córpus ainda é característico de certos grupos e de certo tipo de trabalho, e os dados da lingüística também não podem se limitar àquilo que um córpus pode fornecer. Os psicolingüistas e sociolingüistas também dão importância ao treinamento dos pesquisadores em uma metodologia específica. A falta é aguda, porém, na área gramatical.

Uma das razões para isso, acredito, é que os dados são **selecionados**. Como não há metodologia que garanta a imparcialidade da obtenção de dados, o lingüista confia em sua intuição e na de seus colegas, arranja alguns dados convenientes e não faz um esforço ativo no sentido de procurar dados inconvenientes. Com certa frequência, livra-se de dados embaraçosos com desculpas como “não estou interessado nesses fenômenos”, ou então “é preciso dar à teoria espaço para respirar”. Com essa (pseudo-)metodologia, a confirmação é quase automática. O resultado é não a corroboração da teoria (no sentido de Popper, 1959) mas um aumento da confiança que o próprio lingüista tem em sua teoria escolhida. Uma sensação de segurança, psicologicamente satisfatória, mas sem valor real para o desenvolvimento da ciência.

Advogo, para o momento atual, o cultivo de uma lingüística **descritiva**, assessorada por uma metodologia adequada de obtenção de dados e caracterizada por uma extrema cautela e um alto grau de exigência na elaboração de teorias. Como bem aponta Gross, essa lingüística descritiva não pode ser a única meta da ciência da linguagem; mas é um passo necessário à elaboração de teorias mais gerais.

O problema da elaboração de uma metodologia era considerado central pelos estruturalistas, que produziram diversas propostas muito específicas: pode-se citar Longacre (1968), e principalmente Harris (1951). Essas propostas, no entanto, não são utilizáveis atualmente porque se vinculam a uma concepção muito estreita do que se entende por “fato lingüístico”. A partir da ascensão do gerativismo, o problema parece ter caído em relativo esquecimento; algumas propostas muito interessantes, como a de Itkonen (1974) e a de Schütze (1996) não tiveram quase impacto nenhum sobre a massa dos lingüistas. Como resultado, a pesquisa continua sendo baseada em dados pinçados ao acaso da lembrança e inspiração de cada um, e hoje colhemos o fruto na forma de teorias altamente elaboradas mas que têm pouca relação com a realidade das línguas – teorias de vida útil tão curta que não chegam a desempenhar uma das funções primordiais das teorias, que é a de orientar o trabalho empírico de descrição (veja-se a passagem de Robinson citada acima).

1.7. Em que basear uma metodologia?

A esta altura, vale perguntar o que foi que os lingüistas descobriram sobre a estrutura das línguas desde que os gramáticos helenísticos iniciaram o trabalho de descrever as línguas, por volta do terceiro século a. C. Acredito que existe um grande corpo de conhecimentos que orienta o trabalho de todos os lingüistas atuais, muito embora em geral não se tenha consciência plena desses pontos de convergência. De outro modo, não seria possível fazer lingüística hoje.

Vou fazer um apanhado de alguns desses pontos – mas uma lista completa é algo que nunca se fez. O importante aqui é deixar claro que esse corpo de conhecimentos comuns aos pesquisadores da linguagem é a base que temos para começar a construir teorias mais satisfatórias e uma metodologia para dirigir e controlar nossos esforços de descrição das línguas. Assim, como um simples exemplo do que se precisa fazer, dou a seguir uma lista (parcial!) de pontos em que os lingüistas estão de acordo. Não se trata de um paradigma, porque esses pontos não se integram em uma teoria da linguagem; trata-se

de hipóteses de trabalho, que orientam o levantamento de dados e facilitam o diálogo entre os pesquisadores (quando há diálogo).

Qualquer lingüista ocupado com a descrição gramatical conhece e utiliza em seu trabalho as noções seguintes, com extensões toleravelmente uniformes:

Níveis de análise:

palavra
lexema
morfema
constuinte
oração
período
[...]

Funções:

sujeito (gramatical)
núcleo
modificador
[...]

Classes:

artigo
verbo
preposição
pronome pessoal
[...]

Categorias morfológicas:

palavra variável x invariável
tempo verbal
caso (explícito, representado morfológicamente)
número
gênero
pessoa
sufixo, prefixo, infixos
radical
derivação
composição
flexão
[...]

Relações estruturais:

concordância nominal
concordância verbal
regência, valência
coordenação

subordinação
[...]

Categorias semânticas:

referência
modificação
agente
paciente
aspecto
sinonímia
paráfrase
polissemia
anáfora
dêixis
[...]

Categorias discursivas:

tópico
foco
contraste
[...]

As definições variam, naturalmente: um linguísta vê a concordância verbal como um fenômeno sintático, resultado de regras puramente formais, enquanto outro a vê como um fenômeno complexo, parte formal, parte semântico, parte discursivo. Mas como rótulo de um grupo de fenômenos, “concordância verbal” é eficiente; quando alguém vai fazer uma comunicação sobre esse tema, qualquer profissional do ramo sabe com bastante aproximação do que ele vai tratar. Ou seja, esses rótulos têm, antes de tudo, um valor extensional, delimitando domínios dentro do campo geral da gramática: apesar das controvérsias sobre como analisar os verbos, há muito pouca dúvida sobre se determinada palavra é ou não é um verbo.

Algumas das noções acima são questionadas, mas apenas por pesquisadores isolados ou muito minoritários. Por exemplo, Martinet (1979) questiona a validade da palavra como nível de análise; mas, que eu saiba, muito pouca gente o segue neste particular. Outras, embora de uso geral, são muito mal definidas, como a maior parte das funções sintáticas: sujeito, modificador etc. No entanto, incluo-as na lista porque são de uso geral, e os linguístas costumam concordar nas análises (por exemplo, dada uma frase, identificam o mesmo sintagma como sujeito), ainda que por razões diversas.

Me parece que essa coincidência extensional é principalmente baseada em definições semânticas; ou seja, quando se fala do “sujeito” em latim, em guarani ou em inglês, fala-se de algo como “o elemento que exprime o agente de verbos que significam ‘comer’ (na voz ativa) e o paciente de verbos que significam ‘morrer’”.¹⁴ Isso é de se esperar, já que a semântica é um dos pontos “concretos”, onde a língua por assim dizer encosta na realidade. Já a sintaxe é basicamente relacional, composta principalmente de hipóteses, e por conseguinte é também o campo das divergências mais profundas.

¹⁴ Ver a respeito minha crítica de Keenan (1976), em Perini (1985).

Quando se fala de “substantivo”, portanto, em geral se está falando não da palavra que pode ser precedida de artigo, ou que varia em caso e número, mas da palavra que “nomeia os seres”. Acredito que é isso que permite o limitado diálogo que se processa entre lingüistas de diferentes tendências.

Mas o campo de consenso não se limita a noções delimitadas como as da lista acima. Como vimos, um lingüista interessado nos fatores que determinam a força ilocucionária de um enunciado não pensaria em levar em conta particularidades da fonética, nem da morfologia, nem o gênero dos substantivos, etc., porque sabe que tais elementos não são relevantes para seu objetivo.

Aqui temos um grande campo a explicitar: ninguém tenta descrever a estrutura da oração em termos fonológicos, por exemplo. Não se procura prever a probabilidade de ocorrência de um fonema particular no início de uma oração, pois se sabe que não há nada de gramaticalmente interessante na distribuição ou no significado das orações começadas por /f/. Sabe-se que a exigência de um –s no final de *grandes* em *casas grandes* tem a ver com o –s de *casas*, e que não se trata de fenômeno ligado ao fonema /s/, mas antes ao morfema {-s “plural”}. Aceita-se universalmente que morfemas podem ter alomorfes, ou seja, em determinado nível de análise o –s de *casas* e o –es de *rapazes* são a mesma coisa.

E por aí vai. É fácil ver que a área de consenso é muito grande; falta, e seria muito útil, delimitá-la e explicitá-la. A partir daí seria possível estabelecer o elenco das hipóteses de trabalho que norteiam o trabalho de pesquisa em lingüística e que devem inclusive fazer parte do treinamento dos lingüistas, assim como a lista dos períodos geológicos faz parte do treinamento dos geólogos e paleontólogos, assim como a classificação periódica dos elementos faz parte do treinamento dos químicos, e assim como a diferença entre Idade Média e Antiguidade faz parte do treinamento dos historiadores.¹⁵

1.8. Guias de descrição

1.8.1. Observações preliminares

Vamos tratar agora de algumas guias para a descrição gramatical. Não falo de um “modelo descritivo” porque não me parece que seja possível oferecer um conjunto estruturado, nem toleravelmente completo, nem muito menos teoricamente motivado como um todo. Podem ser entendidas como hipóteses de trabalho, elas próprias constantemente sob escrutínio. Basicamente, essas guias derivam das considerações sobre forma e conceito feitas na seção 1.1, levadas, como disse, a seus extremos metodológicos.

¹⁵ Alguns colegas podem achar que sou excessivamente otimista quanto à possibilidade de diálogo no âmbito de uma disciplina que carece de um paradigma. É claro que não posso ter certeza, mas estou disposto a pagar para ver, e apostar que, havendo boa vontade, o diálogo é possível em muito maior medida do que se verifica atualmente. Creio que o próprio Kuhn, no posfácio que acrescentou a seu livro em 1969, dá a entender que a troca de dados e de idéias pode ser útil mesmo na ausência de um paradigma comum (pp. 223-225 da tradução brasileira).

Nos aspectos que nos interessam, concebemos o funcionamento de uma língua como a interpretação semântica de seqüências sonoras (ou gráficas). Por exemplo, diremos que a frase

[8] Daniela vendeu o carro.

se interpreta através de regras semânticas¹⁶ parcialmente diferentes das que interpretam a frase

[9] O carro, Daniela vendeu.

Nos dois casos o sintagma *o carro* representa o “paciente”, e isso será analisado como resultando de duas regras independentes: uma que (em uma frase com o verbo *vender* na forma chamada “ativa”) atribui o papel de “paciente” ao SN imediatamente seguinte ao verbo, e outra que atribui o mesmo papel ao primeiro SN da sentença, antes do sujeito. Pode perfeitamente ser que essas duas regras tenham conexões, e quanto a isso serei neutro. As duas regras propostas descrevem adequadamente os fatos tais como são observados, e isso é suficiente para nossos objetivos imediatos. Não podemos, por ora, ter a preocupação de procurar a formulação mais econômica possível, e muito menos a de procurar a formulação que melhor se coaduna com eventuais resultados da comparação entre as línguas, ou da pesquisa psicolinguística.

Em outras palavras, nossas guias deixam de lado, em parte, a questão da economia (ou simplicidade), concentrando-se no objetivo mais modesto de descrever os fatos sistematicamente, ainda que à custa de repetições, e ainda que deixando escapar generalizações. Em certos casos pode ser inclusive prudente **evitar** generalizações, porque elas implicam em previsões que podem ser incorretas. É necessário ter sempre em mente que nosso objetivo é **descritivo**; tentamos atingir o que se chama “adequação observacional”. Queremos um dispositivo que nos permita enumerar todas as estruturas da língua, e apenas elas.

Um dos objetivos desse esforço é o de estabelecer uma **linguagem** (isto é, um sistema de nomenclatura) capaz de exprimir os traços principais da estrutura gramatical de uma língua, de forma a permitir a listagem de seus padrões gramaticais – uma linguagem que permita comunicar os resultados de pesquisas empíricas sem se vincular estritamente a pressupostos teóricos muito particularizados. Em última análise, uma linguagem apta a veicular relatórios descritivos. Sustento que o estabelecimento dessa linguagem é uma tarefa urgente, para evitar um grande desperdício de esforço na área do levantamento de dados das línguas naturais. E acredito que a comunicação entre os lingüistas é possível em muito maior grau do que o que se observa atualmente, mas não se realiza, muitas vezes, justamente por causa da diversidade de linguagens utilizadas.

A coincidência de noções apontada na seção precedente não significa coincidência de nomenclatura, que varia enormemente e com freqüência atrapalha o diálogo. Vejo com preocupação a tendência a utilizar os mesmos termos em acepções radicalmente diferentes, justamente porque dificulta as já precárias condições de comunicação entre os lingüistas de diferentes tendências. Por exemplo, *tema* significa, para alguns, o tópico; para outros, é um papel temático que inclui o paciente de uma ação. Os gerativistas têm

¹⁶ Mais precisamente, regras simbólicas.

muita culpa a expiar neste particular: por exemplo, freqüentemente se referem ao “sujeito” de um sintagma nominal, ou ao “determinante” de uma oração, levados pelas implicações da Teoria X-barras, que afinal de contas nem todos aceitam.

Proponho, então, a realização de trabalhos de levantamento e sistematização de dados das línguas naturais e a progressiva elaboração de um sistema de unidades e relações gramaticais de aceitação geral entre os lingüistas; e insisto na construção paralela de uma linguagem para exprimir esses dados. Isso pode parecer muito pouco, principalmente se comparado com os objetivos muito mais ambiciosos de boa parte da pesquisa gramatical, que pretende atingir a “adequação explicativa”. Mas sem adequação observacional, isto é, sem dados adequadamente descritos e sistematizados, os estágios ulteriores não podem ser atingidos. E acrescento (assumindo o risco de levar pedradas de mais de um lado) que os resultados de boa parte da pesquisa gramatical atual não alcançam mais do que uma mera ilusão de adequação descritiva ou explicativa.¹⁷

Mesmo o modesto objetivo aqui proposto só é alcançável em termos relativos. Como Hockett (1968) já apontou, uma língua não é um sistema bem definido, e freqüentemente é impossível decidir com clareza se uma seqüência sonora pertence ou não ao elenco das seqüências bem formadas da língua. Por ora vamos ignorar essa observação, porque não há falta de problemas. Mas, naturalmente, ela terá que ser levada em conta na eventual elaboração de uma metodologia da pesquisa lingüística.

Repito, correndo o risco de parecer redundante, que as limitações acima se aplicam ao trabalho descritivo concebido como “história natural”, isto é, como tentativa de retratar uma realidade. Para mim é evidente que qualquer teoria gramatical adequada precisa adequar-se de alguma forma aos resultados de outras áreas que se ocupam da linguagem, em especial a psicolingüística, e não se pode escapar dessa tarefa; mas tudo a seu tempo. O trabalho descritivo, além de seu interesse intrínseco, é também um dos degraus necessários para conseguir essa compreensão abrangente e integrada do fenômeno da linguagem.

Ele também condiciona uma parte importante do treinamento profissional dos lingüistas, a saber a coleta e interpretação preliminar de dados. Um aspecto fundamental da formação de cientistas de muitas áreas é a análise de dados; assim, um geólogo aprende a identificar minerais, e um biólogo se capacita a classificar plantas e animais, através de critérios uniformes baseados em traços concretos dos espécimes. Ninguém, nessas áreas da ciência, pensaria em dispensar esse tipo de treinamento, que é visto como essencial ao pesquisador praticante. No que pesem as importantes diferenças entre a lingüística e as ciências naturais, a análise preliminar de dados é igualmente crucial na formação do pesquisador; mas ela depende do estabelecimento de uma linguagem basicamente conhecida e aceita pela maioria dos profissionais da área. Essa linguagem, como já deixei claro acima, é possível em grande medida, e é preciso que seja explicitada o quanto antes.

1.8.2. Partindo de exemplos

¹⁷ É interessante notar que Chomsky, em 1965, parecia concordar com isso, a julgar pela passagem seguinte: “Evidentemente, seria utópico esperar conseguir adequação explicativa em larga escala no atual estado da lingüística.” [Chomsky, 1965: 26]. À luz dessa opinião, seria talvez de esperar que os gerativistas fossem mais cuidadosos em suas tentativas de atingir a adequação explicativa.

1.8.2.1. Sujeito “oculto”

Nosso princípio básico, não custa relembrar, é partir dos dois níveis diretamente acessíveis aos falantes: a forma e o significado. Aqui, para simplificar, a forma será representada pela ortografia. Seja o exemplo seguinte:

[10] Comi uma pizza.

Essa é a forma da seqüência. De seu significado, que é complexo, vamos selecionar alguns pontos para estudo (represento elementos do significado entre aspas):

“o agente sou eu”

“a ação é a de comer”

“a ação ocorreu no passado”

“o paciente é a pizza”

“a pizza é apenas uma”

“a pizza não foi mencionada anteriormente no discurso”¹⁸

Todos esses elementos (e outros) estão presentes na representação semântica que o receptor constrói em sua mente a partir de sua percepção da frase [10].

Agora vamos considerar a representação formal de alguns desses ingredientes semânticos. A informação de que a pizza ainda não foi mencionada está codificada na presença do artigo indefinido: *uma pizza*, em vez de *a pizza*, que sinalizaria a informação de que a pizza já foi mencionada. A informação de que a pizza é apenas uma está codificada no número (singular) de *pizza* e de *uma*.

A informação de que “o agente sou eu” está codificada na desinência verbal *-i* (que, cumulativamente, também sinaliza outras informações, como o tempo). Creio que isso será aceito por qualquer pessoa que pense um momento: em [10], a desinência *-i* é a **única** pista de que o receptor dispõe para obter a informação de que “o agente sou eu”.

Na verdade, nossa decisão de relacionar, em um primeiro momento, apenas os dois níveis observáveis nos impede de partir de elementos abstratos tais como sujeitos ou objetos ocultos (elípticos, vazios...), vestígios (*traces*), casos gramaticais não explícitos na morfologia (como os que marcariam os sintagmas nominais na análise gerativa) e estruturas de constituintes (árvores). No primeiro momento, o receptor não tem acesso a nenhum desses elementos, e portanto precisamos partir de seqüências que não os contenham. O fato de o receptor processar frases a partir apenas de informação sensorialmente perceptível¹⁹ precisa ser incluído na análise.

Isso não quer dizer que tais elementos abstratos não possam ser necessários em algum estágio da descrição. Ninguém tem ilusões acerca da complexidade extrema da relação som/significado. Em princípio, estou perfeitamente disposto a aceitar elementos abstratos na análise; mas insisto em dois pontos: (a) esses elementos não podem ser **ponto de partida** da análise. E (b) a postulação de um elemento abstrato requer

¹⁸ Mais exatamente, o emissor acredita que o conceito da pizza não está presente na “consciousness” do receptor; ver a respeito Chafe (1976); Liberato (1980).

¹⁹ Utilizando, é claro, sua informação prévia sobre a língua e o mundo.

fundamentação cuidadosa: é essencial demonstrar, para cada elemento abstrato introduzido na análise, que não há mesmo alternativa não-abstrata para ele. Os elementos abstratos são, por assim dizer, um mal necessário.²⁰

A conclusão é que temos que postular uma diferença de análise entre as frases [10] e [11]:

[10] Comi uma pizza.

[11] Eu comi uma pizza.

porque existe uma diferença formal, a presença de *eu* em [11] e sua ausência em [10]. Essa diferença formal não acarreta, no caso, diferença semântica, mas ainda assim não se pode dizer que as duas frases são totalmente equivalentes. [11] é redundante em um sentido em que [10] não é – em [11] há duas fontes para a informação de que “o agente sou eu”, e em [10] há apenas uma.

Uma maneira adequada de exprimir essa diferença é dizer que [11] tem sujeito (*eu*), mas [10] não tem sujeito. Se distinguirmos entre sujeito (uma noção formal) e agente (uma noção semântica), deve ser fácil compreender por que [10] não tem sujeito – muito embora sinalize adequadamente o agente como sendo “eu”²¹. Essa conclusão é necessária se aceitarmos nosso postulado de que a análise deve, em determinado nível, basear-se nos dados imediatamente acessíveis ao receptor – pois não há dúvida de que, em determinado estágio, o que ele tem em mãos é um verbo (*comi*) não acompanhado de nenhum SN que possa ser analisado como seu sujeito.

Muitos modelos de análise lingüística parecem negar, implícita ou explicitamente, essa verdade tão simples. Um desses modelos é a análise tradicional, que postula um sujeito oculto *eu* na estrutura sintática de [10]. Mas um sujeito oculto não faz parte do nível formal (no sentido em que o tomamos), porque não é imediatamente perceptível ao receptor. O sujeito oculto teria que ser inferido pelo receptor a partir da desinência verbal – que, esta sim, é imediatamente acessível, por ter representação material, fonética ou gráfica.

Note-se que se postularmos um sujeito oculto em [10], então a desinência *-i* de *comi* não será a única pista de que o receptor dispõe para saber que o agente sou eu. Isso, a meu ver, simplesmente nega a evidência; tire-se a desinência e fica impossível, na frase [10] assim mutilada, saber qual é o agente. A situação seria outra se a frase fosse

[11] Eu comi uma pizza.

Aqui, sim, existe um sujeito – mas não oculto – e o receptor dispõe de **duas** pistas que lhe indicam o agente; ou seja, a indicação do agente é redundante em [11], mas não em [10].

²⁰ Essa posição representa a face descritiva da hipótese da “sintaxe simples” de Culicover e Jackendoff (C&J: 5).

²¹ Resta o problema de saber com que *comi* concorda em [10], se não há sujeito. Minha resposta é que não concorda com nada: o sufixo *-i* é um elemento introduzido livremente pelo falante, sendo portanto significativo.

Semanticamente, portanto, existe uma relação entre o sufixo de pessoa-número (-i em *comi*) e o eventual sujeito que o acompanhe: a desinência sistematicamente denota o mesmo papel temático do sujeito. Assim, em *eu comi* o papel temático “agente” é atribuído a *eu* duas vezes: uma vez pela desinência, outra pelo sujeito. Em *Lana comeu*, o sufixo atribui “agente” a uma pessoa singular, diferente do falante²² (a chamada “terceira pessoa do singular”), e o sujeito elabora essa informação, especificando que se trata de *Lana*. Como se vê, a redundância na terceira pessoa é parcial, mas na primeira é total. Nos casos de redundância total, o sujeito é dispensável, e com efeito é omitido muitas vezes.

No contexto da gramática tradicional, acho que a postulação de sujeitos ocultos provém da confusão feita tradicionalmente entre sujeito e agente – elementos que, embora tenham alguma conexão, não se identificam de maneira nenhuma. Essencialmente a mesma análise ocorre na maioria das versões da gramática gerativa. Aqui já não se trata, creio, de confusão entre sujeito e agente, mas de um conjunto de tendências, entre as quais a procura da simetria e da simplicidade da análise mesmo à custa de certa cegueira para com os dados imediatos da experiência fornecidos pelo uso da língua – o que C&J, em sua crítica da posição gerativista, chamam o princípio da Uniformidade da Interface Sintaxe-semântica. O sujeito abstrato (“vazio”, *pro*) postulado pelos gerativistas para frases como [10] é parte da “estrutura-S”, e ainda precisa ser submetido a regras que relacionem essa estrutura com a forma fonética, observável; mas nota-se pouco interesse em formular as regras (ou o que sejam) que realizam esse relacionamento. Os gerativistas tendem a partir de seqüências incluindo pró-formas, vestígios (*traces*) etc., como se fossem dados da experiência, o que certamente não são. O problema de relacionar tais estruturas-S com a forma fonética não é em absoluto trivial.

Um exemplo, entre outros, do tipo de metodologia empregado pelos gerativistas ao se confrontarem com esses problemas se encontra em Haegeman (1991). Tendo chegado à conclusão de que existem sujeitos explícitos marcados com os traços [- anafórico, + pronominal], ela pergunta se não haverá igualmente sujeitos ocultos (isto é, representados por “categorias vazias”) com os mesmos traços:

Não há nenhuma razão para não existir uma categoria vazia caracterizada pelos traços [- anafórico, + pronominal].

[Haegeman, 1991: 413]

E dessa observação Haegeman parte para tentar encontrar uma categoria vazia com os traços mencionados. Se ela encontrar tais categorias, então poderá dizer que os sujeitos explícitos e os implícitos têm traços idênticos, o que em última análise permite simplificar a análise. As categorias procuradas são rapidamente encontradas, na verdade praticamente sem discussão.

No entanto, a argumentação de Haegeman tem pontos perigosamente fracos. Mesmo deixando de lado a questionável postulação de categorias vazias, pode-se observar o seguinte: parece que não se cogita da possibilidade de que não existam categorias vazias sem os traços mencionados. Isso levaria a uma complicação da teoria, mas quem nos garante de partida que as coisas são simples? Haegeman diz que “não há razão” para não existir o elemento procurado; mas não apresenta evidência em favor

²² Não do ouvinte, por causa de *você comeu a pizza* etc.

dessa afirmação. Em princípio, antes de examinar os dados, seria preciso admitir que as duas possibilidades são igualmente possíveis, e que pode haver uma razão para a inexistência de categorias vazias com aqueles traços: se as procurarmos e não as encontrarmos, teremos uma razão suficiente.

A metodologia utilizada por esses autores (implícita, mas muito presente) procura demonstrar a validade da teoria, e não questioná-la com base em dados empíricos. Não surpreende que os dados sejam sempre favoráveis, já que os dados desfavoráveis são deixados na sombra. Não havendo controle de obtenção de dados, estes demonstram aquilo que se quer demonstrar.

1.8.2.2. Topicalização

Pares de frases como o seguinte são muitas vezes analisados como transformacionalmente²³ relacionados:

[15] O Sérgio rasgou meu casaco.

[16] Meu casaco, o Sérgio rasgou.

Em um modelo transformacional, *meu casaco* em [16] é transportado a partir de uma posição “normal” de objeto direto logo após o verbo, deixando um vestígio, e resultando em uma “estrutura-S” mais ou menos como

[17] [meu casaco_i o Sérgio rasgou t_i]

O vestígio (t) é coindexado com o antecedente, e essa estrutura forma a base para a interpretação da forma lógica.²⁴

Por mim, tudo isso pode ser correto e necessário (embora falte mostrar que é). Mas definitivamente não é suficiente: nem o vestígio nem os índices são representados de maneira sensorialmente perceptível, e portanto devem ser (se tanto) resultado de um trabalho de análise realizado pelo receptor.

É possível analisar [16] de maneira mais concreta, sem ter que recorrer a transportes ou vestígios. O receptor certamente conhece o verbo *rasgar* e sua valência: por exemplo, ele sabe que esse verbo pode ocorrer com um SN que exprime seu “paciente”. Em uma oração com *rasgar*, portanto, há lugar para dois SNs: um é o “agente”, e este é identificado basicamente por sua posição imediatamente antes do verbo. O outro é o que exprime o “paciente”, e este tem mais de uma possibilidade aceitável de posicionamento: pode ocorrer depois do verbo ou então no início da frase, antes do sujeito.

Assim, em [15] interpretamos *meu casaco* como “paciente”, não apenas em virtude de sua posição pós-verbal, mas porque a valência do verbo determina que o SN

²³ O termo “transformacional” caiu de moda, mas análises baseadas em transformações continuam sendo adotadas; a operação chamada “mover alfa” é uma transformação.

²⁴ Ver uma crítica desse tipo de análise em C&J, cap. 1 a 4.

que não é o “agente” deve ser o “paciente”. Ou seja, se não o fizermos, a estrutura não vai “fechar”, porque um sintagma (*meu casaco*) vai ficar sem papel temático.

O sintagma *meu casaco* em [16] vai também ser marcado como “paciente”, igualmente por duas razões. Primeiro, ele está em uma posição em que é possível a atribuição desse papel temático; e, depois, porque de outro modo teríamos a situação anômala de o verbo *rasgar* ocorrer sem “paciente”, e um dos SNs da oração, *meu casaco*, vai ficar sem função semântica possível, o que é intolerável.

É necessário levar em conta a posição do SN (o “paciente” depois do verbo ou em primeiro lugar na frase), porque se tivermos

[18] ?? O Sérgio meu casaco rasgou.

não será possível atribuir corretamente os papéis temáticos aos diferentes SNs. Aqui, somos obrigados a atribuir o papel de “agente” a *meu casaco*, porque está imediatamente antes do verbo, e “paciente” a *o Sérgio*, que é o SN inicial da frase – o resultado, claro, é semanticamente anômalo. Como se vê, a aplicação da valência do verbo *rasgar* não é suficiente para garantir a interpretação correta da frase.

Capítulo 2: Posicionamento teórico

2.1. Teoria e descrição

Encontramos com frequência em introduções a trabalhos de análise lingüística a afirmação de que a análise é feita “à luz” desta ou daquela teoria. Isso é uma coisa que não pretendo fazer aqui; e a razão principal é que não confio suficientemente nas teorias lingüísticas estabelecidas para acreditar que alguma delas possa lançar luz sobre dados ou sua análise. Antes, eu diria que o estágio de desenvolvimento em que se encontra a lingüística permite, nas situações mais favoráveis, que o exame dos dados lance luz sobre a adequação das teorias.

Isso não quer dizer que eu defenda, ou ache possível, uma posição “não teórica”. O que quero dizer é que o desenvolvimento da teoria e o exame dos dados devem se processar paralelamente; e que o lingüista corre o perigo permanente de teorizar sem base empírica suficiente ou então, conversamente, de acumular dados sem saber realmente como interpretá-los por falta de pontos de referência em alguma teoria. Tenho uma simpatia especial pelo modelo de análise proposto em C&J, mas não estou preparado para realizar pesquisa “à luz” dessa teoria. No entanto, minhas observações me permitem certo grau de confiança na proposta de Culicover e Jackendoff, de modo que me arrisco a tomá-la como uma guia para o trabalho de descrição. Além disso, me reservo o direito de só aceitar parte de sua proposta, e de acrescentar elementos derivados de minha experiência de pesquisa. A cada momento, as idéias de Culicover e Jackendoff (assim como as minhas) estão sob o escrutínio impiedoso da necessidade de adequação empírica.

Se o conjunto de crenças que norteia este trabalho não chega a compor uma teoria comparável a algumas atualmente disponíveis, é principalmente porque nossa compreensão da sintaxe e da semântica não chegou ao ponto em que valha a pena construir teorias tão estruturadas e específicas quanto, por exemplo, a teoria gerativa da regência e ligação. É prematuro formalizar teorias antes de atingir uma compreensão suficientemente segura dos fenômenos estudados. Precisamos de teorias, é certo; mas no momento precisamos acima de tudo de estudos empíricos amplos, tendo como objeto fatias relativamente grandes da gramática de línguas específicas, e como evidência primária dados, numerosos e não selecionados.

2.2. Por uma gramática descritiva

2.2.1. *Necessidade de descrições*

Precisamos, portanto, de descrições das línguas naturais, descrições que sejam relativamente livres de pressupostos teóricos controversos. Essas descrições funcionarão como fontes de dados, para fundamentar, entre outras coisas, o trabalho de elaboração das teorias. Neste capítulo dou alguns princípios que me têm sido úteis em meu trabalho descritivo, e que podem servir de guia para o estudo empírico do

português. Eles constituem hipóteses de trabalho a que cheguei no decorrer da minha pesquisa sobre a língua portuguesa, mas não formam uma teoria, e não devem ser entendidos como verdades permanentes. Minha simpatia para com o modelo de C&J é motivada pelo exame de problemas encontrados durante minhas tentativas de descrever áreas da estrutura do português (em particular, a estrutura interna do sintagma nominal e as valências verbais). Trabalhando em relativo isolamento durante os anos 90, meus colegas e eu elaboramos muitas soluções análogas às encontradas e publicadas desde o início dos anos 80 por Jackendoff (1983, 1990) e agora publicadas em forma mais acabada e completa em C&J.¹

Não incluo na discussão abaixo o papel do contexto textual e extratextual; tais como se formulam adiante, os princípios valem para um trabalho descritivo em linhas gramaticais, ou seja, fonológico, morfossintático e semântico. Estou convencido, no entanto, de que o contexto é um fator necessário na descrição gramatical, e creio que os princípios terão que ser eventualmente adaptados para incluir esse fator. Isso fica evidente na análise do SN de Perini *et al* (1996).

Em resumo, pois, um dos objetivos fundamentais da lingüística é descrever as línguas naturais. Isso quer dizer prever a aceitabilidade e inaceitabilidade de seqüências formais e de suas associações a representações semânticas. Esse certamente não é o **único** objetivo do trabalho lingüístico; mas se ele não for realizado tudo o mais fica comprometido.

2.2.2. Descrição e teoria(s) da linguagem

A forma mais elementar de uma descrição é uma lista das seqüências possíveis, com as representações semânticas que lhes são respectivamente associadas. É evidente que tal lista seria um recurso muito pouco prático, e provavelmente impossível de elaborar, por ser aberta em princípio. Ela incluiria frases individuais como por exemplo

[1] Helena comprou um CD.

[2] Meu irmão conserta máquinas de lavar.

etc.

Qualquer tentativa de compactar a lista já envolve alguma teorização, de modo que a descrição lingüística está longe de ser uma atividade livre de opções teóricas, e mesmo de especulação teórica. Por outro lado, existe um grande corpo de resultados teóricos que é de aceitação geral, e pode ser considerado “terreno conquistado”, no sentido de que não está sujeito a grandes controvérsias (ver 1.7). Embora seja sem dúvida um risco aceitar como dado esse corpo de resultados, é algo a que não se pode escapar. Voltando às frases [1] e [2] acima, é bastante seguro notá-las como representando ambas a mesma estrutura, ou seja,

[3] SN V SN

¹ Esse trabalho foi relatado em Perini *et al* (1996), para o SN; e em Perini (2005) e na segunda parte deste livro, para as valências verbais.

ou, incluindo parte de sua interpretação semântica,

[4] SN V SN
 Agente Paciente

Ao contrário da lista de frases individuais, uma lista composta de representações esquemáticas como [4] é possível de elaborar, e tem grande utilidade. Como se verá, é justamente uma lista desse tipo que se usa como base para a pesquisa de valências verbais (cap. 8-13).

O presente trabalho se baseia em um grande número de hipóteses, e se houve a preocupação de evitar as mais inseguras, com frequência foi preciso lançar mão de algumas que suscitarão objeções por parte de muitos colegas. Mas é preciso lembrar que aqui se coloca como objetivo fundamental do estudo a representação dos fatos, e não a justificação das teorias. Entenda-se: não estou acusando ninguém de desprezar em princípio a necessidade de fundamentar as teorias nos dados, o que seria equivalente a negar o caráter empírico da lingüística. Mas encontro com certa frequência na literatura uma tendência a selecionar dados, ou mesmo forçá-los, de maneira que se adaptem à teoria em discussão. É uma questão de ênfase: estamos todos lidando com teorias e com dados, mas a relação entre teoria e dados é uma questão delicada. Há uma relação ideal de simbiose entre a descrição e a teorização, de modo que não se pode trabalhar com uma sem lançar mão da outra. Tanto o desenvolvimento da teoria depende de resultados do trabalho de descrição quanto a elaboração de gramáticas descritivas deve basear-se nos resultados disponíveis do trabalho teórico. Retomando a citação de Chomsky (1965) vista na seção 1.5.2, podemos perguntar: onde estão essas gramáticas descritivamente adequadas? Quem está se ocupando em elaborá-las? Acho que existe aqui uma lacuna, que o presente estudo procura preencher, em pequena parte.

Este é portanto um trabalho de descrição. Baseia-se, como não poderia deixar de ser, em uma série de pressupostos teóricos, mas não faz grandes esforços no sentido de construir uma teoria otimizada, nem de encaixar os dados observados em uma tal teoria. O resultado é que a descrição pode ser redundante em muitos pontos, e muitos fenômenos recebem um tratamento heterogêneo. Não nego o interesse de procurar teorias elegantes e econômicas; mas insisto em que a elegância e a economia de nada valem se divorciadas da adequação empírica. No presente trabalho procuro antes de mais nada fornecer bases para a verificação empírica das teorias; e, por outro lado, procuro retratar alguns fatos gramaticais tais como se apresentam, digamos assim, à primeira vista.

2.3. Regularidade e anomalia

2.3.1. *Anomalias*

A língua não se compõe apenas de regularidades; inclui também, em proporção desconhecida, casos particulares, exceções, irregularidades etc. Ou, para usar uma nomenclatura antiga, a língua contém um componente **anomalístico**. Essas anomalias são tradicionalmente colocadas no léxico; assim, um item pode ser marcado como exceção a

uma regra (como o plural de *mal*, que é *males* e não o esperado **mais*), e isso é parte da informação contida no léxico a respeito desse item.

Qualquer levantamento empírico amplo da gramática de uma língua revela uma grande proporção de fenômenos idiossincráticos; Gross sempre insistiu muito nisso (ver Gross, 1975 e 1979). Dentre esses fenômenos pode-se mencionar: irregularidades morfológicas (do tipo *mal / males*); diferenças de transitividade difíceis de atribuir a traços semânticos (como o caso de *roubar*, que pode ocorrer na construção *Roberto roubou a União em mais de 10 milhões*, ao passo que *furtar* não ocorre nessa construção); diferenças de classe flexional (como o fato de que *andar* pertence à primeira conjugação e *vender* à segunda); diferenças de comportamento sintático (como o uso obrigatório da preposição *em* antes do complemento de certos verbos: *confio em você*) e assim por diante. Em princípio pode-se esperar que haja anomalias em qualquer parte da estrutura da língua. Qual é a proporção entre anomalias e regularidades é algo que só a investigação poderá responder.

No caso particular do sintagma nominal, as anomalias são relativamente numerosas e de natureza variada. Assim, temos casos em que um item se posiciona segundo regras próprias (*ruim* só pode vir posposto, nunca anteposto, ao contrário de seu sinônimo *mau*, que ocupa as duas posições); casos em que uma seqüência se fixa em decorrência de uma convenção aparentemente arbitrária (*doce ilusão*, e não ** ilusão doce*); casos em que o item é isento de concordância nominal (como *cada*, ou então certos nomes de cores como *laranja*, que não variam nem em gênero nem em número). Esses casos não podem, aparentemente, ser relacionados com traços semânticos dos itens em questão, constituindo verdadeiras anomalias, a serem aprendidas uma a uma.²

Uma anomalia sintática é exemplificada pela frase

[5] Há dez mil anos o homem caçava mamutes.

O problema aqui está no constituinte *há dez mil anos*. Esse elemento ocupa o lugar sintático e semântico de um advérbio de tempo (pode ser substituído por *na pré-história*). Mas, como inclui um verbo, deve ser analisado como uma oração subordinada. Toda oração subordinada em português é marcada como tal – seja através de uma conjunção subordinativa (*que, se, quando*), seja através de marca no próprio verbo (como quando o verbo está no infinitivo). Mas o constituinte *há dez mil anos* tem a estrutura de uma oração principal: verbo no indicativo, e nenhuma conjunção. No entanto, seu papel semântico é o de uma subordinada, e sua inserção sintática no período é **sui generis**, sendo simplesmente justaposto à oração *o homem caçava mamutes*. Tudo isso constitui uma anomalia, limitada a orações com os verbos *haver* e *fazer* (*está chovendo faz dois dias*). Para complicar um pouco mais as coisas, quando o verbo é *fazer* e a oração com *fazer* vem em primeiro lugar, é possível marcar a segunda oração como subordinada, introduzindo a conjunção *que*; mas com *haver* isso não é possível:

[6] ?? Faz oito anos ele trabalha aqui. / Faz oito anos que ele trabalha aqui.

[7] Há oito anos ele trabalhava aqui. / * Há oito anos que ele trabalhava aqui.

² Essas anomalias são estudadas em detalhe em Perini *et al* (1996: cap. 5).

Apenas mais um exemplo. Quando se utiliza o nome de um objeto ao lado da designação genérica deste, a designação genérica vem sempre em primeiro lugar. Por isso dizemos *o cantor Djavan: cantor* é o termo genérico, e *Djavan* é o nome do cantor específico que se quer mencionar. Mas com nomes de hotéis (e só com nomes de hotéis) a ordem pode ser invertida: *o Normandy Hotel*, ao lado de *o Hotel Normandy*. O receptor, ao processar *o Normandy Hotel*, precisa estar informado de que com nomes de hotéis ele pode ter que procurar o núcleo depois do termo restritivo.

Há grande número de anomalias também ligadas a fenômenos complexos; por exemplo, fenômenos que podem ser descritos em grande parte, mas não totalmente, em termos semânticos. A ordem dos termos internos do sintagma nominal é em grande parte decorrência da semântica dos itens envolvidos. Mas há itens que parecem ser marcados idiossincraticamente, sem que se conheça motivação semântica para isso – é o caso da posição inicial dos artigos *o* e *um*, por exemplo, ou o posicionamento especial da palavra *certo*, que ocupa uma função sintática exclusiva. Essa complexidade de muitos fenômenos tradicionalmente considerados como processos unos é reconhecida por Culicover e Jackendoff, por exemplo quando afirmam que

As restrições sobre a ordem linear vêm em duas variedades. A primeira variedade corresponde às regras de estrutura sintagmática tradicionais [...] A segunda variedade consiste de restrições semânticas sobre a ordem de palavras ou sintagmas [...]

[C&J: 170]

A maioria das teorias sintáticas e semânticas correntes superestima o papel das regularidades, a ponto de dar a impressão de que elas quase esgotam o que há para estudar na língua. As anomalias seriam uma espécie de resíduo, numericamente insignificante e pouco interessante. Essa posição é apriorística, e só subsiste por causa da extrema escassez de estudos voltados para as irregularidades (um trabalho pioneiro, mas que não teve muita seqüência, é Lakoff, 1970). No presente trabalho, tenho a preocupação de só encaixar um caso em uma análise geral (tornando-o assim uma regularidade) na presença de evidência suficiente. Os casos duvidosos são tratados à parte, como anomalias. O resultado é uma análise pouco “elegante”, cheia de exceções, que pode não ser a definitiva, mas que nos é ditada pela prudência e pela consciência de nossa ignorância.

2.3.2. *Léxico e gramática*

Seguindo uma tradição amplamente difundida, distingue-se na língua dois componentes, a **gramática** e o **léxico**. A gramática seria o repositório da informação generalizada, enquanto o léxico seria o repositório da informação particularizada, ligada a itens individuais ou a pequenos grupos de itens.

Essa dicotomia deriva da idéia de que os fatos da língua se diferenciam de maneira mais ou menos nítida em regulares e idiossincráticos. Mais exatamente, suspeito, deriva de uma posição que subestima a importância dos traços idiossincráticos no funcionamento da língua. Aqui se verifica uma situação muito freqüente: os extremos são

claros, mas a delimitação entre as duas situações é muito pouco clara. Só um levantamento amplo dos fatos da língua nos poderá fornecer um panorama realista da importância relativa dos fatos gerais e dos fatos particulares. Por minha parte, quanto mais investigo mais me convenço de que há de tudo: regras inteiramente gerais, que não comportam uma única exceção; regras com poucas exceções; regras com muitas exceções; itens totalmente idiossincráticos, diferentes (sob algum aspecto) de todos os outros da língua; grupos (grandes, médios, pequenos) de itens ligados por traços comuns... Em resumo, não parece ser possível distinguir o léxico da gramática nos termos acima indicados.

Em vez disso, temos uma rede extremamente complexa de semelhanças e diferenças entre os itens (que, por falta de melhor designação, continuarei a chamar **léxicos**). As semelhanças podem ser chamadas **regras** quando valem para grande número de itens, em especial quando as mesmas áreas são afetadas por outros processos. Mas não são menos propriedades de itens léxicos: conhecer o item *comer* implica não apenas em saber seu significado específico ou o fato de que se conjuga pela segunda conjugação, mas também saber que cabe em determinados ambientes, por exemplo com objeto direto (*comi a pizza*), ou sem objeto nenhum (*ele já comeu hoje*), mas não com *a + SN* (**comi ao pernil*). É igualmente saber que pode ocorrer em construções passivas (*Pierre foi comido pelos canibais*). Dessa forma, o conhecimento léxico se integra intimamente com o conhecimento gramatical, e a distinção entre eles muitas vezes não é nada clara.

Essa é uma posição que vem se firmando entre os lingüistas de muitas vertentes; o pioneiro, me parece, foi Gross (1975), mas hoje encontramos basicamente a mesma posição em Langacker (1987, 1991), Kac (1992) e também em C&J, que colocam assim sua posição:

Há um contínuo de fenômenos gramaticais, dos idiossincráticos (inclusive palavras) até as regras gerais da gramática.

[C&J: 15]

No mínimo, temos que reconhecer que a gramática e o léxico são intimamente relacionados. Assim, a valência de um verbo dá informação sobre os ambientes em que esse verbo pode ocorrer:

engordar / SN(Agente) V SN(Paciente) /// SN(Paciente) V

Por outro lado, pode-se também dizer (sem referência direta aos itens léxicos) que o português tem as estruturas

SN V SN
Agente Paciente (*Gracinha derramou o leite*)

e

SN V
Paciente (*o leite derramou*)

mas não, digamos,

* SN	V	SN
Meta		Agente ³

– e essa é uma afirmação geral e, portanto, gramatical. Agora, é verdade que uma estrutura “existe” quando há pelo menos um item léxico que se encaixa nela, caso contrário “não existe”.

Mesmo se distinguirmos léxico de gramática, vai ser impossível estabelecer o limite de forma precisa. Existe uma gradação entre fenômenos muito gerais (gramática) e fenômenos idiossincráticos, limitados a um ou uns poucos itens léxicos (léxico). Isso pode ter importância teórica, mas em termos descritivos não vejo necessidade de jogar uma das perspectivas fora. Acho que se resolvermos não falar de gramática, vamos deixar muitas generalizações de fora. Talvez seja possível simplesmente definir uma “perspectiva gramatical”, que seleciona os traços que mostram ocorrência muito ampla dentro do léxico da língua. Ou seja, em vez de definirmos a “gramática” e o “léxico” como componentes da descrição à maneira tradicional, definiremos dois pontos de vista coexistentes e complementares. Dentro dessa perspectiva, alguns fenômenos são mais convenientemente descritos em termos gramaticais, outros em termos léxicos.

Por isso, me permitirei usar neste texto termos como **regra, léxico, regularidade e irregularidade** (ou **anomalia**) sempre que não resultar daí ambigüidade grave. São noções aproximativas, talvez inevitáveis no atual estado do conhecimento, aceitáveis se usados com cautela.

2.4. Postulados da análise

Como já observei acima, nenhuma análise pode prescindir de postulados. O presente trabalho se baseia em um conjunto de pressupostos teóricos, que nortearam o trabalho de descrição. Desses, quatro são especialmente importantes, a saber:

- (a) Descrição superficial, e conseqüente interpretação semântica das estruturas superficiais; ou seja, não aceitação de estruturas múltiplas relacionadas transformacionalmente;
- (b) consideração estritamente separada de traços de forma e traços de significado;
- (c) sintaxe delimitada residualmente, segundo a hipótese da Sintaxe Simples;
- (d) descrição não dividida em componentes compactos, homogêneos e ordenados (sintaxe, semântica, pragmática). Em particular, aceitação da possibilidade de que regras que estipulam aspectos formais podem lançar mão de informação semântica.

³ O papel temático Meta representa o final de um movimento, como em *cheguei em Campinas*.

Esses quatro pontos se harmonizam com a proposta de C&J, de modo que, na medida em que este trabalho tem uma orientação teórica, é a do modelo ali proposto. Os quatro pontos são detalhados a seguir.

2.4.1. Descrição superficial

O primeiro dos pressupostos que orientam este estudo é que a descrição não deve considerar estruturas múltiplas relacionadas transformacionalmente; ou seja, a relação simbólica se define como a interface entre a estrutura superficial e a semântica, sem níveis intermediários. Culicover e Jackendoff exprimem isso afirmando que

Não há “níveis ocultos” feitos de unidades sintáticas.

[C&J: 15]

O termo **superficial** tem sido utilizado nos últimos tempos em mais de um sentido, e por isso será necessário explicitar o que se entende por descrição da estrutura superficial. Em primeiro lugar, não se trata da **estrutura-S** proposta na teoria da regência e ligação (*government and binding*; Chomsky, 1981). A estrutura-S, nesse modelo, representa um nível bastante abstrato de análise, incluindo, por exemplo, índices de correferência e categorias vazias. Estrutura superficial, no presente trabalho, designa um nível onde se representam, basicamente, os itens léxicos e sua seqüência, tais como se realizam fonologicamente. É um nível extremamente concreto de análise, se comparado com as representações “superficiais” admitidas em certas teorias atuais. Langacker e muitos cognitivistas levam essa tendência a uma posição bastante extrema:

“[A gramática cognitiva] afirma que a estrutura gramatical é quase totalmente expressa [*overt*]: as coisas são realmente como parecem ser, desde que as saibamos interpretar corretamente.”

[Langacker, 1987: 27]

Por minha parte, prefiro uma posição mais prudente: deixo aberta a possibilidade de que a análise das complexidades da língua venha a exigir a postulação de categorias abstratas. Mas, se tal acontecer, o ônus da prova recairá sobre quem propuser as abstrações. Na ausência de evidência clara e abundante do contrário, as estruturas da língua devem ser analisadas tais como se apresentam aos nossos sentidos e à nossa compreensão.

Assim, a **estrutura superficial** tal como aqui se entende não inclui constituintes ou categorias vazias, concebidas como sintaticamente presentes, mas sem representação fonológica; e não participa de relações sintáticas entre estruturas formalmente diversas (isto é, relações de natureza transformacional, como as transformações do modelo gerativo padrão ou o processo de mover-alfa das variedades mais recentes).

Os fatos usualmente capturados através de categorias vazias são analisados como relações simbólicas, que conectam entidades formais e entidades semânticas. Assim, não figuram na estrutura superficial os elementos entre colchetes das frases seguintes:

[8] Silvinha pretende [Silvinha] ir ao concerto.

[9] Zé toca oboé e Leonor [toca] tambor.

A estrutura formal dessas sentenças, base para a análise sintática e para a interpretação semântica, atém-se às formas observáveis:

[10] Silvinha pretende ir ao concerto.

[11] Zé toca oboé e Leonor tambor.

Essas estruturas sofrem a ação de regras de interpretação que atribuem ao agente “Silvinha” a ação expressa por *ir*, e a ação “toca” ao agente “Leonor”. Por exemplo, no segundo caso essas regras são provavelmente sensíveis à presença do coordenador *e*, que normalmente une seqüências semanticamente paralelas. A partir daí, as regras colocam uma cópia da representação semântica de *toca* na segunda metade da frase, com as mesmas conexões encontradas na primeira metade. O resultado é que a representação semântica “toca” terá um agente, representado por “Leonor” (paralelamente ao agente “Zé” da primeira metade), e analogamente para o paciente “tambor”. As condições de aplicação dessas regras já foram em grande parte formuladas, mas concebidas como regras transformacionais que eliminam uma segunda ocorrência de *toca*. Trata-se agora de reformulá-las em termos da interpretação semântica de seqüências superficiais.⁴

Passando à questão das relações transformacionais, nossos pressupostos não admitem a relevância sintática de relações entre estruturas formalmente diferentes. Alguns modelos de análise definem uma relação sintática entre as frases

[12] A cozinheira queimou o feijão.

[13] O feijão, a cozinheira queimou.

Fala-se, nesses casos, de “transporte” do constituinte *o feijão* de uma posição pós-verbal para o início da frase. Naturalmente, isso só faz sentido se se admite uma relação formal entre [12] e [13]. Já em outros casos essa relação não é estabelecida, como em

[14] A cozinheira desacatou o patrão.

[15] O patrão desacatou a cozinheira.

Nunca se diz que *o patrão* foi transportado para o início da frase em [15]. Em vez disso, [14] e [15] são encaradas como sentenças geradas independentemente, não tendo relação transformacional entre elas.

Já o modelo descritivo aqui utilizado não diferencia os dois casos dentro da perspectiva formal: tanto em [12]-[13] quanto em [14]-[15] temos frases independentes, que ilustram estruturas distintas. A evidente relação de significado que existe entre [12] e [13] fica a cargo da interpretação: as regras relevantes atribuirão a *o feijão* tanto em [12]

⁴ Ver a respeito C&J, cap. 7.

quanto em [13] o papel temático de “paciente” da ação expressa por *queimou* e praticada pela cozinheira. O fato de não se poder acrescentar um objeto a [13]

[16] * O feijão, a cozinheira queimou o almoço.

se explica pela saturação da grade temática (a **valência**) de *queimar*, que comporta apenas um agente e um paciente. Um dos argumentos expressos em [16] fica sem lugar. Trata-se de algo semelhante, em seus efeitos, ao princípio expresso por Chomsky (1982) sob o rótulo de **critério-teta**; entretanto, note-se que aqui a grade temática e suas decorrências se analisam explicitamente como fenômenos semânticos.

A conveniência da descrição superficial é algo a ser demonstrado através das análises de muitos exemplos, e aqui não se pode dar mais que uma amostra. Como ilustração adicional, vou considerar brevemente um caso clássico, o da relação entre passivas e ativas. Sabemos que as passivas e as ativas são freqüentemente analisadas como sendo relacionadas transformacionalmente, de modo que uma é derivada da outra, ou, mais freqüentemente, ambas são derivadas de uma estrutura subjacente comum.

Mas seria necessário considerar mais cuidadosamente o que é que nos leva a relacionar uma estrutura passiva com uma estrutura ativa em particular, e não com outras que envolvem os mesmos itens léxicos. Ou seja, por que é que se analisa [17] como passiva de [18], e não de [19]?

[17] O gato foi devorado pelos ratos.

[18] Os ratos devoraram o gato.

[19] O gato devorou os ratos.

As semelhanças formais entre [17] e [19] são tão estreitas quanto as que existem entre [17] e [18]. Por que, então, deveríamos relacionar de preferência [17] e [18]?

Acredito que a razão é a semelhança **semântica** entre [17] e [18], em especial o fato de que os diferentes papéis temáticos são preenchidos pelos mesmos itens léxicos em ambas (*gato* é “paciente”, *rato* é “agente”), ao passo que em [19] encontramos os mesmos itens léxicos com os papéis trocados.

Ora, essas semelhanças e diferenças semânticas são fatos, e portanto devem inevitavelmente figurar em uma descrição completa da língua. Se for possível descrever as relações observadas entre passivas e ativas como decorrência desses fatos, não se justificará a postulação de uma relação formal (isto é, transformacional) entre esses dois tipos de estrutura. Quem propuser a análise das passivas e das ativas como sintaticamente relacionadas deverá, primeiro, mostrar que a relação percebida entre [17] e [18], mas não entre [17] e [19], **não é derivável** de seu parentesco semântico; e isso não se fez, que eu saiba.

O que justificaria a necessidade de estruturas subjacentes seria a descoberta de uma relação sintática entre passivas e ativas; se essa relação sintática não for necessária, segue-se que tampouco é necessário postular estruturas subjacentes. A opção por uma descrição superficial nas linhas acima esboçadas nos leva, naturalmente, a estabelecer

que a relação entre cada estrutura superficial e sua representação semântica se faz diretamente, sem passar por estruturas formais intermediárias.

Voltando aos exemplos [12] e [13],

[12] A cozinheira queimou o feijão.

[13] O feijão, a cozinheira queimou.

temos que interpretar as duas frases independentemente, sem postular que *o feijão* desempenha a mesma função sintática (“objeto direto”) em ambas. Em vez, reconhecemos que *o feijão* tem uma função em [12], representada apenas por sua posição na seqüência linear, e outra função (igualmente linear) em [13]. Por conseguinte, relacionamos a interpretação “paciente de *queimar*” com (pelo menos) duas funções sintáticas, formalmente distintas – o que se harmoniza com a observação ao nível mais concreto, já que é inegável que *feijão* se coloca formalmente de uma maneira em [12] e de outra em [13]: depois do verbo em [12], antes do sujeito em [13].

2.4.2. *Fatos de forma e fatos de significado*

No item precedente me referi às semelhanças entre certas frases como “fatos”. Isso merece algum comentário, e, como veremos, nos levará a algumas considerações relevantes.

Uma observação da maior importância metodológica é que os falantes podem ter acesso intuitivo à diferença entre **traços de significado** e **traços de forma**. Isso não significa que um falante, ou mesmo um lingüista treinado, possa dizer com segurança, em todos os casos, se determinado fenômeno deve ser colocado neste ou naquele componente da gramática; essa é, novamente, uma decisão que depende de teorização. Mas significa que, em geral, podemos distinguir, em um enunciado, traços relativos à organização formal de traços relativos ao significado do enunciado.

Vejamos um exemplo: considerando os sintagmas

[20] Escola pública federal

[21] ?? Escola federal pública

podemos observar que os dois termos *pública* e *federal* têm uma relação especial de extensão: as escolas federais são um subconjunto próprio das públicas. Essa é uma observação inevitável, para quem conhece os princípios da organização escolar brasileira. Não depende de teorização prévia, e faz parte da interpretação normal do sintagma, estando portanto necessariamente à disposição do falante. Esse é um **traço de significado** relativo aos sintagmas [20] e [21].

Acabamos de considerar [20] e [21] do ponto de vista do significado, e depreendemos um de seus traços de significado, a saber, a relação de inclusão existente entre os dois termos *pública* e *federal*. Vamos agora considerar esses sintagmas do ponto de vista da forma. Uma observação imediata é que [20] difere de [21] quanto à ordem de

seus elementos: as palavras *pública* e *federal* trocam de lugar. Essa é uma observação formal que se pode fazer acerca desses sintagmas, e não depende de teorização. Temos aqui um **traço de forma**.

Pode-se ir um pouco adiante: a partir do traço de forma e do traço de significado presentes em [20] e [21], podemos ainda observar que o termo de significado mais abrangente (*pública*) ocorre antes do de significado mais restrito (*federal*), sob pena de inaceitabilidade. Aqui relacionamos um traço de forma com um traço de significado, e obtivemos um **traço de simbolização**.⁵ Novamente, temos um fato relativo a [20] e [21]: um fato particular, não generalizado, porque se limita à observação dos exemplos de [20] e [21], mas inegavelmente um fato.

Qualquer análise que negue, ou negligencie, os traços acima é inadequada ou incompleta. A representação desses fatos de maneira mais generalizada poderá variar segundo a teoria adotada; mas os traços acima expressos podem ser considerados como **fatos**, e deverão necessariamente decorrer de uma análise adequada da língua.

Vou agora exemplificar uma afirmação gramatical que depende de teorização. Parece que existe um princípio segundo a qual determinados elementos do sintagma nominal se ordenam segundo uma condição de restritividade crescente: quando a extensão de um deles é um subconjunto da de outro, o mais restritivo vem sempre depois do menos restritivo.⁶ Isso explica por que [20] é aceitável, mas [21] não é.

Agora já não estamos lidando com uma observação direta dos fatos. Em vez disso, formulamos uma hipótese (a condição de restritividade crescente) para descrever em termos mais gerais o fenômeno exemplificado em [20]-[21]. Seria possível explicar os mesmos fenômenos de outra maneira, e qual delas é a melhor é algo que precisa ser discutido. Para repetir: os fatos devem ser observados, mas não são, eles próprios, passíveis de discussão – pode-se discutir a exatidão da observação, mas nunca os próprios fatos. Por outro lado, as análises podem, e devem, ser discutidas, pois para cada conjunto de fatos muitas análises são possíveis, e umas são melhores do que outras.

Em resumo, ao considerarmos uma forma lingüística dispomos de um grupo de observações imediatas sobre as quais devemos basear a análise. Essas observações se dividem em duas categorias: **traços de significado** e **traços de forma**, com uma terceira categoria, derivada da relação entre as duas primeiras, e muitas vezes tão evidente quanto elas, os **traços de simbolização**. Acredito que muitas vezes é possível distinguir intuitivamente os traços de significado dos traços de forma. Os traços de significado são, evidentemente, os mais difíceis de observar, e principalmente de exprimir, em parte por causa de nosso deficiente conhecimento da semântica; isso não faz com que sejam menos concretos, ou menos essenciais para a análise.⁷

2.4.3. *Sintaxe simples* (“*simpler syntax*”)

⁵ Adaptação da nomenclatura de Langacker. “Simbolização” pode ser entendido como “referente à interface forma / significado”.

⁶ Esta e outras condições são estudadas em Perini *et al* (1996).

⁷ C&J formulam bem concisamente a necessidade de distinguir, em um primeiro momento, fatos de forma e fatos de significado: “Para elaborar a interface entre a sintaxe e a estrutura conceptual, precisamos saber que tipos de estrutura a interface está conectando.” [C&J: 153].

O lingüista descritivo tem necessidade de pontos de referência que guiem seu trabalho e que funcionem como indicações sobre a maneira de formular as regras, categorias e princípios com que deverá lidar ao descrever os fatos da língua. Nesta seção farei algumas sugestões nesse sentido, sugestões que devem ser tomadas como instrumentos metodológicos, úteis para a sistematização dos fatos observados, e não como compromissos teóricos. Aqui o que me interessa primariamente é a elaboração de uma linguagem que permita a apresentação dos dados de maneira sistemática e compacta.

O princípio básico é o seguinte: já que dos fatos não há maneira de escapar, e que nossa escolha está limitada às hipóteses, **uma análise é tanto mais simples quanto mais uso fizer da descrição dos fatos.**

Como sabemos, os fatos são de natureza formal, semântica ou simbólica – esses são os três aspectos diretamente observáveis da língua.⁸ Não há fatos sintáticos, nem morfológicos, nem fonológicos – o que não significa que não deva haver regras sintáticas, morfológicas ou fonológicas. Só significa que esses tipos de regras exprimem hipóteses, em vez de representarem fatos de maneira direta. E significa que a existência de cada uma precisa ser justificada.

Vou explicitar um pouco melhor esse ponto, que é importante. Uma afirmação formal, semântica ou simbólica é acessível à observação direta – observação nem sempre fácil, como sabemos, mas sempre possível em princípio. Exemplificando: a afirmação de que “a palavra *barriga* começa com uma bilabial” pode ser verificada empiricamente de forma muito direta; pode-se apurar se essa afirmação é ou não verdadeira sem referência a nenhuma teoria lingüística. O mesmo vale para outras afirmações formais, como “a palavra *um* pode ocorrer antes da palavra *cavalo*, formando um constituinte, mas não depois dela”; ou “a seqüência *o ataque brasileiro* é aceitável, mas **o brasileiro ataque* não é”. Passando ao campo simbólico, temos afirmações como “*barriga* designa uma parte do corpo”; ou “na seqüência *o cavalo estava correndo* temos a expressão de uma ação” e assim por diante. Qualquer análise que negue, ou deixe de lado, tais observações de fato é inadequada ou, pelo menos, incompleta.

Agora consideremos as hipóteses. Diante do fato de que **o brasileiro ataque* não é aceitável, podemos procurar uma generalização que explique essa inaceitabilidade. Uma possibilidade seria:

Regra 1: A palavra *brasileiro* é marcada como não podendo ocorrer antes do núcleo do sintagma nominal (vamos admitir, para simplificar, que o núcleo do SN já foi definido).

Essa seria uma regra da língua, baseada em uma marca idiossincrática da palavra *brasileiro* (marca formal, pois não faz referência ao significado da palavra). A regra funciona a contento, pois impede a formação da seqüência que se observa ser inaceitável. Essa regra pode ser generalizada, observando-se que há muitas palavras que não podem ocorrer antes do núcleo – por exemplo, *francês*, *russo* ou *pernambucano*. Essas palavras podem ser reunidas em uma classe, que se define por não poder ocorrer antes do núcleo, só depois.

⁸ Deixo de mencionar outros fatores igualmente importantes, ligados ao contexto lingüístico ou situacional, porque têm a ver com o uso da língua, mas não com sua estrutura enquanto código.

Agora consideremos outra possibilidade de analisar os mesmos fatos. Observamos que as palavras que funcionam como *brasileiro* têm um traço semântico em comum, ou seja, todas exprimem proveniência. E observamos que na língua nunca pode ocorrer uma palavra de proveniência antes do núcleo – ou ocorrem elas próprias como núcleos (*um brasileiro*), ou então depois do núcleo.

Podemos utilizar essa observação para formular uma regra, a saber:

Regra 2: A interpretação “proveniência” só pode ser atribuída a palavras em função de núcleo ou a palavras colocadas depois do núcleo.

Diante dos dados considerados, essa regra funciona tão bem como a regra 1, portanto as duas se equivalem em poder descritivo. No entanto, vou argumentar que a regra 2 é preferível à regra 1. Primeiro, note-se que a regra 1 depende da colocação de certas palavras em uma classe especial, definida em termos da própria regra: a classe das palavras que não podem ocorrer antes do núcleo. Essas palavras, segundo esta hipótese, só têm isso em comum; o fato de que são semanticamente semelhantes não é relevante para o funcionamento da regra 1.

Mas a semelhança semântica de palavras como *brasileiro*, *francês*, *russo* e *pernambucano* é um fato inevitável da língua. Isto é, nenhuma descrição do português será completa se não incluir informação sobre o fato de que essas palavras todas exprimem “proveniência”. Se adotarmos a regra 2, já não será necessário inventar uma classe especial para encerrar essas palavras. Elas já estão automaticamente reunidas na classe das palavras que exprimem “proveniência” – uma classe necessária por razões independentes, que têm a ver com a descrição da semântica da língua. Desse modo, a regra 2 nos oferece uma vantagem importante: evita a necessidade de postular uma classe *ad hoc*, cuja única utilidade é possibilitar o funcionamento da própria regra. Isso se torna possível porque a regra 2 faz uso mais intensivo da expressão de fatos, minimizando o recurso a mecanismos hipotéticos.⁹

Generalizando a partir desse exemplo (e de muitos outros), sustento que só se deve aceitar a descrição sintática (morfológica, fonológica) de um fato quando não for possível derivar esse fato de outros fatos, de maneira mais ou menos direta. Conseqüentemente, a descrição de base semântica é sempre preferível à de base sintática, desde, é claro, que ambas dêem conta igualmente dos dados. Outra maneira, talvez mais adequada, de dizer isso, é a seguinte: dentre duas análises que dão conta igualmente dos fatos, a menos abstrata deve ser preferida.

C&J chegaram exatamente à mesma conclusão, e a formularam da seguinte maneira:

Hipótese da Sintaxe Simples (*Simpler Syntax Hypothesis (SSH)*)¹⁰

⁹ Há outra razão para adotar a regra 2, preferivelmente à regra 1: alguns itens que podem denotar proveniência ocorrem antepostos, mas somente quando têm outro significado. É o caso de *britânico*, que pode ocorrer anteposto se tiver acepção qualificativa, sem denotar proveniência: *a britânica pontualidade de Joaquim*. Isso mostra que a possibilidade de anteposição se vincula não ao item léxico como um todo, mas ao item em determinada acepção.

¹⁰ A tradução deveria ser “sintaxe mais simples”, mas como isso não soa bem em português, estou usando “sintaxe simples”.

A teoria sintática mais explicativa é a que atribui o mínimo de estrutura necessário para mediar entre a fonologia e o significado.

[C&J: 5]¹¹

Isso nos fornece um meio de delimitar o campo de ação do componente formal frente ao componente semântico: um fenômeno recebe descrição formal quando não se conhece maneira de derivá-lo diretamente de algum fato semântico (ou fonético, se for o caso). Consequentemente, os limites entre a parte formal e a parte semântica da descrição de uma língua são em princípio provisórios. Um fenômeno é descrito formalmente porque **ainda** não se conseguiu relacioná-lo com algum fato semântico que o preveja. Note-se que isso não implica na crença de que **toda** a descrição da língua pode ser derivada de fatores concretos; a sintaxe, a morfologia e a fonologia podem ser, e certamente são, necessárias.¹²

Voltemos aos exemplos

[20] Escola pública federal

[21] * Escola federal pública

A inaceitabilidade de [21] se deve, segundo a hipótese aqui aceita, ao fato de que a ordenação de *federal e pública* desobedece a uma condição de restritividade crescente: o modificador mais abrangente deve ocorrer antes do mais restritivo. Isso significa que em [20]-[21] um traço de forma observado (a ordenação das duas palavras) é previsível a partir de um traço de significado (a relação de extensão entre *pública e federal*); isto é, conhecendo a relação de restritividade dos dois elementos podemos prever sua ordenação.

Confrontemos essa situação com outra. Sejam as frases

[22] Carolina gosta de Jarbas.

[23] Carolina ama Jarbas.

Um dos traços de forma que distinguem as duas frases é a presença, em [22], da palavra *de* antes de *Jarbas*; sabemos que esse *de* é obrigatório ali, e não pode ser omitido. Já em [23] não se pode colocar *de*:

[24] * Carolina gosta Jarbas.

[25] * Carolina ama de Jarbas.

No entanto, não foi possível até hoje determinar nenhum traço de significado que corresponda à exigência ou não da presença da preposição *de* com certos verbos. No caso

¹¹ A diferença é que C&J oferecem a hipótese como um componente de sua teoria da linguagem; no contexto deste livro, entretanto, ela é apenas uma guia para o trabalho descritivo. Uma coisa não exclui a outra, evidentemente; estou apenas me resguardando.

¹² C&J também consideram brevemente esse problema, e concluem que a sintaxe é necessária (C&J: 22).

de [22]-[23], inclusive, pode-se dizer que as duas frases têm significados muito próximos (são quase sinônimas). Temos aqui, até onde se pode ver, um traço de forma ao qual não corresponde nenhum traço de significado. Em [22]-[23] o traço de forma observado **não** é previsível a partir de nenhum traço de significado – situação oposta à que se verificou com os exemplos de [20]-[21].

Para dar outro exemplo, ainda no campo da distribuição das preposições: nenhum traço semântico conhecido ajuda a prever as estranhas restrições que regem o uso das preposições com as divisões do dia. Temos o seguinte:

<i>de manhã</i>	<i>pela manhã</i>	* <i>à manhã</i>
<i>de dia</i>	* <i>pelo dia</i>	* <i>ao dia</i>
<i>de tarde</i>	? <i>pela tarde</i>	<i>à tarde</i>
<i>de noite</i>	* <i>pela noite</i>	<i>à noite</i>

Aqui, novamente, o único recurso parece ser o de dar conta dessas diferenças através de uma análise formal, desprovida de motivação semântica.

Os exemplos vistos acima mostram a situação em que se encontra o lingüista descritivo diante de seus dados: parte deles decorre, de maneira mais ou menos evidente, de fatos semânticos; outra parte não parece ter correlato semântico nenhum. A idéia é que estes últimos se colocam na sintaxe – ou, para ser mais preciso, são descritos através de regras sintáticas, pois a “sintaxe” não precisa existir como um componente especial. Já os primeiros podem ser analisados como decorrência de regras de caráter simbólico, independentemente motivadas. No caso de [20]-[21], podemos propor uma regra que atribui o significado “proveniência” a um elemento pós-nuclear do SN; na posição pré-nuclear, entretanto, essa acepção não estaria disponível.

O critério de delimitação entre sintaxe e semântica acima proposto tem como conseqüência que a qualquer momento um fenômeno da língua inicialmente analisado como sintático poderá ser posto em correlação com fatos de significado que permitam sua previsão. Nesse momento, esse fenômeno deverá passar a ser analisado como resultado de regras semânticas: a análise formal de qualquer fenômeno será sempre provisória em princípio. Isso equivale a dizer que a delimitação entre sintaxe e semântica é uma questão empírica, não um postulado teórico.

Esse caráter provisório não prejudica a expressividade das descrições. O que se procura com o presente modelo é uma maneira de sistematizar fatos (de forma e de significado); a maneira pela qual isso se realiza é secundária, dados os objetivos primordialmente descritivos deste trabalho.

Acho importante observar que a Hipótese da Sintaxe Simples, ou algo muito semelhante, circulou mais ou menos informalmente entre os lingüistas durante muitos anos. Foi expressa com clareza pelo próprio Jackendoff (1990), e foi proposta independentemente por Langacker (1987: 26), Perini *et al.* (1996) e provavelmente outros. A formulação mais antiga que conheço, entretanto, é anterior às de Jackendoff em cerca de 15 anos, e foi formulada por Maurice Gross:

Partimos de um modelo da língua tão limitado quanto possível e pouco formalizado, e só introduzimos as abstrações que se mostrem absolutamente necessárias. Na verdade, em muitos casos, preferimos deixar um problema no

estado de conjectura antes que dar a ele uma solução que requer mecanismo abstrato novo e não fortemente motivado.

[Gross, 1975: 46]

Gross apenas não mencionou a distinção entre sintaxe e semântica, que aliás tem pouco lugar em seu modelo. Mas, tanto no espírito quanto na aplicação, seu princípio corresponde à Hipótese da Sintaxe Simples de Culicover e Jackendoff.¹³

2.5. Tipologia das regras

Não considero plausível a teoria de que as regras de uma gramática se distribuem em componentes distintos, homogêneos e ordenados como os da gramática gerativa padrão, segundo a qual as regras sintáticas se aplicariam todas em bloco, produzindo um “output” que seria por sua vez “input” para a aplicação das regras fonológicas, por um lado, e semântica por outro. Essa posição é comum hoje a lingüistas de muitas tendências; em Perini *et al.* (1996) pode-se encontrar alguns argumentos em favor dela, baseados na análise do português.

Notemos que essa questão não se identifica com a da distinção entre traços de forma e traços de significado, abordada na seção 2.4.2. Essa distinção é em grande medida acessível à intuição direta; e é adotada como resultado de uma posição metodológica previamente assumida, motivada por uma concepção particular dos objetivos da investigação lingüística. Trata-se da idéia de que a lingüística estuda primariamente a relação entre forma (as “imagens acústicas” de Saussure, 1916) e significado (os “conceitos”). Já a separação entre os componentes morfossintático e semântico (ou entre sintaxe e semântica) como partes da gramática é uma questão teórica a ser debatida – não é algo que possa ser resolvido através da simples observação de fatos. Depende da elaboração e avaliação de teorias sobre esses fatos.

Assim, neste trabalho não é aceita a separação dos componentes da gramática em sintático, semântico etc. Mas isso não quer dizer que as afirmações gramaticais sejam todas de um só tipo. Existem afirmações puramente formais, ou seja, que prevêm, a partir de informação independente de significado, fatos igualmente independentes de significado. Por exemplo, seja a exigência da preposição *de* antes do complemento de certos verbos: *gostar de*, *lembrar(-se) de* etc. Primeiro, os verbos que exigem a presença dessa preposição não formam um grupo semanticamente coerente; depois, a própria preposição não contribui em nada para o significado da frase. Essas afirmações gramaticais podem, a meu ver, ser chamadas **sintáticas** com propriedade. Elas formam um conjunto que pode ser chamado de **sintaxe**, sem violentar o termo; no entanto, a sintaxe, assim definida não é um “constituente” à maneira da gramática gerativa clássica – mesmo porque a concepção clássica depende de ordenação extrínseca das regras ou princípios, recurso que não é adotado aqui.

Existem, por outro lado, afirmações que determinam traços formais das estruturas, mas o fazem como subproduto do processo de interpretação, ou seja, de relacionamento de formas com significados. Um exemplo é a regra, vista acima, que estipula que no SN nunca ocorre um elemento pré-nuclear com o significado de “proveniência”, como

¹³ Ou, melhor, o oposto: a hipótese de Culicover e Jackendoff corresponde ao princípio de Gross.

francês, croata, brasileiro ou *britânico*. Pode-se descrever isso através de uma regra de interpretação, que restringe a acepção “proveniência” à posição pós-nuclear. Ora, essa regra de interpretação traz como consequência que um item cuja única acepção possível seja a de “proveniência” não pode, *ipso facto*, ocorrer antes do núcleo do SN. Isso explica a impossibilidade de antepor o item *croata*:

[26] * As croatas exportações

embora [27] seja aceitável:

[27] As exportações croatas

Explica também por que alguns itens que normalmente exprimem “proveniência” podem ocorrer antes do núcleo, como em

[28] A britânica pontualidade de Carlinhos

mas nesses casos a acepção não é mais “proveniência”, mas “qualidade”. Ou seja, em [28] a pontualidade não é algo que provém da Grã Bretanha, mas algo a que atribuímos a **qualidade** de “britânica”.¹⁴

¹⁴ O superlativo, que tem significado qualificativo e não proventivo, pode ser anteposto com muita facilidade: *uma brasileiríssima caipirinha*.

Capítulo 3: Categorização

3.1. Importância da taxonomia de formas

Categorização é o ato de classificar objetos, considerando alguns semelhantes a outros de acordo com certos critérios selecionados. Essa é uma característica geral do nosso sistema de conhecimento, e não se limita à classificação de objetos lingüísticos. Na verdade, pode-se dizer que categorizar é um imperativo de sobrevivência: por exemplo, uma vaca precisa distinguir entre as plantas que podem lhe servir de alimento e as que não podem, ou mesmo que a podem envenenar. Nesse caso, o animal selecionou o critério “servir ou não para comer” para distinguir as plantas em duas categorias; de outro ponto de vista, claro, a mesma categorização pode não ser válida. Assim, nossa primeira observação (que vai ser importante mais adiante) é que **toda categorização se faz em função de um critério, ou seja, de um objetivo descritivo.**

Nenhum sistema de conhecimento pode funcionar sem um subsistema de categorização. Uma razão para isso é que o mundo é complexo demais para caber literalmente na memória de uma criatura. Eu, no momento, estou sentado em um objeto que chamo de *cadeira*, e há outras cadeiras pela casa, no meu gabinete e em vários outros lugares que frequento. No entanto, há nítidas diferenças entre essas cadeiras: podem ser estofadas ou não, feitas de madeira, de metal ou de plástico, com braços ou sem braços, confortáveis ou desconfortáveis, claras ou escuras e assim por diante. No entanto, para mim são todas cadeiras, se meu objetivo é distingui-las de outros objetos que não servem para sentar; ou servem, mas não têm encosto; ou são feitos de cimento e fixos no chão. Seria totalmente impossível ter uma palavra para designar cada tipo de cadeira – ou, pior, cada cadeira individual, já que nunca há duas realmente idênticas em todos os detalhes. E, acima de tudo, seria inútil, porque meus objetivos comunicativos não exigem essas distinções todas. Em vez disso, criou-se uma **categoria**, atribuindo-se a ela um nome, *cadeira*. Dizemos então que a noção de “cadeira” é **esquemática**, no sentido de que representa uma multidão de conceitos particulares, que são **elaboraões** do conceito esquemático básico.

Ao considerarmos a estrutura de uma língua, seja em seu aspecto formal, seja em seu aspecto semântico, lidamos a todo momento com categorias. Algumas dessas categorias têm nomes bem conhecidos: verbos, preposições, sintagmas nominais, orações. Outras não são em geral mencionadas nas gramáticas, mas são igualmente importantes: verbos de ação, nomes de objetos inanimados, nominalizações etc. Não é possível descrever uma língua sem lançar mão de grande número dessas categorias; por exemplo, não seria prático dar a lista de todas as palavras que aceitam o artigo *a*, mas não aceitam o artigo *o*; em vez disso, usamos as categorias “substantivo feminino” e “substantivo masculino”. Quase todas as afirmações gramaticais se formulam, de alguma forma, com referência a categorias formais ou semânticas.

Disso se conclui que a categorização é algo a que os lingüistas deveriam dar muita atenção. No entanto, isso nem sempre acontece, e em muitos casos se constroem teorias

gramaticais utilizando categorizações tradicionais seriamente deficientes – o que resulta em teorias igualmente deficientes.

Nem sempre foi assim. Os estruturalistas dos anos 40 e 50 davam muita atenção à categorização das formas (embora nem tanto à dos significados). O exemplo paradigmático é Harris (1951), que chegou ao ponto de propor uma gramática de listas, quase que reduzindo a descrição gramatical à categorização de formas:

“Em uma de suas formas mais simples de apresentação, a descrição sincrônica de uma língua pode consistir inteiramente em um conjunto de listas.”

[Harris, 1951: 376]

As listas de Harris incluíam o que hoje chamaríamos de “construções”, ou seja, as possibilidades de distribuição de cada item. Essa concepção, se levada a suas conseqüências lógicas, reduz a descrição ao léxico, sem gramática, o que me parece inadequado, pois identifica idiossincrasias e generalidades.¹

3.2. Correlação entre forma e significado

Harris, como muitos lingüistas de seu tempo, se concentrava na distribuição formal, sem tentar relacioná-la com traços semânticos dos itens envolvidos. Os cognitivistas de hoje adotam a posição oposta, e tendem a ver na categorização gramatical um subproduto da categorização semântica. Essa é uma hipótese interessante, mas necessariamente parcial, pois não se pode negar que certas classificações gramaticalmente relevantes carecem de correlato semântico. Ou seja, a categorização gramatical, como a gramática em geral, é em parte independente da categorização semântica.

Langacker, um dos fundadores da gramática cognitiva atual, reconhece isso quando diz

Obviamente, [...] a lista de membros de muitas classes gramaticais não é totalmente previsível com base em propriedades semânticas ou [formais]. Por exemplo, a classe dos substantivos que sonorizam *f* em *v* no plural (*leaf / leaves*, mas *reef / reefs*).

[Langacker, 1991: 19; ver também 1987: 420 sqq]

No entanto, outros autores parecem não concordar, e assumem uma posição extrema que me parece inadequada:

¹ Não que a distinção entre léxico e gramática seja nítida; mas não há dúvida que certas afirmações são de validade geral, e devem ser expressas como válidas para grandes conjuntos de itens, ao passo que outras são altamente particularizadas (ver 2.3.2 acima). Só que há também muitos casos intermediários: ver, logo adiante, o caso dos **clichês**, que precisam ser considerados ao mesmo tempo construções gramaticais e itens léxicos.

A expectativa é que os fatos sintáticos (e morfológicos) de uma língua sejam *motivados* por aspectos semânticos e que possam ser exaustivamente descritos através de estruturas simbólicas.

[Taylor, 2002: 29]

Não sei até que ponto Taylor leva a sério essa “expectativa”. Mas o exemplo de Langacker, e muitos outros, mostram que nem como expectativa essa afirmação pode valer: os fatos morfossintáticos não podem ser **exaustivamente** descritos em termos simbólicos. O que vejo aqui, no máximo, é uma hipótese de trabalho, que sugere que se deve sempre procurar correlatos semânticos para cada fato formal observado.² Mas a extensão dessa correlação é uma questão empírica, e só será conhecida depois de levantamentos muito amplos de fatos individuais.

Basilio (2004) prefere assumir uma posição de prudência, ao afirmar que

existe uma relação geral óbvia entre as propriedades semânticas e gramaticais das classes de palavras

[Basilio, 2004: 23.]

Ou seja, posso aceitar que existe “uma relação geral”, desde que fique entendido que não se trata de uma regra sem exceções. Se é isso o que Basilio quer dizer, concordo com ela. Já o número e a importância das exceções, como disse, é coisa a verificar.

Temos, portanto, duas perguntas básicas a responder, e nenhuma delas está respondida a não ser em medida muito limitada:

(a) Como se classificam as formas lingüísticas (de maneira linguisticamente interessante)?

(b) Até que ponto uma classificação linguisticamente interessante se correlaciona com uma classificação semântica?

Essas questões devem ser atacadas indutivamente, empiricamente, e não a partir de idéias preconcebidas.

3.3. Insuficiência da classificação tradicional

3.3.1. O que está faltando

Como apontei acima, qualquer descrição gramatical utiliza crucialmente um sistema de categorizações, em geral na forma de classes de palavras e de sintagmas. É o caso da gramática tradicional, que merece atenção porque as teorias gramaticais modernas tendem a aceitar a classificação tradicional sem suficiente crítica.³

Sem entrar em detalhes (para o que existe alguma literatura; ver, para o português, Haug, 1983), eu diria que as deficiências principais da classificação de palavras

² Ou seja, aplicando-se a Hipótese da Sintaxe Simples, vista em 2.4.7.

³ Novamente, os estruturalistas são exceção; ver, por exemplo, a análise do inglês de Fries (1952).

encontrada na gramática tradicional são: o uso de sistemas simples de classificação; a falta de critério nas subclassificações; classes do tipo “cesta de lixo” (por exemplo, advérbios e pronomes); e rejeição da categorização múltipla. Tomando uma por uma:

(a) Uso de sistemas simples de classificação.

Aqui, como em muitos outros pontos, a análise tradicional subestima a complexidade da língua. Um exemplo impressionante, tratado nos capítulos 7 a 13 do presente livro, são as valências verbais, que na visão tradicional se resumem a cinco categorias (verbos **de ligação, intransitivos, transitivos diretos, transitivos indiretos e transitivos diretos e indiretos**) – mas Levin (1993) não conseguiu descrever os tipos de valência dos verbos do inglês em menos do que 49 categorias principais, a maioria com muitas subdivisões.

Em geral se entende um sistema de classificação como um conjunto de classes que se subdividem cada uma em certo número de subclasses e assim por diante. Assim, na gramática tradicional os advérbios se subdividem em: de tempo, de modo, de lugar etc.; os substantivos podem ser próprios ou comuns, e os comuns por sua vez são concretos ou abstratos etc. Mas esse tipo de sistema de classes dentro de classes não é adequado para descrever o comportamento gramatical dos itens léxicos e demais unidades. Os fatos são mais complicados, e envolvem o que se pode chamar “classificação cruzada”. Para explicar o que vem a ser a classificação cruzada, vamos pegar o exemplo de sete verbos: *ser, derreter, abrir, dar, desmaiar, perder e comer*, e classificá-los quanto a diversos traços gramaticalmente relevantes.

De um ponto de vista eles ficam todos juntos: pertencem a lexemas que comportam relações de tempo, pessoa, modo etc. – ou seja, são todos **verbos**. Mas mesmo no que diz respeito à organização dos lexemas, podemos distinguir aí os regulares (*desmaiar, derreter, comer, abrir*) dos irregulares (*ser, perder, dar*).

Por enquanto, temos uma classe e duas subclasses, à maneira clássica. Mas agora, ainda nos atendo à organização morfológica dos lexemas, temos que dividi-los em conjugações: *perder, comer, derreter* e *ser* pertencem à segunda conjugação, *desmaiar* e *dar* à primeira, *abrir* à terceira. Note-se que aqui já não podemos falar de subclasses, porque a classificação por conjugações cruza a classificação por regularidade: *ser* e *dar* ficam juntos como irregulares (e aí se opõem a *derreter* e *desmaiar*), mas separados quando se considera a conjugação: aí *dar* fica com *desmaiar*, e *ser* fica com *derreter*.

As coisas se complicam ainda mais quando passamos a traços descritivos do comportamento sintático e semântico. Por exemplo, *ser* e *desmaiar* se opõem a todos os outros por não admitirem objeto direto.

Não é preciso ir mais adiante; já deve estar claro que a classificação não pode ser feita em termos de classes e subclasses; precisamos lançar mão de traços autônomos, e cada item se classifica em função dos traços que possui. O leitor pode facilmente verificar que, em termos dos quatro traços vistos acima, os sete verbos se distinguem todos entre si. Naturalmente, se ampliarmos o número de itens, encontraremos verbos de comportamento gramatical idêntico (quanto a esses quatro traços): por exemplo, *derreter* é igual a *ferver*, e *desmaiar* é igual a *soluçar*. O problema é que quatro traços são muito pouco para descrever o comportamento dos verbos do português; e se consideramos

também os outros itens (substantivos, advérbios, preposições etc.) a complexidade do sistema cresce significativamente.

A classificação cruzada é extremamente freqüente em gramática, embora não tenha lugar no sistema tradicional. Essa questão será mais desenvolvida mais adiante (ver as seções 3.4 a 3.7). Um recurso para exprimir a classificação cruzada é o uso de traços distintivos, usual desde os anos 30 em fonologia, e originariamente proposto por Jakobson, Trubetzkoy e os lingüistas da Escola de Praga. No entanto, embora também tenha sido usado em sintaxe (por exemplo, por Chomsky, 1965), seu emprego até o momento tem sido curiosamente limitado.

Essa complexidade não é seriamente considerada nas classificações atualmente disponíveis. Praticamente todos os modelos se baseiam em classificações extremamente simples, o que a meu ver prejudica crucialmente sua adequação.

(b) Falta de critério nas subclassificações.

O seguinte exemplo mostra a falta de critério coerente na base da classificação tradicional: os substantivos são divididos em próprios e comuns e também em primitivos e derivados. A primeira divisão tem algum significado gramatical, porque os substantivos próprios ocorrem com artigos em condições diferentes dos comuns. Mas a segunda divisão não tem nenhuma importância gramatical, porque substantivos primitivos e derivados funcionam exatamente da mesma forma na língua – sua diferença é, se tanto, histórica, mas uma gramática descritiva não é a história da língua.

(c) Classes do tipo “cesta de lixo” (advérbios e pronomes).

Esses casos são notórios, e não faltam tentativas de reformular essas classes, que são extremamente heterogêneas (para os advérbios do inglês, há entre outras a de Greenbaum, 1969). O impacto dessas tentativas tem sido menor do que o desejável, e muitos trabalhos de lingüística, assim como a totalidade das gramáticas escolares, se limitam a repetir a classificação tradicional.

(d) Rejeição da categorização múltipla.

Um aspecto da categorização que até hoje não recebeu a atenção devida é a possibilidade de categorização múltipla, isto é, de que um item seja armazenado na memória como duas ou mais unidades ao mesmo tempo. O caso principal são os **clichês**, seqüências sintática e semanticamente regulares que são utilizadas com especial freqüência e que são, ao que tudo indica, armazenadas em bloco pelos falantes, que as recuperam sem passar pelo intermédio das regras da língua.⁴ Por exemplo, quando o falante diz *estou sem camiseta*, ele monta a frase regularmente, utilizando as regras morfossintáticas da maneira usualmente aceita. Mas para dizer a seqüência *estou sem tempo*, que é estruturalmente paralela à primeira, mas de freqüência muito maior, ele tem a alternativa, freqüentemente adotada, de utilizar o armazenamento em bloco, sem necessidade de “gerar” a seqüência cada vez que a utiliza. *Estou sem tempo* é o que se

⁴ Os clichês foram denominados *stems* por Pawley e Syder (1983), em um dos primeiros estudos do assunto.

chama um **clichê**. Há alguma evidência de que os clichês são realmente armazenados em bloco como se fossem itens léxicos, e de que são muito numerosos: Pawley e Syder (1983) falam de “centenas de milhares”, o que talvez seja um exagero. Como se vê, um clichê pode ser assessorado na memória de duas maneiras: primeiro, através de sua formação regular, por exemplo aplicando as regras relevantes; e, depois, através de um armazenamento especial, como se fosse um item léxico complexo.

O estudo dos clichês está ainda na infância, e apresenta problemas metodológicos sérios. No entanto, a meu ver não é razoável negar que existam, e sua existência levanta questões importantes, relativas à propalada criatividade da linguagem – até que ponto as construções utilizadas na fala são realmente “criadas” no momento pelo falante? ⁵

3.3.2. Gerativistas e estruturalistas

A classificação tradicional, com todos os defeitos apontados acima, é muitas vezes adotada sem crítica pelas teorias lingüísticas atuais. Os estruturalistas, que a questionavam, foram prejudicados em suas tentativas de reclassificação por uma perspectiva muito estreita do que é a gramática de uma língua, e do que é uma classificação – excluía considerações semânticas e limitavam a classificação a fatores de ordem distribucional.

Os gerativistas tratam a classificação das palavras muito de passagem, e geralmente mal vão além das noções da gramática tradicional – nesse aspecto, ficam muito aquém de estruturalistas como Fries ou Harris. Por exemplo: os traços $[\pm N]$ e $[\pm V]$ são usados para a classificação básica dos itens léxicos, resultando em apenas quatro classes. Qualquer estudo mais aprofundado mostra que isso é simplificar demais as coisas. É claro que os próprios gerativistas reconhecem (presumivelmente) que cada uma dessas classes deve ser subdividida, porque o comportamento das palavras da língua é muito mais variado. Mas não costumam entrar realmente no assunto, deixando-o para alguma data futura, como se tivesse importância secundária.

Por exemplo, Haegeman (1991: 34) classifica o verbo inglês *meet* como “transitivo”, ou seja, pode ocorrer no ambiente “antes de SN”, comentando que essa fórmula “mostra o quadro sintático em que o verbo *meet* pode e precisa ser inserido”. Ela não diz mais nada, apesar de sua afirmação ser impugnada por frases absolutamente banais como *they never met again*, ou *our class meets on Mondays*. Essa superficialidade frente à questão da classificação das formas é típica de muitos trabalhos realizados na linha gerativista.

3.3.3. Posição dos cognitivistas

⁵ Outro caso possível de duplo armazenamento (e portanto categorização dupla) seria, segundo alguns autores, palavras formadas com afixos regulares, mas que apresentam semântica idiossincrática. Langacker (1987: 28) cita casos do inglês paralelos aos seguintes: *caçador* “pessoa que caça”, *ensacador* “pessoa que ensaca” (ação); *sofredor* “pessoa que sofre” (não é ação); *grampeador* “aparelho para grampear”, *elevador* “aparelho para elevar”. Quanto a esses exemplos, prefiro adotar uma posição de prudência, porque não estou convencido de que haja realmente uma regra que interpreta de maneira “regular” as palavras com sufixo *-dor*.

Os cognitivistas têm o mérito de reconhecer a necessidade da categorização múltipla, e não negam a complexidade fundamental das classes; mas, que eu saiba, não têm realizado trabalhos amplos de mapeamento do léxico. Isso talvez explique uma relativa falta de atenção à importância dos critérios formais (ao lado dos semânticos) na classificação. Langacker reconhece a necessidade de utilizar tais critérios:

Às vezes uma classe distribucional coincide com uma [classe] determinada por uma propriedade intrínseca [...] Frequentemente, entretanto, tal coincidência é só aproximada, e muitas vezes não há nenhuma razão intrínseca para determinar os membros de uma classe distribucional.

[Langacker, 1987: 421]⁶

No entanto, ele parece ter esperanças de se livrar desses casos sem realmente lançar mão de traços formais, isto é, analisando os casos sem correlato semântico como unidades convencionalizadas semelhantes às expressões idiomáticas:

Dizer que um dado elemento participa imprevisivelmente de uma construção equivale a dizer que a estrutura assim definida está estabelecida como uma unidade convencional.

[ibid., p. 422]

Isso pode funcionar para casos como os plurais irregulares do inglês, como *oxen*, *sheep*, *women* etc. Mas a imprevisibilidade (semântica) se estende a casos muito mais ricos e complexos; por exemplo, os verbos *vencer* e *derrotar* são sinônimos e ocorrem no mesmo contexto em

[1] O Avante venceu o Unidos por dois a zero.

[2] O Avante derrotou o Unidos por dois a zero.

Não obstante, mostram comportamento diferente, sem mudança perceptível de significado, em outra construção:

[3] O Avante venceu a partida de sábado.

[4] * O Avante derrotou a partida de sábado.

Isso talvez possa ser reduzido a um fator de significado, mas como fazê-lo sem cair em circularidade – ou seja, sem dizer que os verbos não são, afinal, sinônimos, porque um deles aceita *a partida* como objeto e o outro não? E como encaixar essa afirmação em alguma generalização da língua? Pelo menos no atual estado do conhecimento, é inevitável marcar esses verbos como distintos sem vincular essa distinção a traços do significado próprio de cada um. Por outro lado, seria arbitrário considerar *vencer a partida* uma expressão idiomática, ainda mais que o objeto aí não é fixo: pode-se vencer (mas não derrotar) uma batalha, uma guerra, uma discussão, uma

⁶ Acho que nesta passagem de Langacker *intrínseco* significa *semântico*.

briga, um jogo, um bate-boca etc. Casos como esse são frequentes na língua, e militam contra a solução de Langacker.⁷

3.4. Algumas questões básicas

3.4.1. O que classificar?

A categorização vale para todas as unidades utilizadas na análise lingüística: fonemas, morfemas, palavras e sintagmas. Como exemplos de classificação, temos fonemas vocálicos e consonantais; prefixos e sufixos; verbos e preposições; sintagmas nominais e sintagmas adverbiais – além das orações, que são igualmente uma classe de formas. Temos verbos de ação e verbos de estado, nomes concretos e abstratos, adverbiais de lugar, de tempo e de causa e assim por diante. Todas essas classificações, e muitas mais, são relevantes, em algum ponto e de alguma maneira, para a descrição lingüística.

Em outro nível, agrupamos as palavras em lexemas, e um lexema é também uma classe de palavras, que têm em comum determinado tipo de morfema (tradicionalmente chamado “radical”): assim, as **palavras** *falo, falavam, falando, falar* pertencem ao mesmo **lexema**, ao qual chamamos “o verbo *falar*”. Tanto palavras quanto lexemas se classificam, mas não exatamente pelas mesmas razões. Vamos considerar primeiro a classificação das palavras. É bom notar que a exposição que se segue é um tanto distorcida pelo fato de tomar a palavra como unidade mínima para a classificação; a rigor, deveríamos partir do morfema, e algumas das afirmações teriam que ser significativamente modificadas. No entanto, para simplificar a discussão, vou fazer de conta que a palavra é a unidade mínima. Acredito que não se perde nada de essencial.

Pode ser surpreendente a afirmação de que as diversas palavras que compõem um lexema – por exemplo, as diversas formas de um verbo – não pertencem necessariamente à mesma classe gramatical. No entanto, isso se sabe há tempo, e é mesmo codificado, embora de maneira obscura, na terminologia tradicional, quando se chama formas como *falando, falar* de “formas nominais” do verbo.

Uma forma verbal como *falo* ocupa sempre a função sintática de núcleo do predicado, e o mesmo se pode dizer de *falavam* ou *falássemos*. Mas a forma *falando* pode ocorrer com um verbo auxiliar (*estou falando*), ou então em ambientes em que sua função se assemelha à de certas palavras chamadas de “advérbios”:

[5] Ele não consegue dirigir carro falando. (cf. ele não consegue dirigir carro devagar)

Note-se que as formas finitas como *falo, falavam, falássemos* não podem ocorrer em nenhum desses ambientes. Ora, se a classe a que uma palavra pertence se define em termos de seu potencial funcional,⁸ somos obrigados a reconhecer que *falando* pertence a classe diferente de *falo, falavam* etc. Mais ou menos o mesmo vale para as formas *falado* e *falar*.

⁷ O exemplo vem de Lima (2007).

⁸ Isto é, em termos do conjunto de funções que pode desempenhar.

Assim, o lexema que chamamos “o verbo *falar*” se compõe de um grupo de palavras, nem todas elas pertencentes à mesma classe gramatical. Mas isso não quer dizer que o lexema seja um agrupamento arbitrário: de outros pontos de vista, os membros de um lexema têm funções comuns, o que nos autoriza a mantê-lo como uma noção válida em gramática. Por exemplo, todos os membros de *falar* têm o ingrediente semântico “falar”, ou seja, todos eles exprimem diferentes aspectos dessa atividade. Além do mais, as relações semânticas e formais entre os diferentes membros do lexema são sistemáticas: para *falar* existe um *falando*, assim como para *correr* existe um *correndo*, para *ser* um *sendo* etc. Assim, temos que reconhecer que o lexema *falar* é uma unidade coerente.

Quando falamos da valência de um verbo, esta vale para todos os membros do lexema.⁹ Se *falar* aceita objeto direto (*falar espanhol*), *falando*, *falavam* e *falássemos* igualmente aceitam objeto direto, do mesmo tipo semântico e nas mesmas condições. Assim, se é necessário tratar as palavras individualmente, é igualmente necessário, em outros momentos, fazer referência ao lexema como uma unidade.

Além de palavras e lexemas, é necessário classificar unidades menores e maiores do que as palavras. As unidades menores são os morfemas, que freqüentemente se agrupam em classes pequenas, de distribuição muito restrita, o que reflete a maior rigidez da estrutura da palavra em comparação com a da oração. Assim, temos a classe dos sufixos de pessoa, que se manifesta em paradigmas como:

<i>falo</i>	sufixo de pessoa: -o
<i>falam</i>	- ” - -m
<i>falamos</i>	- ” - -mos

etc. Aqui há uma série de problemas de análise, porque os sufixos freqüentemente representam mais de uma categoria ao mesmo tempo (por exemplo, -o não só caracteriza a primeira pessoa do singular, mas também o presente do indicativo, já que no imperfeito se encontra -a e no futuro -ei). Esses problemas são discutidos em estudos de morfologia (ver, notadamente, Camara, 1969 e Pontes, 1972).

As unidades maiores do que as palavras são os sintagmas. Por exemplo, “sintagma nominal” é o nome de uma classe de formas, assim como “sintagma adverbial” e também “oração”. O sintagma nominal é a classe das formas que podem desempenhar determinadas funções na oração, como sujeito, objeto direto e complemento de preposição. Ou seja, o SN também se define por seu potencial funcional. Além disso, ele se caracteriza semanticamente por seu potencial referencial,¹⁰ podendo referir-se a coisas – ao passo que um verbo finito como *falam* se refere a uma ação, e um sintagma preposicionado como *de madeira* denota uma qualidade.

3.4.2. As classes na descrição gramatical

A descrição gramatical faz uso crucial da referência a classes. Assim, não dizemos que as palavras *pedra*, *casa*, *quadro*, *forte*, *hepático*, *ele* (e mais outras dezenas de

⁹ Há uma complicação aparente, relativa à forma tradicionalmente chamada “particípio”. Como se verá adiante (9.6.2), trata-se na verdade de duas formas, apenas uma das quais pertence ao lexema do verbo.

¹⁰ Um outro aspecto do potencial funcional, já que ter referência é uma função semântica.

milhares) fazem o plural acrescentando um *-s*, mas que **os nominais** fazem o plural dessa maneira¹¹. E não dizemos que as seqüências *Vilma, a Vilma, minha amiga Vilma, a Vilma do terceiro andar* podem ser sujeito de uma oração, mas que **os SNs** podem ter essa função.

As razões são bem evidentes: de um lado, seria excessivamente anti-econômico fazer referência às unidades individuais – e às vezes até impossível: quantos SNs existem na língua portuguesa? Mas há outra razão, ainda mais importante: omitir as classes da descrição, mesmo que fosse possível, impossibilitaria a percepção das grandes linhas da estrutura, reduzindo-a a uma multidão de particularidades, o que a língua certamente não é. Essa perspectiva inviabilizaria qualquer explicação plausível do uso da linguagem, pois não se compreenderia que um falante pudesse utilizar corretamente uma forma que ele nunca viu antes – que é o que fazemos constantemente ao utilizar sintagmas e orações. Um sintagma pode ser totalmente novo para o falante, como *aqueles quatro caminhões amarelados*, e ainda assim não apresentar nenhuma dificuldade de uso. Qualquer falante do português sabe que esse sintagma pode ser sujeito ou objeto direto, mas nunca núcleo do predicado ou modificador.

Como se vê, as classes se colocam bem no centro da descrição e da teoria lingüística. Classificar é um pré-requisito indispensável a qualquer análise. Por isso, é necessário ter uma noção bem clara e explícita da natureza das classes e dos objetivos de seu estabelecimento; e é igualmente necessário levar a efeito o trabalho de depreensão e definição cuidadosa das classes em que se dividem as formas de uma língua. A classificação tradicional – amplamente adotada, sem crítica, por algumas vertentes da lingüística moderna – é um ponto de partida, mas tem deficiências seríssimas, que resumi na seção 3.3 acima. É necessário elaborar uma nova classificação; e, antes disso, precisamos de uma explicitação e avaliação dos critérios que devem basear toda classificação gramatical.

3.4.3. *Classificando por objetivos*

Vimos em 3.1 acima que uma classificação não faz sentido por si só, de maneira que não se pode simplesmente perguntar quais são as classes de palavras do português. Uma classificação se faz por objetivos. Mudando o objetivo, muda a classificação; e uma descrição gramatical comporta muitos objetivos, portanto muitas classificações.

Para entender isso, vamos tomar um exemplo da vida diária: a que classe de objetos pertence um livro? Essa pergunta não tem resposta a não ser que especifiquemos os objetivos da classificação. Por exemplo, consideremos as formas *fizera* e *tinha feito*: do ponto de vista semântico, elas ficam na mesma categoria, porque são exatamente sinônimas. Mas se estamos no momento tentando descrever a morfossintaxe, elas têm que ser separadas, porque uma é “composta” (assim como *estou fazendo* e *tenho feito*), e a outra é “simples”.

Dei um exemplo de classificações diferentes para os mesmos itens quando expliquei a classificação cruzada, na seção 3.3.1 acima. Voltando ao tema, digamos que estamos descrevendo os verbos *afundar*, *andar*, *amar*, *comer*, *derreter* e *correr*. Do ponto

¹¹ “Nominais” são uma classe abrangente que inclui várias classes tradicionais: substantivos, adjetivos, pronomes e artigos. Ver Perini *et al* (1996) para uma conceituação detalhada.

de vista morfológico, eles se categorizam em verbos da primeira conjugação (*afundar, andar, amar*) e verbos da segunda (*comer, derreter, correr*). Mas se estamos descrevendo suas valências, temos que fazer uma categorização muito diferente: temos aí verbos (*afundar* e *derreter*), que ocorrem com ou sem objeto direto, mas que quando ocorrem sem objeto direto têm sujeito paciente, como em

[6] A menina derreteu o chocolate / o chocolate derreteu.

[7] A tempestade afundou o barco / o barco afundou.

Já *andar* e *correr* ocorrem normalmente sem objeto direto, e de qualquer maneira sempre que ocorrem seu sujeito é necessariamente agente:

[8] Ele andou / correu de Nova Lima a Belo Horizonte.

Finalmente, *comer* e *amar* ocorrem com ou sem objeto direto, e o sujeito é igualmente agente em todos os casos:

[9] A menina comeu o pudim / a menina comeu.

[10] Essa menina está amando o vizinho / essa menina está amando.

Note-se que ambas as classificações, embora sejam de certo modo contrastantes, precisam ser conhecidas para que o falante utilize corretamente esses verbos. Assim, se alguém perguntar como se classificam esses seis verbos, só é possível responder se se explicitar previamente o objetivo da classificação. Essa situação é típica em gramática, e precisa ser levada em conta a todo momento.

A idéia de classificar por objetivos não é de todo nova; às vezes é mesmo explicitada, embora me pareça que raramente seja levada a sério. Por exemplo, Vilela (1992), falando sobre a classificação dos verbos pela valência, tal como encontrada na literatura, observa:

[...] a própria classificação é contraditória entre diferentes propostas e mesmo dentro de uma só proposta. Que grandezas escolher? Dependerá da finalidade da classificação.

[Vilela, 1992: 25-36]

Essa passagem exprime bem o que acabei de dizer.

A classificação por objetivos não significa que as classes, uma vez definidas, sejam variáveis em função do objetivo perseguido. Ou seja, se uma palavra foi definida como “verbo transitivo” (e a classificação, naturalmente, obedece a determinado objetivo), essa classificação se mantém. O objetivo é relevante no momento de classificar um grupos de formas, mas não afeta a classificação uma vez realizada, nem os termos que a exprimem: estes devem ser sempre unívocos, de maneira que se saiba com clareza o que se quer dizer quando se afirma que determinado verbo é “transitivo”, ou que determinada palavra é um “paroxítono”.

3.4.4. Classes e funções

É preciso distinguir estritamente classes de funções, mas isso nem sempre se faz, o que constitui uma fonte de confusão. O problema se manifesta com frequência em afirmações de que elementos de determinada classe “funcionam” como se pertencessem a outra classe em determinado contexto. Por exemplo, encontram-se referências à “realização não transitiva de um verbo transitivo”, ou a “adjetivo que funciona em certo contexto como advérbio” e assim por diante. Vou tentar esclarecer esse ponto, e para adiantar minhas conclusões, formulo o seguinte princípio:

As funções se definem no contexto em que ocorrem, mas as classes se definem fora de contexto.

Pode-se testar isso formulando perguntas como:

- (a) A que classe pertence a palavra *gato*?
- (b) Qual é a função sintática da palavra *gato*?

Não há grande dificuldade em responder (a) – *gato* é, digamos, um substantivo. Mas a pergunta (b) não tem resposta possível; só pode ser respondida se for fornecido um contexto, porque *gato* **pode** ter diversas funções sintáticas: pode ser sujeito, objeto direto, predicativo do sujeito, complemento de preposição. Ou seja, uma função se caracteriza pela inserção do item em determinado entorno gramatical – *gato* é sujeito¹² em

[6] Gato dá muito trabalho.

Mas a classe se caracteriza pelo **potencial funcional** da forma, ou seja, pelo que a 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 *gato* pode ser sujeito ou objeto direto. Na frase [6], essa palavra é sujeito. Mas ela continua **podendo** ser objeto, em outra frase. Assim, mudando o contexto, pode mudar a função de *gato*; mas sua classe é sempre a mesma, porque a classe se define por potencialidades, não por realidades presentes do contexto. *Gato* é “substantivo” porque pode ser núcleo de um sintagma nominal, pode vir precedido de artigo etc. Essas funções possíveis constituem seu potencial funcional.

Uma conclusão desse raciocínio é que cada forma pertence a uma (e só a uma) classe. Às 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.

¹² Mais precisamente, *gato* é núcleo do SN sujeito.

[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 modificador¹³; *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 núcleo ou modificador constituem uma classe, distinta das que só podem ser núcleo ou só modificador. Conclui-se que essas três palavras pertencem a três grupos de potencial funcional distinto – ou seja, 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. Outro exemplo é o da classificação dos verbos quanto a sua co-ocorrência com objeto direto: temos verbos como *desmaiar*, que nunca têm objeto; verbos como *dizer*, que sempre têm; e verbos como *comer*, que podem ocorrer mais ou menos livremente com ou sem objeto direto. A solução freqüente é afirmar que *comer* é “transitivo” em certas frases e “intransitivo” em outras. Mas isso esconde a diferença fundamental que distingue *comer* de *dizer* e *desmaiar*. Aqui, novamente, temos três classes (três tipos de potencial funcional), e não apenas duas.

Na morfologia também se encontra essa confusão, por exemplo no conceito de “derivação imprópria” ou mudança de classe. Encontramos afirmações como

As palavras podem mudar de classe gramatical sem sofrer modificação de forma. [...] A esse processo de enriquecimento vocabular pela mudança de classe das palavras dá-se o nome de **derivação imprópria** e por ele se explica a passagem:

[...]

c) de adjetivos a substantivos:

circular *persiana* *veneziana*

[...]

h) de adjetivos a advérbios:

(*ler*) *alto* (*falar*) *baixo* (*custar*) *caro*

[Cunha, 1976: 120-121]

Aqui há muito o que criticar, mas vou me limitar aos casos de *circular* e *alto*, que se ligam diretamente a nosso assunto do momento. É verdade que encontramos *circular* como núcleo de um SN, em função semântica referencial (*ela enviou uma circular*), e também como modificador, em função qualificativa (*ela enviou um ofício circular*, ou então *tem uma praça circular perto da minha casa*). Mas isso não significa (em termos descritivos, sincrônicos), que a palavra tenha “mudado” de classe. Primeiro, o que é que nos autoriza a dizer que o uso como modificador é o básico, e que *circular* “se tornou” núcleo de um SN por efeito de um processo gramatical? Por que não exatamente o contrário? É claro que a decisão é arbitrária, e se tem alguma motivação, deve ser diacrônica – suspeito que é isso por causa da inclusão de casos como *persiana* na

¹³ “Adjunto adnominal”, na nomenclatura tradicional.

mudança de classe, sendo que essa palavra, no português atual, só é usada como nome de um objeto.

Depois, como vimos, a afirmação de que uma forma pertence a mais de uma classe não faz sentido. O que temos aí são palavras (por exemplo, *circular*, *amigo*) cujo potencial funcional é mais amplo do que outras (por exemplo, *mesa* ou *administrativo*); as primeiras podem ser núcleos de SN e modificadores, ao passo que as segundas podem ser apenas núcleos de SN (*mesa*) ou apenas modificadores (*administrativo*).¹⁴ E podemos inclusive acrescentar que essa maior versatilidade sintática de *circular* provém diretamente do fato de que tem um potencial semântico mais amplo: *circular* denota uma qualidade e também é o nome de uma coisa, *mesa* é apenas o nome de uma coisa, *administrativo* é apenas uma qualidade.

O mesmo vale para exemplos como *alto* em oposição a *dourado*: *alto* pode modificar um núcleo de SN (*barulho alto*) ou um verbo (*falar alto*), porque, semanticamente, pode denotar uma qualidade ou um modo; já *dourado* só denota uma qualidade, e correspondentemente não pode modificar um verbo. Mas isso equivale a dizer que *alto* e *dourado* pertencem a classes diferentes (semelhantes em muitos pontos, mas não em todos), não que são “basicamente” de uma classe e “passam” para outra. Note-se, aliás, que esse uso em diversas funções (nunca diversas classes) não constitui inovação por parte do falante, mas está codificado na língua atual, e é aprendido como parte da estrutura léxico-gramatical da língua.

Isso vale para todas as classes: cada palavra (ou forma em geral) 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 contextos. Ou seja, “classe” é uma relação paradigmática, e “função” é uma relação sintagmática, conforme a distinção estabelecida por Saussure (1916).

É importante conceber as classes como existentes na língua fora de contexto¹⁵ para relacionar as classes com o uso que o falante faz da língua. A oposição entre relações sintagmáticas (funções) e paradigmáticas (classes) se correlaciona, respectivamente, com o produto da atividade lingüística (sintagmas, frases, enunciados) e com as regras e princípios que governam a construção desse produto. Um falante, para usar a língua, dispõe de uma “receita” de como construir frases. Evidentemente, essa receita (a gramática mais o léxico) precisa estar presente na mente do falante antes que este construa seus enunciados. Um aspecto da receita são as classes de formas; ou seja, ele sabe, de saída, que uma palavra como *gato* pode ser núcleo de um SN, mas não de um predicado, ao passo que *chegou* pode ser núcleo de um predicado, mas não de um SN. Essas potencialidades definem a classificação dessas formas, e estão presentes na mente do falante, sem especificação de contexto. Elas são, antes, instruções sobre como construir contextos – onde inserir *gato* e onde inserir *chegou*, por exemplo.

Vimos na seção precedente que existe a possibilidade de classificar uma mesma forma de várias maneiras (correspondentes a objetivos descritivos distintos). Isso não quer dizer que uma forma possa pertencer a mais de uma classe, pelo menos não no sentido tradicional. A posição tradicional é de que a classificação depende do contexto, e

¹⁴ Essas palavras têm outras funções possíveis, que estou ignorando porque não são relevantes para o ponto em discussão.

¹⁵ Elas se definem em termos de contexto, como “pode ser núcleo de um SN”, mas isso é outra coisa. Uma forma, uma vez classificada, pertence àquela classe e a nenhuma outra.

que uma forma pode pertencer a uma classe em um contexto e a outra em outro contexto. Isso, como acabamos de ver, é incorreto. Uma palavra pode pertencer a mais de uma classe quando muda o objetivo descritivo, e com ele de certo modo todo o quadro de referência da classificação. Dessa maneira, *estávamos* se classifica junto com *estavam* por ser uma forma verbal finita, e nisso se distingue de *pérola*. Mas é claro que se o objetivo descritivo for fonológico, podemos classificar *estávamos* juntamente com *pérola* (por serem ambas palavras proparoxítonas). Isso não viola o princípio de que a classificação se faz independentemente de contexto, por se tratar de uma relação paradigmática.

Essa é a diferença entre as classes e as funções, e é por essas razões que elas não podem ser confundidas, como tantas vezes acontece.

3.4.5. Exclusão do contexto anafórico

Quando se classifica uma forma gramatical, é necessário levar em conta sua ocorrência em contexto não-anafórico. A razão é que em contexto anafórico encontramos estruturas reduzidas, que podem ser analisadas como contendo lacunas; e as lacunas são recuperáveis a partir do contexto (lingüístico ou extralingüístico). Assim por exemplo encontramos frases como

[a] Fiz.

Essa frase só ocorre em contextos que permitam a recuperação do sujeito e do objeto (ou seja, do Agente e do Paciente), como em

[b] – Você fez as empadinhas?
– Fiz.

Como se vê, o contexto anafórico tende a nivelar os verbos, de maneira que, por exemplo, qualquer verbo pode ocorrer sem objeto direto, e o comportamento de *fazer* fica igual ao de *comer* desse ponto de vista. Em outras palavras, o contexto anafórico mascara algumas das diferenças que se pretende descrever ao estudar as valências verbais.

Por outro lado, as frases reduzidas típicas de contextos anafóricos podem ser analisadas, semanticamente, em termos de suas versões plenas. Em [b], é necessário entender *fiz* como tendo um Agente “eu” e um Paciente “as empadinhas”, e isso se faz através de um processamento supra-sentencial (ou seja, discursivo) que se coloca fora do estudo da gramática propriamente dito. As estruturas reduzidas são invariavelmente identificáveis como estruturas sintáticas normais, amputadas de um ou mais termos. Por exemplo:

[c] – Que roupa você achou mais bonita?
– A da Regina.

[d] Serafim toca flautim, e Ivone trombone.

É bastante claro que estruturas como *a da Regina* ou *Ivone trombone* representam reduções de estruturas completas.

Acrescente-se que esse processamento se faz em termos de contexto, de modo que se levarmos em conta a ocorrência de *fazer* sem objeto teremos que especificar que isso só se dá em determinados contextos – ou, melhor dizendo, que a diátese sem objeto de *fazer* só ocorre em determinados contextos. Mas o contexto não é usado para definir diáteses, porque estas, como aspecto que são da categorização dos verbos, definem relações paradigmáticas. Em termos gramaticais, portanto, as diáteses são autônomas (embora sua ocorrência possa ser condicionada a fatores discursivos e comunicativos em geral).

Por tudo isso, a pesquisa das valências verbais, assim como a pesquisa sintática em geral, deve ser feita com exclusão do contexto anafórico. No caso das valências, isso significa considerar apenas períodos simples. Essa é a posição universalmente adotada, embora eu nunca tenha visto uma discussão explícita do problema.¹⁶

3.5. Calibrando regras

Vamos considerar a relação entre as regras gramaticais¹⁷ e os traços categorizadores que marcam cada unidade lingüística. Vou primeiro exemplificar essa relação na fonologia, onde o uso de traços tem uma longa tradição.

O segmento representado por [ĩ] (isto é, um *i* nasal) tem, falando foneticamente, diversos traços: é uma vogal alta, não-retraída, nasal etc. Mas quando uma regra precisa referir-se a ele, não precisa levar em conta todos os traços. Por exemplo, há uma regra no português do Brasil (região sudeste) que palataliza um /t/ diante de vogais altas e não-retraídas: é a regra responsável por nossa pronúncia da palavra *tinta* com “tch” inicial. Para formular essa regra, temos que mencionar os traços “alta” e “não-retraída”, mas não é necessário mencionar o traço “nasal”, porque a regra se aplica igualmente diante de nasais ou de orais. Podemos dizer que a regra é calibrada para esses dois traços, desprezando os outros. Mas a regra que nasaliza qualquer vogal diante de uma vogal nasal (e que produz a pronúncia da palavra *cães* com *a* e *e* nasais)¹⁸ não precisa levar em conta o caráter “alto” e “não-retraído”, calibrando-se exclusivamente em função do traço “nasal”. Ou seja, uma unidade como “a vogal [ĩ]” não se classifica de maneira simples como “vogal nasal”, ou como “vogal não-retraída”, mas como um conjunto de traços; e desse conjunto somente alguns serão relevantes em uma situação dada. Essa situação é reconhecida em todas as teorias fonológicas modernas.

O mesmo ocorre na morfossintaxe e na semântica, mas a situação costuma ser obscurecida por uma concepção inadequada de “classe”, concepção essa que persiste em virtude da falta de atenção ao tema, no que pese sua importância. Se tomarmos uma

¹⁶ Essa posição, embora universalmente aceita, não é a única possível. Seria possível, em princípio, analisar as estruturas reduzidas como novas construções, e conseqüentemente novas diáteses dos verbos. Meu palpite é que vai complicar muito a gramática (mas pode haver ganhos em outras áreas).

¹⁷ Ou “princípios”, ou “relações entre estruturas”, ou “restrições de boa formação”, pouco importa. Qualquer que seja o recurso descritivo utilizado, deve funcionar com base em traços, da maneira aqui explicitada.

¹⁸ Essas regras são dadas aqui apenas como exemplos; não estou necessariamente defendendo a análise de *cães* com uma vogal oral /a / subjacente antes da semivogal.

unidade gramatical como, por exemplo, a palavra *mão*, podemos associá-la a um grande número de traços. Dependendo da afirmação gramatical que queremos fazer sobre essa palavra, contudo, somente alguns desses traços serão relevantes. Assim, por exemplo, para atribuir à palavra *mão* uma função sintática na oração, é importante saber que se trata de um nominal – por isso pode ser precedida de um artigo, por exemplo. Para descrever a forma desse artigo (assim como de outros elementos do SN), é necessário saber que essa palavra é feminina. Mas seu gênero não tem importância para se saber se ela pode ou não ser núcleo de um SN – para isso o importante é saber que é uma palavra com potencial referencial, ou seja, é o nome de uma coisa. Como não tem potencial qualificativo, não pode ocorrer como modificador:

[11] A Carol machucou a mão.

[12] A Carol detesta trabalhos manuais / * trabalhos mão.

Ou seja, para cada ponto de vista gramatical que se adote, *mão* se classifica de uma maneira diferente. Essa situação é típica, e requer, como vimos, uma noção adequada de classificação.

3.6. Taxonomia e descrição gramatical

A taxonomia das formas gramaticais depende, pois, do estabelecimento adequado de objetivos de análise; e esses objetivos obedecem à necessidade de descrever a estrutura da língua. Isso acarreta uma necessidade de colaboração estreita entre a pesquisa léxica e a gramatical. Um item léxico pode, em princípio, ser classificado de muitos pontos de vista, e nem todos são relevantes. A descrição gramatical é a fonte principal onde o taxonomista vai buscar informações sobre a relevância de cada traço possível. Por exemplo, o item léxico (lexema) *urso* pode ser classificado, no que diz respeito a seu significado, como “animado”, “não-humano”, “grande” (em oposição, digamos, a *rato*), “concreto” (em oposição a *importância*), “comum” (em oposição a *Ivan*) etc. Todos esses traços são importantes de algum ponto de vista: se alguém não sabe que um urso é grande, ou não-humano, não sabe direito o que é um urso. Mas a pergunta que nos interessa é: quais deles são necessários para a descrição gramatical?

O levantamento taxonômico precisa progredir em constante comunicação com a descrição gramatical. Esta pode nos dizer, a respeito de *urso*, que o traço semântico “animado” é provavelmente importante, porque condiciona a possibilidade de esse item ocorrer no papel temático de “agente”. O traço “comum” também é importante porque os nomes comuns ocorrem com ou sem artigo em condições diferentes dos nomes próprios: quando sem artigo, um nome comum tem referência genérica, como em

[13] Urso não suporta o calor excessivo. (*urso*: “todo e qualquer urso”)

Mas os nomes próprios ocorrem sem artigo com referência específica:

[14] Ivan não suporta o calor excessivo. (*Ivan*: “pessoa específica chamada Ivan”)

Por outro lado, o traço “grande” não tem utilidade nenhuma para a descrição das estruturas gramaticais da língua, podendo ser omitido da caracterização taxonômica de *urso*. Dessa maneira a descrição gramatical e a classificação dos itens léxicos caminham (ou deveriam caminhar) lado a lado, interagindo constantemente.

3.7. Protótipos

Existe uma maneira alternativa de lidar com a complexidade das classes, que analiso acima basicamente em termos de traços formais e semânticos: trata-se dos chamados **protótipos**, bastante freqüentes na literatura; os protótipos lançam mão, freqüentemente, de uma classificação por contínuos, em vez de postular classes discretas. Não favoreço o uso desse recurso na análise, por razões que preciso adiar por ora.

Vou apenas adiantar que vejo dois tipos principais de problemas no uso de protótipos: primeiro, acredito que há evidências de que as categorias gramaticais (em oposição a algumas categorias conceptuais) são discretas, e não contínuas; depois, vejo dificuldades metodológicas incontornáveis no uso de protótipos na análise. Esses problemas são de tal monta que inviabilizam o uso de protótipos com um mínimo de confiabilidade e rigor, razão pela qual tenho que descartá-los. Assim, o sistema de classes do português, embora certamente muito mais complexo do que geralmente se pensa, é composto por grupos definíveis em termos discretos. Remeto o leitor interessado à seção 8.4, onde retomo a questão dos protótipos com os devidos detalhes.

Capítulo 4: Funções sintáticas: sujeito e objeto direto

4.1. Funções sintáticas

4.1.1. *A Hipótese da Sintaxe Simples revisitada*

Neste capítulo vou considerar a noção tradicional de “função sintática”, focalizando em especial as de sujeito e de objeto direto e discutindo a maneira de incluir essas relações gramaticais na definição das construções. Conforme se verá, a representação das funções sintáticas deve ser feita de maneira significativamente diferente da que é geralmente aceita, tanto na gramática tradicional quanto em teorias recentes. A proposta se baseia na Hipótese da Sintaxe Simples, já exposta no capítulo 2, e que repito aqui:

A teoria sintática mais explicativa é a que atribui o mínimo de estrutura necessário para exprimir a mediação entre fonologia e significado.

[Culicover e Jackendoff, 2005: 5]

É para dirigir o lingüista em suas escolhas (de hipóteses, não de fatos) que se elaboram as teorias. A HSS faz isso minimizando o conteúdo hipotético das análises – não negando a necessidade de hipóteses, mas, digamos, colocando-as em seu devido lugar, que é o menor possível. Com isso, ela pretende evitar duplicações: descrições sintáticas ou morfológicas que simplesmente replicam informações disponíveis na fonologia e na semântica.¹

4.1.2. *Repensando as funções sintáticas*

Não é possível realizar uma descrição lingüística sem utilizar, de alguma maneira, algo que se pode chamar de “funções sintáticas”. Por exemplo, é inevitável considerar a ordem dos elementos na frase, e isso é legitimamente uma função; nas duas frases abaixo, *Patrícia* e *o cachorro* têm funções diferentes:

[1] Patrícia mordeu o cachorro.

[2] O cachorro mordeu Patrícia.

Essa diferença de função é indispensável para descrever tanto a forma quanto a semântica dessas sentenças. No caso, podemos dizer que a função se identifica com a posição linear na oração, antes ou depois do verbo.

¹ Ver justificação da Hipótese em 2.4.3.

Mas não é isso o que se entende em geral por “função sintática”. As funções seriam relações de natureza mais abstrata, no caso relacionando um sintagma nominal com o verbo (ou um argumento com o predicado). A diferença entre essas duas concepções de função se vê nas frases

[1] Patrícia mordeu o cachorro.

[3] O cachorro, Patrícia mordeu.

Segundo a primeira concepção (concreta), *o cachorro* teria funções formais diferentes em [1] e em [3], já que sua relação de ordenação difere. Mas segundo a segunda concepção (abstrata) *o cachorro* seria objeto direto nas duas frases. Isso não se deve, ou não se deve exclusivamente, ao fato de *o cachorro* receber o mesmo papel temático (Paciente) nos dois casos: entende-se que existe uma relação abstrata entre esse SN e o verbo, relação essa que se mantém mesmo quando a ordenação varia. Dessa relação abstrata é que derivariam os aspectos formais e semânticos observáveis, tais como o papel temático de Paciente, a posição do sintagma, e em certos casos sua forma (*me* em vez de *eu* etc.). Essa segunda noção de “função sintática” pode ser considerada dominante na lingüística atual, e é a que pretendo contestar, para o caso específico do objeto direto.

Procurarei mostrar que é possível atribuir o papel temático correto ao SN geralmente analisado como objeto direto sem lançar mão das funções sintáticas tradicionais. Segundo essa hipótese, a atribuição funciona apenas com base em fatores básicos observáveis, como a ordem dos elementos, mais categorizações prévias – no caso, apenas duas: a distinção entre SN e não-SN e a identificação da classe valencial do verbo. Se essa posição for correta, não há necessidade de postular a função de “objeto direto” tal como usualmente se entende. Já o caso do sujeito é diferente, de modo que não se pode reduzir essa função a uma relação posicional, embora ainda assim seja possível defini-la com base em fatores concretos, observáveis.

Segundo a posição tradicional, as funções sintáticas (sujeito, objeto direto, adjuntos...) são etapas intermediárias no relacionamento entre os espaços formal e semântico. Por exemplo, a rotulação de um sintagma como “sujeito” é normalmente parte do processo de identificação do papel temático desse sintagma. Isto é, o que justifica a eventual postulação de funções sintáticas é que elas permitem estabelecer o relacionamento entre formas e significado. A indagação deste capítulo é até que ponto a noção de “função sintática” é necessária para isso.

Vou discutir aqui as funções de objeto direto e de sujeito, e minhas conclusões só se aplicam a essas duas funções. A partir daí se levanta imediatamente a pergunta de se a mesma argumentação não se aplicaria igualmente às demais funções sintáticas tradicionais. Essa é outra questão, evidentemente, e terá que ficar para outro momento.²

² Bhat (1991) sustenta que, para certas línguas, as funções sintáticas são supérfluas, e que a descrição pode ser feita em termos mais concretos. Não tive acesso ao texto de Bhat, mas a julgar pela breve resenha de Ichihashi-Nakayama (1994) parece que ele tem uma idéia bastante próxima da que defendo aqui para certos casos.

4.1.3. Para que funções sintáticas?

Vou começar discutindo a noção de “sujeito”. Mas antes vamos considerar a questão de para que precisamos de funções sintáticas na descrição da língua.

Em português, a motivação para se postular a função de sujeito vem basicamente da necessidade de descrever os seguintes fenômenos:

- (a) a atribuição de papéis temáticos aos diversos SNs da oração;
- (b) a concordância verbal; e
- (c) a distribuição de itens como *eu*, em oposição a *me*.

Se pudermos dar conta desses fenômenos sem necessidade de postular uma relação abstrata entre um dos SNs da oração e o verbo (a função sintática de “sujeito”), essa solução deve ser preferida, em virtude da Hipótese da Sintaxe Simples. Ou seja, a função sintática, em si, não tem utilidade nenhuma; falamos do “sujeito” para podermos descrever os fenômenos acima que, eles sim, são necessários em si mesmos. A interpretação de *o cachorro* em

[1] O cachorro me mordeu.

como Agente, a correlação desse sintagma com a terminação do verbo (*-eu*, não *-i* ou *-emos*), e a interpretação de Paciente, não de Agente, da forma *me* são fatos da língua que precisam figurar na descrição. Se pudermos descrevê-los sem lançar mão de noções abstratas, tanto melhor.

Passando ao caso do objeto direto, observamos que em uma língua sem casos morfológicamente marcados, como o português, a importância de identificar o objeto direto, distinguindo-o do sujeito, se vincula fundamentalmente à necessidade de atribuir o papel temático correto a um dos SNs da oração; no caso de [1], *me* precisa receber o papel de Paciente.

4.1.4. Redundância nos papéis temáticos

Como se sabe, há uma assimetria entre as funções de sujeito e de objeto direto (que vou de agora em diante chamar simplesmente de “objeto”). O sujeito é envolvido em uma relação chamada de “concordância”, da qual o objeto não participa. Para descrever essa diferença, vou partir da seguinte observação:

Um dos papéis temáticos da oração pode ser indicado por um SN presente na oração e, redundantemente, pelo sufixo de pessoa-número do verbo.

É o que se verifica em

[2] Antônio beliscou as meninas.

onde *Antônio* está em uma posição tal que recebe o papel de Agente; e o SPN *-ou* indica, redundantemente, que o Agente deve ser de 3ª pessoa do singular (semanticamente, isso só quer dizer que não se trata do falante sozinho, porque o sufixo pode denotar *você*, por exemplo, ou ainda *a gente*, onde o falante figura em companhia de outras pessoas).

A redundância, nesse caso, é parcial, porque o SN *Antônio* elabora a informação fornecida pelo sufixo. Em outros casos, verifica-se redundância total, como em

[3] Eu belisquei a menina.

e é justamente nesses casos que o SN pode ser omitido livremente. Compare-se

[4] Belisquei a menina.

[5] * Beliscou a menina.

[5] só é aceitável em contexto anafórico. O mesmo ocorre sempre que o sufixo é excessivamente ambíguo, como em

[6] Eu beliscava a menina.

[7] * Beliscava a menina.

Comparando-se [4] com [7] vemos que a omissão do sujeito não depende propriamente da pessoa gramatical, mas das condições de identificação do referente ao qual se deve atribuir o papel temático indicado pelo sufixo de pessoa-número. Em *belisquei*, trata-se de “eu”, sem margem de dúvida; mas em *beliscava* pode ser “eu” ou então qualquer das outras opções oferecidas pela terceira pessoa: “ele”, “ela”, “você” etc. A ambigüidade é intolerável, e a frase é mal formada.

Há ainda casos em que a indicação redundante não se verifica por falta de sufixo de pessoa-número no verbo; isso ocorre em construções de gerúndio como na oração marcada em

[8] Antônio encontrando a menina, ela vai ficar de castigo no quarto.

Note-se como nesses casos, apesar da ausência total de concordância, não há ambigüidade.

O SN que partilha seu papel temático com o sufixo de pessoa-número, elaborando-o, é o que chamamos de **sujeito**.

4.1.5. Identificação do sujeito

A próxima pergunta é: como é que o receptor, confrontado apenas com uma seqüência de palavras, sabe qual dos constituintes da oração é o que se relaciona com o sufixo de pessoa-número dessa maneira?

Em uma frase como

[9] Antônio beliscou a menina.

é claro que o fator decisivo é a ordem dos sintagmas: o sintagma relacionado com o sufixo de P-N é o que aparece imediatamente antes do verbo. É fácil testar isso trocando de lugar os SNs, o que acarreta troca de papéis temáticos:

[10] A menina beliscou Antônio.

É evidente que essa regra simples não funciona para todos os exemplos. Em muitos outros casos, o sufixo de PN se relaciona com sintagmas em outras posições. O exame de todos os casos conhecidos nos fornece uma regra mais complexa, que é a seguinte:

Regra de identificação do sujeito

Condição prévia: O sujeito é um SN cuja pessoa e número sejam compatíveis com a pessoa e número indicados pelo sufixo de PN do verbo.

(i) Se na oração só houver um SN nessas condições, esse SN é o sujeito.

(ii) Se houver mais de um SN, então o sujeito é o SN que precede imediatamente o verbo.

(iii) Mas se o SN em questão for um clítico (*me, te, nos, se*), ele não conta, e o sujeito é o SN precedente.³

Cada item da regra se baseia exclusivamente em informação disponível ao usuário da língua, a saber, a seqüência formal⁴ e conhecimentos prévios sobre a estrutura da língua. Vamos examinar cada um deles, sob essa perspectiva.

Começando pela condição prévia: naturalmente, o falante possui meios de identificar um SN e diferenciá-lo de sintagmas de outras classes (sintagmas adverbiais, adjetivos ou verbais). Ele faz isso com base em regras de estrutura sintagmática de formato tradicional. Essa condição é necessária para explicar por que constituintes como *antigamente* ou *sem a menor dúvida*, mesmo que ocorram logo antes do verbo, não recebem o papel temático reservado ao sujeito. Trata-se, evidentemente, em última análise, de uma condição semântica: o sujeito tem que ser um constituinte com potencial

³ A regra funciona para orações do tipo das que estudamos aqui. Para períodos compostos e casos de redução anafórica tornam-se necessárias outras restrições.

⁴ Em última análise, fonética. Estou saltando os estágios de processamento que relacionam a forma fonética com a seqüência de palavras que constitui o ponto de partida tradicional nos estudos de sintaxe.

referencial -- o que significa, em termos sintáticos, um SN.⁵ Mas é claro que o receptor precisa partir de informação presente no sinal, ou seja, informação de caráter léxico e sintático. Mais adiante (4.1.8) veremos alguns detalhes importantes desse processo de identificação dos SNs.

O item (i) da regra não apresenta dificuldade de aplicação, já que depende apenas da contagem dos SNs (que, como veremos mais adiante, é uma operação necessária em mais de um ponto do processo) e da verificação da compatibilidade em pessoa e número entre um SN e o sufixo de PN.

Quando há mais de um SN na oração, há mais de um candidato a sujeito, e aí em princípio temos um problema. O problema entretanto, se resolve pela posição dos SNs (item (ii)), de modo que, no português falado atual, não se conhece nenhum caso em que haja dúvida quanto à identificação do sujeito. Este precede imediatamente o verbo, exceto no caso dos itens léxicos de uma lista pequena e fechada, a saber, *me*, *te*, *nos* e *se* (item (iii)).⁶ Ou seja, a identificação do sujeito se faz sem deixar resto.

Por outro lado, pode-se sustentar que a função de “sujeito” é necessária, embora não seja básica, mas antes derivada da aplicação da regra de identificação. Isso será desenvolvido mais abaixo. Antes, vamos examinar a função tradicional de “objeto”.

4.1.6. Identificação do objeto

Se imaginarmos que o processo de identificação do objeto se dá em conexão com o de identificação do sujeito, não encontramos complicações maiores, pois o objeto se caracterizaria por sua posição pós-verbal.

Mas, como já adiantei, o importante não é identificar o objeto, mas atribuir o papel temático correto ao SN em questão. Se a cadeia sonora corresponder a

[11] Sônia comeu uma pizza.

será necessário atribuir a *Sônia* o papel de Agente, e a *uma pizza* o papel de Paciente. Nesse caso, e já que a forma do verbo não identifica o elemento com o qual ele concorda, isso se faz através da posição de cada sintagma – digamos, em relação ao verbo. O SN imediatamente antes do V será Agente, e o SN que vier após o V será Paciente. Note-se que não há necessidade de atribuir a *uma pizza* uma função sintática abstrata (a de “objeto”) distinta de sua posição, já que esta basta para caracterizar o sintagma como candidato único ao papel de Paciente. Isso se faz com base no sinal fonético e no conhecimento de que se o verbo é *comer* (e não está na construção passiva)⁷, o SN pós-verbal é necessariamente Paciente.

Já quanto ao SN pré-verbal, é o Agente na frase acima; mas em certos casos ocorrem dois SNs antes do verbo, como em

⁵ Isso será eventualmente reduzido às condições de atribuição de papéis temáticos como Agente, Paciente etc., que ocorrem com o sujeito, e que só podem ser atribuídos a sintagmas com potencial referencial.

⁶ Se considerarmos que *lhe* não é propriamente um SN; mas em certos dialetos o *lhe*, tomado como sinônimo de *te*, também entra na lista -- trata-se dos dialetos onde se diz *ele conseguiu lhe agarrar*?

⁷ A passiva não é uma exceção, como veremos na seção 9.6.3.

[12] A pizza, Sônia comeu com muito gosto.

Esse exemplo sugere que o papel de Agente é atribuído ao SN pré-verbal mais próximo do verbo, caso haja dois SNs pré-verbais; e quando só há um, ele recebe o papel de Agente -- e o SN inicial recebe o de Paciente. Novamente, a função de “objeto” é supérflua. Em particular, não há necessidade de marcar *a pizza* como tendo a mesma função sintática em [11] e em [12]. As funções são diferentes, como se pode ver pela estrutura da frase (em uma *a pizza* vem depois do verbo, na outra vem no início da frase). Acontece, apenas, que essas duas funções recebem o mesmo papel temático.⁸

Estou vinculando aqui a atribuição dos papéis de Agente e Paciente exclusivamente ao posicionamento pré- e pós-verbal. No caso do sujeito, normalmente se considera também a relação de concordância com o verbo. Mas já vimos que essa relação nem sempre funciona: em [11], por exemplo, não há modo de demonstrar que *comeu* concorda com *Sônia* e não com *uma pizza*. Ou seja, já que não há ambigüidade em [11], é inevitável lançar mão da ordenação para distinguir os dois termos.

Se nos limitamos ao verbo *comer*, os fatos não são excessivamente complexos, e parece possível atribuir os papéis corretamente de maneira bastante simples. Há um grande número de verbos que funcionam exatamente como *comer*: *matar*, *ler*, *escrever*, *comprar* etc. Assim, o sistema acima esboçado tem bastante quilometragem. O falante sabe quais são esses verbos, e sabe que eles fazem parte de um grupo caracterizado por uma valência específica. Há outros grupos de verbos, com propriedades diferentes. Vamos examinar brevemente esse fenômeno, que será o tema da segunda parte deste livro.

4.1.7. Papel das valências

A atribuição dos papéis temáticos depende em parte da classe a que pertence o verbo. Os verbos da classe de *comer* se caracterizam por ocorrerem em frases com sujeito Agente e mais um outro SN Paciente. Essa construção se denomina **transitiva**, e pode ser definida, provisoriamente, assim:⁹

[13] **Construção transitiva**, composta de

(Sujeito-Agente) + verbo + sufixo de PN-Agente + SN-Paciente

A associação de *comer* com essas construções faz parte do conhecimento dos falantes a respeito dos verbos da língua. É parte de sua competência, e portanto pode ser utilizada como um dos pontos de partida para o processamento das orações do português. *Comer*, nesse particular, é igual a muitos outros verbos do português (*matar*, *escrever*,

⁸ Não apenas com o verbo *comer*. Trata-se de uma regra geral da língua.

⁹ A notação de [13] difere da que uso na lista de diáteses do capítulo 13, e que será devidamente exposta na seção 4.5. Para contornar certas discussões que nos levariam longe do ponto principal deste capítulo, acrescentei o rótulo de “sujeito”, incluí o sufixo de PN e seu papel temático (sempre o mesmo do sujeito) e marquei o SN correspondente como opcional.

aprender, rabiscar etc.). Esses verbos compõem uma classe valencial, ou seja, têm valências semelhantes.

Vamos agora examinar outra classe, a que pertence o verbo *esquentar*. Em frases como a seguinte, *esquentar* se comporta exatamente como *comer*:

[14] A cozinheira esquentou o leite.

Ou seja, o SN pré-verbal é o Agente, e o pós-verbal é o Paciente -- trata-se, mais uma vez, de um exemplo da construção transitiva. Mas se o SN pós-verbal não ocorrer, os dois verbos funcionam de maneira diferente:

[15] Sônia já comeu.

[16] O leite esquentou.

Com *comer*, as coisas ocorrem como previsto, ou seja, retirou-se o Paciente e sobrou o Agente. Mas em [16] o SN pré-verbal (sujeito) não é Agente, mas Paciente.¹⁰ Isso acontece com grande número de verbos; e descrevemos o fenômeno associando *esquentar* à diátese ergativa, ou seja,

[17] **Construção ergativa**, composta de

(Sujeito-Paciente) + verbo + sufixo de PN-Paciente

Ou seja, diremos que a valência de *comer* inclui a diátese transitiva (exemplificada em [11]), mais uma outra diátese, idêntica à transitiva mas sem o SN pós-verbal -- chamo-a **construção transitiva de objeto elíptico** (exemplificada em [15]). Mas a valência de *esquentar* inclui, além da diátese transitiva, a diátese ergativa (exemplificada em [16]). Essa associação entre verbos¹¹ e listas de diáteses constitui sua valência, e é parte do conhecimento da língua dominado pelo usuário. Os verbos se classificam segundo suas valências, e é necessário conhecer a classe do verbo para processar corretamente uma frase. Com verbos como *comer* o SN pré-verbal é sempre Agente; mas com verbos como *esquentar* será Agente ou Paciente, conforme as circunstâncias.

O receptor, ao processar uma oração, deve identificar a construção que a oração representa. Isso é importante porque, com o mesmo verbo, podemos ter sujeito Agente ou sujeito Paciente, dependendo da diátese que a frase realiza. Há naturalmente muitas maneiras de representar isso na descrição da língua. A maneira que prefiro é a de associar cada verbo a um conjunto de construções onde ele pode ocorrer. Cada uma dessas construções é uma **diátese** desse verbo, e o conjunto de todas as diáteses de um verbo é a sua **valência**.

Esse sistema descreve satisfatoriamente o fenômeno (a) da lista acima, ou seja, **a atribuição de papéis temáticos aos diversos SNs da oração**. E note-se que faz isso

¹⁰ A rigor, frases como [16] são ambíguas, e o sujeito também pode ser Agente: comparar [16] com *conhaque com mel esquento*.

¹¹ Não apenas verbos, como veremos a seu tempo; trata-se de pares verbo + predicado.

partindo exclusivamente de informação disponível ao receptor: o sinal sensorialmente perceptível fornecido pelo emissor e o conhecimento da língua que ambos dominam previamente à realização do ato de fala. Em particular, o sistema não depende de elementos abstratos tais como categorias vazias, funções abstratas ou marcas de coindexação.

4.1.8. SNs e não-SNs

Vamos agora desenvolver alguns pontos importantes que ficaram mais ou menos em suspenso na discussão que precede. Esses pontos têm a ver com a forma de aplicação do mecanismo esboçado de atribuição dos papéis temáticos dentro da oração.

O procedimento de atribuição do papel temático desenvolvido acima depende não apenas da posição do constituinte na oração, mas também de uma categorização do mesmo como sendo um SN. Essa categorização nos impediria de atribuir o papel de Paciente ao sintagma *com muito gosto* em

[18] Sônia comeu com muito gosto.

Mas, como já vimos, há outra razão para que *com muito gosto* não possa ser Paciente: esse sintagma não denota uma coisa (pessoa, animal, objeto), e um Paciente é sempre uma coisa. O sintagma *com muito gosto* exprime uma circunstância de modo, sendo nisso paralelo a *rapidamente*, *à força* etc. Assim, podemos dizer que o papel de Paciente (com o verbo *comer* na voz ativa) é atribuído ao constituinte pós-verbal **que seja semanticamente apto a receber esse papel**. Esses constituintes são, normalmente, SNs típicos. Mas nem sempre; em alguns casos (raros, é verdade) elementos que não se analisam geralmente como SNs aparecem com o papel de Paciente, como em

[19] Eu vou apagar hoje da minha memória.

Aqui o Paciente (a coisa que vai ser apagada) é expresso pela palavra *hoje*, que em geral se analisa como advérbio, e não como nome (logo, não poderia ser núcleo de um SN).

Mas o problema é apenas aparente, e decorre da visão simplista (embora dominante) da categorização das formas lingüísticas que venho criticando no presente trabalho. Rótulos como “SN”, e também “substantivo”, “conjunção” etc., que denotam classes de formas, são apenas abreviaturas para matrizes de traços, cada traço descrevendo uma propriedade gramatical. Há grandes concentrações de itens em certas matrizes – por exemplo, os substantivos típicos¹² são dezenas de milhares em português – mas há também itens que apresentam um grupo de propriedades mais ou menos idiossincrático. Assim, sabemos que *alto* tem propriedades que o aproximam dos adjetivos (*ele é alto*) e dos substantivos (*ele gritou lá do alto*), mas também propriedades

¹² Isto é, as palavras que designam ou podem designar coisas. No Aurélio há cerca de 80 mil, de acordo com estimativa de Luís Antônio Marcuschi (comunicação pessoal). A noção de “coisa” como termo técnico de semântica é discutida por Langacker (1987: 183 – 213).

que o aproximam de certos advérbios (*ele fala alto*). Esse fenômeno é muito freqüente e bastante importante na língua.

A palavra *hoje* é um desses casos limítrofes. Funciona em geral de maneira paralela à de advérbios como *atualmente*, mas também (e ao contrário de *atualmente*) pode aparecer em funções geralmente reservadas a SNs, como em [19] ou em

[20] Hoje foi um dia ótimo para mim.

Essa dualidade provém, em última análise, de uma maior riqueza semântica da palavra *hoje*: além de exprimir uma circunstância de tempo (como *atualmente*), *hoje* também se refere a uma **coisa**, o dia de hoje. Para nos referirmos, então, ao tempo atual, temos que dizer *o tempo atual*, *os dias atuais* etc. Mas para nos referirmos ao dia de hoje, podemos usar simplesmente a palavra *hoje*, porque ela tem propriedades que o permitem.

Um SN se identifica a partir de conhecimento léxico e gramatical. Quando se trata de apenas uma palavra (*Fernando*, *hoje*) tudo o que é preciso fazer é verificar se a palavra em questão tem os traços semânticos que lhe permitam ser núcleo de um SN e, eventualmente, ser o SN inteiro, caso não tenha determinantes nem modificadores. Mas quando o SN é composto de várias palavras o processo é mais complicado. Digamos que o receptor se defronta com a seqüência

[21] O tio de Fernando ronca.

De partida, ele sabe que há dois itens aí que poderiam ser núcleos de um SN: *tio* e *Fernando*. Mas qual deles é núcleo do SN sujeito de *ronca*? O usuário da língua pode escolher *Fernando* como o núcleo de um SN composto apenas dessa palavra, e atribuir a esse termo o papel temático de Agente. Sabemos que essa solução é incorreta, mas como é que o usuário a evita?

A resposta está na noção de **fechamento** (*closure*): todos os itens (palavras e morfemas) precisam ser incluídos na estrutura, seguindo as regras da língua, sem sobras nem faltas. E acontece que se *Fernando* for o sujeito da frase, ficamos com *o tio de*, que não pode encaixar-se em lugar nenhum sem violar as regras da sintaxe. Consultando essas regras, vemos que há outra possibilidade de preencher o lugar do sujeito, sem deixar resto: *o tio de Fernando*; e esse SN sujeito se soma ao verbo *ronca*, formando uma oração que se conforma às regras da língua. Fechamos, portanto, a estrutura desta maneira:

[22] ([o tio de Fernando]_{SN} ronca)_o

Como parte desse processo, identificamos *o tio de Fernando* como um sintagma; e as regras da língua, juntamente com a categorização da palavra *tio* como um nominal passível de ser núcleo de um SN, nos permitem identificá-lo como um sintagma nominal.

4.1.9. Verbos

Outra tarefa que o receptor precisa realizar é a identificação do verbo. Isso é necessário porque falamos de SN pré- e pós-verbal, o que pressupõe o reconhecimento prévio do verbo.¹³

Nos casos mais favoráveis, a forma do verbo é suficientemente característica, tanto morfológica quanto sintaticamente, para que a identificação não apresente problemas. Assim, em

[23] Nós corríamos pela beira da lagoa.

é muito pouco provável que *corríamos* seja outra coisa que não um verbo, tanto por sua forma (com a terminação característica *-íamos*) quanto por sua posição logo depois de *nós*, item que funciona tipicamente como sujeito.

Em casos menos claros entra em jogo novamente a noção de fechamento, já mencionada acima. Um exemplo é

[24] Gripe nunca mata.

As palavras *gripe* e *mata* são, quando consideradas isoladamente, ambíguas: *gripe* pode ser forma do verbo *gripar* ou o nome da doença; *mata* pode ser forma de *matar* ou sinônimo de *floresta*. E como não há na frase nenhuma marca de subordinação (como a conjunção *que*, por exemplo), é de esperar que só haja uma oração e portanto só um verbo. Mas qual dessas palavras é o verbo?

O acesso à informação gramatical nos dá a resposta: a palavra *nunca* não pode aparecer depois do verbo e antes de um SN, ou seja, a estrutura V + *nunca* + SN não é aprovada pelas regras da gramática. Mas se *gripe* for o nominal, poderá ser um SN, já que um SN pode ser composto de apenas um nominal; nesse caso *mata* será o verbo, e teremos uma estrutura bem formada, a saber, SN + *nunca* + V. Logo, o verbo é *mata*, e *gripe* é um SN (o sujeito).

O exemplo acima ilustra apenas uma pequena parte das informações que podem ser utilizadas para identificar o verbo em um enunciado. Há certamente muitos caminhos alternativos, mas até hoje ainda não foram sequer listados, quanto mais verificados em casos concretos.

4.2. Fenômenos gramaticais ligados ao sujeito

4.2.1. Concordância

Na seção precedente tratamos do fenômeno da atribuição dos papéis temáticos, o que nos levou a uma conceituação mais concreta de sujeito e objeto. Vamos agora passar ao segundo dos fenômenos que devem ser descritos, a saber, a chamada **concordância**

¹³ Não é preciso levar o termo “prévio” a sério. O que quero dizer é apenas que uma tarefa depende da outra, sem que uma tenha que ser realizada necessariamente **antes** da outra – já que temos acesso simultâneo a todo o nosso conhecimento da língua (ver a seção 5.8).

verbal. Isto é, precisamos explicitar os fatores que marcam como inaceitáveis orações como

[25] * A menina quebramos a janela.

Basta aplicar o sistema proposto para verificar por que essa frase é mal formada: o SN pré-verbal (*a menina*) não pode ser marcado como sujeito porque não é compatível com a pessoa e número indicados pelo sufixo do verbo. Mas a valência de *quebrar* só admite dois SNs na oração se um deles for o sujeito; logo, se *a menina* não pode ser sujeito, vai ficar sem papel temático. E há uma condição que impede a ocorrência de constituintes destituídos de papel temático na oração.¹⁴

No caso de

[26] * Morri um gato.

a explicação é semelhante, levando-se em conta que *morrer* admite um sujeito pós-verbal (na construção ergativa), e nenhum outro SN. Como *um gato*, nessa oração, não pode ser sujeito, fica sem papel temático, e conseqüentemente a oração é mal formada.

A oração

[27] Comi uma pizza.

é sintaticamente semelhante a [26], mas é aceitável. A razão é que o verbo *comer*, ao contrário de *morrer*, ocorre na construção transitiva, e portanto *uma pizza* pode receber o papel de Paciente (será o segundo SN da fórmula dada em [13]).

Ou seja, no caso de [27] é como se houvesse duas análises em competição: numa delas *uma pizza* seria sujeito, na outra seria objeto. Ambas as possibilidades se coadunam com a valência do verbo, mas a segunda análise vence a competição porque é a única compatível com os dados. Note-se que isso requer que todo o mecanismo gramatical esteja disponível simultaneamente, e que o usuário possa lançar mão deste ou daquele recurso conforme for mais conveniente. Explicito esse mecanismo mais detalhadamente na seção 5.8 abaixo.

Essa idéia de processos mentais em competição parece ser muito antiga. O neurofisiologista William H. Calvin relata uma sugestão feita por William James em 1874, no sentido de que

as idéias poderiam “competir” de alguma forma dentro do cérebro, deixando apenas as melhores ou “mais aptas”. Assim como a evolução darwinista modelou um cérebro melhor em 2 milhões de anos, um processo semelhante agindo dentro

¹⁴ Certamente se trata de uma condição de base semântica. Não é exatamente o **critério teta** dos gerativistas, porque este impede a atribuição de mais de um papel temático ao mesmo argumento, o que me parece inevitável em certos casos. Aliás, a formulação usual do critério (Chomsky, 1982: 36, por exemplo) me parece pouco significativa, em vista da ignorância em que nos encontramos a respeito dos papéis temáticos: como, exatamente, podemos saber se um argumento tem um ou mais de um papel temático? Basta lembrar a formulação de Dowty (1991), em que ele propõe a análise dos papéis tradicionais em feixes de “proto-papéis”.

do cérebro modelaria soluções inteligentes para problemas na escala do pensamento e da ação.

[Calvin, 2007: 89]

Essa passagem lembra muito o fenômeno que estamos examinando.

Haveria dificuldade com a análise vista com orações que tivessem sujeito posposto e mais de um SN -- mas, sintomaticamente, tais orações não parecem existir em português. Me refiro a frases como

[28] * Comeu um gato o mingau.

[29] * O mingau comeu o gato. (com *o gato* Agente)

que, no que pese seu uso ocasional na literatura tradicional, são inaceitáveis em português moderno (tanto falado quanto escrito).

Já em

[30] Quebraram a janela.

não aparece a ambigüidade que se observa em

[31] Quebraram alguns pratos.

justamente pela impossibilidade de rotular *a janela* como sujeito, dada a incompatibilidade de pessoa e número com o sufixo verbal. Em [31] a incompatibilidade não existe, e por isso a frase é ambígua: pode ser entendida como uma construção ergativa (= *alguns pratos quebraram*) ou uma transitiva com Agente indeterminado (= *alguém quebrou alguns pratos*).¹⁵

Ou seja, os chamados erros de concordância são na verdade casos de má formação semântica -- mais especificamente, casos de ocorrência de argumentos destituídos de papel temático. Essa análise se harmoniza com o espírito da Hipótese da Sintaxe Simples, dispensando mecanismos específicos para tratar do fenômeno da concordância, e reduzindo-o a simples consequência de outros fatores, independentemente motivados.

4.2.2. Clíticos

Ainda temos um fenômeno a explicar, a saber, a distribuição dos pronomes oblíquos em relação a seus congêneres retos (*eu x me* etc.). Na regra de identificação, coloquei simplesmente uma condição que impede os oblíquos de serem rotulados como sujeito; na prática, isso impede que eles recebam o papel temático correspondente, de

¹⁵ O sujeito indeterminado ainda tem que ser integrado no sistema, certamente como parte das condições de introdução de argumentos não expressos; um primeiro passo nessa direção é Moreira (2005).

modo que em termos descritivos o problema está resolvido. Ninguém negaria que os pronomes oblíquos são formas excepcionais dentro da língua.¹⁶

Mas em algum momento teremos que explicitar também as possibilidades de posicionamento, para evitar frases como

[32] * Toninho ajudou me.¹⁷

[33] * Me ajudou Toninho. (com *me* Agente)

[34] * Me ajudei Toninho. (com *me* Agente)

O caso de [32] precisa ser resolvido através de uma marca idiossincrática dos pronomes oblíquos: ao contrário de todos os outros SNs da língua, eles ocorrem em posições fixas em relação ao verbo.¹⁸ Essa marca é inevitável, e precisa aparecer, de uma maneira ou de outra, em qualquer descrição da língua.

Em [33] e [34] não é possível atribuir a *me* o papel temático associado ao sufixo de PN, a saber, Agente. Disso dá conta a condição (iii) da regra de identificação, que impede que *me* seja sujeito.

Ou seja, os chamados erros de concordância são na verdade casos de má formação semântica. Essa análise se harmoniza com o espírito da Hipótese da Sintaxe Simples, dispensando mecanismos específicos para tratar do fenômeno da concordância, e reduzindo-o a simples consequência de outros fatores independentemente motivados.

4.2.3. Sujeito posposto

Nas frases

[35] Pulou um gato no meu colo.

[36] Caiu uma bandeja.

temos que analisar o SN que vem após o verbo como sujeito, pois ele se comporta de maneira idêntica, no que nos interessa, à dos SNs pré-verbais de

[37] Um gato pulou no meu colo.

[38] Uma bandeja caiu.

¹⁶ Eventualmente, poderemos chegar a uma restrição que impeça os pronomes oblíquos de elaborar o papel temático indicado pelo sufixo de PN. Não podemos discutir essa questão aqui, mas ela pode vir a ser interessante, pois reduziria a restrição às propriedades semânticas de itens como *me*, *se*, *nos*.

¹⁷ Em português brasileiro falado. Não estou considerando aqui os fatos do português padrão ou do europeu, onde é possível uma frase muito próxima dessa, *Toninho ajudou-me*. Para essa variedade, claro, as regras podem ser diferentes.

¹⁸ A saber, sempre antes do verbo principal: *ele me ajudou*, *ele tem me ajudado*, *me ajuda aqui*. Essa é a posição do clítico em brasileiro falado (cf. Perini, 2002, cap. 29).

Sabe-se que a posição do sujeito tem a ver com seu “status” funcional e sua carga informacional, mas não com sua relação temática com o verbo.

Nesses casos, o SN em questão será rotulado como sujeito pela regra de identificação, porque só há um SN na oração. Em [35], por exemplo, o sufixo de PN indica o papel de Agente, e o mesmo papel será atribuído ao SN *um gato*, o que é correto.

Outro fato que confirma que as restrições sobre a posposição do sujeito são independentes da valência do verbo é o seguinte: o sujeito posposto, que como vimos ocorre com verbos ergativos, sofre restrições aparentemente relacionadas com a ocorrência de pronomes pessoais. Assim, se o sujeito for um SN pleno, o sujeito posposto ocorre normalmente. Mas com sujeito pronominal o sujeito posposto é muito menos freqüente, e só é plenamente aceitável se for contrastivo:

[39] Chegou o guitarrista.

[40] Ele chegou.

[41] ?? Chegou ele.

Uma frase como [41] só se enuncia em contextos contrastivos como

[42] O Ricardo disse que ia chegar junto com a Mirtes. Mas só chegou ele.

O mesmo vale para pronomes de primeira pessoa, para *você*, etc. Os itens tradicionalmente chamados “pronomes” mas não “pessoais”, como *isso*, *tudo*, etc. não são sujeitos a essa restrição; assim, [43] e [44] podem ocorrer autonomamente, sem problemas:

[43] Começou tudo de novo.

[44] Chegaram todos?

4.2.4. O que é o sujeito

O sujeito seria, então, na melhor das hipóteses uma relação derivada. Mas eu acho que não é nem isso: é apenas um termo conveniente para resumir “o SN que ocupa determinada posição na oração, e que recebe o mesmo papel temático indicado pelo sufixo de pessoa-número do verbo”. A se aceitar essa análise, uma frase como

[9] Antônio beliscou a menina.

se analisa como uma seqüência de **SN + verbo + SN**, sem marca especial que vincule um desses SNs ao verbo. Naturalmente, cada um dos constituintes tem uma estrutura interna, de modo que o receptor na verdade parte de algo mais parecido com

[45] SN V SN

Antônio	belisc-ou	a menina
[3 ^a p. sing]	[3 ^a p. sing]	[3 ^a p. sing]

Note-se que [45] só inclui fatos inegáveis da língua, a saber, unidades foneticamente representadas e marcas gramaticais diretamente relacionadas com a estrutura observável -- informação de que o receptor dispõe previamente. É claro que falta algo na fórmula [45], a saber, qual dos dois SNs é o que elabora a indicação feita pelo sufixo de pessoa-número, já que ambos os SNs são de terceira pessoa do singular. Isso se faz através da regra de identificação dada acima, isto é, com base na posição dos SNs da oração.

Como sabemos, o papel temático atribuído ao sujeito *Antônio* em [45] é o de Agente. Mas isso não acontece automaticamente em orações da forma SN-V-SN, porque há casos como

[46] Antônio viu a menina.

etc., onde o sujeito não é Agente.¹⁹ Isso, naturalmente, depende da valência do verbo. Assim, uma vez identificado o sujeito, podemos analisar a frase como realização de uma das diáteses do verbo que aparece nela. No caso de [45], o verbo é *beliscar*, e podemos analisar a frase como um exemplo da construção (diátese) transitiva, definida acima. Já [46], com o verbo *ver*, não pode ser exemplo da construção transitiva, porque esse verbo não a conta entre suas diáteses.²⁰

Voltando à concepção de sujeito como “o SN que ocupa determinada posição na oração, e que recebe o mesmo papel temático indicado pelo sufixo de pessoa-número do verbo”, vale a pena fazer algumas observações (em parte repetindo o que já se viu em 1.8.2.1). Dada essa definição, é claro que o próprio sufixo de pessoa-número, embora seja uma sinalização de papel temático, não é o sujeito. O sujeito é o SN que replica esse papel temático, e que como vimos ocorre em determinadas posições da oração, detectáveis a partir das valências verbais.

Isso significa que uma frase como

[27] Comi uma pizza.

não tem sujeito. Com efeito, não há aí nenhum SN que corresponda à definição proposta: o único SN presente é *uma pizza*, mas não é compatível com as marcas de pessoa e número do sufixo. Essa é uma consequência necessária da definição adotada para a função de sujeito, e não me parece particularmente complicada. Decorre, pura e simplesmente, dessa definição.

A definição, evidentemente, pode ser contestada, mas não conheço nenhuma alternativa melhor. Se admitirmos que o próprio sufixo é o sujeito, que função terá o SN? Não podemos atribuir a mesma função ao SN e ao sufixo, chamando ambos de “sujeito”, pois isso obscureceria as grandes diferenças que existem entre os dois: o sufixo é uma

¹⁹ Tem o papel de Experienciador.

²⁰ [46] é exemplo da construção **de apreciação**, que se define como tendo sujeito Experienciador e objeto Causador de Experiência. A “apreciação” (*amar*) e a “percepção” (*ver*) se representam nesta construção de maneira idêntica.

unidade morfológica presa, e sua semântica se limita a sinalizar a diferença de número e de pessoa, ao passo que o SN é uma forma livre com estrutura interna complexa, e semanticamente tem um potencial referencial muito mais amplo do que o do sufixo.

Por outro lado, postular um “sujeito” oculto em [27] é desprezar a diferença que existe entre essa frase e

[47] Eu comi uma pizza.

As duas frases, embora sejam proposicionalmente equivalentes, têm uma diferença formal evidente (uma tem *eu*, a outra não tem). Além disso, são diferentes do ponto de vista informacional, pois [47] é redundante onde [27] não é. O que tanto [27] quanto [47] têm é um Agente representado por “eu”, mas o sujeito, de acordo com nossa definição, não é o Agente, mas um SN que pode, eventualmente, mas nem sempre, receber o papel de Agente.

Não me parece possível escapar, portanto, da conclusão de que a diferença gramatical entre [27] e [47] é que a primeira não tem sujeito, e a segunda tem.

4.2.5. *O que é a concordância*

A análise acima explica o desempenho do usuário da língua, que parte de um sinal material (a “imagem acústica” de Saussure) e, com base em seu conhecimento da língua, estabelece certas relações simbólicas -- no caso, em especial a relação entre os papéis temáticos e os elementos formais aos quais se associam.

Nesse contexto, o sujeito já não pode ser definido como o termo com o qual o verbo concorda. Isso funciona apenas em casos como o de

[2] Antônio beliscou as meninas.

[35] Pulou um gato no meu colo.

onde a relação de concordância é clara. Mas em frases como

[48] Antônio encontrando a menina, ela deve ficar no quarto.

[9] Antônio beliscou a menina.

não há nada que mostre a concordância do verbo com um ou outro dos SNs, e em

[49] Quebrei a janela.

não é possível captar uma relação de concordância entre o verbo e o único SN da oração. Não obstante, em todos esses casos a atribuição do papel temático é clara, o que mostra que o receptor parte de algum outro sinal que não a concordância para realizar essa atribuição.

O fenômeno que chamamos “concordância verbal” é, portanto, um corolário do fato de que um dos papéis temáticos da oração é sinalizado redundantemente em alguns casos. Em particular, não é resultado de um processo de harmonização, ou de transferência de traços (como nas análises gerativas, cf. Haegeman, 1994; Ouhalla, 1999; Napoli, 1993).

Uma saída amplamente adotada é a de lançar mão de constituintes abstratos (categorias vazias, vestígios, sujeito oculto), mas essa hipótese não faz mais do que transferir o problema para uma instância ulterior. Se, por exemplo, existe um sujeito oculto em [49], ele precisa ser suprimido na superfície; e como o receptor só tem acesso à superfície, ele precisa se basear em alguma coisa para restaurar corretamente o sujeito. Ou seja, o problema continua de pé, e vai ter que ser resolvido, inevitavelmente, através da presença do sufixo de pessoa-número. O que faço aqui é simplesmente eliminar o estágio intermediário (a postulação, e posterior supressão, do sujeito oculto), que não traz nenhum benefício nem no plano sintático nem no semântico, além de ser parcial, pois com [9] teremos que imaginar outro recurso que permita ao receptor identificar *Antônio*, e não *a menina*, como sujeito (isto é, no caso, *Agente*).

4.3. Complexidade: fatores semânticos

Como vimos, o fenômeno a que normalmente chamamos “concordância verbal” é essencialmente o resultado de fatores semânticos. Digo “essencialmente” porque parece haver também fatores não-semânticos em jogo, de modo que teríamos aqui um fenômeno gramatical complexo. A existência de tais fenômenos, compostos de fatores heterogêneos convergindo para um resultado único, ou tradicionalmente considerado como único, não tem sido objeto de investigação -- mas a maioria dos modelos gramaticais correntes, como em especial os gerativos, parecem negar sua existência. Como não posso desenvolver a questão no âmbito deste capítulo, vou expor os fatos da maneira mais clara possível, como preliminar ao exame do problema em profundidade.

Vou começar dando alguns argumentos adicionais em favor da análise semântica da concordância, exposta acima.

4.3.1. Nós = eu e ele

Para efeitos de concordância, *nós* é equivalente a *eu e ele* e seqüências semelhantes: *nós chegamos*, *eu e ele chegamos*.

As análises tradicionais, que tratam a concordância como fenômeno formal, precisam descrever esse fato como uma coincidência – ou seja, é necessário um recurso independente para fazer com que *eu e ele* dispare as mesmas marcas de concordância que *nós*.

Na presente análise, entretanto, a identidade de *nós* com *eu e ele* deriva automaticamente da semântica dos elementos: *nós* é definido no léxico como “eu mais outra(s) pessoa(s)”, o que inclui “eu e ele”, que é precisamente a representação semântica do sintagma *eu e ele*. Entendendo a concordância como fenômeno semântico, o efeito idêntico de *nós* e *eu e ele* fica explicado de maneira natural, sem necessidade de

dispositivos *ad hoc*. É fácil verificar que o mesmo raciocínio explica por que SNs compostos do tipo *João e Maria* se comportam, para efeitos de concordância, como plurais.

4.3.2. O núcleo do SN como controlador da concordância

Dentro da perspectiva tradicional, o fato de a palavra que constitui o centro de referência do SN ser igualmente o controlador da concordância verbal é uma coincidência a explicar. Assim, em

[50] Os gatos da vizinha são barulhentos.

é a palavra *gatos* que determina tanto a forma do verbo *são* quanto a do predicativo do sujeito *barulhentos*; a palavra *vizinha*, por exemplo, é irrelevante para ambas as coisas.

Sabemos que há boas razões semânticas para isso, mas a análise tradicional não permite expressá-las. Segundo essa posição, os dois fenômenos, a referência de *gatos* e a relação formal entre *gatos* de um lado e *são / barulhentos* de outro lado precisa ser feita indiretamente, através de uma relação especial que estipule que o centro de referência é também o núcleo formal do SN (o controlador da concordância).

Na análise aqui proposta, tudo decorre do caráter de centro de referência de *gatos*, sem necessidade de explicitar na gramática que esse mesmo elemento “controla” a concordância. Essa relação decorre de dois fatos: o sufixo de pessoa-número do verbo indica um sujeito de terceira pessoa do plural, e o sintagma *os gatos da vizinha* é igualmente marcado como sendo de terceira pessoa do plural.

4.4. Complexidade: fatores sintáticos

Agora vejamos alguns fatores não-semânticos que parecem influenciar a concordância. A meu ver, são fatores locais, de menor importância do que os fatores semânticos -- funcionam, poderíamos dizer, em contextos marcados.

4.4.1. Coletivos

4.4.1.1. Coletivos e plurais

A hipótese de que a concordância é primariamente um processo semântico parece inadequada quando se lida com coletivos. Um SN como *os mecânicos da equipe* é plural para efeitos de concordância, mas pode parecer semanticamente idêntico a *a equipe dos mecânicos*, que é singular.

A isso se pode responder que, embora tais pares de sintagmas possam ser correferentes no que diz respeito aos membros da equipe, eles não são sinônimos, e a diferença pode ser evidenciada em alguns contextos. Por exemplo, quando dizemos

[51] A equipe dos mecânicos ganhou 200 reais.

é necessário entender que 200 reais foi a quantia total envolvida, ao passo que em

[52] Os mecânicos da equipe ganharam 200 reais.

pode-se entender não apenas que eles ganharam 200 reais para todo o grupo, mas também (e preferivelmente) que ganharam 200 reais cada um (este é o significado **distributivo**).

A diferença semântica aparece também nas possibilidades posicionais de certos itens que têm conexão com o significado distributivo, como *todos* e *cada*. Podemos dizer

[53] Os mecânicos da equipe ganharam todos 200 reais.

[54] Os mecânicos da equipe ganharam 200 reais cada um.

Mas se o sujeito for um coletivo, o significado distributivo já não é disponível, e *todos* e *cada* não podem ocorrer como acima:

[55] * A equipe dos mecânicos ganhou / ganharam todos 200 reais.

[56] * A equipe dos mecânicos ganhou 200 reais cada um.

Isso mostra que os coletivos não são semanticamente idênticos a sintagmas no plural, mesmo quando são extensionalmente equivalentes. A equivalência extensional não é co-referência completa, porque o coletivo tem um significado mais rico: um sintagma como *os mecânicos da equipe* se refere a um conjunto de pessoas; *a equipe dos mecânicos* se refere ao mesmo conjunto, mas também a algo mais, a equipe propriamente dita. Isso é provavelmente resultado do que os cognitivistas chamam **figuração** (*imagery*):

[...] a gramática tem uma função “figuracional”, e [...] grande parte dela tem caráter figurativo. A gramática (como o léxico) incorpora **figurações** convencionais. Com isso quero dizer que ela estrutura uma cena de certa maneira para efeitos de expressão lingüística, enfatizando algumas facetas em detrimento de outras, encarando-a de um certo ponto de vista, ou entendendo-a em termos de alguma metáfora. Sustento que duas sentenças aproximadamente sinônimas, com as mesmas palavras de conteúdo, mas estruturas gramaticais diferentes [...] são semanticamente distintas por causa de suas organizações gramaticais diferentes.

[Langacker, 1987: 39]

O que Langacker diz acerca das sentenças se transfere facilmente para o SN *os mecânicos da equipe* frente a *a equipe dos mecânicos*: o primeiro enfatiza (figura) o caráter plural de uma equipe, que é necessariamente composta de várias pessoas. Já o segundo figura seu caráter unitário. Os exemplos acima mostram que a figuração é relevante para determinar as condições de boa formação das frases.

A conclusão é que há uma razão semântica para considerar *a equipe dos mecânicos* um singular para efeitos de concordância. Obviamente, a forma sintática do sintagma contribui para isso, porque *equipe*, núcleo do SN, é morfologicamente singular. Mas não se pode dizer que o SN é singular meramente porque seu núcleo é singular; é também semanticamente singular em um sentido em que *os mecânicos da equipe* não é.

4.4.1.2. Silepse

O fato de *a equipe dos mecânicos* ser singular para efeitos de concordância ilustra a complexidade do fenômeno, que não pode ser rotulado como simplesmente sintático nem semântico. Como vimos, a referência desse sintagma é dupla, referindo-se ao mesmo tempo a várias pessoas e a um conjunto visto como um todo.

Seria de esperar, então, que o verbo fosse livre para ocorrer tanto no singular quanto no plural. Essa possibilidade é, de fato, admitida pelas gramáticas, pelo menos em certos casos; diz-se que frases como [57] ocorrem, sendo chamadas casos de **silepse**:

[57] A equipe dos mecânicos lotaram a sala.

O exemplo [57] seria provavelmente rejeitado pela quase totalidade dos falantes, mas parece que existem exemplos em que a silepse dá resultados aceitáveis. Nos exemplos geralmente citados o que está envolvido não é o número, mas a pessoa. Assim, lê-se às vezes

[58] Os brasileiros tendemos a desconfiar das autoridades.

Tais frases não são comuns, e só parecem ocorrer em textos escritos de natureza especialmente formal. Mas é interessante ver que não são totalmente desconhecidas, sendo mencionadas em todas as gramáticas. Sua eventual aceitabilidade não pode ser explicada pela análise tradicional da concordância: não haveria modo de explicar por que o sufixo é *-mos* quando não há na oração nenhum SN de primeira pessoa do plural.

Por outro lado, de acordo com minha hipótese, a frase [58] deixaria de ser filtrada desde que ficasse claro que quem fala é brasileiro, e portanto *os brasileiros* inclui o falante em seu conjunto referencial. Assim, pode-se entender que o sufixo e o sujeito indicam a mesma entidade, e não há contradição na frase. Prefiro, porém, não levar muito adiante esse argumento, porque frases silépticas são certamente marginais.

4.4.2. Outros fatores não-semânticos

4.4.2.1. Marcados e não-marcados

Além dos exemplos vistos acima, com sujeito coletivo, há outros casos em que a concordância parece depender de fatores não-semânticos. Alguns deles se relacionam com o caráter **não marcado** dos sufixos de terceira pessoa do singular. Será necessário,

portanto, considerar essa noção de “não-marcado”, corriqueiramente utilizada mas nem sempre claramente definida.

Sem entrar em uma discussão do que vem a ser “marcação” (*markedness*) em geral, vou fazer apenas a seguinte observação: os sufixos de primeira pessoa ocorrem em uma situação facilmente definível – ou seja, quando o sujeito é marcado como de primeira pessoa, o que significa que é *eu*, *nós* ou alguma combinação que inclua *eu* ou *nós*, ou então sem sujeito – ao passo que os sufixos de terceira pessoa ocorrem em uma grande variedade de ambientes, o que indica que são os que aparecem quando não há nenhuma razão especial para que apareçam os de primeira pessoa. Por exemplo, a terceira pessoa ocorre com todos os SNs da língua, exceto os que incluam *eu* ou *nós*: *ela*, *o senhor*, *Nagib*, *meus amigos paulistas*, etc.²¹ Ocorre também quando o verbo não remete a nenhum argumento que tenha um papel temático específico, como em frases “impessoais”:

[59] Choveu um pouco no dia 18.

[60] Está cheio de cupins nesse armário.

A terceira pessoa também ocorre quando o sujeito é representado por uma oração:

[61] Você fumar desse jeito vai acabar te matando.

Ou seja, a terceira pessoa é compatível com uma variedade muito grande de sujeitos – é o caso geral, ou seja, **não-marcado**. Essa concepção da terceira pessoa como caso não-marcado não é novidade, e já se encontra com bastante clareza em Jespersen (1924: 212) e outros autores antigos. Dentre os autores que trabalharam com o português, pode-se citar Martin (1975):

[...] a forma que é tradicionalmente chamada de ‘terceira pessoa do singular’ [...] nada mais é [...] que o predicado que porta um valor de T[empo] sem outra marcação qualquer.

[Martin, 1975: 58]

No entanto, a análise de Martin considera a concordância como um processo autônomo, ou seja, ele adota a posição tradicional. Apenas acrescenta que a concordância se processa nos casos chamados de primeira e segunda pessoa, mas não nos casos de terceira (do singular).

No contexto da presente análise, temos que ver as coisas de outro ângulo, já que estou sustentando que o verbo nunca realmente “concorda” com o sujeito. Para nós, a ocorrência ampla do sufixo de terceira pessoa decorre de sua própria complexidade referencial: enquanto o de primeira pessoa indica o papel temático de um conjunto facilmente delimitado de entidades (todas as que incluem “eu”), o de terceira pessoa se refere a qualquer conjunto de entidades não definido como de primeira pessoa.

²¹ Os sufixos de segunda pessoa, nas variedades em que ocorrem, são semelhantes aos de primeira, ocorrendo apenas com *tu*, *vós* e combinações incluindo esses pronomes.

É importante notar que isso vale para o sufixo de terceira pessoa **do singular**. O de terceira pessoa do plural tem complexidades próprias, e além de se referir a entidades plurais que não incluem “eu” ou “nós”, ele também se refere ao Agente não especificado, como em

[62] Depredaram o meu carro.

4.4.2.2. Impessoais

Agora vamos examinar as frases

[63] Há muitas praias na Bahia.

[64] Existem muitas praias na Bahia.

Essas frases são basicamente sinônimas, mas em [63] o verbo está no singular e em [64] está no plural. Aqui, acredito, nenhum fator semântico justifica a forma dos verbos: em termos tradicionais, como nos termos desta análise, [64] tem sujeito e [63] não tem. Temos que analisar as duas frases como realizações de duas construções distintas, que chamo **de apresentação de existência e de apresentação de existência com concordância**, respectivamente. Trata-se de construções de extensão muito limitada, compostas de um SN (sujeito no caso de [64]) com o papel temático de Apresentando, isto é, o elemento cuja existência é asserida. A primeira construção vale só para o verbo *ter* e, na língua escrita, *haver*. É interessante notar que se verifica uma tendência a reduzir a primeira à segunda, dizendo por exemplo *tinham dois atacantes na frente do goleiro*. Neste caso, portanto, a concordância é resultado de marca formal, e não decorre de fatores semânticos.

4.4.2.3. *Nós = a gente*

Um fenômeno difícil de analisar em termos semânticos é a ocorrência como sujeitos dos itens *nós* e *a gente*, um dos quais elabora informação de um sufixo de primeira pessoa do plural, e o outro de um de terceira do singular. Por exemplo,

[65] Nós já vimos esse filme.

[66] A gente já viu esse filme.

Pelo menos em termos de referência, *nós* e *a gente* se equivalem, e ambos incluem em sua extensão o falante, mais pelo menos uma pessoa. Ao que tudo indica, isso nos impede de descrever semanticamente essa diferença de sufixos. Por conseguinte, será necessário marcar o item *a gente* como uma exceção – embora seja semanticamente “primeira pessoa”, é formalmente “terceira”. Talvez haja analogia com casos como

[67] Agora a mamãe vai descansar.

com a diferença de que *a mamãe* só excepcionalmente inclui o falante, ao passo que *a gente* o faz normalmente. De qualquer forma, aqui temos algo ainda a investigar, e deve ser difícil escapar de alguma marca formal interferindo no fenômeno da concordância.

4.4.2.4. Casos de “atração”

Restam alguns casos pouco estudados (embora bastante freqüentes) como

[68] A maioria dos comentaristas foi / foram favoráveis ao treinador.

[69] Um dos analistas que trabalhou / trabalharam no estudo...

onde parece haver uma espécie de conflito entre o critério semântico e o formal. Nessa frase o sintagma sujeito é formalmente singular; mas semanticamente plural (ou coletivo), o que ocasiona a dualidade de formas do verbo -- ambas comuns, e inclusive admitidas pela norma. Já em

[70] A solução desses problemas graves e urgentes tiveram que esperar o novo governo.

parece que o que funciona é algum princípio de proximidade, talvez aliado ao caráter mais marcado, e portanto cognitivamente mais saliente, do plural em comparação com o singular. Digo isso porque se invertermos os números o efeito parece desaparecer:

[71] As fábricas da cidade são / ?? é de calçados.

Nesses casos é arriscado confiar em julgamentos pessoais; certamente a aceitabilidade de frases como [71] merece alguma verificação, nem que seja através de testagens. Já frases como [68], [69] e [70] são amplamente atestadas.²²

4.5. Como definir as construções: a notação ‘X’

Nas seções precedentes me estendi bastante propondo uma análise da concordância verbal que difere marcadamente da tradicional, a saber, a hipótese de que os fenômenos chamados de “concordância (verbal)” não são o resultado de um processo gramatical homogêneo, mas antes a convergência de diversos fatores bastante distintos, alguns sintáticos, outros semânticos.

Embora eu tenha chegado a essa proposta independentemente, descobri depois que não é totalmente nova. Uma análise bastante parecida foi proposta por Lapointe (1980). No entanto, esse autor a abandonou, ao que parece em razão de uma tentativa de formular uma análise universal da concordância (Lapointe, 1988). Reid (1991) também

²² Há um estudo a respeito, de Elizabeth Saraiva, realizado nos anos 70 na UFMG.

fez uma proposta semelhante para o inglês. Meu compromisso é com o português, e para essa língua estou convencido de que essa é a melhor análise.

As conseqüências para a conceituação de sujeito são que essa função, tradicionalmente analisada como puramente sintática, é na verdade uma relação de ordem simbólica. O sujeito, segundo essa hipótese, é o SN da oração que é indicado na diátese verbal como devendo receber o mesmo papel temático já indicado, redundantemente, pelo sufixo de pessoa-número.²³

Essa análise tem conseqüências para a maneira como vou descrever as diáteses verbais. Até o momento, as descrições incluíram o SN sujeito, colocado arbitrariamente na posição pré-verbal -- que, como sabemos, é apenas uma das posições possíveis. Assim, defini as diáteses segundo o modelo seguinte:

[72] Construção transitiva

(SN-sujeito)	V-SPN	SN
Agente	Agente	Paciente

[73] Construção ergativa

SN-sujeito	V-SPN
Paciente	Paciente

No entanto, vimos que: (a) o sujeito pode ocorrer posposto; (b) o sujeito pode não ocorrer; e (c) o sufixo de pessoa-número ('SPN' nas fórmulas acima) pode não aparecer. Para exprimir esses fatos de maneira mais eficiente, vou mudar o formalismo, introduzindo uma variável 'X', que deve ser lida como "sufixo de pessoa-número e/ou SN rotulado como "sujeito" pela regra de identificação", ou seja, a regra dada na seção 4.1.5.

Agora as construções acima podem ser definidas assim:

[74] Construção transitiva

X	V	SN
Agente		Paciente

[75] Construção ergativa

X	V
Paciente	

Tomando [75] como exemplo: entende-se que o Paciente é atribuído ao sufixo de pessoa-número e/ou ao SN indicado como sujeito. O fato de o papel temático ser

²³ Com as ressalvas vistas para casos de orações sem sujeito ou com formas verbais não flexionadas.

Paciente nesse caso é governado pelo verbo, isto é, [75] é um ambiente simbólico onde ocorrem os verbos que têm a construção ergativa em sua valência.

Uma nota final sobre a convenção “X”: uma fórmula como **X V** representa, na verdade, três estruturas, a saber

SN	V-SPN	(<i>o gato engordou</i>)
SN	V	(<i>o gato engordando...</i>)
V-SPN		(<i>engordei</i>)

Como as três definem exatamente o mesmo recorte dentro do universo dos verbos da língua (ou seja, qualquer verbo que ocorre em uma delas ocorre também, necessariamente, nas outras duas), poderíamos considerar apenas uma delas como sendo “a” diátese, deixando as outras por conta de regras gramaticais – tal como fazemos com as construções topicalizada, negativa, interrogativa etc. Essa seria a solução mais coerente. No entanto, vou preferir a notação “X” por causa da importância do sujeito dentro da oração; dessa maneira, ele fica marcado como um SN com propriedades particulares, em oposição aos demais SNs da oração.

Há ainda estruturas invertidas como

V	V-SPN	(<i>chegou um gato</i>) etc.
---	-------	--------------------------------

cujas condições de ocorrência são mal conhecidas, mas parecem ter a ver com a estrutura informacional da frase, e não com a subclasse do verbo.

Capítulo 5: Outras funções sintáticas

5.1. Objeto direto

5.1.1. Problemas de definição

De todas as funções sintáticas tradicionais, o objeto direto é provavelmente a de conceituação menos satisfatória. Sua definição apresenta problemas tanto na face formal quanto na semântica; e embora haja bastante consenso quanto a que sintagma em cada oração deve ser analisado como objeto direto, a base para essa decisão continua obscura. Os artigos reunidos na coletânea de Plank (1984), assim como a resenha de Vilela (1992), apresentam uma variedade de propostas que reflete a desorientação reinante.

Em geral se procura uma definição universal de objeto direto, mas essa tentativa me parece prematura. Sustento que se deve partir, previamente, do exame do problema em línguas particulares, de modo a permitir uma comparação fundamentada entre o que se chama “objeto direto” em uma língua ou em outra. Podem ser coisas bem diferentes: por exemplo, em latim o objeto direto (tradicional) é um termo marcado pelo caso acusativo, e sua posição é bastante livre na oração, embora talvez a mais comum seja depois do sujeito e antes do verbo. Em português, como veremos, o objeto direto se define posicionalmente, e a marca de caso é marginal, só aparecendo em alguns pronomes (*me, te, nos, se*). Em náhuatl, o termo que exprime o Paciente de certos verbos, e que seria presumivelmente analisado como objeto direto, é um dos participantes da concordância verbal.

As comparações interlingüísticas, em geral, se fazem com base no papel temático. Assim, ao se comparar o objeto direto na língua A e na língua B, muitas vezes o que realmente se está fazendo é comparar os recursos que A e B possuem para exprimir funções semânticas como Paciente – o que definitivamente não é a mesma coisa.

A abordagem gerativista oferece uma definição formal de objeto direto. Por exemplo, Haegeman observa que

[...] a função gramatical de “sujeito” não é uma noção **primitiva** na teoria da Regência e Ligação: não é um conceito simples ou inalisável da teoria. Ao contrário, “sujeito” (assim como “objeto”) se define em termos de configurações sintáticas; é uma noção derivada baseada em conceitos mais elementares da teoria.

[Haegeman, 1991: 105]

O sujeito, segundo essa teoria, é o “SN imediatamente dominado pelo IP¹, isto é, o SN na posição de especificador de IP”; e o objeto é o SN imediatamente dominado por V-barra. Essas relações podem ser representadas em uma árvore sintagmática do tipo bem conhecido.

¹ “Inflectional phrase”, correspondente ao que outras teorias denominam “oração”; “V-barra” é o que outros chamam de sintagma verbal.

O problema é que a base para a definição dessas funções não é a seqüência concretamente observada, mas configurações sintáticas abstratas. Uma árvore representa alguma coisa relacionada com os dados, mas ela própria não é parte dos dados. Quero dizer que o receptor não pode partir da árvore; esta precisa ser construída por ele com base em informações disponíveis no sinal fonético recebido, processado através de seu conhecimento prévio da estrutura da língua e de outros eventuais aspectos da realidade. Em algum momento, pois, será necessário estabelecer relação entre a configuração mostrada na árvore e esses recursos materiais (o sinal fonético) e cognitivos (conhecimentos prévios). Portanto, a árvore, se necessária, também não pode ser tomada como um elemento primitivo; e a posição gerativa deixa uma lacuna entre a análise e os dados.

A gramática funcional (na variedade de Dik, 1980) distingue três níveis de análise, a saber: os papéis temáticos (agente, meta, receptor); as funções sintáticas (sujeito, objeto); e a ordem dos constituintes (estabelecida, esta, através de **regras de expressão**). Isso se deduz com facilidade da exposição de Dik (1980: 13 e 19-20), onde fica claro que **sujeito** e **objeto** não são nem ingredientes semânticos nem designações de posições na sentença.

Vou tentar mostrar que, para efeitos de definição do objeto direto, é possível lidar com apenas dois desses níveis, a saber, o semântico e o de ordenação – ao qual podemos chamar **sintático**, entendendo-se contudo que a sintaxe não se resume à especificação da ordem dos constituintes.

Um ponto de partida possível são as perguntas seguintes:

- (a) Quais são as bases formais para distinguir o objeto direto dos outros constituintes da oração, e em particular dos que são realizados por sintagmas nominais?
- (b) Que papéis temáticos podem ser expressos através do objeto direto?

A segunda pergunta só pode ser respondida a partir da resposta dada à primeira. Acho importante essa observação porque, primeiro, a caracterização formal é específica de cada língua, e como já disse é preciso iniciar o estudo do objeto direto pelas línguas particulares. E, depois, mantenho o princípio metodológico básico de descrever separadamente (em um primeiro momento) os aspectos formal e semântico, para só então relacioná-los em estruturas simbólicas.

Como resultado da pesquisa aqui relatada, entretanto, vou acabar chegando a uma reformulação das perguntas acima, de modo que meu ponto de chegada será uma resposta à pergunta seguinte:

- (c) Quais são as bases formais para atribuir cada um dos papéis temáticos usualmente associados à função de objeto direto?

Ou seja, chego a uma relação direta entre bases formais (no caso, a ordem dos constituintes) e o papel temático, sem o intermediário da “função sintática” concebida como uma relação abstrata entre constituintes. Ao contrário do sujeito, que estudamos no capítulo precedente, o objeto direto não tem uma relação especial com outro elemento da

oração – é apenas um SN em determinada posição. Isso nos permitirá dispensar totalmente a noção de objeto direto enquanto função gramatical.

Um aspecto importante do objeto direto, mas que não vai ser levado em conta aqui, é o da chamada “incorporação”. A incorporação ocorre quando a relação entre o objeto direto e o verbo é especialmente estreita, do ponto de vista semântico e morfossintático. Em algumas línguas, como em maia de Yucatan, o objeto se integra morfologicamente em um verbo transitivo, e o conjunto funciona como um verbo intransitivo:

Os verbos transitivos podem passar para a forma intransitiva quando o significado combinado do verbo e de seu objeto representam uma ação habitual:

[Tozzer, 1977: 65]

Em outras línguas, a incorporação não chega a esse ponto, mas ainda assim tem conseqüências gramaticais detectáveis, conforme foi mostrado por Saraiva (1997) para o português. Saraiva estudou construções como *alugar apartamento*, que se comporta diferentemente de *alugar o apartamento*: no primeiro caso, por exemplo, não é possível topicalizar o objeto, e há restrições quanto aos elementos que podem ser colocados entre o objeto e o verbo, ao passo que essas restrições não funcionam para *alugar o apartamento*, *alugar este apartamento* etc.

No que diz respeito aos objetivos deste estudo, o objeto incorporado do português não se distingue do não-incorporado porque os papéis temáticos são os mesmos, e a topicalização e a possibilidade de inserir adjuntos não são relevantes para a definição das valências.

5.1.2. Semântica do objeto direto: papéis temáticos

Já foi mostrado que é possível identificar o sujeito utilizando apenas a ordem dos constituintes, sua categoria (SN ou não) e a valência do verbo da oração; ora, uma vez identificado o sujeito, é possível rotular corretamente outros eventuais SNs da oração, atribuindo-lhes os papéis temáticos de acordo com sua ordenação e a valência do verbo. Ou seja, até onde se pode ver, seria supérfluo marcar esses SNs com funções abstratas, já que a atribuição do papel temático de cada um, objetivo final da análise, pode ser feita em termos de informações concretas (ordenação), mais outras informações que precisam ser, por razões independentes, incluídas no equipamento cognitivo do usuário da língua.

Para dar um exemplo, os dois SNs pós-verbais da frase

[1] Daniela considera [Ronaldo]_{SN} [o maior jogador do mundo]_{SN}

recebem seus respectivos papéis temáticos com base em sua ordenação relativa (*Ronaldo* antes de *o maior jogador do mundo*) e em seu conteúdo semântico (*o maior jogador do mundo* tem potencial qualificativo, *Ronaldo* não tem). Não há portanto necessidade de rotular esses dois SNs através de funções abstratas (“objeto direto”, “predicativo do objeto”), já que é possível atribuir a eles os papéis temáticos a partir de informações concretas de caráter sintático (ordenação) e semântico (potencial qualificativo). Ou seja, a

seguirmos a Hipótese da Sintaxe Simples, não há necessidade de usar rótulos funcionais abstratos nessa frase. A questão é se isso vale para as frases do português em geral; assim, vamos voltar ao problema de como caracterizar o que se chama tradicionalmente “objeto direto”.

A base semântica a partir da qual se computa o significado final de um objeto direto não precisa ser necessariamente uniforme. As tentativas de estabelecer um significado único “básico” para o objeto direto não são muito convincentes. Sanders (1984) propõe o seguinte:

a função semântica prototípica dos objetos diretos é exprimir **os argumentos relativamente menos ativos, menos controladores, ou de menor iniciativa dos predicados de dois lugares.**

[Sanders, 1984: 224-225]

Mas mesmo Sanders lista imediatamente algumas exceções, como os verbos *receive* ‘receber’, *hear* ‘ouvir’, onde o objeto direto não é perceptivelmente menos ativo do que o sujeito. Prefiro portanto manter provisoriamente a posição de que cada verbo, ou cada grupo de verbos semanticamente aparentados, estipula a relação semântica exata que tem com seu objeto direto. É impossível “quebrar” alguma coisa sem mudar seu estado (se o estado não mudar, não houve “quebra”), ao passo que é possível “tocar” alguma coisa sem mudar o estado dessa coisa.

5.1.3. Comportamento sintático

Apesar da grande variedade de papéis temáticos observada com os complementos geralmente analisados como objeto direto, seu comportamento sintático, formulável essencialmente em termos de posição, é relativamente simples. Isso já foi observado por Sanders, que comenta que

Todas as variedades de objetos [...] têm essencialmente a mesma gama de posições de ocorrência nas sentenças de uma língua, apesar de sua ampla variedade de papel ou função semântica.

[Sanders, 1984: 232]

A estrutura interna dos objetos diretos é igualmente consistente, sendo composta sempre de um SN. Sanders contrasta a situação do objeto com a do adjunto adverbial, que além de ser semanticamente complexo, apresenta excessiva variedade de estrutura interna e de comportamento sintático (a razão, a meu ver, é que o que chamamos de “adjunto adverbial” é uma falsa função, composta na verdade de diversas funções distintas).

A relativa consistência sintática do objeto direto pode sugerir que se trata de uma função única. No entanto, a posição que adoto me obriga a distinguir pelo menos duas funções, ambas tradicionalmente chamadas de objeto direto, a saber, a pós-verbal, exemplificada por

[2] Patrícia vendeu o cachorro.

e a inicial, exemplificada por

[3] O cachorro, Patrícia vendeu.

Do ponto de vista aqui adotado, não é possível atribuir a mesma função a *o cachorro* nas duas frases, porque sua ordenação difere. O que *o cachorro* tem em comum nas duas frases acima é sua composição interna (SN) e seu papel temático (Paciente). Mas nenhum desses fatores é relevante para a definição da função: a composição interna determina a **classe** do sintagma (não sua **função**), e o papel temático, sendo um traço semântico, é automaticamente irrelevante para a caracterização da função sintática (isto é, formal). O que temos, portanto, são duas funções formais, que se assemelham por serem preenchidas pela mesma classe de formas e por receberem o mesmo papel temático.² Sua semelhança se coloca no espaço simbólico, mas do ponto de vista formal elas são nitidamente diferentes.

5.1.4. O objeto direto como relação simbólica

Limitando-nos ao verbo *comer* e seu Paciente, vimos que esse papel temático é atribuído ao constituinte que seja semanticamente apto a recebê-lo e que esteja convenientemente posicionado na oração. A aptidão semântica se cifra na possibilidade de ser usado referencialmente; é o que *hoje*, *Fernando* e *o tio de Fernando* podem fazer, mas *atualmente* e *com muito gosto* não podem. E a posição conveniente é determinada pelas diáteses associadas ao verbo – no caso, *comer* se associa às diáteses **transitiva** e **transitiva de objeto elíptico**, além da construção topicalizada; esta não é considerada diátese porque não subcategoriza o verbo, já que vale para qualquer verbo.

O papel de Paciente de *comer* é portanto atribuído ao SN em posição pós-verbal ou ao SN colocado no início do período. O resultado é que *o sanduíche* será Paciente nas frases

[4] A menina comeu o sanduíche.

[5] O sanduíche, a menina comeu.

Note-se que com isso conseguimos atribuir o papel temático ao SN não sujeito sem necessidade de rotulá-lo segundo uma função sintática abstrata – sua função sintática se limita a sua posição na seqüência.

Como explicar, agora, a impossibilidade de ocorrência de um SN pós-verbal ao lado de um antes do sujeito, ou seja, por que a frase seguinte é inaceitável?

[6] * O sanduíche, a menina comeu as batatas.

² Essas duas características, na verdade, se reduzem a uma só, pois os papéis associados a essas funções são tipicamente desempenhados por SNs: Paciente, por exemplo.

A explicação tradicional é de que uma oração não comporta dois objetos diretos. Mas essa é uma restrição *ad hoc*, não tendo motivação independente. Agora podemos dar uma explicação mais fundamentada: dada a valência do verbo *comer*, tanto o SN pós-verbal quanto o inicial teriam que receber o papel de Paciente, mas isso resultaria em uma diátese inexistente.

5.1.5. O que é um objeto direto?

5.1.5.1. Regras de atribuição do papel Paciente

Como fica a nossa análise, então? Vamos admitir que partimos de estruturas formais já montadas (orações). A montagem dessas estruturas pode ser tarefa de regras de estrutura sintagmática ou outros recursos sintáticos, mas certamente pelo menos parte do trabalho será realizado por regras de boa formação semântica. Neste trabalho não tentarei entrar em detalhes a respeito; me contento com uma lista de construções, que certamente não é a maneira mais econômica de representar as possibilidades de estruturação sintática, mas que é conveniente como ponto de partida, principalmente por sua relativa neutralidade, sendo compatível com uma grande gama de teorias diferentes. Vamos passar diretamente à atribuição de papéis temáticos.

Dada a lista de diáteses de cada verbo, a atribuição dos papéis temáticos de uma oração se processa através do esquadrinhamento da lista, depois de localizado o verbo e os SNs presentes, e tomando em consideração a valência do verbo. No caso de

[7] Sônia comeu uma pizza.

esse processo nos dá a análise preliminar seguinte:

[8] Sônia comeu uma pizza
 SN V SN

Como *comer* pode ocorrer na construção transitiva, o papel de Agente será atribuído ao SN pré-verbal, *Sônia*; e o de Paciente ao SN pós-verbal, *uma pizza*.

Em certos casos o SN pós-verbal pode receber o papel de Agente; mas isso nunca acontece quando o verbo é *comer*, nem quando há outro SN na oração. Pode acontecer em

[9] Pulou um menino no meu colo.

Agente

Esses casos de Agente pós-verbal ocorrem com verbos característicos (monovalentes) em estruturas características, em especial estruturas com um só SN. O fato é que nunca resultam em ambigüidade do papel temático do SN. Só conheço casos de ambigüidade na língua padrão tradicional, mesmo assim só em textos bastante

antigos.³ No português brasileiro falado, que é o que nos interessa aqui, a atribuição dos papéis de Agente e Paciente é sempre unívoca: o sistema de atribuição de papéis temáticos pode ser complexo, mas funciona.

Um traço importante da formulação das construções é que ela só precisa levar em conta a classe de certas formas (SN, a subclasse do verbo) e a posição relativa dessas formas (logo antes do verbo, depois do verbo). Não é necessário mencionar a “função sintática” de nenhum dos constituintes, se por “função” se entende outra coisa que não apenas a posição linear do constituinte em relação ao verbo. Ou seja, essa análise dispensa a noção de objeto direto enquanto função sintática autônoma. E ela é capaz de atribuir o papel temático corretamente aos constituintes usualmente chamados de objeto direto sem lançar mão da noção de objeto direto propriamente dita.

5.1.5.2. A representação sintática do Paciente

Sabemos que o Paciente é considerado o papel temático “típico” do chamado objeto direto. A representação do Paciente nas diáteses do português é relativamente simples. São conhecidas, até o momento, 21 diáteses que incluem um Paciente.⁴ Dessas, o Paciente é o SN pós-verbal em 11 diáteses; é o sujeito em 2 diáteses; é um SPrep em 6 diáteses (as preposições que ocorrem são *de*, *em*, *com*); e um caso especial, C21, parece ter dois Pacientes, um SN pós-verbal, outro com preposição (*a*, *em*, *de*, *com*). Isso significa que o emissor tem poucas opções quando quer exprimir um Paciente.

Agora, considerando o problema do lado do receptor, vemos que o SN pós-verbal, associa-se a uma grande variedade de papéis temáticos, a saber: Paciente (as mesmas 11 diáteses mencionadas acima); Tema (8 diáteses); Meta (5 diáteses); Causador de Experiência (3 diáteses); e Possuído, α Ref, Medida, PRS, Fonte, Possuidor e Apresentando (1 diátese cada).⁵

Isso pode parecer uma dificuldade para o receptor, mas temos que lembrar que o verbo está sempre presente. Assim, digamos que a frase a ser interpretada seja

[10] A professora viu o gato.

O receptor sabe que o verbo *ver* pode ocorrer na diátese de Apreciação, que se define como

[11] Construção de Apreciação:
 X V SN
 Exp CausExp

³ Por exemplo, “essa exposição ou relatório publicou Cândido Mendes de Almeida” (Varnhagen, 1854) onde *Cândido Mendes de Almeida* é Agente, mas isso só se percebe porque o oposto seria implausível. Esses casos são raros, mesmo nesse tipo de texto. Outro estilo em que ocorrem sujeitos pospostos mesmo na presença de objeto é em narrações esportivas: “afasta o perigo o zagueiro do Atlético”.

⁴ Não incluí na lista os casos em que o papel de Paciente é atribuído a um complemento nominal, como em *a destruição da ponte*. Excluo também a construção passiva, pelos motivos explicitados em 9.6.3.

⁵ Ver a definição desses papéis temáticos no capítulo 8. Aqui só nos interessa o número de papéis associados ao SN pós-verbal.

Assim, não terá dificuldade em atribuir o papel temático de Causador de Experiência ao SN pós-verbal.

Nunca é demais repetir que a informação contida em [11] não foi incluída na gramática (ou no léxico) meramente para possibilitar a atribuição correta dos papéis temáticos em frases como [10]. Trata-se, ao contrário, de informações ligadas ao verbo *ver* das quais não podemos escapar, a saber, que pode ocorrer com SN pós-verbal e que, nesse caso, o SN pós-verbal tem o papel de Causador de Experiência. Quem não detém essas duas informações não conhece suficientemente o verbo *ver* para usá-lo corretamente. Aqui o que estamos fazendo é apenas utilizar essas informações, independentemente necessárias, para atribuir o papel temático adequado a *o gato* em [10]. E para isso não é preciso rotular esse SN de “objeto direto”.

Note-se que é necessário identificar o verbo e usar suas diáteses para levar a efeito essa tarefa. É verdade que um SN pós-verbal é, com muita frequência, interpretado como Paciente, mas essa informação não tem utilidade em casos como o de [10] e muitos outros. O receptor não pode confiar na hipótese de o SN pós-verbal ser **sempre** um Paciente, porque o risco de erro é excessivo; não sei exatamente qual seria em termos de texto corrido, mas em termos de número de diáteses⁶ apenas cerca de um terço das que incluem um SN pós-verbal atribuem a ele o papel de Paciente (11 em 34).

O resultado é um sistema bastante complexo, envolvendo o conhecimento e manipulação de dezenas, talvez centenas, de diáteses, cada uma associada a uma lista de verbos. Essa complexidade pode assustar, mas temos que lembrar que a simplicidade de certas alternativas disponíveis é ilusória. C&J chamam repetidamente a atenção para esse fato: aplicar o critério de simplicidade (a chamada “navalha de Occam”) localmente é usar de uma contabilidade falsa; o que se simplifica em um setor é pago com complexidade em outro. A meu ver, creio que aplicar sistematicamente a navalha de Occam às descrições lingüísticas é prematuro, justamente porque não dispomos de descrições abrangentes que nos permitam ter uma visão geral do fenômeno. Nessa dúvida, opto por salvar as aparências, utilizando o mais possível noções concretas como a de ordem, seguindo um princípio informal assim expresso por C&J:

Dada a prioridade epistemológica da ordenação linear – ela é imediatamente disponível ao aprendiz em uma medida em que a estrutura não é – parece-nos que a abordagem natural seria ver quanta quilometragem se pode tirar da ordenação linear [...]

[C&J: 52]

A proposta que acabo de apresentar resolve (ou, talvez melhor, dissolve) o problema de “quais são os papéis temáticos do objeto direto (do sujeito...)”, porque associa diretamente o papel temático à posição do SN em função da valência do verbo da oração. Já não precisamos procurar o que há de comum entre o papel temático do objeto de *comer* e o de *tocar*, o de *criar*, o de *ver* etc., como faz Sanders (1984) na passagem citada acima. Se houver (como talvez haja) duzentos papéis temáticos para duzentos verbos, o modelo dá conta da descrição dos fatos sem dificuldade.⁷

⁶ Limitando-nos às que constam da lista, evidentemente. Há muitas outras.

⁷ Embora talvez um tanto prolixamente.

Isso não quer dizer, obviamente, que não valha a pena procurar as generalizações; mas permite que a descrição seja elaborada sem ficar à espera dos resultados dessa procura. Ou seja, a conveniência da proposta é basicamente descritiva, contemplando antes de tudo o trabalho de coleta e sistematização de dados que se impõe no presente momento. Trata-se de quebrar um círculo vicioso no seu ponto mais vulnerável.

5.1.6. Complementos de medida

A concepção de objeto direto desenvolvida acima pode nos ajudar a dirimir algumas dúvidas da análise tradicional. Aqui vou apenas examinar sumariamente um caso, como exemplo do que se pode fazer.

Há alguma controvérsia sobre se o complemento de verbos como *medir* e *pesar* seria um objeto direto em frases como

[12] A mesa pesa 40 quilos.

[13] Meu quintal mede 30 metros de comprimento.

Huddleston (1984), discutindo casos semelhantes do inglês, afirma que

Com alguns verbos fica longe de claro se o complemento é [objeto] ou [predicativo]. Entre esses estão alguns verbos que tomam SNs de ‘medida’ como complementos – [*custar, pesar, medir*] etc. [...] Esses complementos carecem das propriedades positivas tanto do [objeto] quanto do [predicativo]: não podem tornar-se sujeitos mediante passivização [...], não concordam com o sujeito, e não estão em contraste paradigmático com sintagmas adjetivos. Não é de espantar que algumas gramáticas os considerem [objetos], outras [predicativos].

[Huddleston, 1984, p. 189]⁸

Ele deixa a questão por aí, sem tentar uma solução.

Allerton (1982), debatendo-se com a mesma dificuldade, acaba colocando esses complementos, com base em seu comportamento sintático, entre os exemplos da função a que chama “objóide”. No caso de [12] e [13], ele analisa o complemento como “objóide de medida” (*measure objoid*), que se define pelas seguintes características:

- (a) não ocorre como sujeito da passiva;
- (b) não ocorre na construção de alçamento,⁹ como *esse livro é difícil de Maria ler* – compare-se com * *40 quilos são difíceis da mesa pesar*.

⁸ Huddleston usa as abreviaturas ‘O’ e ‘PC’, que substituí pelos equivalentes portugueses para facilitar a leitura. Traduzi também os verbos, pois nesse particular o inglês e o português são semelhantes.

⁹ Esse era o termo usado pelos gerativistas nos anos 70. Vou mantê-lo porque essa construção não tem designação tradicional em português.

Vou argumentar que a própria questão está mal colocada, e que o problema (se é que é um problema) não tem nada a ver com o estatuto de objeto ou não objeto dos sintagmas envolvidos.

O que Allerton tentou fazer foi encontrar uma diferença de funções (objeto x objóide) para explicar as diferenças observáveis entre a construção exemplificada em [12] e [13] e a construção de

[14] O motorista ligou o carro.

[15] O cachorro comeu meu chinelo.

Mas essa diferença de funções é apenas uma maneira indireta de usar as valências dos verbos em questão para explicar as diferenças de comportamento. Seguindo o modelo de Allerton, diríamos que os objóides não podem ocorrer como sujeito de uma passiva, nem de uma construção de alçamento.

Mas talvez haja fatores mais concretos em jogo. Para começar, digamos que a valência de *pesar* e *medir* inclua a diátese seguinte:

[16] **Construção de objeto de medida**¹⁰

X	V	SN
Localizando		Medida

Ou seja, temos um sujeito que é Localizando (coloca-se em determinada posição em uma escala de mensuração) e um SN pós-verbal, que exprime a Medida.

À parte a delimitação dos papéis temáticos, isso não é uma hipótese: é bastante evidente para qualquer usuário do português que *40 quilos* é a medida usada, e que *a mesa* é o elemento cuja medida é afirmada. O tipo de medida é elaborado através do significado desse SN pós-verbal: é um peso em [12], porque o SN se refere a quilos; e é uma medida de comprimento em [13], porque o SN é *30 metros* etc.¹¹

Note-se que o SN pós-verbal não designa uma “coisa”, ou seja, não é referencial. Se colocarmos naquele lugar um SN referencial, o verbo terá que assumir outro significado:

[17] Eva pesou o arroz.

Aí temos outra diátese, pois o SN pós-verbal não aceita o papel temático de Medida.

Até o momento, temos uma análise para as frases [12] e [13], sem menção da função sintática; ou seja, não temos que decidir se esse SN é “objeto”, “objóide”, “predicativo” ou seja o que for.

Agora, passando às demais características usadas por Allerton, observa-se que o sujeito de uma frase passiva é necessariamente referencial. Por exemplo, podemos dizer

¹⁰ O termo “objeto” no nome da construção tem função mnemônica. Como veremos, não há propriamente “objeto” aí.

¹¹ Os verbos também são, em parte, especializados quanto o tipo de medida: *medir*, *custar*, *pesar*.

[18] O rato foi comido pelo gato.

que seria a passiva de

[19] O gato comeu o rato.

Mas a passiva de

[20] Gato come rato.

só poderia ser

[21] * Rato é comido por gato.

mas essa frase é mal formada, justamente porque contém um sujeito não referencial (“atributivo”, no dizer de Donnellan, 1966) em uma construção passiva. Essa é a verdadeira razão da inaceitabilidade de passivas com sujeito de medida: não sendo referencial, esse sujeito não cabe na passiva.

Já a construção de alçamento aceita sujeito não referencial:

[22] Rato é difícil de matar.

Aqui o que parece estar em jogo é um traço semântico do verbo: ele precisa denotar uma ação. Por isso é que não se diz

[23] * Cachorro é difícil de ter.

[24] * Uma pessoa honesta é difícil de ser.

apesar da aceitabilidade de

[25] É difícil ter cachorro.

[26] É difícil ser uma pessoa honesta.

Isso explica a impossibilidade de “alçar” o SN pós-verbal de frases como [12] e [13], já que o verbo não denota ação.

Se essas observações forem válidas em geral (e reconheço que não foram verificadas amplamente), teremos uma explicação para os fenômenos citados por Allerton sem necessidade de lançar mão de funções sintáticas abstratas, utilizando apenas relações concretas e independentemente motivadas. E a questão de se o SN pós-verbal de [12] e [13] é ou não é objeto direto é fica esvaziada.¹²

¹² Pode-se citar as condições de retomada pronominal como argumento para distinguir o caso de [12] – [13] do caso de [14] – [15]. Por exemplo, temos *o cachorro comeu ele* [= meu chinelo], mas * *meu quintal mede eles* [= 30 metros]. A meu ver isso pode ser facilmente explicado em termos das restrições referenciais do próprio pronome; cf. também *ele dormiu o dia inteiro* mas * *ele dormiu ele (inteiro)*.

5.1.7. Conclusão

Não é possível tirar conclusões de grande generalidade desse exame um tanto restrito de fatos da gramática portuguesa. Mas o caso do objeto direto pelo menos sugere que as chamadas funções sintáticas deveriam ser reexaminadas, para ver se muitas delas não seriam supérfluas para a análise da língua.

Langacker sustenta que nas descrições lingüísticas não devem nunca aparecer unidades ou relações sem representação fonológica ou interpretação semântica. Diz ele:

Como a linguagem é simbólica, a realidade das estruturas semântica e fonológica dificilmente pode ser questionada, mas o valor de construtos que não têm conteúdo semântico nem fonológico é evidentemente menos seguro, e a gramática cognitiva sustenta que tais construtos nunca são válidos.

[Langacker, 1987:26]

[...] não há um “espaço gramatical” distinto, nem unidades gramaticais que não possam ser descritas em termos de unidades semânticas e fonológicas, juntamente com suas associações simbólicas.

[ibid., p. 422]

Já C&J assumem uma posição que me parece mais prudente, e admitem a necessidade de definir certas relações sintáticas não relacionadas com a ordem dos sintagmas :

[...] a ordem linear, a marcação de caso e a concordância são recursos gramaticais independentes que podem ser usados para relacionar a estrutura fonológica ao significado; nenhum deles é dependente dos outros, mas em algumas línguas eles podem ocorrer redundantemente.

[C&J : 190]

Não sei se o português é uma das línguas mencionadas. De qualquer maneira, estou disposto a explorar a possibilidade de que a concordância (verbal) não é um fenômeno autônomo (fiz isso na seção 4.2). E quanto à marcação de caso, vou me esforçar por substituí-la por **ordenação + categorização** + (em certos casos) **filtragem semântica**, aliados à presença eventual de uma preposição. O sistema resultante é complexo, mas utiliza apenas mecanismos independentemente motivados.

Só a pesquisa extensa vai poder mostrar se essa tentativa é ou não suficiente para dar conta de todos os detalhes da relação entre a estrutura fonológica e o significado. Ou seja, pode ser que em certos casos haja necessidade de relações abstratas. O que se pode

dizer é que o caso do objeto direto em português fornece apoio à proposta de Langacker.¹³ Resta a tarefa de generalizá-la para outras áreas da gramática.

5.2. Notas sobre a construção ergativa

Antes de terminar o capítulo, vou fazer duas observações sobre a construção ergativa, que definimos acima.

5.2.1. Sujeito ou tópico?

A primeira observação tem a ver com a possibilidade de que a posição de *o tanque* em

[27] O tanque encheu rapidamente.

se deveria não a ser esse sintagma o sujeito de uma construção específica (a ergativa), mas à topicalização de um objeto direto. [27], portanto, seria mais ou menos paralela a

[28] O tanque nós enchemos rapidamente.

onde *o tanque* normalmente se analisa como objeto topicalizado.

O teste mais imediato da hipótese seria a concordância verbal: se *encheu* em frases como [27] concorda com *o tanque*, então não se trata de objeto topicalizado, mas de sujeito. De acordo com esse teste, trata-se realmente de um sujeito:

[29] Eu encho rapidamente (com essa conversa do Dedé).

Outro argumento em favor da análise como sujeito é que a ocorrência na construção ergativa é lexicamente condicionada; isto é, tal construção pode ocorrer com certos verbos, mas não com outros:

[27] O tanque encheu rapidamente.

[30] * O tanque pintou facilmente.

Mas a topicalização, pelo que se sabe, não é sujeita a condicionamento léxico, sendo um processo controlado apenas por fatores estruturais e discursivos – por isso, justamente, não é considerada uma diátese verbal.

Além disso, a topicalização, nos casos claros, é bloqueada justamente quando não há sujeito; assim, podemos dizer

¹³ Partindo de uma argumentação diferente, Gross (1969) também chegou à conclusão de que a noção de “objeto” é completamente inútil para a descrição gramatical, por não corresponder a nenhum fenômeno lingüístico preciso.

[31] Esse filé eu comprei congelado.

mas não

[32] ?? Esse filé comprei congelado.

Logo, não é provável que frases como [27] sejam exemplos de topicalização do objeto direto.

Outro argumento¹⁴ é que se *o tanque* em [27] fosse um objeto direto topicalizado, seria de esperar que pudesse também ocorrer depois do verbo, isto é, na posição habitual, não-topicalizada, do objeto; mas uma frase como

[33] ?? Enche rapidamente o tanque.

não é natural, o que mostra que *o tanque* não deve ser objeto direto.

Esses argumentos mostram que o que temos em [27] não é uma topicalização de objeto, mas uma construção ergativa em que o Paciente aparece como sujeito.

5.2.2. Ambigüidades

Não é totalmente correta a afirmação feita acima de que o sujeito das frases dadas como exemplos da construção ergativa se interpreta como Paciente. É verdade que ao ouvirmos

[34] Esse lombo assa rapidamente.

a única acepção que nos vem à mente é a de que o lombo fica assado rapidamente. Mas isso se deve aos traços semânticos de *lombo*, que se refere a um ser incapaz de provocar eventos. A acepção de Agente não aparece nesse caso, mas permanece latente, e pode vir à tona se houver condições favoráveis. Com o mesmo verbo *assar* pode-se ter uma frase como

[35] Esse forno assa muito bem.

Agora a interpretação mais plausível, de longe, é a que atribui a *esse forno* o papel de Agente.¹⁵

A ambigüidade de [35] nasce do fato de que em português os objetos podem ser omitidos com muita facilidade. Conseqüentemente, [35] pode ser interpretada de duas maneiras: ou o sujeito é Paciente, ou então é Agente e há um Paciente esquemático. Assim, devemos reformular a caracterização semântica de verbos como *assar*, dizendo que seu sujeito em estruturas da forma SN + V é interpretado como Paciente ou Agente (e em geral outros fatores se encarregarão de desambiguar a frase).

¹⁴ Que devo a Glória Figueiredo de Souza.

¹⁵ Relembro que não se inclui no Agente o componente semântico “voluntário”.

Mas não temos que dizer que a construção ergativa é ambígua; segundo o modelo de análise aqui adotado, uma construção nunca é ambígua no que diz respeito aos traços que a definem.¹⁶ Em vez disso, direi que a **frase** [34] é ambígua porque representa formalmente duas **construções**, a saber: a ergativa, que se define como SN + V com sujeito Paciente, e a transitiva de objeto elíptico, definida como

[36] Construção transitiva de objeto elíptico

SN	V	∅
Ag		Pac

Como as construções se definem simbolicamente, duas construções podem ser sintaticamente idênticas sem se confundirem. Obviamente, a ambigüidade acima mencionada só se verifica se o verbo for da classe dos que cabem tanto na construção ergativa quanto na transitiva de objeto elíptico. Nem todos os verbos que ocorrem na construção ergativa ocorrem na transitiva de objeto elíptico (por exemplo, *morrer*); e nem todos os verbos que ocorrem na transitiva de objeto elíptico ocorrem na ergativa (por exemplo, *gritar*). Portanto, a ambigüidade do sujeito nas frases ergativas depende da valência do verbo – só aparece quando o verbo é transitivo-ergativo. No caso de *o coelho morreu*, ou no de *o menino chorou*, não há ambigüidade.

5.3. Outras funções sintáticas

Vimos acima que a função de objeto direto é desnecessária, ou melhor, não se distingue de um conjunto de posições específicas dentro da oração; e vimos que o sujeito se caracteriza por elaborar uma indicação temática fornecida inicialmente pelo sufixo de pessoa-número. Vou agora examinar brevemente outras funções de nível oracional, mostrando que em todos os casos é possível reduzi-las a formulações concretas em termos de posição, classe e presença de preposições.

Deixo de lado algumas funções tradicionais por não serem relevantes para a definição de diáteses verbais: o vocativo e o agente da passiva, assim como o aposto e as funções internas do sintagma nominal.

5.4. Adjuntos e complementos adverbiais

O chamado adjunto adverbial é, como se sabe há muito tempo, um grupo heterogêneo de relações, com propriedades semânticas e sintáticas extremamente variadas. A única característica que os une é sua representação através de sintagmas preposicionados ou através de palavras invariáveis chamadas “advérbios” -- outra classe heterogênea.

Assim, à primeira vista parece que temos aqui um problema de maiores proporções a resolver. No entanto, e felizmente para nós, é possível evitar esse problema, pois de qualquer maneira a formulação das diáteses verbais precisa ser feita com

¹⁶ Uma *frase* pode ser ambígua; uma construção não pode, porque se expressa em termos esquemáticos.

referência às preposições específicas que ocorrem com cada verbo. Não há, por exemplo, nenhum caso de complemento que possa ser introduzido por toda e qualquer preposição; embora muitos possam ocorrer com diversas preposições, há sempre restrições (ver abaixo o caso da construção de residência). Fica o problema da subclassificação dos advérbios, que ainda está por realizar.¹⁷

Esses constituintes são às vezes considerados complementos (ver Kury, 1985, entre outros), por causa de frases como

[37] O Rafael mora em Aracaju.

A questão do estatuto de constituintes como *em Aracaju* será discutida no capítulo 6. Aqui nos interessa a maneira de representar a diátese para efeitos de subcategorização do verbo *morar*. A construção ilustrada em [37] foi incluída na lista do capítulo 13 da seguinte maneira:

[38] **Construção de Residência**

X	V	<i>em</i> SN ~ SAdv
Localizando		Lugar

Essa definição está incompleta em dois pontos. Primeiro, podemos ter outras preposições, desde que expressem Lugar:

[39] O Rafael mora fora da cidade / debaixo da ponte / ao lado da padaria...

E, depois, não é qualquer “SAdv” que cabe ali:

[40] O Rafael mora aqui / bem...

[41] * O Rafael mora imensamente / sem dúvida...

É possível suprir em parte essa deficiência lembrando que a definição é simbólica, e em sua face semântica exige a presença de um sintagma que expresse Lugar. Isso explica por que não se pode complementar a frase com *imensamente* ou *sem dúvida*. Já a possibilidade de *bem* talvez tenha que ser introduzida como marca idiossincrática, porque nem todo advérbio “de modo” cabe com *morar*:

[42] * O Rafael mora devagar / bondosamente / regularmente...¹⁸

Em alguns casos, a preposição é determinada individualmente, como com o verbo *gostar*, que exige *de* em seu complemento; em outros, como vimos com *morar*, parece

¹⁷ Que eu saiba. Há trabalhos pioneiros como o de Greenbaum (1969), certamente seguido de outros. Aqui se faz necessário realizar um levantamento bibliográfico preliminar a uma proposta de subclassificação dos diversos elementos denominados “advérbios” e “adjuntos (complementos) adverbiais”.

¹⁸ Esses itens são dados como “advérbios de modo” por Cunha (1981). De qualquer maneira, uma decisão a respeito teria que esperar uma classificação aceitável dos advérbios, como sugeri na nota 31**.

que predomina o papel temático. De qualquer maneira, um problema que não se coloca é o de definir a função sintática relevante para a definição da diátese. A discussão aqui se dirige a pontos da estrutura concreta, como a identidade das preposições, a classe dos sintagmas e o papel temático que veiculam.

Naturalmente, sobram problemas a resolver, o que me forçou a lançar mão de formulações provisórias em muitos casos; os problemas são discutidos detalhadamente na segunda parte do livro. Aqui basta observar que a função tradicional de “adjunto (ou complemento) adverbial” não ajuda em nada a esclarecer a situação. Temos que falar de seqüências de *de* + SN, de advérbios que veiculam o papel temático de Lugar, e coisas desse gênero. Portanto, para nossos objetivos, a função (ou funções) de “adjunto adverbial” é supérflua.

5.5. Objeto indireto

O mesmo se diga da função de “objeto indireto”, outro cesto de despejo da sintaxe tradicional. É fácil ver que casos como o de

[43] Meu avô gostava de vinho.

podem ser descritos simplesmente como casos de complemento regido de preposição. A única diferença que consigo ver entre os chamados “objetos indiretos” e os “adjuntos (complementos) adverbiais” é que com os primeiros a preposição costuma ser fixa, e com os segundos há uma preponderância do papel temático, admitindo-se certa variação na preposição desde que o ingrediente expresso (como Lugar, por exemplo) esteja presente.

Mas mesmo isso é apenas uma aproximação, por causa de casos como

[44] Entregamos o certificado a / para mais de cem alunos.

onde a preposição não é individualizada, mas a análise tradicional é de objeto indireto.

Assim, na formulação das diáteses não distingo casos de adjuntos ou complementos adverbiais de casos de objeto indireto, incluindo apenas uma especificação concreta do tipo de sintagma que deve ocorrer, mais o papel temático que cada um exprime. A construção exemplificada em [44] é, portanto,

[45] Construção dativa

X	V	SN	Prep SN
Agente			Beneficiário
Fonte		Tema	Meta

Notar a ocorrência de mais de um papel temático para certos sintagmas; isso é discutido na seção 8.7.

5.6. Predicativos

Para nossos objetivos, o predicativo do sujeito exemplificado em

[46] Meu avô era italiano.

se reduz a um sintagma (SN ou sintagma adjetivo) com um papel temático -- no caso, vários papéis, mas pode ser que se trate na verdade demais de uma construção. Por ora, defino a construção como

[47] **Construção Estativa**

Definição:

X	V	SAdj ~ SAdv
Localizando		(*)

(*) O papel temático é algo a discutir (ver 8.6.2); mas em todos os casos se trata de uma característica atribuída ao sujeito.

O que nos interessa aqui é que não é necessário fazer menção da função do complemento: basta delimitar sua classe e seu papel temático.

O mesmo vale para o predicativo do sujeito com verbos não de ligação, como em

[48] A diretora apareceu furiosa.

O constituinte *furiosa* pode ser analisado, para efeitos da expressão das valências, como um sintagma adjetivo (que se realiza como um nominal ou uma preposição +SN), com o papel temático de Qualidade.¹⁹ Isso basta para a expressão da diátese; permanecem outros problemas, como a possibilidade de concordância nominal com o sujeito, mas isso não parece ter nada a ver com a valência verbal.

A mesma argumentação vista para o predicativo do sujeito vale para o predicativo do objeto, que aparece em frases como

[49] O presidente nomeou Sara ministra da Educação.

Trata-se da construção de nomeação, definida assim:

[50] **Construção de Nomeação**

X	V	SN	SAdj ~ SN
Agente		Tema	Qualidade

Há, a bem dizer, um problema a resolver em relação com essas frases: a qualidade expressa se refere ao sujeito em [46] e em [48], mas em [49] se refere ao objeto direto.

¹⁹ “Qualidade” é às vezes analisado como um predicado. Aqui o que me interessa é que é uma relação semântica entre um sintagma da oração e o verbo, e portanto funciona, para os atuais objetivos, como um papel temático. A categoria lógica precisa dessa relação é algo que fica fora da minha área de competência.

Como representar isso na análise? E isso é ou não é relevante para a subcategorização do verbo? No caso de [46] e [48], evidentemente, não há propriamente um problema: o complemento qualifica o único SN possível, que é o sujeito. Mas comparemos [49] com

[51] O presidente descascou o abacaxi furioso.

Aqui *furioso* poderia qualificar tanto *o presidente* quanto *o abacaxi*. No caso, a segunda possibilidade não se realiza por razões pragmáticas (abacaxis em geral não ficam furiosos). Mas é possível imaginar um caso em que ambas as acepções sejam plausíveis:

[52] O presidente encontrou o ministro furioso.

Nesses casos surge uma ambigüidade, pois se pode entender a qualidade de “furioso” atribuída tanto ao presidente quanto ao ministro. A acepção mais saliente me parece a segunda, mas a primeira é certamente possível.²⁰

Digamos que essa situação seja de validade geral para tais estruturas. Nesse caso, a ligação semântica do qualificativo com este ou aquele SN da oração não será relevante para a subcategorização dos verbos. Podemos dizer que o sintagma qualificativo se vincula a um dos SNs precedentes; quando houver dois, pode vincular-se a qualquer um deles, desde que não acarrete violação de ordem pragmática. Vou adotar essa hipótese, reconhecendo que precisaria ser mais investigada, e não utilizarei a vinculação a este ou aquele SN como fator de subcategorização. Conseqüentemente, defino aqui apenas duas construções, a saber, [50], com sujeito, objeto e sintagma de valor qualificativo, e [47], com apenas sujeito e qualificativo.

5.7. Conclusão: funções sintáticas

Aqui há ainda muito a pesquisar. Mas por ora já se pode tirar algumas conclusões.

A principal é que em nenhum dos casos examinados há necessidade evidente de postular funções sintáticas de natureza abstrata. Para efeitos da descrição das valências verbais, todas as relações podem ser expressas em termos de (a) classe do sintagma (SN, SAdj etc.); (b) presença e identidade de preposições; (c) posição do sintagma na oração, em particular em relação ao verbo; e (d) relação semântica do sintagma com o verbo, expressa basicamente em termos de papel temático.

O sujeito é, até certo ponto, uma exceção, porque requer a explicitação de sua relação com o sufixo de pessoa-número em geral presente no verbo. Mas mesmo aqui a relação se define em termos concretos -- isto é, em termos sintáticos (posição) e semânticos (identidade de papéis temáticos).

Note-se como isso leva a uma concepção das funções sintáticas bastante diferente da tradicional. Essa concepção é válida para efeitos de formulação das valências verbais, um fenômeno reconhecidamente central na estrutura da oração. Se vale igualmente para outros objetivos dentro da descrição gramatical é coisa que fica a investigar.

²⁰ Refiro-me a acepções em que *o ministro furioso* não constitui um SN.

5.8. Funcionamento e interação das regras

A esta altura vale a pena parar um momento para nos perguntarmos como funciona esse complicado sistema de dispositivos.

Alguns modelos lingüísticos lançam mão da ordenação de regras. Em geral, a ordenação tem a utilidade de possibilitar uma notação mais econômica das regras. No entanto, esse recurso não é muito mais do que uma conveniência notacional, sem verdadeira significação teórica. Para compreender como funcionam e interagem as regras gramaticais, temos que recorrer a uma concepção mais realista do funcionamento da memória – dado que as regras gramaticais são, em última análise, parte do nosso conhecimento do mundo, ao lado de coisas como nosso conceito de “cachorro” e a disposição dos cômodos de nossa residência. Todos esses conceitos se organizam no que alguns autores chamam de **esquemas**,²¹ e uma das características fundamentais dos esquemas é que eles são todos interligados, podendo ser acessados quando necessário.

Assim, para compreender uma afirmação como

[53] Não dormi direito porque teve uma festa no apartamento de baixo.

precisamos fazer diversas ligações, de diversas naturezas. Temos que usar o conhecimento de que festas podem ser barulhentas; que festas às vezes duram até tarde da noite; que geralmente as pessoas dormem de noite; que o barulho atrapalha o sono; que o apartamento de baixo é muito próximo do meu; que quanto mais próxima a fonte do barulho, maior o incômodo; que o som se propaga mais facilmente entre apartamentos do que entre casas.

Não faz muito sentido perguntar qual dessas informações precisa ser posta em jogo primeiro. Elas estão todas disponíveis simultaneamente, e nossa mente faz as conexões de acordo com as necessidades do momento; certamente, deve haver várias maneiras de abordar essa tarefa. Não sabemos grande coisa de como a nossa mente faz isso, mas é bem claro que não é um processo linear, e que o resultado é uma rede complexa de conexões entre vários esquemas presentes em nossa memória, que colaboram para construir nosso entendimento de [53].

Outro exemplo é a tarefa de reconhecer uma pessoa que encontramos. Sabemos que Márcia é loura, e Sandra, Felipe e Zé são morenos; que Márcia e Felipe geralmente se vestem de maneira formal, mas Sandra e Zé costumam vir à escola de jeans etc. Certamente não passamos por cada um desses itens, fazendo eliminações uma a uma, no estilo: (a) a pessoa é morena, logo não pode ser Márcia; (b) é homem, logo não pode ser Sandra; (c) está vestido formalmente, logo deve ser Felipe. Em vez disso, nossa apreciação é global, e todos os detalhes são levados em conta mais ou menos simultaneamente. Ordem, caso haja, é necessariamente flexível, porque não podemos prever o que vamos notar primeiro: se se trata de homem ou mulher, que roupa está usando, ou a cor do cabelo. Qualquer que seja o ponto de partida, precisamos estar prontos para decidir quem é; uma ordenação rígida impediria o sistema de funcionar com eficiência.

²¹ **Esquemas** é o termo utilizado por Rumelhart & Ortony (1976). Designações alternativas: **frames** (Minsky, 1975); **scripts** e **plans** (Schank & Abelson, 1977) e outras.

Essa concepção do processamento da informação é praticamente um consenso entre os psicólogos cognitivos, e isso há muito tempo, conforme atesta a afirmação de que

Se o organismo leva um “modelo reduzido” da realidade externa e de suas ações possíveis dentro da cabeça, ele consegue tentar várias alternativas, concluir qual é a melhor delas, reagir diante de situações futuras antes que elas ocorram, utilizar o conhecimento de eventos passados ao lidar com os presentes e futuros, e de todas as maneiras reagir às emergências que tem que enfrentar de maneira muito mais completa, segura e competente.

[Craik, 1943, *apud* Johnson-Laird, 1981: 149]

A flexibilidade é uma condição para o bom funcionamento do sistema cognitivo em um mundo em grande parte imprevisível.

Acredito que o processamento gramatical funciona basicamente da mesma forma. O receptor tem acesso simultâneo às regras, categorias e demais componentes de seu conhecimento da língua, e faz uso de tudo isso (ao lado de seu conhecimento do mundo) para construir a compreensão das frases. As regras, itens léxicos e demais recursos são também esquemas, e compreender uma frase é estabelecer as conexões certas entre os esquemas certos.

Essa concepção tem relevância para certas controvérsias da lingüística moderna. Uma dessas controvérsias é a da proposta separação dos componentes: sintático, semântico etc., e sua ordenação, tal como se encontra na maioria dos modelos derivados da gramática gerativa clássica. Existe evidência suficiente em favor da necessidade de regras (assim como princípios, formas, unidades e traços) de base formal; de regras (etc.) que descrevem formas com base parcialmente semântica; de regras semânticas que funcionam a partir de formas; de regras semânticas que funcionam a partir do conhecimento do mundo (mais as condições de produção de sentido no texto). E, por outro lado, há casos claros de regras puramente sintáticas, assim como puramente fonológicas, sem correlato semântico. Ou seja, não dá para negar que o fenômeno da linguagem é fundamentalmente complexo.

No entanto, me parece inadequado argumentar sobre a separação ou não dos diversos componentes, sua ordenação (ou a das regras dentro dos componentes), ou discutir se uma estrutura em particular deve ou precisa ser analisada em termos semânticos, formais ou mistos. A evidência disponível sempre desautoriza essas especulações. Creio que o quadro que emerge é mais de um conjunto de estratégias dominadas pelo usuário, que as usa dependendo do momento, da forma mais conveniente – um modelo “oportunístico” do uso da língua.

Para pegar um exemplo, seja a seguinte frase:

[54] Se todos fossem iguais a você, que maravilha viver!

[Morais, 1956]

Aqui temos um problema sintático: uma oração subordinada típica (*se todos fossem iguais a você*) sem uma oração principal completa. Como não se trata de situação anafórica, podemos nos perguntar como é que o ouvinte processa essa frase,

presumivelmente inserindo o elemento semântico “seria” na segunda parte. Se isso for feito explicitamente, o resultado vai ser

[55] Se todos fossem iguais a você, que maravilha seria viver!²²

Sabemos que, a se inserir ali um verbo, tem que ser *seria*: não cabe *foi*, nem *é*, nem *tem sido*, nem *será* – isso em virtude de uma regra de correlação temporal que leva em conta a forma do verbo da subordinada (*fossem*). Mas o fato é que o verbo não está ali, e é introduzido em [55] a partir de quê? A regra sintática de correlação não nos ajuda, porque ela não introduz verbos, apenas os correlaciona.

A melhor saída é considerar um “espaço contrafactual” evocado pela subordinada, que nos obriga a entender determinada situação na principal – e, se for o caso, introduzir a forma verbal adequada, como está em [55]. O que é que o usuário faz, exatamente? Ao receber a primeira oração, estabelece o espaço contrafactual. Isso lhe possibilita, por exemplo, saber que *seria* fica bem como verbo de uma eventual oração principal subsequente. Mas ele certamente não espera que o verbo apareça para julgá-lo – antes, a partir do espaço estabelecido, faz previsões. Na ausência do verbo, mas na necessidade de entender uma relação entre *que maravilha* e *viver*, ele entende algo como “seria” – o conceito, não a forma verbal propriamente dita.

Aqui temos um exemplo do uso oportunístico da linguagem. Antes de verificar a seqüência que vem depois da subordinada, ele fica preparado para:

- (a) ler uma forma verbal como *seria*; ou
- (b) não encontrar nenhuma forma verbal ali onde é esperada.

Ele tem estratégias para enfrentar qualquer das situações, e essas estratégias são de diversas naturezas: sintáticas, simbólicas, semânticas, pragmáticas etc. Só o que não se aceita é que a seqüência não forme um texto.

No exemplo acima, existe um contexto lingüístico estabelecendo o espaço mental. Mas isso não é necessário. Se meu interlocutor e eu estamos vendo uma grande atleta na TV, posso dizer

[56] Que maravilha ter essa agilidade!

e aqui o que se entende é novamente “seria” – porque, entre outras coisas, meu interlocutor sabe muito bem que eu não tenho grande agilidade.

Agora, se eu chego para alguém e digo

[57] Que maravilha trabalhar na PUC!

isso pode ser entendido como “que maravilha é trabalhar na PUC”, ou então “seria” – vai depender do que se sabe a respeito da situação. Se se sabe que eu trabalho na PUC, a interpretação é “é”; se se sabe que eu estou pensando em fazer concurso, a interpretação é “seria”. Em outras palavras, o significado “é” ou “seria” tem pelo menos duas fontes possíveis: pode ser depreendido da forma verbal explícita, presente na frase, ou pode ser

²² Com minhas desculpas ao poeta.

suprido através das restrições fornecidas pelo espaço mental relevante. O significado é o mesmo nos dois casos, apesar da origem radicalmente diferente. Esse é um exemplo da disponibilidade de diferentes estratégias para atingir o mesmo fim.

Para que isso funcione, é preciso que as estratégias estejam disponíveis mais ou menos em pé de igualdade. Não é plausível que o usuário tenha que tentar primeiro uma, e só recorrer à outra quando a primeira não funciona – ou seja, a noção habitual de ordenação não se aplica aqui.²³

²³ É claro que o sistema inclui restrições, que não são conhecidas e têm que ser pesquisadas eventualmente. Aqui sou obrigado a pular por cima do problema, que é acessível à pesquisa empírica mas não à simples especulação.

PARTE II: VALÊNCIAS VERBAIS

Capítulo 6: Resenha da literatura

Antes de atacar a questão das valências verbais, vou examinar brevemente alguns trabalhos que são de particular importância para o estudo desse fenômeno: Gross (1975), Allerton (1982), Levin (1993), Goldberg (1995), Vilela (1992) e Vilela & Koch (2001). Além disso, vou fazer uma breve crítica do tratamento das valências (“regência verbal”) em gramáticas tradicionais e dicionários. Alguns trabalhos não são resenhados neste capítulo porque contêm afirmações importantes sobre problemas específicos que são discutidas diretamente nas passagens que tratam desses problemas; um exemplo é Whitaker-Franchi (1989).

6.1. A tradição gramatical

6.1.1. Gramáticas

As gramáticas disponíveis não oferecem nenhuma exposição ou discussão do sistema de valências, à parte a sumaríssima visão tradicional em cinco classes: transitivos diretos, transitivos indiretos, transitivos diretos e indiretos, intransitivos, de ligação. Alguns autores, como Kury (1985) questionam a posição tradicional, mas de forma muito limitada.

6.1.2. Dicionários

Os dicionários de regência (Fernandes, 1940; Borba *et al.*, 1990; Luft, 1999) são bem mais minuciosos que as gramáticas, e constituem fonte útil de informação, desde que manejados com as devidas precauções. Todos eles descrevem a língua escrita. Os mais estritamente tradicionais (que são os de Fernandes e de Luft; o de Borba foge bastante aos moldes usuais, e deve ser examinado à parte) apresentam diversos defeitos que limitam sua utilidade enquanto trabalhos de descrição da língua. Assim, examinando Luft (1999), pode-se apontar problemas como os seguintes:

(a) A informação é atomizada, sem um esforço no sentido de obter generalizações: a informação a respeito dos verbos é repetida para cada um deles, seja ou não idêntica à referente a outros verbos. Desse modo, um verbo idiossincrático como *ser* é tratado da mesma maneira que verbos como *comer* e *encher*, que representam grandes grupos de verbos de valência semelhante.

(b) Seguindo a tendência geral, categorias como **transitivo**, **intransitivo** etc. são tratadas ora como classes, ora como funções (como quando se diz que um verbo “pode ser” transitivo direto ou intransitivo, o que reduz essas categorias a funções).

(c) Não há notação clara dos papéis temáticos, o que dificulta a percepção de generalizações. As relações semânticas são captadas indiretamente, através do uso de paráfrases ou sinônimos. Por exemplo, para exprimir os papéis temáticos possíveis do sujeito de *entortar* (Agente ou Paciente), Luft oferece as paráfrases “(fazer) ficar torto”, “tornar(-se) torto”, “entortar(-se)”.

(d) A nomenclatura é pobre, incapaz de cobrir todos os casos gramaticalmente relevantes. Luft, cuja terminologia é bem mais ampla do que a tradicional, utiliza cerca de uma dezena de tipos: **intransitivo, transitivo direto, transitivo indireto, transitivo direto e indireto, predicativo** (=de ligação), **auxiliar, transitivo direto predicativo e transitivo indireto predicativo** e mais uma ou duas. Fernandes, que se limita mais de perto à nomenclatura tradicional, tem apenas as cinco categorias usuais.

O trabalho de Borba *et al.* (1990) é muito mais rico, e tenta descrever a semântica ao lado da regência, indicando os papéis temáticos e utilizando uma variedade maior de descrições estruturais. Assim, ele distingue verbos que indicam “ação-processo”, verbos que indicam “processo” e verbos que podem indicar ambas as coisas, sem no entanto incluir uma tentativa de análise dessas construções.

A informação fornecida para cada item é copiosa, mas às vezes omissa ou incorreta em alguns pontos (defeito talvez inevitável em uma lista de 1362 páginas!). Por exemplo, ele descreve o sujeito de *beber* em *Machado de Assis bebeu seu humor nos ingleses* como “paciente”; diz que o paciente de *ressecar* é um nome “concreto não-animado” (mas podemos dizer *o calor resseca as pessoas*); e representa de maneira inutilmente difusa a semântica do complemento de *gostar*, como “nome indicativo de **algo comestível / bebível**” + “nome **animado**” + “nome não animado”, quando a relação semântica (o papel temático) é sempre a mesma. Isso, naturalmente, dificulta o trabalho de estabelecer generalizações.¹

A estrutura sintática das construções em que cada verbo se insere pode ser deduzida da informação fornecida, mas as relações entre construções não são explicitadas. A necessidade de exprimir as diferentes acepções semânticas tende a obscurecer as oposições diatéticas, que são as que nos interessam aqui. Assim, Borba *et al.* distinguem em subverbetes separados acepções como as encontradas em *a velha caiu no chão, os raios do sol caíram sobre a terra e um abatimento grande caiu sobre a população*,² que para nós certamente não vale a pena distinguir. E, provavelmente por não ocorrerem em seu cópula, deixa de consignar diáteses de certo interesse, como a ilustrada na frase *a fábrica caiu de produção*.

Um problema do dicionário de Borba *et al.*, para nossos objetivos, é o de se concentrar exclusivamente nos fatos da língua escrita; por isso, deixa de consignar construções coloquiais como, por exemplo, a “pseudo-ergativa” exemplificada em *ele cortou o cabelo*. E, naturalmente, sendo um dicionário, não inclui discussão das questões teóricas e metodológicas envolvidas no trabalho de investigação de valências; este se realizou em grande parte com bases intuitivas, decisão inevitável no momento em que foi

¹ A bem da justiça, observo que o objetivo da obra de Borba não é servir de base a generalizações gramaticais; e que tal falta de sistematicidade é traço comum aos dicionários em geral.

² Exemplos do autor.

elaborado.³ Quanto ao mais, é uma obra pioneira, construída sobre dados reais tirados de um corpus escrito, constituindo importante fonte de idéias e dados para qualquer pesquisa que se venha a realizar sobre a valência dos verbos do português.

6.2. Tesnière (1959)

Tesnière (1959) foi quem introduziu o termo “valência” na teoria lingüística, e seu trabalho tem o mérito de ter chamado a atenção para o fenômeno e estimulado muitas pesquisas. No entanto, apresenta muitos dos defeitos da abordagem tradicional, de modo que hoje seu valor é principalmente histórico.

Tesnière distingue os **actantes** (*actants*, que correspondem aos nossos complementos) e **circunstantes** (*circonstants*, isto é, adjuntos); mas a distinção é tão confusa quanto nos trabalhos anteriores (e posteriores). Assim, ao tratar do verbo *changer* ‘trocar’, Tesnière comenta que ele

tem um valor todo especial quando se concebe como verbo de um actante. Nesses casos ele se constrói com um circunstante precedido de *de* e, enquanto que com dois actantes dizemos *l’horloger change le ressort de ma montre* [‘o relojoeiro troca a mola do meu relógio’], temos, ao contrário, com um só actante: *Alfred change de veste* [‘Alfred troca de paletó’].

[Tesnière, 1959: 241]

Fica bastante claro como a distinção entre actante e circunstante é exatamente tão nebulosa quanto a que se estabelece em geral entre complemento e adjunto. Note-se que em *Alfred change de veste* o chamado circunstante é inclusive obrigatório, pois presumo que **Alfred change* deve ser tão anômalo quanto o português **Alfredo troca*.

A sintaxe de Tesnière é pouco rigorosa, e escorrega em afirmações como a de que

Os **actantes** são sempre **substantivos**, ou equivalentes de substantivos.

[p. 102]⁴

quando o próprio Tesnière cita como actantes sintagmas como *à Charles*, em *Alfred donne le livre à Charles* [p. 102]; este é claramente um sintagma preposicionado e, portanto, não pode ser considerado o “equivalente” sintático de um substantivo (ou seja, na nomenclatura atual, um SN).

Os actantes (que seriam no máximo três por oração) se definem simbolicamente, mas sem deixar espaço para a grande variedade de papéis temáticos que ocorrem na realidade. Por exemplo,

Do ponto de vista **semântico**, o primeiro actante é o que **pratica a ação**. Por isso, o primeiro actante é conhecido na gramática tradicional com o nome de **sujeito**, que vamos manter.

³ Borba (2003) discute alguns problemas teóricos relativos à elaboração de dicionários, com um resumo útil do problema das valências verbais (pp. 219-228).

⁴ Mantenho o uso de negritos como no original.

[p. 108]

Por outro lado, Tesnière já exprime algumas noções fundamentais, certamente defensáveis. Ele reconhece o papel central do verbo na estruturação da oração:

[...] o verbo fica no centro do nódulo verbal e por conseguinte da frase verbal. Ele é portanto o **regente** de toda a frase verbal.

[p. 103]

E observa a possibilidade de não-ocorrência de actantes, ou seja, o problema da opcionalidade dos chamados complementos:

nunca é necessário que as valências de um verbo sejam inteiramente preenchidas pelo actante e que o verbo seja, por assim dizer, saturado. Algumas valências podem ficar desocupadas [...]

[p. 238]

No entanto, essa observação não se harmoniza com a concepção de Tesnière do que vem a ser um actante, pois ele afirma em outro lugar que

o **primeiro actante** se encontra em princípio em todas as frases de um, dois ou três actantes; [...] o **segundo actante** se encontra em princípio nas frases de dois ou três actantes [...]

[p. 108]

Não sabendo o que ele quer dizer exatamente com “em princípio”, ficamos exatamente na situação de incerteza em que nos deixam as teorias tradicionais.

Sem levar adiante a crítica, concluo que o trabalho de Tesnière representa o marco inicial dos estudos valenciais, no que pesem suas evidentes deficiências. Além disso, gostaria de observar que é um livro cheio de *insights* importantes, embora o modelo adotado em geral tenha impedido Tesnière de desenvolver essas percepções em termos de análise da língua.

6.3. Gross (1975) e Levy (1983)

Os trabalhos de Gross (1975) e de Levy (1983) não tratam das valências verbais no sentido aqui adotado; ambos tratam das completivas (orações subordinadas), que ficam fora do âmbito do presente trabalho. Mas a discussão da metodologia e dos problemas encontrados nessas obras tem validade bem ampla. Ambas incluem discussões extremamente lúcidas de problemas metodológicos importantes para quem quer que pretenda fazer estudos de âmbito lexical. Aqui não será possível entrar em pormenores dessas discussões, mas é importante deixar claro que elas constituem boa parte das bases da abordagem adotada neste trabalho.

A seguinte passagem resume bem as preocupações de Gross relativas à gramática gerativa dos anos 70, e que a meu ver se aplicam a muitos trabalhos atuais, gerativos ou não:

Os estudos transformacionais só afetam pequenas quantidades de exemplos. Eles descobriram grande número de fenômenos novos, mas não permitem avaliar a extensão desses fenômenos para uma língua dada. Por “extensão”, não entendemos frequência de ocorrência nos textos, já que essa noção não tem praticamente nenhum sentido estatístico para a sintaxe, mas frequência de ocorrência no léxico da língua.

[Gross, 1975: 20]

Embora reconheça a contribuição da gramática gerativa para a delimitação rigorosa dos estudos e para a apreensão de regularidades numerosas e importantes, Gross critica a falta de estudos abrangentes e suas conseqüências:

Depois de um momento em que os sucessos [da análise gerativa] levaram a crer que o uso das transformações nas descrições ia regularizar consideravelmente estas últimas, ficou claro que as novas regras continuavam a comportar “exceções” em número apreciável. Tornou-se pois crucial verificar essas teorias levando a efeito a descrição de pelo menos uma língua [...]

[Gross, 1975: 20]

Levy (1983) chega às mesmas conclusões de Gross a respeito da complexidade dos fatos, muitíssimo maior do que dá a entender a maior parte da literatura. Assim, observa ela:

Como é bem sabido, o objetivo da lingüística gerativa não é a descrição particular de uma língua, mas antes a caracterização formal do que é uma língua natural. Por isso, para os gerativistas bastam uns tantos exemplos-chave que mostrem certas propriedades formais. Em geral, as decisões teóricas tomadas por muitos autores dentro dessa corrente se baseiam em estudos muito detalhados, mas de um número muito reduzido de exemplos lingüísticos. [...] Me parece que antes de entrar nos detalhes sutis de um fenômeno lingüístico, é muito importante ter uma idéia do que é o geral para esse fenômeno e o que é o particular.

[Levy, 1983: 15]

Como se vê, é exatamente a preocupação de Gross, e também a que norteia o presente trabalho. Se for preciso resumir as preocupações que partilho com Gross, Levy e outros autores, eu diria que a maior parte da lingüística moderna (principalmente no que diz respeito aos estudos de sintaxe e semântica) subestima radicalmente o componente anomalístico da linguagem, passando a imagem de uma estrutura altamente regular, que não corresponde à realidade. Uma das saídas para essa situação será a realização de estudos que cubram áreas amplas da gramática e do léxico de línguas particulares.

6.4. Allerton (1982)

Allerton (1982) realizou um levantamento amplo dos verbos do inglês e elaborou uma lista das construções relevantes para sua subcategorização. Como veremos, a utilidade de seu trabalho é limitada por alguns aspectos do modelo empregado, mas de qualquer maneira é necessário reconhecer o caráter pioneiro da obra, que inclui discussões bastante lúcidas de muitos problemas ligados à formulação de diáteses.

Vejo dois problemas fundamentais no estudo de Allerton. O primeiro é que ele usou um modelo transformacional, distinguindo uma estrutura subjacente e definindo as diáteses em termos dessa estrutura subjacente. Assim, para ele ativas e passivas representam a mesma construção, porque ambas seriam derivadas da mesma estrutura subjacente. Em *Fagin was seen by Oliver* ‘Fagin foi visto por Oliver’ o sujeito seria *Oliver* para efeito de diátese:

[em *Fagin was seen by Oliver*] encontramos que, embora o sujeito superficial seja agora *Fagin*, o sujeito de valência com seu papel temático de “experenciador” pode ser mantido como *Oliver* [...]

[Allerton, 1982: 43]

O segundo problema é que Allerton não leva em conta o plano semântico para a definição das diáteses, de maneira que para ele só existe uma construção da forma SN + V, exemplificada por *Fido barked* ‘Fido latiu’ [p. 145]. Ora, isso mascara diferenças importantes, como a que existe entre o verbo *comer*, que pode ocorrer sem objeto, mas com sujeito Agente e Paciente subentendido (na construção que chamo “transitiva de objeto elíptico”) e o verbo *assustar*, que ocorre sem objeto com sujeito Paciente e sem Agente subentendido (na construção “ergativa”). Para Allerton, portanto, *comer* e *assustar* ocorreriam nas mesmas construções (SN + V e SN + V + N), sem maiores comentários. Aqui, entretanto, sustento que são verbos de subclasses diferentes. O modelo de Allerton nos obrigaria a repetir a subclassificação dentro do componente semântico, desta vez, presumivelmente, distinguindo os dois verbos *comer* e *assustar*, acrescentando ainda a relação função sintática / papel temático de cada um.

Esses problemas devem ser imputados ao modelo utilizado mais do que ao trabalho pessoal de Allerton. Apesar deles, a lista de construções que ele apresenta (páginas 145-147) é útil como ponto de partida para novos levantamentos.

6.5. Levin (1993)

O livro de Levin (1993) é o trabalho mais importante na área do estudo de valências verbais. É uma obra fundamental, repleta de idéias e exemplos fecundos, e constitui um ponto de partida indispensável para quem quer que deseje se aprofundar no assunto. Assim, as críticas que se seguem pretendem apenas aperfeiçoar certos aspectos da perspectiva adotada por Levin, sem desmerecer a importância de sua pesquisa.

Levin aparentemente pressupõe que há uma relação íntima entre o significado e o comportamento sintático dos verbos (no que diz respeito às diáteses, pelo menos). Por exemplo, diz ela:

Se o comportamento distintivo das classes de verbos em relação às alternâncias de diáteses decorre de seu significado, qualquer classe de verbos cujos membros funcionam paralelamente quanto às alternâncias de diáteses deve ser uma classe semanticamente coerente: seus membros devem compartilhar pelo menos algum aspecto de seu significado.

[Levin, 1993: 14]

Isso pode valer como uma hipótese a ser investigada, e eventualmente confirmada total ou parcialmente.

Levin não explicita um fator essencial na conceituação das diáteses, que é a capacidade ou não de uma construção de subclassificar os verbos. Ou seja, o fato de se detectar uma diferença semântica ou formal entre duas frases não significa automaticamente que estas sejam representantes de duas diáteses diferentes – é o caso das frases topicalizadas, que apesar de se distinguirem formalmente de suas versões não topicalizadas, não podem ser consideradas base para uma diátese verbal, porque os verbos não se dividem entre os que admitem e os que não admitem topicalização. É verdade que, de uma maneira ou outra, Levin evita incluir em sua lista de diáteses construções claramente não-classificadoras (como frases topicalizadas, clivadas, negadas ou interrogadas), o que mostra que, pelo menos implicitamente, ela estaria aplicando o princípio de que relevância para a subclassificação é critério para a postulação de uma nova diátese.

Levin tampouco leva em conta o que chamei de “nível de detalhamento” (Perini, 1989), o que a leva a fazer distinções de utilidade questionável. Ou seja, tenho a impressão de que as distinções entre as diáteses de Levin são, por vezes, muito finas, a ponto de serem gramaticalmente irrelevantes. Tudo isso resulta em um número de diáteses maior do que o desejável. Assim, Levin (seguindo aqui a tendência geral da literatura) distingue a construção “média” da “ergativa”,⁵ sendo que a diferença radica em fatores estranhos à subclasse do verbo (como mostro no capítulo 10). Distingue, também, a construção sem objeto, exemplificada em

[1] O marido dela nunca bebe.

da de “objeto elíptico parte do corpo”, como em

[2] A menina piscou para mim.

No entanto, essa diferença me parece nascer de informações pragmáticas estranhas à classe do verbo: assim como em [1] se entende um Paciente “bebida alcoólica” (o marido dela certamente bebe água), em [2] se entende que a menina piscou o olho (e não a boca). Coloco essa diferença no nível de detalhamento representado pelo objeto de *comer* e *beber* discutido no capítulo 9.⁶ Isto é, há uma subclassificação dos

⁵ *Inchoative* na terminologia de Levin; essas construções são muitas vezes chamadas “inacusativas”.

⁶ Mostro ali que a diferença de objeto direto desses dois verbos pode ser prevista facilmente a partir de fatores estranhos ao comportamento gramatical de cada um: bebe-se líquidos, come-se sólidos.

verbos aqui, mas é excessivamente “fina” para ser de interesse em uma descrição gramatical com pretensões a generalização.

Admito que a calibragem do nível de detalhamento é algo intuitivo, e não posso certamente demonstrar que minha abordagem é superior à de Levin neste particular. No entanto, prefiro apostar em uma descrição mais ampla, desprezando certos detalhes. O limite do detalhamento é inevitável, e o problema que cada lingüista precisa enfrentar é onde o colocar. Uma decisão mais objetiva só poderá ser atingida quando dispusermos de descrições suficientemente abrangentes; por ora, terei que confessar que minha intuição difere da de Levin. Trata-se, no fundo, da decisão teórica de deixar certos fenômenos a cargo de uma descrição de nosso conhecimento do mundo (“regras de bom senso”), excluindo-os da descrição da língua que falamos.⁷

Outra coisa importante que pode ser dita a respeito do trabalho da Levin é que a homogeneidade das classes semânticas é desmentida, em parte, pela ocorrência muito grande de particularidades, que ela apenas comenta, sem tirar conclusões. Assim, por exemplo ao comentar as propriedades de cada grupo (semântico) de verbos, ela constantemente acrescenta notas como “a maioria dos verbos” ou “alguns verbos”, mostrando que a homogeneidade tem limitações. Ora, ela não explora essas limitações, e por isso, sua lista (que é aliás preciosa como fonte de indicações) é imprecisa demais para servir como base de teste para a hipótese. Creio que o problema aqui é a falta de critérios prévios de relevância, o que obriga Levin a fazer distinções ou não com base em uma avaliação arbitrária e um tanto subjetiva.

Levin agrupa as diáteses em “alternâncias”, isto é, grupos de construções consideradas como tendo alguma coisa em comum. Ela não discute explicitamente o que vêm a ser as alternâncias, mas até onde posso ver trata-se de relações de natureza transformacional. Assim, diz ela ao tratar de “alternâncias de transitividade”:

Esta seção inclui alternâncias que envolvem uma mudança na transitividade de um verbo. Essas alternâncias têm a forma de ‘SN V SN’ alternando com ‘SN V’ ou então ‘SN V SN’ alternando com ‘SN V SPrep’.

[Levin, 1993: 25]

Ou seja, existiria uma relação privilegiada entre certas construções, o que configura um modelo transformacional – não necessariamente o modelo clássico de Chomsky (1965), já que Levin leva em conta a supressão de um objeto direto, por exemplo.

Não vejo vantagens em aceitar essas relações privilegiadas. A meu ver, uma frase como

[3] Os meninos tocaram a campainha.

⁷ A determinação do nível de detalhamento me parece uma questão metodológica importante, no que pese a impossibilidade em que nos encontramos de colocá-la em terreno relativamente firme. A respeito desse contínuo de pormenores, é interessante a afirmação do matemático Benoît Mandelbrot (apud Gould, 2000: 346) de que “a costa do Maine não tem extensão absoluta, mas depende da escala de mensuração” (isto é, da decisão de contar ou não todos os pequenos desvios da linha costeira), que mostra que o problema não é só dos lingüistas.

se relaciona com

[4] A campainha tocou.

não em virtude de operações gramaticais específicas, mas por causa de sua semelhança semântica: ambas têm uma ação (“tocar”), sendo que em [3] o agente é expresso (“os meninos”), e em [4] este fica não especificado. Isso é decorrência automática da interpretação que se atribui a essas frases e do fato de que elas são preenchidas, em parte, com os mesmos itens léxicos.

Aqui adoto uma posição estritamente não-transformacional:⁸ cada construção é gerada e interpretada independentemente, e quaisquer semelhanças entre elas são resultado de semelhanças no processo de geração e interpretação. Isso exclui a noção de “alternância”, de maneira que falarei simplesmente de “construções” (em geral) e “diáteses” (quando uma construção é relevante para a formulação da subcategorização dos verbos).

Um problema com a lista de Levin, que limita sua utilidade, é a ausência de qualquer tentativa explícita de análise das diversas construções utilizadas. Assim, ela dá um nome a cada uma de suas “alternâncias” e acrescenta exemplos; mas não inclui uma análise sintática e semântica, embora esta esteja certamente implícita. Fica a cargo do leitor deduzir do que é exatamente que ela está falando em cada caso. Um exemplo típico é:

Material/Product Alternation (intransitive)

[...] Alternating verbs:

GROW verbs: develop, evolve, grow, hatch, mature

[...] a. That acorn will grow into an oak tree.

b. An oak tree will grow from that acorn.

[Levin, 1993: 57]⁹

Além disso, Levin em geral acrescenta comentários sobre a extensão da alternância (exceções e ocorrência de verbos nesta e em outras construções), sobre as relações semânticas encontradas (no caso, ela observa que “o sujeito exerce o papel que o objeto direto exerce na forma transitiva desta alternância”). Acabamos tendo muita informação, mas apresentada de maneira um tanto informal e às vezes assistemática.

Para fazer uma crítica global do trabalho de Levin, eu diria que ela procedeu ao levantamento e classificação dos verbos sem antes desenvolver devidamente um modelo de análise adequado à tarefa empreendida. Por isso, a classificação resultante é um tanto

⁸ Note-se que o modelo transformacional, introduzido por Harris (1957) e desenvolvido por Chomsky (1957), não foi abandonado pela maioria dos modelos gerativos, que continuam admitindo operações como a de “Mover α ”. A posição não-transformacional aqui adotada se harmoniza no essencial com a dos cognitivistas (Langacker, 1987), e também com a teoria mais recente de Jackendoff, tal como expressa em C&J.

⁹ Neste caso preferi não traduzir, para manter os fatos do inglês tais como apresentados por Levin.

heterogênea, pouco sistemática e menos elucidativa do que poderia ser. Neste trabalho me concentro na elaboração do modelo descritivo que falta em trabalhos como o de Levin.

6.6. Goldberg (1995)

Goldberg (1995) defende a importância da formulação de construções como base para a descrição da estrutura simbólica das orações simples. Essa é basicamente a posição que adoto, e que é resumida na passagem seguinte:

Uma tese central deste trabalho é que as sentenças básicas do inglês são realizações de *construções* – correspondências formas-significados que existem independentemente dos verbos individuais. Isto é, sustento que as próprias construções veiculam significados, independentemente das palavras da sentença.
[Goldberg, 1995: 1]

Eu só acrescentaria que as “sentenças básicas” de Goldberg podem ser entendidas como períodos simples (unioracionais), com a possível exclusão dos adjuntos.¹⁰ Pode-se dizer, portanto, que neste ponto concordo com Goldberg.

Ela propõe um sistema de notação das construções, mas neste particular faço objeção. A notação de Goldberg, como a minha, é simbólica, incluindo traços sintáticos e semânticos. Assim, para a construção inglesa representada por [5],

[5] Pat faxed Bill the letter. ‘Pat mandou por fax a carta para Bill’

ela dá a seguinte formulação:

[6] X CAUSES Y TO RECEIVE Z Subj V Obj Obj₂
[Goldberg, 1995: 3]

Isso corresponde bem de perto à minha notação

[7] X V SN SN
 Agente Meta Tema

Mas há uma diferença: Goldberg inclui traços do significado do verbo (CAUSE TO RECEIVE), ao passo que eu coloco os papéis temáticos. Acho que a minha notação é mais clara, porque a de Goldberg não explicita que complementos sintáticos correspondem ao X, Y e Z da representação semântica; por isso, utilizo neste livro a notação exemplificada em [7].

Goldberg define uma **construção** da seguinte maneira:

¹⁰ “Adjunto” é uma noção muito pouco clara, aliás; ver o capítulo 9.

C é uma CONSTRUÇÃO se e somente se_{def} C for um par forma-significado $\langle F_i, S_i \rangle$ tal que nenhum aspecto de F_i nem de S_i seja estritamente previsível a partir das partes componentes de C ou de outras construções previamente estabelecidas.

[Goldberg, 1995: 4]

Essa definição corresponde de perto a minha noção de **diátese**; mas prefiro chamar de construção todo e qualquer par forma-significado definível dentro da língua, mesmo se sua existência puder ser prevista a partir de outro par forma-significado.

Para dar um exemplo, podemos tomar o caso de

[8] O Sérgio rasgou meu casaco.

[9] Meu casaco, o Sérgio rasgou.

Em português, para cada oração da forma (e significado básico) de [8] existe uma da forma de [9]; isso não depende do verbo, mas da própria estrutura. Por conseguinte, não vale a pena definir aqui duas diáteses independentes, já que todo verbo que cabe na estrutura exemplificada em [8] cabe também em [9], e vice-versa. No entanto, [8] e [9] não são idênticas, e precisam ser descritas como estruturas diferentes em algum lugar da gramática. Faço isso considerando-as **construções** distintas, embora só uma delas seja **diátese** do verbo *rasgar* (volto ao problema na seção 9.5.1.1).

Segundo Goldberg (1995: 4), “as construções são consideradas as unidades básicas da língua”. No entanto, não vejo inconveniente em formular regras de estrutura sintagmática que enumerem de maneira mais econômica as construções possíveis de uma língua. Ou seja, em vez de dar simplesmente uma lista de construções, pode ser vantajoso dar uma “receita” que defina o que é e o que não é uma construção em determinada língua. Não vou investigar essa possibilidade no presente trabalho, porque ela depende do estabelecimento prévio da lista de construções; mas sugiro que essa é uma das tarefas importantes que nos esperam no futuro próximo. Ela se insere na investigação mais ampla das dimensões dos dois grandes componentes da estrutura da língua, a gramática e o léxico, ou seja, as regularidades e as idiosincrasias.

No que pesem algumas objeções, o trabalho de Goldberg (1995) é teoricamente muito próximo das idéias expressas neste livro, e se move na direção de objetivos basicamente idênticos. É uma obra muito rica em propostas de interesse, que infelizmente não podem ser examinadas como merecem dentro das limitações do presente estudo.

6.7. Vilela (1992)

Em Vilela (1992) encontramos uma série de ensaios, alguns dos quais tratam da questão das valências verbais. O primeiro capítulo oferece uma resenha de propostas, mostrando que a maioria delas inclui na subcategorização traços do próprio significado do verbo, o que é contra-indicado porque impede que se verifique a correção da hipótese da motivação semântica das valências. Ora, a meu ver a testagem dessa hipótese é justamente um dos produtos mais interessantes do levantamento das classes valenciiais (ver discussão do problema em 8.6.4).

Sem realmente optar por um dos modelos resenhados, Vilela assume uma posição um tanto pessimista, exprimindo um ideal certamente inatingível, de

uma classificação conjunta e completa, em que se tome em consideração o morfológico, o sintático, o semântico e o pragmático (no verbo) [...] [Vilela, 1992: 26]¹¹

Isso é sem dúvida impossível, e, inclusive pouco interessante, dadas as condições sob as quais toda classificação é elaborada (ver a respeito o capítulo 3). Vilela termina afirmando que

Uma proposta de classificação completa que seja capaz de abarcar todos os verbos ... pertence, na melhor das hipóteses, apenas ao mundo dos “possíveis”. [ibid.]

Essa afirmação é curiosa, dado que o próprio Vilela cita trabalhos muito extensos, como os do grupo do *lexique-grammaire* (iniciado por Gross), cujos levantamentos abrangem milhares de verbos – cerca de 3000 no estudo de Willems (1981) para o francês.

6.8. Concluindo

Como se vê, ao reconhecer deficiências em todos os sistemas de notação das valências disponíveis (em função dos objetivos descritivos aqui adotados), me vi obrigado a criar um sistema próprio. Este é exposto e discutido no capítulo 9; no capítulo 13 sumario as convenções adotadas e dou uma lista (parcial) das construções encontradas no português.

Evidentemente, meu sistema tem uma grande dívida para com os pesquisadores que se ocuparam anteriormente do problema das valências verbais. Dentre essas fontes, saliento Gross, que foi o primeiro linguista a não só apontar o problema da extensão dos fenômenos idiossincráticos no léxico e na gramática, mas ainda a enfrentar a tarefa de explicitá-los extensivamente;¹² Levin, que realizou um trabalho comparável (embora muito diferente em detalhes) para o inglês; e Borba *et al.*, cujo dicionário se mostrou uma fonte útil de idéias e exemplos.

¹¹ Não sei exatamente o que Vilela quer dizer com uma classificação “pragmática” dos verbos. Talvez tenha algo a ver com a “valência pragmática”, de Růžička (1977; *apud* Welke, 2005); mas não me parece que fatores contextuais possam interferir na valência (que é um sistema de classes), embora certamente afetem as possibilidades de ocorrência ou não de complementos em casos específicos.

¹² Com seus colaboradores do *Laboratoire Automatique de Documentation Linguistique*.

Capítulo 7: Papéis temáticos

7.1. Preliminares

Os papéis temáticos apresentam um problema espinhoso, que requer discussão. Não só se ignora a lista completa dos papéis temáticos, como não se propôs até agora um conjunto confiável de critérios que levem a sua elaboração. Acredito que a investigação das valências verbais pode trazer algum subsídio a essa tarefa. Nas páginas seguintes vou tratar do problema, em uma exposição mais programática do que decisiva.

Na oração, cada sintagma, exceto o verbo, é associado a um papel temático.¹ Por exemplo, a frase *Antônio comeu a pizza* ilustra a construção seguinte:

<i>Antônio</i>	<i>comeu</i>	<i>a pizza</i>
SN	V	SN
Agente		Paciente

Os papéis temáticos aí são Agente e Paciente. O papel temático (PT) é uma relação semântica entre um verbo (e seu predicado) e um complemento (ou adjunto).² Mas, como se verá mais adiante, as relações possíveis são muito variadas, de maneira que freqüentemente nos vemos perdidos quanto à identificação de um PT particular.

Isso em geral é reconhecido, como atestam as passagens seguintes:

a teoria dos papéis temáticos é ainda muito imprecisa. Por exemplo, no atual estágio da teoria não há consenso sobre quantos papéis temáticos específicos existem e quais são seus rótulos.

[Haegeman, 1991: 41]

Não há talvez nenhum conceito na teoria sintática e semântica moderna que seja tão freqüentemente envolvido em uma ampla variedade de contextos, mas sobre o qual haja tão pouco consenso sobre sua natureza e definição quanto o **papel temático**.

[Dowty, 1991: 547]

Anos depois dessas afirmações, a teoria continua imprecisa.

A lista inicial de papéis temáticos foi proposta por Gruber (1965), e adotada por Jackendoff (1972); até hoje, a maioria dos lingüistas se louva nessa lista, com as respectivas definições, como base para seu trabalho. Posteriormente, houve algumas tentativas de refinar o sistema, das quais as mais importantes são a de Dowty (1989, 1991) e a de Jackendoff (1990), além de C&J. Mas me parece que esses autores deixam

¹ E às vezes mais de um, como se verá. Há restrições: a negação verbal (*ele não veio*) por exemplo, não recebe nada que se possa chamar “papel temático”. Mas não conheço razão válida para restringir os papéis temáticos aos chamados “complementos”; ou seja, o papel temático Lugar ocorre tanto em *Sérgio morou em Recife* (tradicionalmente um complemento) quanto em *Sérgio morreu em Recife* (tradicionalmente um adjunto).

² A diferença entre complemento e adjunto é discutida no capítulo 6.

escapar alguns aspectos importantes da questão, pelo menos quando ela é encarada do ponto de vista descritivo.

O papel temático não se identifica com a relação que estabelecemos entre um evento ou estado e uma coisa; é, antes, uma relação codificada na língua de forma esquemática, eventualmente elaborada pelo falante em situações concretas. Assim, ao ouvir uma frase como *a menina abriu a porta*, o receptor estabelece uma função para a menina, incluindo o elemento “ação voluntária”, ao contrário de *o vento abriu a porta*, onde não se pode falar de “ação voluntária”, porque sabemos que o vento não é um ser vivo. No entanto, tudo indica que o papel temático do sujeito das duas frases é o mesmo, definido esquematicamente como “causador direto”. Embora as duas relações sejam nitidamente distintas, a gramática reúne as duas em uma relação única, mais esquemática. A elaboração da relação fica por conta do conhecimento de mundo do usuário, e não é codificada nas formas sintáticas da língua.

De agora em diante, vou me referir às relações elaboradas chamando-as **relações conceptuais temáticas**, ou **RCTs**, para distingui-las dos **papéis temáticos**, que são as unidades gramaticalmente relevantes. Correspondentemente, usarei aspas para falar das relações conceptuais “agente”, “paciente”, “causador” etc. Os papéis temáticos serão distinguidos por inicial maiúscula: Agente, Paciente etc. Podemos definir um papel temático como um grupo de (uma ou mais) RCTs que se comportam de maneira gramaticalmente semelhante. Espera-se, evidentemente, que o número de papéis temáticos seja limitado, e muito provavelmente seu elenco varia de língua para língua. Já as RCTs são muito mais numerosas, representando possivelmente pontos em um contínuo; além disso, como não são dependentes de codificação lingüística, devem ser em grande medida as mesmas para falantes das diferentes línguas.

Aqui, mais do que em qualquer outra passagem deste livro, é preciso ter em mente o caráter descritivo, e portanto provisório, da proposta. A concepção de “papel temático” aqui adotada é bastante superficial, e deixa de lado muitas possibilidades interessantes. Em particular, acho que a análise dos papéis temáticos em traços (“controle”, “causação” etc.), como se encontra nas propostas de Dowty (1991) e de Cançado e Franchi (2003), deverá ser eventualmente explorada. Mas também acho que essa exploração vai depender da disponibilidade de dados sistematizados e em grande quantidade, o que só pode ser obtido admitindo-se um conceito provisório de “papel temático” (e outros conceitos) e partindo daí para levantamentos do léxico.

Uma palavra sobre os termos adotados: estou adotando a nomenclatura implícita em Jackendoff (1990: 48), onde ele distingue a “estrutura argumental” da “estrutura conceptual”. Os papéis temáticos pertencem à primeira, as relações conceptuais à segunda. A extensão do termo “relação conceptual” é tremendamente ampla, mas aqui só nos interessam as relações temáticas, que definem os participantes do “pequeno drama” que, no dizer de Tesnière, é expresso por cada oração. Por outro lado, é claro que os papéis temáticos são também relações conceptuais (embora esquemáticas); mas como tenho que escolher uma maneira simples e direta de me referir a esses dois tipos de relação, fiquem os termos **RCT** e **papel temático**, entendidos como se definiu acima.

7.2. Por que e para que papéis temáticos?

Vou começar elencando as razões pelas quais a definição dos papéis temáticos é de importância crucial para os estudos de gramática.

Entre as tarefas básicas da gramática estão as de

(a) explicitar as estruturas formais possíveis na língua; e

(b) relacionar cada uma dessas estruturas (e cada detalhe de cada estrutura) com interpretações semânticas correspondentes.

Isso, hoje, não é controverso, sendo aceito em todas as teorias gramaticais; e decorre diretamente da idéia (igualmente não controversa) de que o objetivo da gramática é explicitar a relação som / significado (ou conceito / imagem acústica).

Assim para dar um exemplo, o português comporta uma estrutura que, formalmente, se expressa assim:

[1] SN V

ou seja, um sintagma nominal seguido de um verbo, e nada mais. Essa é uma das estruturas sintáticas existentes da língua, e sua explicitação preenche uma parte da tarefa (a) acima.

Muitas descrições, principalmente as elaboradas pelos estruturalistas de linha bloomfieldiana, se limitam a fornecer a lista dessas construções definidas formalmente. Um exemplo é Allerton (1982), que dá a estrutura [1] incluindo a função sintática, assim:

[2] SUBJECT + V

[Allerton, 1982: 145]

Tratando apenas de estruturas oracionais – isto é, deixando de lado os sintagmas menores (SN, SAdj, SAdv) e os períodos compostos, Allerton lista um total de 30 estruturas. Adaptando-as para o português, teríamos desde estruturas só com o verbo (*choveu*) até estruturas complexas com o verbo e três complementos (o sujeito, mais outros dois: *o menino pagou 10 reais pelo carrinho*). Esse tipo de análise nos diz algo sobre a estrutura da língua, evidentemente. Mas, como veremos, é insuficiente.

Note-se que fórmulas como [1] e [2] não nos dizem nada a respeito do significado dessas construções – ou seja, deixam de lado a tarefa (b), tão importante que a tarefa (a) pode ser encarada como um preliminar a ela – isto é, estudamos a forma **para** podermos descrever seu relacionamento com o significado. Observe-se, por exemplo, que a construção formal descrita em [1] (e em [2], na notação de Allerton) pode ser associada com representações semânticas bem diferentes, como

[3] O gato fugiu.

[4] O gato morreu.

[5] O gato cegou.

Em [3], falamos de um gato que praticou uma ação. Em [4] e [5], o gato sofreu um processo. Ou seja, a relação conceptual temática (RCT) é diferente em [3] e em [4] – [5], e também é diferente o papel temático.

Essa associação de sintagmas com RCTs é, evidentemente, essencial para que essas frases sejam compreendidas; ela faz parte da representação semântica que o receptor constrói a partir das formas que o emissor lhe passa. Por isso, ao estabelecermos a lista das construções do português, não nos limitamos à informação formal, como fez Allerton, mas acrescentamos também informação de carácter semântico – não apenas os papéis temáticos, mas principalmente eles. Desse modo, e admitindo que as RCTs “agente” e “paciente” representam papéis temáticos distintos, [1] deve ser representada como (pelo menos) duas construções, a saber:³

[6] X V (“construção intransitiva”; exemplo, [3])
 Agente

[7] X V (“construção ergativa”; exemplos, [4] e [5])
 Paciente

Dizemos então que as construções do português são definidas em termos **simbólicos** – isto é, em termos de relação forma / significado. Até onde se pode ver atualmente, os papéis temáticos são o ingrediente principal na distinção semântica entre as diversas construções da língua. Existem outros, mas não nos interessam no momento.

Encarando o problema sob outro ângulo, note-se que o conhecimento das diversas construções, definidas como se ilustrou acima, é uma parte essencial do conhecimento que o falante precisa ter da língua, a fim de poder utilizá-la corretamente. Mas não basta conhecer as construções: se uma pessoa ouve uma seqüência analisável como SN + V, ela precisa ainda saber se esse SN vai ser “agente” ou “paciente” (ou outra coisa qualquer). Por exemplo, ao ouvir

[5] O gato cegou.

temos que saber se o gato cegou alguém ou se ficou cego – note-se que ambas as alternativas são possíveis, em termos de mundo real.

Aqui, o falante se vale de seu conhecimento das subcategorias verbais – ou seja, das **valências** dos verbos. Ao aprender um verbo, temos que aprender uma grande quantidade de informações – a que conjugação pertence, se é regular ou irregular, seu significado e assim por diante. E também temos que conhecer sua valência – isto é, em quais **construções** ele pode aparecer.

Ora, uma coisa que aprendemos é que *cegar* pode ocorrer na construção **ergativa**, definida em [7]; por isso, quando aparece apenas com sujeito, esse sujeito

³ Quando me refiro a construções em termos gerais, adoto a notação exposta em 4.5, onde ‘X’ representa o sujeito e/ou o sufixo de pessoa-número.

deve ser entendido como “paciente”.⁴ Como resultado, interpretamos [5] como significando que o gato ficou cego.

Já o verbo *fugir* não pode aparecer na ergativa; aparece, em vez, na construção intransitiva, o que nos faz entender

[3] O gato fugiu.

como exprimindo uma ação praticada pelo gato. É claro que com *fugir* é de qualquer maneira impossível entender o gato como “paciente”; mas mesmo com o verbo *matar* (que tem “agente” e “paciente”), ainda temos que entender o sujeito como “agente”, porque *matar* não ocorre na construção ergativa:

[8] Essa onça já matou.

A diferença de interpretação entre [5] e [8] depende de nosso conhecimento da diferença que existe entre os verbos *cegar* e *matar*. Essa diferença se exprime em termos das construções em que ocorrem, e elas diferem em termos dos papéis temáticos que cada uma atribui ao sujeito. Por isso é que a definição dos papéis temáticos é tão importante para o estudo gramatical.

Sabe-se que os papéis temáticos não se vinculam apenas a verbos. Temos sintagmas como

[9] A invasão americana

onde o nominal *americana* exprime o “agente” de uma ação expressa por outro nominal, *invasão*. Já em

[10] Destruição ambiental

ambiental exprime o “paciente” de um evento, *destruição*.

Além disso, há outras RCTs que se verificam dentro do sintagma nominal e que podem ser consideradas papéis temáticos; por exemplo, em

[11] Um engenheiro simpático

simpático exprime uma “qualidade”, atribuída ao engenheiro. A meu ver, temos aqui uma RCT funcionalmente análoga às que vimos examinando. Neste livro nos ocupamos exclusivamente das RCTs (e dos papéis temáticos) associados a verbos.

A lista das construções (definidas em termos de classes + funções sintáticas + papéis temáticos e provavelmente alguns outros fatores semânticos) nos fornece uma boa parte da gramática da língua. A lista nos diz coisas como:

⁴ Mais precisamente, **pode** ser entendido como paciente. A construção ergativa parece ser sempre ambígua em princípio, e em muitos casos a outra acepção aparece com clareza: por exemplo, *essa doença cega*, com sujeito “agente”.

(a) a ordem dos elementos na oração é limitada: podemos ter SN “**agente**” + verbo + SN “**paciente**” (*a menina comeu os bombons*), ou então SN “**paciente**” + SN “**agente**” + verbo (*os bombons, a menina comeu*) mas não * SN “**agente**” + SN “**paciente**” + verbo (* *a menina os bombons comeu*);

(b) o SN pós-verbal (objeto direto) pode ter o papel de “paciente” (*a menina comeu os bombons*), de “meta” (*o alpinista atingiu o pico*), de “coisa possuída” (*a menina tem um cachorrinho*), mas nunca de “agente”;

(c) o “agente” pode ser expresso por um SN sujeito (*a menina comeu os bombons*) ou por um SN precedido de *de* (*a menina apanhou da mãe*);

(d) o primeiro caso (“agente” sujeito) é muito mais freqüente (ocorre com um número muito maior de verbos) do que o segundo (“agente” regido por *de*).

Ou seja, a lista nos fornece uma grande fatia da estrutura da língua – em especial no que diz respeito às estruturas sintáticas e às relações simbólicas. Aprender uma língua é, em grande parte, aprender suas construções.⁵

7.3. Conceito e codificação

7.3.1. Como a sintaxe vê o significado

Uma vez estabelecida a importância dos papéis temáticos para a análise gramatical, vamos considerar o problema de como defini-los. Em particular, temos que estabelecer critérios para distinguir relações conceptuais e papéis temáticos.

Jackendoff expressa a distinção entre relações conceptuais e papéis temáticos de maneira bastante clara, ao afirmar que

a “estrutura argumental” pode ser concebida como abreviatura da parte da estrutura conceptual que a sintaxe consegue “ver”.

[Jackendoff, 1990: 48]

Portanto, a tarefa que se coloca é a de definir com a maior precisão possível as indicações de como a sintaxe “vê” a estrutura conceptual.

Quando falamos de Agente, Paciente ou Lugar, não estamos nos referindo a relações tais como observadas na realidade, mas tais como representadas na língua. Isso já foi reconhecido por Jackendoff (1990), que define, por exemplo, *angry* em

[12] John stayed angry. ‘John continuou com raiva’

⁵ Não quero dizer que as construções devam ser aprendidas (e enumeradas) como uma lista, uma a uma. Certamente existem regras que abreviam essa tarefa, enumerando as construções de maneira mais econômica. Neste trabalho, entretanto, vou apenas lidar com listas, entendidas como uma maneira preliminar de enumerar as construções possíveis.

como um Lugar (*location*) abstrato – isto é, embora a relação conceptual seja claramente diferente de em *John continuou em casa*, a língua trataria os dois casos como se fossem a mesma coisa. Em outra passagem, Jackendoff (1990: 25-26) menciona um critério para a identificação gramatical de relações que são distintas no mundo real. Na versão de Taylor, o critério se formula assim:

Consideremos o seguinte:

- (1)a. I went from the hotel to the airport. [‘Eu fui do hotel para o aeroporto’]
- b. The inheritance went from George to Philip. [‘A herança passou (*lit.* ‘*foi*’) de George para Philip’]
- c. The light went from green to red. [‘O semáforo passou (*lit.* ‘*foi*’) do verde para o vermelho’]

Essas expressões designam eventos em três domínios diferentes: movimento no espaço, transferência de posse e mudança de estado. No entanto, o mesmo verbo, *go*, é usado em todas as três frases, em associação com as mesmas preposições, *from* e *to*. [...]

Poderíamos dizer que (1b) e (1c) ilustram como a mudança de posse e a mudança de estado são entendidas metaforicamente em termos de um domínio experiencialmente mais básico, a mudança de lugar.

[Taylor, 2003: 506]

Essa intuição de Jackendoff necessita de uma validação empírica ampla, que até hoje, que eu saiba, não foi feita.

A distinção entre a estrutura argumental (lingüística) e a estrutura conceptual (não-lingüística), que é clara no texto de Jackendoff, é negada por muitos cognitivistas atuais. Me parece, entretanto, que estes se deixam levar por uma posição assumida *a priori*, a de que os princípios que regem o funcionamento da língua são semelhantes aos que regem o funcionamento da cognição em geral – haveria um isomorfismo entre os dois planos ou, melhor dizendo, uma falta de distinção entre eles, que seriam na verdade um único plano. Mas existe, a meu ver, evidência contra essa hipótese: tudo indica que existem distinções discretas, que se superpõem às distinções naturais (que são talvez contínuas). Veremos mais adiante vários exemplos de categorias gramaticais discretas que correspondem a categorias conceptuais contínuas. Logo, nada nos impede de imaginar que o mesmo se dê no caso particular dos papéis temáticos; acho que a tentativa é válida, e pode dar resultados positivos.

7.3.2. Diferenciando os papéis temáticos

Um problema metodológico que se plantea de imediato é: que critérios nos autorizam a definir um papel temático, diferenciando-o de outros papéis gramaticalmente relevantes na língua? Por que, e quando, podemos ajuntar diferentes relações conceptuais em um papel temático único? Ou, colocando a questão em termos informacionais: o que é que permite ao receptor depreender as relações conceptuais no sinal recebido?

Para começar o processo, o receptor dispõe do seguinte:

- (a) o verbo e sua classificação;
- (b) os eventuais complementos do verbo, com suas classes e funções sintáticas.

A classe do verbo inclui sua valência; e a função de cada complemento se define através de fatores concretos como, por exemplo, a posição que ocupa na oração.

Podemos chamar a isso “informação gramatical”, que está presente no sinal formal e no conhecimento da língua; daí será necessário retirar, de algum modo, os papéis temáticos que ligam o verbo a seus complementos. Por exemplo, se a frase for

[13] Zé matou uma onça.

o receptor tem o verbo *matar* (que ele conhece e sabe como se categoriza); tem *Zé*, identificado como um SN com a função de sujeito (pelo procedimento exposto no capítulo 4), e tem *uma onça* como SN pós-verbal. Falta, no que nos interessa, estabelecer os papéis temáticos desses dois SNs.

Mesmo sabendo que *matar* exige, por seu significado, um “agente” e um “paciente”, não podemos estabelecer quem é quem com base apenas nessa informação, porque é perfeitamente plausível que a onça mate Zé, em vez do contrário: isso não contradiz as propriedades nem da onça nem de Zé. Mas a frase [13] não comporta ambigüidade, e tem que ser entendida com Zé como Agente. Ou seja, nesse exemplo, não podemos prescindir das funções sintáticas.⁶ Fora as funções sintáticas, nada em *Zé matou uma onça* nos diz quem é “agente” e quem é “paciente”. De alguma forma temos que especificar que com o verbo *matar* o sujeito é Agente e o SN pós-verbal é Paciente, o que se faz incluindo esses papéis temáticos na valência do verbo, donde a necessidade de incluir as funções semânticas Agente e Paciente como parte da estrutura da língua, e não apenas como relações conceptuais, parte do nosso conhecimento do mundo.

Para ver um caso diferente, vamos comparar [13] e [14]:

[13] Zé matou uma onça.

[14] O incêndio matou uma onça.

Aqui temos, no plano conceptual, uma diferença evidente: em [13] o “agente”, Zé, é entendido como o causador voluntário do evento descrito, mas em [14] o incêndio não realizou o evento por vontade própria; podemos dizer que ele é o “causador”, não propriamente o “agente”. Alguns autores distinguem essas duas relações também na gramática, propondo dois papéis temáticos distintos, o de Agente e o de Causador. Como veremos, essa solução não é adequada.

Voltemos à pergunta feita acima: dadas as frases [13] e [14], o que é que permite ao receptor distinguir o “agente” do “causador”? Aqui as funções sintáticas não podem ajudar, porque ambos têm a mesma função (sujeito); e o verbo é o mesmo. Mas há um fator extragramatical em jogo, que é o decisivo: um incêndio **não pode** ser “agente”, porque (pelo que sabemos dos incêndios) não tem vontade própria. Note-se que essa

⁶ Relembro que as funções sintáticas incluem, por exemplo, a simples ordem do sintagma em relação ao verbo.

informação não se encaixa nas categorias (a) e (b) vistas acima; na verdade não tem nada a ver com nosso conhecimento da língua portuguesa, mas antes com nosso conhecimento do mundo.

Ou seja, a informação em questão é fornecida por fatores independentes de uma presumível distinção de papéis temáticos. O resultado é que neste caso não há necessidade de definir dois PTs distintos: se o sujeito de [13] e o de [14] forem ambos analisados como tendo o papel de Agente, isso não causará perda de informação, porque o sentido final depende em parte de fatores pragmáticos, e estes nos informam que o incêndio causou a morte da onça de maneira diferente da de Zé. Naturalmente, devemos entender Agente como uma relação mais esquemática, incluindo tanto a relação conceptual de “agente” (voluntário) quanto a de “causador” (inanimado). A diferença entre as RCTs “agente” e “causador” não corresponde a uma oposição formalmente codificada pela língua; ela é apreendida, quando necessário, de fatores extragramaticais, referentes basicamente a nosso conhecimento do contexto imediato e do mundo. Marcar essa diferença na estrutura gramatical é ser desnecessariamente redundante.

Chegamos a essa conclusão a partir do exame de apenas dois exemplos, e naturalmente é preciso validá-la para a totalidade da língua. Mas já temos um roteiro para a investigação, que pode nos ajudar a obter uma resposta final. Se se verificar que a situação examinada acima é típica, então concluiremos que a distinção entre “agente” (volitivo) e “causador” (não-volitivo) não é gramaticalmente relevante em português. Na seção 8.6 voltarei ao problema para formular os critérios com mais cuidado e exemplificar mais amplamente sua aplicação.

7.3.3. Relações conceptuais como elaboração dos papéis temáticos

A diferenciação que estamos fazendo entre “relações conceptuais” e “papéis temáticos” corresponde à oposição entre um contínuo de relações minimamente distintas (as relações conceptuais) e um conjunto de relações discretas, esquemáticas (os papéis temáticos), das quais as relações conceptuais representam elaborações.⁷ Assim, o papel temático Agente é realizado, no espaço conceptual, como “agente volitivo”, “causador” e certamente muitas outras relações; além de [13] e [14],

[13] Zé matou uma onça.

[14] O incêndio matou uma onça.

podemos citar muitos outros casos em que as relações se desviam minimamente umas das outras; em

[15] Zé escreveu um poema.

⁷ As RCTs são elaborações dos PTs, e estes são representações esquemáticas daquelas. Não se deve entender essas relações como direcionais -- trata-se apenas de formular uma relação de esquematicidade entre dois grupos de conceitos.

a ação de Zé é diferente tanto da de Zé em [13] quanto da do incêndio em [14]: aqui o Agente causa o aparecimento de algo que não existia antes. Outros exemplos são

[16] Cientista muda teoria.

[17] Sarcófago muda teoria.⁸

É claro que em [16] entendemos *cientista* como “agente”, e em [17] *sarcófago* é apenas o causador de alguém (“agente”) mudar a teoria: um causador inanimado, mas indireto, ao contrário do incêndio em [14].

Isso pode ser multiplicado quase ao infinito, fornecendo um número ilimitado de relações. Uma tese aceita neste trabalho é que não é necessário representar todas essas distinções na descrição gramatical, e que um número talvez relativamente grande, mas limitado, de papéis temáticos discretos pode dar conta dos fenômenos observados. Como vimos, os papéis temáticos fornecem uma delimitação esquemática; e outros fatores, em particular fatores pragmáticos, elaboram a relação até o nível de detalhe necessário para a comunicação do momento. Por exemplo, em [16] – [17] a diferença semântica nos é fornecida pelo conhecimento que temos sobre cientistas e sarcófagos; e note-se que em [16] seria possível entender *cientista* de maneira paralela à de *sarcófago* em [17], caso o cientista fosse apenas o objeto de uma observação (a teoria é de que todo alemão tem pé chato; Schmidt, cientista alemão, não tem pé chato; logo, temos que mudar a teoria).⁹

7.3.4. Caráter esquemático dos papéis temáticos

O grande problema atual relativo à delimitação dos papéis temáticos é, como já aponte, o da validação empírica. Vou a seguir fazer algumas sugestões sobre como se pode começar a enfrentar essa tarefa. Devo enfatizar que se trata apenas de algumas indicações sucintas do que deve constituir um programa de pesquisa muito amplo.

Jackendoff (1983), seguindo a intuição básica de Gruber (1965), levantou a hipótese de que as relações temáticas são representadas na língua de forma esquemática, e que fica por conta do usuário da língua elaborá-las em relações conceituais específicas. Assim, as relações de tempo seriam codificadas da mesma maneira que as de lugar; elementos contextuais funcionariam para discriminar “tempo” de “espaço”. Existe evidência em favor dessa hipótese, a saber, o conhecido paralelismo entre as expressões de tempo e de espaço:

[18] Machado morreu em Niterói. / em 1908.

[19] A estrada vai de Niterói a Teresópolis. / A festa foi das 8 às 11 horas.

[20] O asfalto foi estendido até Barbacena. / O contrato foi estendido até 2009.

⁸ Manchete do *Estado de Minas*, 11-10-2004.

⁹ Essa concepção esquemática de Agente se opõe à posição de muitos linguistas; ver por exemplo Givón (1990: 565), onde o Agente é definido como “volicional”.

etc.

No entanto, em todos os casos existem formas especializadas, de maneira que a expressão do tempo não é propriamente idêntica à de espaço. Pode-se citar palavras como *aqui*, *lá*, que só exprimem espaço, e *hoje*, *agora*, que só exprimem tempo. Isso, entretanto, não significa automaticamente que seja necessário distinguir “tempo” e “lugar” como papéis temáticos diferentes. Pode-se argumentar que a diferença decorre de informações estranhas à valência dos verbos. Ou seja, assim como não se entende *em Niterói* como um sintagma de significado temporal, mas sempre como “lugar” (dado o significado locativo da palavra *Niterói*), só se entende *aqui* como “lugar” igualmente por força da semântica dessa palavra. Em outras palavras, mesmo se atribuirmos o mesmo papel temático, Lugar, a *em Niterói*, *aqui*, *em 1908* e *hoje* nas frases acima, continuará sendo possível distinguir “tempo” de “lugar” sem margem a dúvida, com base no significado interno dos itens léxicos e no nosso conhecimento do mundo.

O paralelismo sintático entre as expressões de “tempo” e de “lugar” vai muito longe; assim, as mesmas funções sintáticas são usadas para ambas as relações conceptuais nos seguintes casos:

complemento de preposição:

[21] Ele morreu em Belém.

[22] Ele morreu em 1908.

e assim com muitas outras preposições:

[23] Eu só acordei depois das duas horas.

[24] Eu só acordei depois de Juiz de Fora.

[25] Essa reunião vai até as duas horas.

[26] Essa estrada vai até Juiz de Fora.

Existem, é verdade, algumas preposições especializadas: *perante*, *sob*, *debaixo de*¹⁰ são exclusivamente locativas.

objeto direto:

[27] Os alpinistas atingiram o pico.

[28] Meu avô atingiu os noventa e cinco anos.

Um caso talvez menos claro é o da relação conceptual de “posse”, que, segundo alguns, é outra das elaborações da relação temática de Lugar. Essa relação inclui um

¹⁰ Mas não *em cima de: ele chegou em cima da hora*.

possuidor e uma coisa possuída; segundo a hipótese, o possuidor seria expresso pelo papel de Lugar, e a coisa possuída por Tema (ou Localizando). Seguindo essa hipótese, analisa-se a frase [29]

[29] Beth tem uma boneca

atribuindo-lhe uma estrutura semântica que, no essencial, coloca *Beth* com o papel temático de Lugar e *uma boneca* com o de Localizando;¹¹ ou seja, é como se disséssemos que a boneca está “em (na posse de) Beth”: a frase é processada elaborando os papéis temáticos dentro do campo conceptual da “posse”. Como veremos, não há propriamente sinalização formal que contenha essa instrução – o receptor deverá inferir o campo conceptual adequado a partir de coisas como a implausibilidade de alguém dizer que a boneca está “em” Beth.¹²

Esse processo de base pragmática é necessário porque mesmo a presença do verbo *have*, ou o português *ter*, não é suficiente para sinalizar o campo conceptual a ser ativado, por causa de frases como

[30] Ipanema tem 250 restaurantes.

onde certamente o campo conceptual é o de “lugar” mesmo.

As frases [29] e [30] mostram que há pelo menos uma estrutura comum às duas interpretações (“lugar” e “posse”). Agora temos que perguntar até onde vai esse paralelismo – e quanto maior for a área de coincidência sintática, mais plausível será a identificação de “lugar” e “posse” como um papel temático único.

Temos, em primeiro lugar, estruturas sintaticamente idênticas, inclusive com os mesmos verbos e / ou preposições. Além de [30] e [31],

[30] Ipanema tem 250 restaurantes.

[31] Beth tem cinco bonecas.

podemos citar o modificador regido da preposição *de*, que não só expressa a “posse”, como também o “lugar”:

[32] A boneca de Beth

[33] Os restaurantes de Ipanema

O paralelismo entre [32] e [33], a bem dizer, não impressiona muito, em vista da imensa variedade de acepções veiculadas pela preposição *de*. Mas mesmo assim constitui um argumento extra: podemos dizer que há na língua pelo menos duas estruturas que expressam “posse” **ou** “lugar”.

Uma terceira estrutura é apenas parcialmente idêntica, porque as preposições são diferentes:

¹¹ Localizando é o elemento cuja localização é expressa.

¹² Não haveria problemas com a interpretação locativa se a frase fosse *Beth tem uma úlcera*.

[34] A boneca é de Beth.

[35] O restaurante é em Ipanema.

Ambas as orações têm a estrutura **SN + *ser* + SPrep**, mas a primeira tem *de* e a segunda *em*. Não deixa de haver paralelismo, mas nessa construção as preposições são especializadas.

Não vou prosseguir com a exemplificação, porque aqui só pretendo dar sugestões. Naturalmente, será necessário procurar outras construções que expressem “posse” e “lugar”, submetendo-as a um exame comparativo como o que fiz acima.¹³

Além disso, será útil comparar a expressão dessas duas relações conceptuais com outras relações, o que nos poderá dar uma perspectiva mais adequada da importância do eventual paralelismo sintático encontrado entre as expressões de “posse” e de “lugar”. Peguemos, por exemplo, a relação conceptual “agente”; é fácil mostrar que há menos paralelismo entre a realização sintática do “agente” e a de “lugar” do que o que existe entre “lugar” e “posse”. Assim, o “agente” se codifica sintaticamente em construções como

[36] Beth comeu a pizza.

[37] A pizza foi comida por Beth.

[38] A destruição da cidade pelos invasores.

[39] O assassinato do presidente.

Dessas quatro construções (três, se resolvermos identificar [37] e [38]), apenas uma é formalmente semelhante a uma das anteriores (a saber, [39] e [32] / [33]). Nesse caso único, o “agente” é realizado sintaticamente da mesma maneira que o “lugar”. Será que vale a pena identificar “agente” e “lugar” como elaborações do mesmo papel temático?

Certamente não: há razões sintáticas e semânticas contra essa análise. Sintaticamente, a área de coincidência é pequena, e envolve justamente a suspeitíssima preposição *de*, de valor semântico extremamente vago. E semanticamente é difícil imaginar analogias conceptuais entre a idéia de “agente” e a de “lugar”. Temos que concluir que “agente” e “lugar” são elaborações de papéis temáticos distintos.

É claro que essa conclusão não vale grande coisa enquanto não se fizer um levantamento amplo – começando, digamos, por uma lista de todos os recursos sintáticos oferecidos pelo português para exprimir “posse”, “lugar” etc. Essa lista nos fornecerá meios de apurar o grau de semelhança sintática e léxica entre as codificações das relações em causa, e de avaliar a plausibilidade de que a diferenciação conceptual (a elaboração que resulta nos conceitos de “lugar”, “posse” etc.) se realiza utilizando o conhecimento

¹³ Como se verá, mantenho a distinção entre “posse” e “lugar” para efeitos de notação das diáteses, dentro do princípio de que é melhor fazer distinções de mais do que de menos.

do mundo do receptor. O que temos aqui, portanto, não é propriamente uma resposta – coisa que não é possível dar nas atuais condições – mas um programa de pesquisa.¹⁴

7.4. Contra os protótipos-como-contínuos

7.4.1. Protótipos

Alguns lingüistas se colocam numa posição oposta à que descrevi na seção anterior, lançando mão da noção de uma concepção especial de **protótipo**, que aparece com certa frequência na literatura (ver resenha em Taylor, 1989). Para eles, não haveria papéis temáticos discretos, mas apenas regiões no contínuo conceptual que poderiam ser considerada mais típicas; cada exemplo se colocaria mais ou menos próximo dessa região, sendo, correspondentemente, mais ou menos prototípico. Essa é apenas uma das noções de protótipo; outros autores, que rejeitam o caráter contínuo das categorias, entendem os protótipos como configurações típicas, definidas em termos discretos, porém mais importantes em termos estruturais ou estatísticos. Isto é, uma categoria será prototípica se ocorrer em muitas configurações da língua e/ou se tiver grande incidência no léxico ou na gramática.

Não tenho objeções de princípio contra essa segunda concepção de protótipo, antes tendo a considerá-la adequada, embora ainda careça de confirmação empírica. Entenda-se, então, que a crítica que se segue se aplica apenas aos protótipos entendidos como “pontos de saliência em um contínuo”. Vou argumentar que os protótipos, assim entendidos, não são a maneira adequada para lidar com o problema da delimitação dos papéis temáticos, nem com categorias lingüísticas em geral.

A noção de “protótipo” em questão se refere a contínuos, limitados por zonas preenchidas por “uma penumbra de exemplos não tão típicos” (Taylor, 2002:177). Trata-se, pois, de uma questão de grau: um objeto pertence a determinada categoria em determinado grau. Isso certamente se aplica a muitos aspectos da cognição propriamente dita, mesmo porque nosso sistema cognitivo precisa incorporar um alto grau de flexibilidade, a fim de dar conta da tarefa de categorizar e interpretar um mundo de informações novas que surgem a todo momento.¹⁵ Esse caráter contínuo certamente reflete a realidade (pelo menos em muitos casos), mas introduz um sério problema metodológico, o de identificar e distinguir os graus de prototipicidade: como distinguir um item que é 40% membro de uma categoria de outro item que é 50% membro da mesma categoria? Aqui só se pode usar um julgamento intuitivo, e o resultado é que o uso da noção de “protótipo” nesse sentido muitas vezes não vai além da observação anedótica, sendo muito difícil de incorporar em análises minimamente rigorosas.

Em casos dependentes de levantamento léxico, entretanto, o problema tem saída. Aqui podemos lançar mão da extensão léxica de cada fenômeno, no espírito da proposta de Gross, que, referindo-se aos estudos gerativos até os anos 70, comenta que estes

¹⁴ Observe-se como aqui (e em muitas outras áreas) o levantamento empírico e o desenvolvimento da teoria precisam caminhar lado a lado.

¹⁵ Como foi mostrado nos trabalhos de Berlin e Kay (1969), Rosch (1973; 1975; 1978) e muitos outros, para a categorização de elementos não lingüísticos.

[...] só levam em conta pequeno número de exemplos. Descobriram um grande número de fenômenos novos, mas não permitem avaliar a extensão desses fenômenos para uma língua dada.

[Gross, 1975: 20]

A extensão dos fenômenos gramaticais não costuma ser levada em conta, mas tudo indica que tem importância para a explicação do conhecimento e uso da língua, e o grande mérito do trabalho de Gross é justamente chamar a atenção para esse aspecto da análise.¹⁶ Ou seja, a extensão de um fenômeno deve ser incluída como parte dos fatos de que uma descrição gramatical precisa dar conta; e, em muitos casos, é possível reduzir a prototipicidade a uma questão de estatística lexical.

7.4.2. *Protótipos e a descrição gramatical*

A noção de “protótipo” foi inicialmente proposta para lidar com a categorização de noções não lingüísticas, tais como “copo” / “caneca”, “azul” / “verde”, “pássaro” etc. Mostrou-se que as pessoas têm uma espécie de imagem prototípica dessas categorias, e conseguem avaliar cada exemplo concreto em termos do grau de semelhança entre ele e a imagem. Aqui, aparentemente, temos que lidar com contínuos, que inclusive apresentam certo grau de elasticidade conforme a pessoa, o momento, o contexto de percepção ou as necessidades comunicativas.

Mas quando consideramos a área gramatical, as coisas podem ser diferentes. Há uma grande diferença entre julgar se um objeto é um caneco ou uma xícara e julgar se uma palavra é substantivo ou adjetivo. Outro problema é que as distinções são claramente discretas em outras partes da lingüística: assim como não existe na fonologia portuguesa um fonema intermediário entre **a** e **o** aberto – embora isso seja foneticamente possível, e ocorra na fala¹⁷ – pode-se argumentar que não existe gradação contínua entre substantivo e adjetivo, por exemplo. Os casos aparentes de gradação contínua, como os “graus de nominalidade” observados por Ross (1972), seriam na verdade um artefato da simplicidade excessiva do sistema tomado como ponto de partida. Se para descrever determinado grupo de itens só temos duas categorias (digamos, “substantivo” e “adjetivo”), então encontraremos um grande número de casos intermediários, que podem dar a impressão de um contínuo. Mas se o sistema for mais complexo, como precisa ser,¹⁸ poderemos descobrir que o intervalo é preenchido por um número maior de categorias discretas.

A evidência disponível, embora esporádica, é toda favorável ao caráter discreto, e não contínuo, das unidades gramaticais. Por exemplo, sabemos que o tempo é um contínuo: não há quebras naturais entre 10.000 A.C. e 10 minutos atrás. No entanto, a língua manipula esse contínuo estabelecendo pontos de referência: em português temos

¹⁶ Tendo a criticar outros pontos, principalmente o fato de Gross não ter hierarquizado as propriedades em termos de importância gramatical, o que resultou em uma matriz excessivamente complexa (600 propriedades x 12.000 verbos). Por outro lado, essa complexidade serve para ilustrar a grande extensão do comportamento idiossincrático dentro do léxico.

¹⁷ Por exemplo, é comum pronunciar a última vogal de *carnaval* arredondada.

¹⁸ Ver Perini *et al.* (1996) para uma proposta nesse sentido.

comeu para uma ação acontecida em qualquer momento do passado; e *tinha comido* para uma ação anterior a outra igualmente passada, e só isso.¹⁹ Não há representação gramatical do contínuo temporal: não existe forma intermediária entre *como* e *comeu*, no que diz respeito à referência temporal. E como fazemos para situar um evento no tempo com precisão? Aqui temos que utilizar outros recursos que não o tempo verbal, tipicamente adjuntos adverbiais como *ontem*, *há mais de dois meses*, *em 1990* etc.

Não é difícil multiplicar os exemplos: os substantivos são todos masculinos ou femininos, os adjetivos podem ou não ocorrer antes do substantivo etc. Há casos em que se poderia ver um contínuo, como no caso das conjugações verbais: *amar* é da primeira, *comer* é da segunda, mas *dar* tem formas que parecem da primeira (*dava*) e formas que parecem da segunda (*desse*); mas essas formas são localizáveis com toda a precisão, e o verbo *dar* pode ser caracterizado sem ambigüidades como tendo um grupo de formas regulares da primeira, um grupo de formas da segunda, mais um grupo de formas que não se encaixam em nenhuma conjugação. Ou seja, não temos aqui um contínuo, mas um conjunto complexo de unidades discretas. Temos que concluir que a evidência disponível na gramática aponta para um sistema de unidades discretas, e não para um contínuo.

Essa posição se contrapõe à adotada por alguns lingüistas que, partindo dos trabalhos de Rosch (1973; 1975; 1978), Berlin & Kay (1969) e outros, estenderam a noção de protótipo-como-contínuo não só às possibilidades referenciais dos itens léxicos (onde é provável que ela se aplique), mas também à definição das unidades formais e semânticas relevantes em gramática. Segundo esses autores,

os falantes das línguas naturais criam categorizações de objetos lingüísticos da mesma maneira que criam categorizações de objetos naturais e culturais.

[Bybee e Moder, 1983: 267]

Ora, isso pode ser verdadeiro ou não – é impossível decidir *a priori*. Baseando-me na evidência disponível, acredito que a melhor hipótese de trabalho é a de que as categorias gramaticais (classes, regras, traços formais e semânticos) são discretos, mesmo quando expressam contínuos do mundo real.

A hipótese que desejo explorar aqui é a de que a língua utiliza, como matéria prima para a formulação de regras, não as distinções naturais, mas distinções demarcadas em termos especificamente lingüísticos – distinções discretas, semanticamente esquemáticas. A situação no campo da gramática seria análoga à que se verifica na fonologia, estabelecida com clareza desde os anos 30. Trata-se, evidentemente, de uma hipótese, que pode ser refutada; mas merece uma chance, porque é tão plausível quanto seu oposto (eu diria até mais plausível, em vista do que se sabe a respeito).

Essa hipótese se harmoniza com a idéia de que a linguagem tem um caráter maximamente econômico. O sinal lingüístico explícito fornece a informação mínima necessária para que o receptor, de posse de seu conhecimento do mundo, possa construir uma paisagem mental que faça sentido. A razão principal dessa economia é a limitação dos recursos que a linguagem coloca à nossa disposição, frente à infinidade de conceitos que a experiência nos oferece.

¹⁹ *Comia* é igualmente passado, e não se distingue, em termos de referência temporal, de *comeu*.

7.4.3. *Perspectiva descritiva*

Devo esclarecer que não há **provas** de que essa hipótese corresponda à realidade da língua, assim como não há provas em contrário. Talvez os limites entre as categorias gramaticais sejam realmente contínuos, mas, primeiro, não há maneira de verificar isso, pelo menos no atual estado do conhecimento; e, depois, uma das maneiras de abordar a questão é justamente partir de uma hipótese de trabalho que lida com distinções discretas, e eventualmente – se for o caso – descobrir que essas distinções se multiplicam, indicando a possível existência de um contínuo. Estou adotando, portanto, uma hipótese de trabalho que me parece a que melhor se presta a um estudo empírico como o que se faz necessário atualmente em gramática. Os contínuos, para serem estudados, requerem algum tipo de quantificação; mas, embora seja fácil estudar a altura dos cavalos (usando uma fita métrica, por exemplo), não se inventou ainda uma maneira de medir com precisão o grau de nominalidade de uma palavra. As propriedades gramaticais discretas (traços) fornecem um padrão de medida; e, além do mais, talvez nem sequer se tenha que ir além deles, no caso, muito provável a meu ver, de que as categorias gramaticais sejam de fato discretas. Em resumo, para evitar uma controvérsia que, nas atuais circunstâncias, não pode levar a nada, adoto uma hipótese de trabalho sem compromisso necessário no plano teórico. Trata-se de um instrumento que nos permitirá colher dados e sistematizá-los: um instrumento de **descrição**, não de explicação ou de teorização.

Nessa perspectiva, a noção de “protótipo”, se for usada, terá que ser definida de maneira diferente (e essa é a opção de muitos lingüistas).²⁰ Em vez de “uma penumbra de exemplos não tão típicos”, teremos uma série de casos que se distinguem por traços discretos; e os que são menos típicos o são apenas pelo fato de se aplicarem a pequeno número de itens léxicos. Para dar um exemplo, dentre os verbos da primeira conjugação há muitos (*andar, amar, fechar, pintar*) que se conjugam de maneira perfeitamente regular; e há um número muito menor de irregulares (*odiar, ansiar, dar*) que apresentam irregularidades mais ou menos extensas. De certo modo, podemos dizer que *amar* é “mais prototípico” do que *odiar*. No entanto, não cabe aqui entender os protótipos como contínuos: não há zona nebulosa entre regulares e irregulares (é possível dar a lista completa de cada grupo), e o próprio “grau” de irregularidade pode ser formulado com toda a precisão: *odiar* só é irregular em algumas formas: *odeio* (e não **odio*); *dar* é irregular na maior parte do paradigma.

A se seguir as sugestões acima, a investigação, partindo de um nível inicial muito concreto, poderá estabelecer papéis temáticos altamente esquemáticos, na medida em que se puder distinguir oposições previsíveis (a partir de outras informações) de oposições imprevisíveis (e portanto portadoras de informação própria).

A questão tem a ver com a função delimitativa dos significantes. Já nos anos 30 Troubetzkoy falava de “regras que recortam o mundo dos significados em fragmentos que elas organizam” (Troubetzkoy, 1939: 2). Os papéis temáticos têm esse papel delimitativo. O problema é como averiguar isso, e como apurar os detalhes com um mínimo de rigor. Em outras palavras, o grau de generalidade com que se define um papel temático vai depender de pesquisa de caso por caso.

²⁰ Devo essa observação a Beth Saraiva.

7.5. Previsibilidade

7.5.1. O que se pode prever?

Um ponto muito discutido, em geral inconclusivamente, é até que ponto a valência de um verbo é previsível a partir de detalhes de seu significado. Essa previsibilidade, desconfio, se aplica muito mais a que papéis temáticos vão ocorrer com cada verbo; a previsibilidade é muito mais baixa no que diz respeito à representação formal de cada papel temático. Assim, *espancar* tem (mais ou menos necessariamente) um Agente e um Paciente; *apanhar* também tem Agente e Paciente – isso decorre de sua semântica; e o mesmo para *bater*. Mas as representações formais são diferentes para os três verbos, e isso é imprevisível, como mostram os exemplos abaixo:

[40] Rosa espancou o vizinho.

Ag Pac

[41] O vizinho apanhou de Rosa.

Pac Ag

[42] Rosa bateu no vizinho.

Ag Pac

Estamos aqui, portanto, pelo menos restringindo a hipótese de Levin de que

qualquer classe de verbos cujos membros funcionam paralelamente quanto às alternâncias de diáteses deve ser uma classe semanticamente coerente: seus membros devem compartilhar pelo menos algum aspecto de seu significado.

[Levin, 1993: 14]

Ao que parece, isso vale para a presença de argumentos (eventualmente realizados formalmente, mas sempre presentes na representação semântica final), mas não para a realização formal desses argumentos. Ou seja, mesmo que um verbo exija, em virtude de seu significado, um Agente e um Paciente, não se pode prever se ambos os papéis temáticos precisam aparecer explicitamente, nem como eles vão ser codificados sintaticamente. Aqui há certamente generalizações, mas o quadro é muito complexo e em grande parte imprevisível a partir da semântica do verbo.²¹ É esse fato que faz com que o estudo das valências verbais tenha que incluir, necessariamente, o levantamento de listas de itens léxicos.

Note-se que os mesmos verbos *espancar*, *bater* e *apanhar* se associam a relações conceptuais como “tempo” etc. Aqui as representações formais são bem mais previsíveis, porque “tempo” se codifica sempre da mesma maneira (advérbio ou sintagma preposicionado). Daí, o verbo não tem o que dizer a respeito, e “tempo” não tem lugar na valência. A valência será tanto mais interessante quanto mais imprevisível for a

²¹ Acredito que as listas de Levin (1993) confirmam amplamente essa afirmação. Note-se que, a rigor, a afirmação de Levin fala de prever a semântica a partir da valência, não o oposto; mas em geral se entende as duas coisas.

representação dos papéis temáticos. O mesmo vale, em princípio, para a presença dos próprios papéis temáticos, caso esta seja imprevisível – mas tendo a achar que é previsível.²²

7.5.2. Arbitrariedade da representação sintática

A dificuldade de prever a forma pela qual cada argumento é codificado merece mais algum comentário. Não parece possível estabelecer correlações categóricas entre papéis temáticos e funções sintáticas – por exemplo, “o agente é sempre representado pelo sujeito” e coisas assim. Isso vai variar para cada caso, e o grau de previsibilidade também varia. Pode-se dar o exemplo de *gostar* e *adorar*, em que o primeiro verbo exige preposição e o segundo não, apesar de o papel temático do complemento ser o mesmo:

[43] Gostei do concerto.

[44] Adorei o concerto.

Outro exemplo: para muitos brasileiros, *alugar* e *emprestar* se constroem diferentemente, embora os papéis sejam aparentemente idênticos. Assim, dizemos²³

[45] Vou emprestar um casaco para a Cristina.

[46] Vou tomar emprestado (*e não *emprestar*) um casaco da Cristina.

mas

[47] Vou alugar um apartamento para a Cristina.

[48] Vou alugar um apartamento da Cristina.

Mais um exemplo: os verbos *embrulhar* e *enrolar* ocorrem em pares de frases como

[49] O coveiro embrulhou o corpo no lençol.

[50] O coveiro enrolou o lençol no corpo.

onde os papéis temáticos de *o corpo* e *o lençol* são os mesmos nas duas frases, embora as funções sintáticas estejam trocadas.

Mais um exemplo vem dos verbos *varar*, *atravessar*, *trespassar*, *cravar*, *fincar*, *espetar*, *enfiar*, que descrevem aproximadamente o mesmo evento, e que certamente se

²² É preciso tomar cuidado com o perigo de circularidade ao se definir a semântica do verbo a partir de seus PTs, e depois dizer que os PTs são previsíveis a partir da semântica. Dowty (1991: 555) chama atenção para isso.

²³ Em Minas.

associam aos mesmos participantes, mas que aparecem em construções distintas. Assim, *espetar* e *fincar* ocorrem em duas construções:

[51] Jorge espetou / fincou a lança no dragão.

[52] Jorge espetou / fincou o dragão com a lança.

Enfiar e *cravar* só ocorrem na primeira dessas construções:

[53] Jorge enfiou / cravou a lança no dragão.

[54] * Jorge enfiou / cravou o dragão com a lança.

E *atravessar*, *trespassar* e *varar* só ocorrem na primeira:

[55] * Jorge varou / atravessou / trespassou a lança no dragão.²⁴

[56] Jorge varou / atravessou / trespassou o dragão com a lança.

Nas duas construções acima os papéis temáticos são atribuídos aos mesmos sintagmas, mas as funções sintáticas variam segundo o verbo.

Isso não exclui a possibilidade de estabelecer correlações entre papel temático e função sintática; mas estas terão que ser de natureza estatística, ou seja, tendências e não regras categóricas. Não há dúvida de que o Paciente é com muita frequência representado por um objeto direto, e o Agente por um sujeito. Por outro lado, a Causa (caso seja mesmo um papel temático) se representa de maneira bastante sistemática através de determinadas preposições, tipicamente *por causa de*.²⁵

7.6. Critérios de delimitação dos papéis temáticos

Agora vou resumir as considerações acima formulando critérios gerais de delimitação de papéis temáticos – critérios que nos digam quando devemos unir duas relações conceptuais (RCTs) em um papel temático único, e quando devemos mantê-las separadas.

7.6.1. Critério de necessidade

O critério principal é o seguinte:

Critério de necessidade

²⁴ Esta frase me parece menos inaceitável do que [53]; é possível que alguns falantes a aceitem.

²⁵ Às vezes apenas *de: o velhinho morreu de gripe*.

Duas RCTs devem ser distinguidas em dois papéis temáticos independentes quando a atribuição dessas RCTs não puder ser derivada de informação extragramatical ou do significado dos itens léxicos envolvidos.

Ou seja, vamos nos apoiar principalmente na procura de imprevisibilidades dos complementos – sua presença, assim como o modo como são codificados formalmente. O objetivo é consignar na descrição gramatical aquilo que não pode ser depreendido nem da semântica dos itens léxicos e das estruturas, nem do conhecimento geral do mundo e do contexto. Seguimos aqui o exemplo já citado da pesquisa fonológica; nessa área, a contribuição de Troubetzkoy e seus colegas foi a de formular com precisão, no âmbito da fonologia, a relação entre os fatos naturais e os estruturais. No fundo, o princípio básico é a previsibilidade: algo como “a estrutura (gramaticalmente) relevante é feita de imprevisibilidades”. Nas seções seguintes veremos como se justifica e como se aplica esse critério de delimitação dos PTs.

7.6.2. *Aplicando o critério de necessidade*

7.6.2.1. Agente e causador

Vamos começar pelos casos já estudados. A diferença entre as RCTs “agente” (entendido como volitivo) e “causador” se realiza em

[13] Zé matou uma onça. (“agente”)

[14] O incêndio matou uma onça. (“causador”)

[16] Cientista muda teoria. (“agente”)

[17] Sarcófago muda teoria. (“causador”)

Em todos esses exemplos há elementos que nos permitem inferir corretamente a RCT (se a dúvida for entre “agente” e “causador”). Digamos, por hipótese, que se juntem essas duas RCTs em um só PT; nesse caso, o sujeito dessas quatro frases seria uniformemente marcado como Agente. No entanto, essa marcação não nos obriga a interpretar as relações como a mesma em todos os casos, porque Agente corresponde a uma relação esquemática, que pode ser elaborada como “causador voluntário”, “causador (não voluntário) direto” ou “causador (não voluntário) indireto” – essas três RCTs são suficientemente semelhantes para serem incluídas em uma só representação esquemática. Assim, em [13] e em [16] os traços de *Zé* e de *cientista* (animados, humanos) nos levarão a elaborar Agente como “causador voluntário”. Já em [14] isso não é possível, porque *o incêndio* não poderia desempenhar um papel voluntário; nesse caso o entendemos como “causador”, em geral direto, pelo que sabemos dos incêndios; e em [17] o sarcófago receberá a RCT elaborada de “causador indireto”. Tudo isso provém de nosso conhecimento dos objetos e das situações envolvidas, não de nossa competência léxica.

Vemos que não há necessidade de distinguir essas diferentes modalidades de “agente” (ou de “causador”) no nível argumental (gramatical); as modalidades serão devidamente produzidas através do processo de elaboração do único papel temático esquemático Agente.

Naturalmente, a conclusão é provisória, pois se baseia no exame de um número limitado de exemplos; temos agora a tarefa de corroborá-la com outros exemplos, e quanto mais e mais variados, melhor. Alguns casos que podem ser úteis são:

[57] A porta foi aberta por Renata.

[58] A porta foi aberta pelo vento.

[59] Os estragos de Renata

[60] Os estragos do vento

Os exemplos [57] – [60] fornecem evidência adicional do paralelismo sintático entre as RCTs “causador” e “agente (voluntário)”.

Temos também contra-exemplos aparentes, como o seguinte:

[61] A professora tentou quebrar a janela.

[62] ?? O vento tentou quebrar a janela.

Mas o que marca [62] como mal-formada é nosso conhecimento do mundo. O verbo *tentar* engloba em seu significado o ingrediente “tentativa voluntária”. Como o vento não é capaz de volição, a frase [62], uma vez cabalmente interpretada, inclui uma contradição. A (relativa) inaceitabilidade de [62] se explica porque o que conhecemos a respeito do vento não se harmoniza com parte do significado do verbo *tentar*, que exige um agente dotado de volição.

Agora, se definíssemos as duas RCTs como PTs independentes, teríamos que atribuir a inaceitabilidade de [62] a **dois** fatores, em princípio não relacionados: o primeiro é o mesmo visto acima (o que conhecemos a respeito do vento não se harmoniza com parte do significado do verbo *tentar*, que exige um agente dotado de volição). Esse fator não podemos evitar, porque não podemos negar as afirmações independentemente motivadas de que (a) o vento é inanimado, logo não tem volição; e (b) o significado do verbo *tentar* inclui o ingrediente “tentativa voluntária”. Mas haverá ainda uma outra razão, a saber, *tentar* exige “agente” e não apenas “causador”; e *vento* só pode ser “causador”, nunca “agente”. Em outras palavras, estaremos representando duplamente as propriedades semânticas do verbo *tentar*: no plano do nosso conhecimento do mundo (saber o que é “tentar” envolve saber que isso só se faz “de propósito”); e no plano léxico, onde estipulamos que o verbo *tentar* entra em construções com “agente”. A informação é a mesma, representada de duas maneiras. A solução que unifica as RCTs “agente volitivo” e “causador não volitivo” em um só papel temático evita essa duplicação.

Conclui-se que o caso de [61] e [62] não constitui contra-exemplo à hipótese de que se deve reunir essas duas RCTs em um PT único. Note-se que, neste caso, não é particularmente difícil encontrar um núcleo semântico comum às duas relações cognitivas: trata-se, nos dois casos, de um elemento causador do evento expresso pelo verbo. Naturalmente, o traço “volição” fica excluído da definição do papel temático. Temos que sustentar que, quando presente na semântica de uma frase, esse traço provém de processamento independente da estrutura argumental – em particular, processamento baseado no conhecimento extralingüístico.

Há uma certa controvérsia na literatura quanto a isso. Alguns autores preferem distinguir o “agente voluntário” do “causador (não voluntário)” – por exemplo Fillmore que, no dizer de Whitaker-Franchi, define o Agente como

a função desempenhada por um ente animado que é responsável, voluntária ou involuntariamente, pela ação ou pelo desencadeamento dos processos [Whitaker-Franchi, 1989: 63].

Já Chafe (1970) assume uma posição idêntica à minha, observando que

[...] somente alguns substantivos são elegíveis para ocorrer como agentes de verbos de ação [ele dá exemplos como *rope* ‘corda’ e *dish* ‘prato’, que não são]. A habilidade de um substantivo de ocorrer como agente depende de sua especificação semântica como uma coisa que tem o *poder* de fazer alguma coisa, uma coisa que tem uma força própria, auto-motivada. Em ampla medida [...] esse conceito de auto-motivação coincide com o conceito de animação [...] Mas parece haver alguns substantivos que não são animados mas que podem ocorrer como agentes:

- (5)a. The heat melted the butter. ‘o calor derreteu a manteiga’
- b. The wind opened the door. ‘o vento abriu a porta’
- c. The ship destroyed the pier. ‘o navio destruiu o cais’

[...] Com base nessas observações, sugiro que um substantivo pode ser especificado opcionalmente como *potente*, no sentido de que tem, ou é imaginado como tendo, potência interna própria.

[Chafe, 1970: 109]

Concordo com essas observações de Chafe, mas acho importante acrescentar um argumento que me parece decisivo, ou seja, não se observa diferença nenhuma de codificação sintática entre o Agente animado e o Agente apenas “potente”.

7.6.2.2. Lugar e tempo

Um par altamente suspeito são as relações conceptuais de “lugar” e de “tempo”, já examinadas em 8.3.4. Como vimos, essas duas RCTs devem ser reunidas sob um papel temático único, que denominei Lugar (com maiúscula: um papel temático, não uma relação conceptual). A diferença semântica entre as relações é nítida, mas por outro lado

há também semelhanças suficientes, que vêm certamente do fato de que tendemos a ver o tempo em termos espaciais: uma data é um “lugar no tempo”, e um período transcorrido é uma “trajetória”.²⁶

O que é mais importante nos exemplos vistos é que a diferença entre “lugar” e “tempo” em cada caso provém de fatores extragramaticais. Assim, a diferença entre [63] (“lugar”) e [64] (“tempo”) vem do que sabemos de um pico e de uma idade:

[63] Os alpinistas atingiram o pico.

[64] Meu avô atingiu os noventa e cinco anos.

7.6.2.3. Lugar e estado

À luz do que se discutiu acima, vamos considerar as relações conceptuais de “lugar” e de “estado” (ou “qualidade”), identificadas no plano argumental por Jackendoff (1990), na citação de Taylor vista em 8.3.1. Segundo Jackendoff, o “estado” pode ser entendido como um “lugar” metafórico, ou seja, posicionamento em um espaço não físico: seria uma extensão da relação “mais básica” de “lugar”. O exemplo de Jackendoff é

[12] John stayed angry. ‘John continuou com raiva.’

onde *angry* ‘com raiva’ seria uma ocorrência do papel temático “lugar”.

Mas o fato de ser possível ligar “lugar” e “estado” através de uma extensão de sentido não nos autoriza a identificar essas duas relações como um papel temático único. O critério não é apenas o da proximidade semântica, mas também, e principalmente, o da maneira como a língua codifica essas duas relações. Elas serão identificadas apenas no caso de serem tratadas de maneira semelhante pela língua. Jackendoff sugere um critério, quando aponta que “o mesmo verbo, *go*, é usado em todas as três frases, em associação com as mesmas preposições, *from* e *to*”. Mas veremos na seção 8.6.4 que a simples ocorrência com os mesmos verbos (e itens léxicos em geral) não vale como critério de identificação – é essencial que as duas relações conceptuais em pauta sejam codificadas sintaticamente de maneira semelhante.

Desse modo, a proposta de Jackendoff deve ser avaliada em termos de formas de codificação formal, não em termos da lista de verbos envolvidos. No caso, podemos dizer que relevante não é o fato de aparecer o mesmo verbo (*stay*, por exemplo), mas o fato de que “lugar” e “estado” se substituem na mesma função: *stay angry*, *stay home*.

Por outro lado, a representação do sintagma não é idêntica, pelo menos em geral. O “estado” é tipicamente representado por nominais: *estou furioso*, *fiquei triste* etc. Só

²⁶ É possível que a semelhança semântica entre tempo e espaço se baseie em uma relação codificada, expressa por Lakoff e Johnson (2002) como “o tempo é um objeto parado e nós nos movemos através dele na direção do futuro” (acrescente-se que podemos também estar parados nesse espaço). Aliás, não concordo com a posição de Lakoff e Johnson de tratar essas conexões codificadas na estrutura da língua paralelamente às metáforas, que prefiro entender como conexões estabelecidas no momento, dependentes de um contexto e válidas apenas para ele.

em raros caso encontramos sintagmas preposicionados: *ele está na fossa, está na maior tristeza*. O oposto se dá com o “lugar”; só dois ou três nominais da língua expressam a RCT “lugar”: *próximo, distante e longínquo* são os únicos que me ocorrem. Isso não deixa de ser uma diferença importante, a ser eventualmente avaliada.

Como se vê, reina a incerteza neste campo. Mas acho que ainda assim há alguma esperança de unificar essas duas relações conceptuais em um papel temático, como propõe Jackendoff.²⁷ Os exemplos abaixo parecem constituir um contra-argumento, mas é só aparente:

[64] O presidente passou por Belo Horizonte.

[65] O presidente passou por louco.

Aqui temos *por* + SN com o mesmo verbo, e até onde posso ver na mesma função sintática. Mas as relações conceptuais são claramente distintas: em [64] é “lugar”, e a representação é típica (Prep + SN); em [65] é “estado”, e a representação é igualmente típica (Adjetivo). A diferença semântica observada é nítida, e pode parecer impossível de representar sem lançar mão de uma diferença de papéis temáticos.

Mas há uma maneira de fazer isso, jogando com a polissemia da preposição *por* e com o conteúdo semântico do SN que a segue. Sabemos que *por* pode significar, entre outras coisas, “lugar por onde” (como em [64]) e “introdutor de uma qualidade” (em [65]). Como *Belo Horizonte* é o nome de um lugar, e não de uma qualidade, necessariamente *por* assume em [64] a primeira acepção; e o oposto para [65]. A partir daí é possível inclusive explicar as duas acepções do próprio verbo *passar*: seguido de “lugar por onde” ele é um verbo de movimento, mas seguido de “qualidade” ele significa “ser considerado”.

Recapitulando, o que vimos é que se “estado” e “lugar” forem representados no nível gramatical por um único papel temático, Lugar, será possível ainda assim derivar corretamente os significados dos exemplos vistos. Por exemplo, temos

[66] João está com raiva.

[67] João está no banheiro.

Em ambas as frases o complemento (*com raiva, no banheiro*) recebe o papel temático de Lugar. Esse papel temático pode ser interpretado, no espaço semântico, como “lugar” ou como “estado” (ou, ainda, como “tempo”). Mas o significado de *com raiva* exclui todas as interpretações, exceto a de “estado”; o significado de *no banheiro* nos obriga a selecionar “lugar” e assim por diante. E no caso de

[68] O jogador nasceu em 1980.

o significado de *em 1980* só nos permite entender esse complemento como “tempo”.

²⁷ Jackendoff, claro, faz a proposta para o inglês; mas nada nos impede de aproveitar a sugestão para tentar a mesma coisa em português: o contexto de descoberta é livre.

A subcategorização é portanto a única maneira de garantir a interpretação correta do sujeito de [69] e [70]. Mas depende da existência dos dois PTs separados – ou seja, a diferença entre esses PTs é gramaticalmente necessária, o que a coloca como um caso positivo do critério de necessidade.

7.6.2.5. Meta e beneficiário

Muitos autores distinguem um papel temático de Beneficiário, que ocorreria em casos como os seguintes:

[72] Maria recebe gorjetas.

[73] Maria ama seus irmãos.²⁹

e também em casos como

[74] Maria deu um presente para sua neta.

Segundo Whitaker-Franchi (que cita Chafe, 1970) o beneficiário seria

o elemento interessado no resultado de uma ação verbal, que altera relações de posse, obtenção, perda e transferência de objetos. De um modo geral, o beneficiário se distingue na estrutura superficial pela presença de uma preposição “dativa” e pela posição sintática de objeto indireto em português.

[Whitaker-Franchi, 1989: 65]

Aqui vou analisar o constituinte marcado nos exemplos [72] e [74] como Meta, e o de [73] como Causador de Experiência, por duas razões: primeiro, porque há certa vizinhança semântica: o “beneficiário” de [72] e [74] pode ser figurado como uma “meta” do presente ou das gorjetas, e o de [73] é bem claramente um “causador de experiência”. E, depois, há um paralelismo estrutural com frases como

[75] Maria recebeu uma pedrada. (cf. [72])

[76] Maria jogou a bola para a neta. (cf. [74])

Quanto a [73], acredito que *seus irmãos* é um “beneficiário” por razões ligadas ao significado específico do verbo *amar* (“ser amado é um benefício”), que desaparece quando se muda o verbo, como em

[77] Maria detesta seus irmãos.³⁰

²⁹ Exemplos de Whitaker-Franchi (1989).

³⁰ Nesses casos, Wenceslau (2003) fala de “maleficiário”.

[78] Maria avistou seus irmãos.

O que essas frases têm em comum é, primeiro, exprimirem uma causação de experiência (que pode ser sentimental ou sensorial, conforme o verbo); e, segundo, uma estruturação sintática idêntica (Experienciador sujeito, Causador de Experiência objeto). Acredito que isso é suficiente para que se analisem de maneira paralela.

Whitaker-Franchi cita a presença de uma preposição “dativa” como evidência para o papel temático de Beneficiário, mas exemplos como [76] mostram que essas preposições (presumivelmente, *a* e *para*) ocorrem em casos em que a RCT de “beneficiário” não está presente. E quanto à posição sintática do objeto indireto, não sei exatamente do que se trata, pois o objeto indireto tradicional se caracteriza não pela posição (que é a mesma do objeto direto), mas pela presença, justamente, de uma preposição.

A análise que adoto pode ser inserida na proposta geral de que esses papéis temáticos seriam basicamente “localísticos” (cf. Anderson, 1971, seguido nesse particular, e para esses papéis, por Jackendoff). Mas note-se que a conclusão vale para os papéis temáticos discutidos, não em geral para todos os papéis temáticos.³¹

7.6.2.6. Sumário parcial

Vamos nos deter um momento para examinar as implicações teóricas do critério de necessidade.

A interpretação semântica de uma oração aceitável inclui certo número de RCTs, cada uma delas associada a um complemento.³² A análise precisa explicar como se produz essa associação, dado que o receptor só dispõe do sinal concreto (em última análise, fonético) e de seu conhecimento prévio da língua e do mundo; ou seja, ele não conhece previamente as relações que unem as diversas partes do sinal, mas tem que construí-las com base nos conhecimentos disponíveis.

As RCTs poderiam, em princípio, ser atribuídas com base em conhecimento do mundo, pelo menos em alguns casos. Assim, se a oração é

[79] Toninho tomou o chope.

o conhecimento do mundo poderia fornecer a informação de que Toninho é “agente”, e o chope é “paciente”, dado que o oposto se chocaria com o que sabemos das pessoas e dos chopos.

Mas não é assim que as coisas se passam. Se o processo fosse o descrito acima, a frase seguinte seria tão normal como [79], e sinônima dela:

[80] * O chope tomou Toninho.

³¹ Whitaker-Franchi não argumenta em favor de sua decisão, limitando-se a dizer que “não adot[a] a generalização que se poderia obter com a hipótese localística” [p. 66].

³² Existe também a possibilidade de um complemento se associar a mais de uma RCT, mas vamos deixar isso de lado por ora.

A inaceitabilidade de [80] mostra que há alguma coisa na estrutura gramatical que governa a atribuição de papéis temáticos – e, conseqüentemente, de RCTs. No caso, é evidente que se trata da função sintática dos SNs envolvidos. Podemos dizer que com *tomar* o sujeito é o Agente – ou melhor, que a valência de *tomar* inclui uma diátese em que o sujeito é marcado como Agente.

Note-se que aqui simplesmente não há outra saída: a diferença formal entre [79] e [80] está nas funções sintáticas dos dois SNs, que são trocados; e semanticamente as RCTs também acabam trocadas. Temos que concluir que as RCTs de [79] dependem da função sintática, sem interferência significativa do conhecimento do mundo ou do contexto.³³

O que acabamos de fazer para a oposição “agente” e “paciente” foi mostrar que se trata de uma oposição necessariamente vinculada a elementos gramaticais. Portanto, itens léxicos como *tomar* e muitos outros devem ser marcados como atribuindo PTs diferentes às diferentes funções em questão. Neste caso, como há verbos que exigem isso, a conclusão é final: a gramática do português precisa incluir a distinção entre “agente” e “paciente” – o que se faz através da postulação dos PTs Agente e Paciente, que são posteriormente elaborados de acordo, agora sim, com base no conhecimento do mundo ou do contexto.

Para representar tudo isso em uma análise formal, precisamos de (a) funções sintáticas; (b) papéis temáticos e (c) relações conceptuais. Como se vê, os PTs representam uma espécie de estágio intermediário entre os fatos fonéticos (as funções sintáticas, entendidas basicamente como relações de ordenação de sintagmas) e os fatos semânticos. Ou seja, a postulação dos PTs aumenta o grau de abstração da análise, e precisa ser justificada.

A justificação está implícita no que já dissemos, mas creio que vale a pena resumi-la aqui, porque os pressupostos da análise descritiva (a Hipótese da Sintxe Simples) requerem que a relação entre forma e significado seja a mais concreta possível. Sabemos que uma relação direta é muitas vezes impossível, mas isso tem que ser mostrado de cada vez.

Para isso, podemos examinar os exemplos

[70] O professor assustou.

[81] Esse professor não perdoa.

Aqui simplesmente não é possível fazer o relacionamento direto entre a forma e o significado. A função sintática de *o professor* é a mesma nas duas frases (sujeito, ou SN pré-verbal), de modo que não podemos derivar dela a diferença de RCTs que se verifica: “paciente” em [70], “agente voluntário” em [81]. E não há mais nada, na representação formal das frases, que possa ser utilizado para isso. A diferença que sobra é a identidade

³³ Em certos casos, pelo menos, oposições gramaticalmente relevantes também podem ser resolvidas pelo contexto; um exemplo é *Hélio arrancou um dente ontem*, onde o papel temático do sujeito muda conforme Hélio seja um dentista ou um paciente. No entanto, como sabemos, a oposição aí presente (Agente x Possuidor) precisa ser distinguida gramaticalmente, por causa de outros exemplos.

dos dois verbos, *assustar* e *perdoar*; evidentemente, teremos que derivar dessa diferença de verbos a diferença de RCTs.

Ou seja, os dois verbos terão que ser marcados de maneira diferente; em termos coloquiais, podemos dizer que “o sujeito de *assustar* é “paciente”, e o de *perdoar* é “agente voluntário”. Aqui estamos estabelecendo uma relação direta entre RCT e verbo.

Para essas frases, a afirmação é válida. Mas no caso de *perdoar* ela falha em

[82] Essa doença não perdoa.

Aqui já não temos “agente voluntário”, mas um “causador”, não dotado de volição. Ou seja, não podemos amarrar o significado do sujeito de *perdoar* à RCT “agente voluntário”, porque aqui já não funciona. Seria o caso de marcar o verbo *perdoar* duplamente, podendo ter sujeito “agente” ou “causador”?

Mas aqui interfere um novo fator: o caráter voluntário ou não do sujeito de *perdoar* é previsível a partir de nosso conhecimento do mundo. Se sabemos o que é uma doença, e o que é um professor, podemos deduzir que o sujeito de [81] é voluntário, mas o de [82] não é. Ou seja, não vale a pena marcar o verbo *perdoar* duplamente. Basta atribuir ao sujeito desse verbo um PT esquemático (Agente), que será eventualmente elaborado em diversas RCTs, conforme nosso conhecimento do referente do sujeito de cada frase.

Em conclusão, parece inevitável que teremos que lidar com o estágio abstrato dos papéis temáticos, intermediário entre a função sintática e a RCT. Para o verbo *perdoar*, o processo de interpretação segue a ordem abaixo:³⁴

função sintática	→	PT	→	RCT
sujeito		Agente		“agente voluntário” “causador” etc.

7.6.3. Critério de semelhança semântica

Além do critério de necessidade, é bem provável que se precise acrescentar um segundo, o **critério de semelhança semântica**. Por exemplo, um argumento que se pode levantar é o de que os papéis temáticos “estado” e “lugar” não deveriam ser identificados porque são muito distantes semanticamente.

Essa objeção tem precedentes em fonologia. Sabe-se que sons que se encontram em distribuição complementar são identificados como realizações de um mesmo fonema.³⁵ Um exemplo clássico do português é a identificação de [a] e [ə], porque o primeiro nunca ocorre em final átono, que é onde ocorre o segundo – ou seja, estão em distribuição complementar. No entanto, [ə] está em distribuição complementar não apenas com [a], mas também com [e], [o] e vários outros segmentos (por exemplo, [p]). Por que então juntar [ə] e [a], e não, digamos, [ə] e [o]? A resposta é que [ə] e [o] são

³⁴ Trata-se de uma ordem abstrata, ou seja, uma relação de implicações; não quero dizer que um receptor realmente perpassa esses estágios em tempo real na ordem dada (ver 5.8).

³⁵ Ou “segmento subjacente”, se se preferir. Acho “fonema” um termo muito conveniente.

mais diferentes do que [ə] e [a]; aqui, então, é a substância fonética que serve de base para a decisão.³⁶ O resultado é que cada fonema (ou segmento subjacente, ou como se chame), embora seja uma unidade estrutural, tem um “núcleo” fonético, de tal modo que todas as suas realizações compartilham certo número de traços.

É verdade que é mais fácil avaliar a semelhança fonética do que a semelhança semântica. O julgamento mencionado acima, de que a diferença entre “estado” e “lugar” é grande demais, é puramente intuitivo; e, dada nossa ignorância dos traços semânticos em geral, vai provavelmente continuar intuitivo por algum tempo. Não obstante, é essencial que a junção de diversas relações conceptuais em um único papel temático seja sujeita a algum controle semântico. Caso contrário, acabaremos identificando “papel temático” e “função sintática”; por exemplo, qualquer objeto direto (independentemente da relação conceptual que exprima) poderá ser associado ao papel de Paciente.³⁷

Por isso, faz sentido procurar o “núcleo” semântico que subjaz a cada papel temático definido. Agente, por exemplo, não pode ser definido simplesmente como o papel atribuído ao complemento de um particípio, regido de *por* (*o carro foi lavado por Joca*). É essencial que se procure um grupo de traços semânticos que caracterizem o Agente, porque é possível que algum *por* + SN não deva ser analisado como tendo o papel de Agente – isso deve ser uma questão empírica, atendendo à finalidade básica do estudo gramatical, que é o de relacionar formas e significados. Corolário dessa posição é que qualquer princípio que relacione automaticamente³⁸ uma forma com um significado, ou vice-versa, é suspeito.

Uma razão para enfatizar a necessidade de semelhança semântica é que um papel temático se entende como um **esquema**, ou seja, como uma unidade de conhecimento, definido por exemplo como

estruturas de dados que representam conceitos genéricos armazenados na memória.

[Rumelhart e Ortony, 1976: 3]

Os esquemas são instrumentos de categorização, permitindo que a mente reconheça estímulos mesmo quando estes, estritamente falando, são inéditos. Assim, ao ver um objeto que nunca vimos, podemos identificá-lo imediatamente como sendo um “pneu”, embora seja em muitos detalhes diferente de todos os pneus que vimos até o momento. Acontece que o objeto em questão tem traços conhecidos suficientes para que o coloquemos na categoria dos pneus. Esses traços constituem justamente o esquema de “pneu”.

Mas é claro que esse processo depende de alguma semelhança entre os diversos representantes da categoria dos “pneus” encontrados no mundo real. Não reunimos em um esquema objetos totalmente diferentes (em seu aspecto, finalidade, forma, cheiro, seja o que for); o esquema (como diz o nome) é uma representação esquemática, que é

³⁶ Em alguns casos há razões morfológicas que motivam a análise; mas isso não acontece em todos os casos, e freqüentemente a decisão precisa ser feita sobre bases exclusivamente fonéticas.

³⁷ Parece que é a esse perigo que Rabin se refere, quando fala de lingüistas que “forçados a fazer ‘funcionar’ os papéis temáticos, [...] gradualmente os esvaziam de conteúdo semântico até que os papéis já não correspondem a nenhum conceito semântico coerente.” (Rabin, 1990: 4)

³⁸ Isto é, **a priori**, independentemente de verificação empírica adequada.

elaborada em cada caso particular pela adição de traços não-essenciais. Ou seja, os esquemas

contêm, como parte de sua especificação, a rede de inter-relações que é aceita como sendo geralmente válida entre os constituintes do conceito em questão. Os esquemas, em certo sentido, representam estereótipos desses conceitos.

[Rumelhart e Ortony, 1976: 3]

Como já vimos em 8.3.4, os papéis temáticos são esquemas, realizados (elaborados) em cada frase individual de maneira às vezes muito claramente diferente, mas sempre incluindo alguns traços em comum, que permitem o reconhecimento dos casos individuais como elaborações daquele papel temático, e não de outro qualquer. Como se vê, a semelhança semântica é um requisito para a análise dos papéis temáticos como esquemas. E essa análise permite a integração da teoria dos papéis temáticos no que se sabe em geral sobre a cognição humana.

Voltando então ao caso do “lugar” e do “estado”, eles poderão ser identificados em um único papel temático dependendo (entre outras coisas, é claro) de se encontrar entre eles alguma semelhança semântica clara, algum núcleo comum de significado. Note-se que essa procura de um núcleo comum precisa ser controlada de alguma forma, porque em princípio não tem limites. É possível ver mudança de lugar no sujeito de *a menina correu*, porque afinal de contas não se pode correr sem mudar de lugar; ou no objeto direto de *o cachorro comeu o bife*, porque o bife se transferiu para o estômago do cachorro. Mas isso não quer dizer, necessariamente, que *a menina* e *o bife* tenham o papel temático de Tema (definido como “o elemento cuja mudança de lugar é afirmada”). Avaliar a semelhança semântica, portanto, não é um problema trivial.

7.6.4. Conteúdo semântico dos sintagmas

Já vimos casos em que o conteúdo semântico dos sintagmas de uma oração é suficiente para eliminar incertezas a respeito da RCT: por exemplo, o caso de

[16] Cientista muda teoria.

[17] Sarcófago muda teoria.

onde apenas em [16] podemos ter uma relação de “agente voluntário”, por causa do significado inanimado de *sarcófago*. A seguir vou expandir um pouco esse ponto, para rebater certas objeções possíveis.

Sugeri acima que as RCTs “lugar” e “tempo” sejam reunidas sob o único papel temático Lugar. As frases abaixo poderiam ser entendidas como contra-exemplo a essa análise:

[83] Eu moro em Campinas / aqui.

[84] * Eu moro hoje.

O problema é que, aparentemente, o verbo *morar* exige um complemento de “lugar”, e não de “tempo”; isso seria razão suficiente para se separar essas duas relações em PTs distintos, já que a exigência deve aparecer na valência do verbo.

No entanto, acredito que os fatos podem ser explicados em termos de filtragem semântica, sem recorrer a uma diferença de PTs entre “lugar” e “tempo”. Para ver isso, temos que examinar o significado intrínseco do verbo *morar* e de palavras como *hoje* e *quando*.

O verbo *morar* existe, por assim dizer, para expressar um tipo de relacionamento entre uma pessoa (geralmente não se aplica a gatos ou aranhas) e um lugar: esse é o núcleo de seu significado. Por isso mesmo, ele co-ocorre quase sempre com um complemento que denota “lugar”. Note-se que aqui não estamos falando de papéis temáticos, mas de relações conceptuais; trata-se não da valência de *morar*, mas de seu significado. Sabemos que o significado dos itens léxicos vai muito além dos papéis temáticos, e que esses detalhes gramaticalmente insignificantes são relevantes para a boa formação das frases. Por exemplo, não podemos dizer

[85] * Entornei o bife no prato.

porque *entornar* se relaciona com um paciente líquido ou em pequenas partículas; mas isso não é **gramaticalmente** relevante, embora seja necessário para explicar a inaceitabilidade de [85].

Voltando então a *morar*: uma exigência desse verbo é a presença de um argumento que exprima “lugar” – a exigência não se refere ao papel temático Lugar, mas à relação conceptual “lugar”, e acontece que *hoje* não tem esse ingrediente em seu significado: é um dêitico de “tempo”, e por isso [85] é semanticamente mal formada.

Uma explicação análoga vale para o caso de *durar*, que exige um argumento que exprima “tempo”; por isso não podemos dizer

[86] * A festa durou aqui.

porque *aqui* não tem o ingrediente semântico “tempo”.

Igualmente, o interrogativo *quanto* tem o ingrediente semântico “tempo”, mas não “lugar”, e *onde* tem “lugar”, mas não “tempo”, o que explica as frases abaixo:

[87] * Quando é que você mora? (*mas Onde é que você mora?*)

[88] * Onde foi que a festa durou? (*mas Quanto foi que a festa durou?*)

Em outras palavras, se uma pessoa sabe o significado de *morar*, não aceita que esse verbo venha acompanhado (apenas) de um complemento de “tempo”, porque nesse caso falta um dos termos da relação de “lugar” que *morar* expressa. Ou seja, aplicamos o critério da necessidade: sempre que for possível explicar a atribuição das RCTs sem utilizar uma diferença de papéis temáticos, essa diferença deve ser dispensada. Concluo que a objeção não se sustenta, e podemos continuar a considerar “tempo” e “lugar” como manifestações conceptuais do mesmo papel temático Lugar.

Resta um possível problema: o complemento de “lugar” e o de “tempo” podem aparecer lado a lado na mesma oração. Isso não seria argumento para distingui-los?

[89] O incêndio aconteceu aqui ontem.

Ainda aqui acho que o contra-exemplo não funciona. Observe-se que, embora não se possa ter na mesma oração mais de um complemento Agente,³⁹ é possível incluir mais de um complemento de “lugar” ou de “tempo” na mesma oração:

[90] O incêndio aconteceu [em Brasília], [perto da Esplanada].

[91] O incêndio aconteceu [ontem], [às duas horas].

Isso só pode acontecer quando os diversos complementos têm referências em parte coextensivas: a Esplanada é parte de Brasília, e às duas horas é parte de ontem. Mas aqui a razão da restrição é evidente: se não fosse assim, a frase daria informação contraditória a respeito do tempo ou do lugar:

[92] * O incêndio aconteceu em Brasília, em Goiânia.

[93] * O incêndio aconteceu ontem, há quatro dias.

Isso mostra que, ao contrário do que acontece com Agente, o papel temático Lugar pode ser atribuído a mais de um complemento na mesma oração. Por conseguinte, a frase [89] não constitui argumento contra a identificação de “tempo” e “lugar” em um só papel temático, que estamos chamando Lugar. Esse papel temático deverá ser definido em termos de **localização no tempo ou no espaço**, no espírito da citação de Jackendoff (1990: 25-26): as relações de “tempo” se codificam em português da mesma maneira que a relação “experencialmente mais básica” de “lugar”.⁴⁰

7.6.5. *Perigos da circularidade*

Um procedimento que deve ser evitado é o de distinguir papéis temáticos com base nas construções nas quais co-ocorrem; por exemplo, definir “paciente” como o papel que ocorre como objeto da construção transitiva e sujeito da ergativa, e “resultativo” como o que ocorre como objeto da transitiva, mas não como sujeito da

³⁹ Para o Paciente, é possível que essa afirmação não seja exata: há casos de diáteses nas quais, ao que parece, mais de um complemento recebe o papel de Paciente. São casos de **construções complexas** (ver 8.7.2).

⁴⁰ Acho justo notar que Jackendoff não foi o primeiro a perceber o papel básico das relações de “lugar” na definição de várias relações gramaticais. Anderson (1971: 6) já propõe uma teoria localística do **caso** (que, para nossos propósitos, corresponde ao papel temático), e cita vários precedentes que remontam até a Idade Média.

ergativa. Isso é o que faz Whitaker-Franchi (1989: 68-69), com base nos seguintes dados:⁴¹

- [94] a. Os vinhateiros quebravam os brotos das videiras. (“paciente”)
 b. Os brotos das videiras quebravam.

- [95] a. Os vinhateiros podavam os brotos das videiras. (“resultativo”)
 b. * Os brotos das videiras podavam.

Isso segundo a autora, mostraria que “paciente” e “resultativo” são papéis temáticos distintos.

Mas esse argumento acarreta distinguir PTs com base nas classes verbais e, ao mesmo tempo, distinguir classes verbais com base nas valências, que por sua vez dependem da distinção de PTs -- incorrendo portanto em circularidade. Ou seja, *os brotos das videiras* seria Paciente em [94a] porque se correlaciona (uma relação transformacional) com o sujeito de [95b], ao passo que o mesmo sintagma seria Resultativo em [95a] porque não existe um correlato ergativo.

A relação transformacional é essencial para o argumento porque, ainda que definíssemos duas construções para [94a] e [95a], a motivação para essa dualidade de construções seria a existência ou não na língua de uma construção ergativa com aquele verbo e com sujeito idêntico ao do objeto da primeira construção (ou construções). Não é possível distinguir as construções de [94a] e de [95a] diretamente porque não há diferença semântica saliente entre os dois casos: Whitaker-Franchi define o paciente como “o objeto afetado ou modificado no processo e nas ações”, e o resultativo como “o objeto resultante do processo ou ação construtivos”. Não vejo essa diferença, nem nenhuma outra, entre os casos de [94] e [95] acima: em ambos os casos *os brotos das videiras* é “o objeto afetado ou modificado”, e os significados dos dois verbos (*quebrar* e *podar*) são muito próximos.

A situação seria outra se houvesse uma diferença nítida entre as relações semânticas entre o SN e o verbo nos dois casos. Aí, poderíamos ter que distinguir diáteses meramente com base nessa diferença -- aplicando o critério de semelhança semântica dado em 8.6.3. Mas não é o caso.

A maneira mais conveniente de representar a diferença observada é subcategorizar diferentemente os verbos: diremos que *quebrar* tem em sua valência as diáteses ergativa e transitiva, e *podar* apenas a transitiva.⁴² O papel temático de *os brotos das videiras* é sempre o mesmo (Paciente) em todos os casos.

7.7. PTs múltiplos

7.7.1. Mais de um PT para o mesmo sintagma

⁴¹ Exemplos de Whitaker-Franchi (1989: 69). Minha crítica ao argumento apresentado pressupõe que o objetivo é descrever as valências verbais.

⁴² Seja ou não essa diferença decorrente de alguma diferença semântica entre os verbos; eu tendo a pensar que não.

Uma possibilidade, já aventada pelo menos desde Jackendoff (1972), é a de que certos complementos tenham mais de um papel temático, cumulativamente. Assim, observa-se que em

[96] O cachorro arrastou para debaixo da mesa.

é necessário entender o cachorro como Agente e, ao mesmo tempo, Tema. Esse fenômeno é muito freqüente, e precisa ser levado em conta.

Em primeiro lugar, gostaria de notar que talvez o problema seja na verdade decorrência de uma concepção deficiente do que vem a ser um papel temático. Dowty (1991) propõe a dissociação dos papéis temáticos em “proto-papéis” (*proto-roles*), de modo que o Agente, por exemplo, seria analisado como um feixe de traços como *controlador, não-afetado, voluntário* etc. Essa possibilidade precisa ser mantida presente, embora eu não possa incorporá-la em minha análise no momento.

Por ora, vamos falar simplesmente de papéis cumulativos. Utilizei a acumulação de papéis para fazer a distinção entre certas construções. Assim, temos a construção **transitiva**, que se define como

Construção transitiva

X	V	SN
Ag		Pac

Exemplo:

[97] A multidão depredou o estádio.

Distingo essa construção da **transitiva de sujeito Tema**, que se define assim:

Construção Transitiva de Sujeito Tema

X	V	SN
Ag		Pac
Tema		Fonte

Exemplo:

[98] A multidão abandonou o estádio.

Como se vê, *a multidão* denota o Agente nos dois casos; mas no segundo temos ainda a informação da movimentação da multidão, informação que não aparece no primeiro caso. Seguindo a prática de Jackendoff (1972: 34 sqq), explico isso atribuindo papéis temáticos duplos aos complementos do verbo no caso da construção de sujeito locado.

Que os fatos são esses não há dúvida; o que se precisa discutir é se a diferença deve ou não ser incorporada à definição das construções. Como optei por maximizar as distinções em caso de dúvida, defini as duas construções como está acima, ou seja, fazendo uso do recurso dos papéis múltiplos.

7.7.2. Os papéis temáticos nas construções complexas

Jackendoff (1972) apresenta casos de papéis temáticos múltiplos realizados de forma mais complexa, organizados em uma espécie de hierarquia de planos semânticos, de modo que na mesma frase podemos ter papéis primários e secundários:

Consideremos o verbo *trade* [‘trocar’], que tem objeto direto, um sintagma opcional com *to* [port. *com*] e um sintagma obrigatório com *for* [port. *por*].

(2.48) Esau traded his birthright (to Jacob) for a mess of pottage.
[Esaú trocou seu direito de primogenitura (com Jacó) por um prato de guisado]⁴³

Essa frase descreve duas ações relacionadas. A primeira é a mudança de mãos do direito de primogenitura de Esaú para Jacó. O objeto direto é Tema, o sujeito é Fonte, e o objeto com *to* é Meta. Há também o que vou chamar de *ação secundária*, a mudança de mãos do prato de guisado na direção oposta. Nessa ação, o sintagma com *for* é Tema Secundário, o sujeito é Meta Secundária e o sintagma com *to* é Fonte Secundária.

[Jackendoff, 1972: 35]

Em casos como esse, naturalmente, não se pode simplesmente atribuir papéis temáticos duplos aos diferentes complementos: Esaú não é simplesmente Fonte e Meta, mas Fonte do direito de primogenitura e Meta do prato de lentilhas. A questão que nos interessa é se esse paralelismo e hierarquização de relações deve ser representado na valência de *trocar*.⁴⁴

Uma análise plausível seria a de analisar estruturas como

[99] Esaú trocou seu direito de primogenitura com Jacó por um prato de lentilhas.

como realização de uma espécie de soma de duas diáteses. Uma delas seria algo como

[100] X V SN *com*SN
 Fonte Tema Meta
 Ag

⁴³ Essa é a versão tradicional em português. A Bíblia ainda especifica que se tratava de um guisado de lentilhas (*Gênesis*, 25: 29-34).

⁴⁴ Outros verbos que ocorrem nessa construção são *comprar*, *vender*, *barganhar* e *permutar*, que podemos denominar “verbos de troca”. Esses são os únicos verbos conhecidos que envolvem essa dualidade de eventos paralelos.

Essa diátese (ou melhor, uma diátese parecida, com preposição diferente) vale, por exemplo, para o verbo *dar*:

[101] O velhinho dava migalhas a / para os pombos.

No entanto, não se conhece nenhuma diátese que corresponda à outra metade, ou seja, uma com a forma

[102]	X	V	<i>com</i> SN	<i>por</i> SN
	Meta		Fonte	Tema

Tenho que concluir que a alternativa de analisar [99] como representante de um cruzamento de duas diáteses simples não funciona. De qualquer modo, um cruzamento de duas diáteses seria, inevitavelmente, uma terceira diátese, já que não se aplica a qualquer verbo.

A segunda saída que vejo seria encaixar parte de [99] em uma diátese simples e derivar o resto da informação da semântica das preposições, mais informação pragmática. Vamos examinar essa possibilidade.

A idéia é argumentar que a primeira parte de [99], a saber,

[103] Esaú trocou seu direito de primogenitura.

é um exemplo da construção transitiva, com *Esaú* como Agente e *seu direito de primogenitura* como Paciente. Os outros sintagmas seriam adjuntos, semanticamente caracterizados através de preposições predicadoras; ou seja, [99] seria tratada de maneira paralela à de

[104] Esaú comeu as lentilhas na cozinha em poucos minutos.

No entanto, é preciso observar que os adjuntos de [104], que exprimem “lugar” e “duração”, são de validade muito geral, e se aplicam a uma gama imensa de eventos e respectivos verbos; sua adição à oração não parece ser governada pelo verbo. Já em [99] os presumíveis adjuntos são característicos dos verbos de troca (listados na nota 45) e de nenhum outro. Por exemplo, o sintagma *com Jacó* só pode exprimir a Meta e a Fonte com esses verbos. Isso parece indicar que a derivação da relação semântica desses sintagmas a partir da semântica da preposição ou da informação pragmática não é possível: eles subcategorizam nitidamente os verbos, e só assumem aquele significado com uma classe claramente definida, a dos verbos de troca.

Ao que tudo indica, portanto, é inevitável analisar [99] como realização de uma diátese própria (que podemos denominar **Transitiva de troca**). Conseqüentemente, temos agora um novo tipo de construção, uma que leva em conta dois eventos diferentes e simultâneos. Vou chamar esse tipo **construção complexa**; atualmente, o único exemplo conhecido é a construção de troca exemplificada em [99]. Ela se define da seguinte maneira:

[105] Construção Transitiva de troca

	X	V	SN	<i>com</i> SN	<i>por</i> SN
Ev.1:	Fonte		Tema	Meta	
Ev.2:	Meta			Fonte	Tema

Entende-se que os dois eventos são simultâneos e interdependentes – de outra maneira, ficariam alguns complementos sem papel temático, o que sabemos que não é permitido. Em geral, o sujeito é interpretado como Agente. Mas prefiro não incluir isso na definição porque tenho dúvidas: talvez seja possível entender [99] tendo *Jacó* como iniciador do evento.

A preposição varia: é *com*, *de*, ou *a / para* segundo o verbo. Ou seja, estritamente falando, temos aqui três construções parecidas:

[106] Esaú trocou um burro com Jacó por uma galinha.

[107] Esaú comprou um burro de Jacó por mil reais.

[108] Jacó vendeu um burro a / para Esaú por mil reais.

Em todos esses casos se observa a ocorrência dos dois eventos simultâneos (no caso de *comprar* e *vender*, um dos Temas é necessariamente dinheiro). Estou analisando todas essas frases como realizações da construção Transitiva de troca, em três subtipos distinguidos pela preposição que governa um dos sintagmas.

7.7.3. Sobre a representação formal dos papéis temáticos

Parece que existem relações entre certos papéis temáticos e suas representações formais; assim, pode-se imaginar que um Agente é sempre representado por um sintagma nominal. No entanto, o exame dos fatos mostra que se trata mais de tendências estatísticas do que de regras categóricas. Assim, o Agente é representado por um sintagma preposicionado quando o verbo é *apanhar*:

[109] Ricardo apanhou de Milena.

Agente

O Lugar, por sua vez, na maioria dos casos é representado por um sintagma com preposições locativas, como *em*, *dentro de*, *debaixo de* etc. Mas há exceções: primeiro, pode ser representado por um elemento adverbial (*aqui*, *lá*); depois pode ser um SN, como em

[110] A multidão lotava o estádio.

Lugar

O Paciente é tipicamente representado por um SN, mas também aqui há casos de representação por sintagma preposicionado, em

[111] Milena batia em Ricardo.
Paciente

Assim, embora haja tendências -- e estas sejam certamente interessantes -- não se pode fazer generalizações. É possível que haja casos de relações categóricas, mas tudo isso depende de levantamentos ainda não realizados.

7.8. Fazendo pesquisa em tempos de incerteza

As observações feitas nas seções precedentes apontam o caminho para se mostrar que uma distinção de PTs é necessária: basta mostrar que ela é necessária em um caso. Por outro lado, não é tão fácil mostrar que uma distinção é desnecessária, porque seria preciso mostrar que ela é dispensável em **todos** os casos; aqui, portanto, há a necessidade de uma pesquisa muito mais ampla, cobrindo uma área considerável do léxico e um número significativo de construções.

Naturalmente, não é possível esperar a elaboração de uma lista definitiva de PTs para iniciar a pesquisa léxica: primeiro, porque o levantamento léxico é urgente para vários objetivos; e, depois, porque a lista dos PTs depende, pelo menos em parte, do levantamento léxico. Portanto, as duas tarefas têm que ser desenvolvidas paralelamente, e em comunicação constante, de maneira a que cada uma delas controle os resultados da outra e ao mesmo tempo se beneficie deles.

Nessa situação, acredito que a posição mais prudente é a seguinte: **quando em dúvida, faça a distinção**. Isso porque, como acabamos de ver, é relativamente fácil demonstrar a necessidade de uma distinção, mas é muito difícil demonstrar a necessidade de uma junção. Desse modo, convém adotar essa estratégia, chegando eventualmente a uma lista um tanto inflada de PTs; a partir dela, será fácil eventualmente reunir PTs tentativamente separados, ao passo que a tarefa contrária nos obriga a refazer toda a pesquisa.

Capítulo 8: Construções, diáteses, valência

8.1. A apresentação dos dados

O trabalho descritivo tem como principal objetivo a apresentação e sistematização dos fatos da língua, de maneira a criar uma base de dados que seja útil não apenas a profissionais interessados na língua em si, mas também a lingüistas teóricos preocupados com a validação empírica de suas hipóteses. Essa apresentação de dados deve ser feita, evidentemente, da maneira mais compacta e sistemática possível. A apresentação direta dos dados tais como são observados é impossível e inconveniente por duas razões principais.

Primeiro, como se sabe, as frases e sintagmas de uma língua são em número ilimitado; qualquer enumeração precisa ser realizada de forma esquemática. Essa enumeração é justamente o objetivo descritivo das gramáticas em geral. Ou seja, não dizemos que a frase

[1] Lucas rasgou o diploma.

existe em português, mas damos uma representação estrutural esquemática, expressa em termos como “sintagma nominal”, “verbo”, “agente”, “paciente” etc. Essa representação se acopla a uma taxonomia dos itens léxicos envolvidos (*Lucas* é um “nominal”, *-ou* é um “sufixo de tempo-pessoa-número”). Assim se torna possível enumerar objetos cuja lista completa é de tamanho ilimitado.

Depois, como já apontei no capítulo 1, existe uma grande área de consenso entre os lingüistas, suficiente para permitir a elaboração de representações estruturais aceitáveis (ou, pelo menos, compreensíveis) para a grande maioria dos profissionais do ramo. Seria insensato ignorar esses recursos, que constituem grande parte do terreno conquistado pelos lingüistas nos últimos milênios, em sua tentativa de compreender a estrutura das línguas.

Assim, com base nessa área de consenso, vou propor aqui um sistema de representação das estruturas oracionais do português, baseado em uma lista de **construções** (uma proposta parecida, embora diferente em seus detalhes, se encontra em Goldberg, 1995). Meu objetivo final é elaborar uma maneira de representar esquematicamente as construções que servem para subcategorizar os verbos – as **diáteses verbais** – e toda a discussão é subordinada a esse objetivo. Acredito que com isso se obterá uma linguagem útil para representar as construções da língua, independentemente de sua relevância para as valências verbais; por isso, a pesquisa aqui sugerida e em parte realizada deve ser de interesse para os lingüistas em geral. Se tive sucesso em minha tentativa é coisa que o leitor crítico deve decidir.

A seguir, exponho a constituição dessas construções e seu papel na representação descritiva das estruturas da língua.

8.2. Construções: papel descritivo

A noção de “construção” é na verdade muito simples, e vem de uma longa tradição. Uma construção é, essencialmente, uma representação esquemática que se realiza concretamente como um conjunto de frases ou sintagmas. Assim, dadas as frases

[1] Lucas rasgou o diploma.

[2] Eu lavei as janelas.

[3] Seu filho beliscou aquele aluno do 4º período.

podemos dizer que todas representam uma só construção, que em termos tradicionais seria analisada (falando em termos de funções) como composta de **sujeito** e **predicado**, sendo o predicado formado de **núcleo** mais **objeto direto**; ou então (falando em termos de classes) como composta de **SN** mais **SV**, sendo o SV composto de **V** mais **SN**. Como se vê, não há nada de realmente novo nessa noção de construção: a construção, definida em termos esquemáticos (gramaticais) se realiza (ou se **elabora**) em termos de palavras e morfemas particulares, de modo a produzir as bases de um enunciado.¹

Boa parte da gramática de uma língua pode ser expressa como uma lista de construções, enumeradas uma a uma ou reduzidas a um conjunto de regras de estrutura. Aqui vamos explorar a conveniência de enumerar as construções uma a uma; mas fique claro que não é essa a única maneira de fazer essa enumeração. A lista é uma espécie de catálogo das construções possíveis na língua; e esse catálogo é extensionalmente equivalente a um conjunto de regras (ou princípios, ou o que seja) que definam adequadamente essas mesmas construções. A opção pela lista é vantajosa, no estágio primitivo em que se encontra a pesquisa, porque evita um mundo de discussões teóricas e, principalmente, notacionais que, a meu ver, têm pouca utilidade para efeitos da descrição preliminar.²

Mesmo se decidirmos simplesmente listar as construções, ainda se coloca uma série de problemas. Um dos principais é o do nível de abstração a ser adotado para representar as construções. Por exemplo, uma construção como [3] pode ser representada em diversos níveis de esquematicidade:

[3] Seu filho beliscou aquele aluno do 4º período.

(a) SN + SV

(b) SN + V + SN

(c) [possessivo + N]_{SN} + [V + SN]_{SV}

(d) [possessivo + N]_{SN} + [V + (determinante + N + SPrep)_{SN}]_{SV}

¹ O enunciado não se limita a isso, naturalmente: inclui também elementos contextuais.

² Isso não quer dizer, evidentemente, que essas discussões não possam se tornar importantes em estágios ulteriores da pesquisa.

(e) SN + V + SN
 Agente Paciente

(f) SN + V + SN
 [+animado] [+ animado]
 [+humano] [+ humano]

etc.

Cada uma dessas representações atende às necessidades de determinado objetivo descritivo. Na exemplificação acima, os esquemas (e) e (f) são diferentes dos demais porque incluem não apenas a seqüência de classes morfossintáticas e sua agrupação em constituintes, mas também elementos de seu significado, a saber, em (e) os **papéis temáticos** dos dois SNs envolvidos, e em (f) alguns traços semânticos dos núcleos dos SNs. Como se verá, lidaremos principalmente com fórmulas do tipo de (e) – mas não do tipo de (f) – ao definirmos as construções que nos interessam para a descrição das valências verbais. Mas essa é apenas uma das possibilidades de representar a construção exemplificada por [3]; se nosso objetivo descritivo fosse outro, poderíamos ter que optar por algum outro formato.

As valências verbais serão descritas, neste trabalho, em termos de construções. Assim, vamos distinguir o verbo *beliscar* do verbo *aparecer* porque *beliscar* pode aparecer na construção dada acima como (e) (chamada **transitiva**), ou seja, precedido de um SN com o papel temático de Agente e seguido de outro SN com o papel de Paciente.³ Já *aparecer* aparece precedido de um SN com o papel de Paciente, sem nenhum SN seguinte – na construção ergativa; mas *aparecer* não pode ocorrer na construção transitiva, nem *beliscar* na ergativa. A **valência** de um verbo é o conjunto de construções em que ele pode ocorrer; a exemplificação acima mostra que as valências de *beliscar* e de *aparecer* são diferentes. Alguns verbos têm valências muito simples (*devorar*, por exemplo, que parece só ocorrer na construção transitiva), outros têm valências muito ricas e complexas (por exemplo, *dar*, que ocorre em grande número de construções). Uma construção que vale para a subcategorização de verbos se denomina **diátese**; assim, dizemos que *quebrar* tem as diáteses transitiva e ergativa, porque pode ocorrer nessas duas construções.

A noção de **diátese** é comum em gramática e em lexicologia (freqüentemente sob os rótulos de **voz**, **transitividade** ou **regência**). O conjunto das diáteses de um verbo é sua **valência**.⁴ O termo “valência” é usado aqui em sentido um tanto ampliado em relação à noção tradicional tal como se encontra em Allerton (1982), Trask (1992: 296) e Lyons (1977: 481). Aqui, seguindo autores como Villela e Koch (2001) e outros, considero não apenas o número de argumentos associados a cada verbo, mas também seu tipo: classe sintagmática, papel temático, posição (função) sintática e eventualmente outras propriedades.

³ Relembro que o sujeito será representado por ‘X’, que simboliza o SN sujeito, o sufixo de pessoa-número ou apenas um dos dois.

⁴ A designação “valência” foi introduzida na terminologia lingüística por Tesnière (1959). Tal como usada neste texto, corresponde de perto à noção de “grade temática” usada em gramática gerativa, e me parece idêntica à *valency* de Lyons (1977: 486).

Assim, o verbo *assustar* pode ocorrer na construção transitiva ou na ergativa:

[4] O professor assustou os alunos. [transitiva]

[5] Os alunos assustaram. [ergativa]⁵

Já o verbo *castigar* só ocorre na transitiva, não na ergativa:

[6] O professor castigou os alunos. [transitiva]

[7] * Os alunos castigaram. [inaceitável em acepção ergativa]

A construção ergativa se caracteriza por atribuir ao sujeito o papel temático de Paciente; e com efeito essa é a interpretação que se verifica em [5] (ou seja, os alunos são a entidade que sofre o susto).⁶ Mas [7] é inaceitável nessa acepção: não é possível entender de [7] que os alunos são a entidade que sofre o castigo. Nessa frase, *os alunos* só pode ser o Agente. Isso, naturalmente, coloca *assustar* e *castigar* em subclasses diferentes de verbos – eles têm valências distintas. Essa diferença é uma informação que o usuário da língua precisa dominar para usar corretamente esses verbos.

Cada construção recebe, pois, uma definição expressa em termos formais e semânticos – ou seja, em termos **simbólicos**, para usar a terminologia de Langacker (1987, 1999). Podemos definir **diátese** como uma construção definida em termos simbólicos, considerada como entorno possível de ocorrência de um verbo. Assim, pode-se falar da “diátese transitiva de *assustar*”, ou dizer que “*assustar* tem a diátese ergativa”. As formulações propostas descrevem cada construção da maneira mais neutra e direta possível; não procuro generalizações dependentes de discussões teóricas que, embora possam ser importantes, nos desviariam de nossos objetivos descritivos imediatos.

8.3. Limitações do estudo

Antes de passar adiante, é necessário explicitar uma importante limitação do presente estudo (e dos estudos de valência em geral): aqui só se consideram estruturas unioracionais. Assim, fica excluído todo o sistema de complementação, com suas grandes complexidades – classe da oração subordinada (nominais, adjetivas, adverbiais), tipos de complementizadores, modo da oração subordinada (indicativo, subjuntivo, infinitivo), sistema de correferências de SNs etc. A meu ver, esses fenômenos podem ser considerados legitimamente como manifestações da valência, pois contribuem para a subclassificação dos verbos; mas incluí-los na pesquisa nos levaria desmesuradamente

⁵ Ou seja, *assustar* funciona como *engordar*. Esse é o uso normal no português de Minas; para muitos brasileiros, entretanto, há diferença entre os dois verbos, e *assustar* exige *se*: *os alunos se assustaram*, mas nunca * *meu cachorro se engordou*. Neste estudo tomo como base os dados do dialeto mineiro.

⁶ Construções da forma SN-V, onde o verbo é da classe de *assustar*, são ambíguas; assim, o sujeito de [5] pode ser entendido como Agente; isso aparece mais claramente em *cobra assusta*. Já com verbos da classe de *castigar* não há ambigüidade, e [7] não pode ser salva entendendo seu sujeito como Paciente (ver 10.1.2).

longe.⁷ Por isso, aqui só se estuda a ocorrência dos verbos nos chamados “períodos simples”.

Além disso, o presente estudo se limita aos verbos, embora seja conhecimento tradicional que nominais, advérbios e certas expressões idiomáticas também podem ter valência. E, por minha conta, acrescento mais uma limitação: não incluo as construções com *se* obrigatório, como *esse fato se deu há mais de cem anos* etc. O estudo detalhado dessas construções requer a consideração de fatores de variação, já que a ocorrência do *se* depende de fatores sociolinguísticos e dialetológicos. Esta pesquisa se limita aos fatos tais como observados na fala coloquial de Belo Horizonte, onde o *se* obrigatório é muito pouco usado.

Esses recortes são impostos pela amplitude do problema geral da valência e pela necessidade de manter a pesquisa em dimensões controláveis. Ainda assim, o tema é muito amplo, e espero que a área delimitada seja suficientemente extensa para fornecer resultados significativos para a compreensão do funcionamento da língua.

8.4. Contando verbos

Outra questão que precisa ser esclarecida é se é válido distinguir verbos homônimos, em especial quando os significados e as valências seriam muito diferentes. Um caso típico é o de *pintar*, como em

[8] Leonardo pintou a janela.

[9] Leonardo pintou lá em casa.

Ou seja, *pintar* ocorre em dois significados nitidamente diferentes, correspondendo a valências também diferentes. A pergunta é se devemos distinguir aqui dois verbos *pintar*, ou apenas um.

Se a decisão for a de separar os dois verbos com base em seu significado, vão surgir problemas bastante sérios (de que trato em outro texto; ver Perini, 1999), isso porque a diferença entre os significados na maioria dos casos é muito pouco nítida, freqüentemente gradual, de modo que faltam critérios para estabelecer os limites. Já a diferença de valências é nítida, mas se separarmos os verbos pelas valências, isto é, se considerarmos cada verbo em cada valência como um item léxico separado, vamos acabar tendo apenas verbos monovalentes, prejudicando assim o objetivo principal do estabelecimento das valências, que é o de subclassificar os verbos.

Minha solução é considerar *pintar* em [8] e [9] um só item, pelo menos para efeitos do estabelecimento da valência. Há boas razões para isso, de modo que os problemas mencionados acima não nos afetam. Vamos lembrar um dos princípios básicos que regem o trabalho descritivo, ou seja, o de que temos que estabelecer uma relação entre **formas** (tais como percebidas pelo receptor) e **significados**. As formas são o ponto de partida necessário para a interpretação, e precisam portanto estar registradas na memória do falante. Ao ouvir ou ler [8] ou [9], o que ele tem é um verbo com a forma

⁷ Há estudos importantes na área do período composto. Entre meus favoritos está o de Levy (1983), para o espanhol; é preciso também citar Gross (1975), para o francês.

pintar.⁸ É tarefa do receptor atribuir a essa forma o significado “cobrir de tinta” ou “aparecer”, com base no contexto da sentença. Ora, parte desse contexto é justamente o entorno sintático – a construção em que a forma ocorre. Se *pintar* tem sujeito e é seguido de um SN, ou seja, se ocorre na construção transitiva, só pode ter a primeira das acepções acima.

Se quisermos relacionar a descrição da língua com seu uso – o que, acredito, todos admitimos que é um dos objetivos da lingüística – teremos que descrever o processo que, a partir de um sinal formal, atribui significados aos enunciados. É claro que temos que distinguir, em algum momento, o item *pintar* que ocorre em [8] do que ocorre em [9], mas isso é resultado de um processamento, não um ponto de partida. No primeiro momento, *pintou* é uma forma fonológica a ser associada a determinados traços morfológicos, sintáticos, semânticos com base no contexto em que ocorre – basicamente, no que nos interessa, no contexto morfossintático.

Por isso, a descrição deve partir do ponto inicial; e, para nós, *pintar* é um verbo único. Vamos estabelecer sua valência, que inclui, como vimos, as diáteses transitiva (como em [8]) e intransitiva (como em [9]); e teremos que descrever igualmente as conexões entre as diáteses e as diferentes acepções semânticas, que no caso podem ser muito diferentes. Para *pintar*, diremos que na construção transitiva a acepção é “cobrir de tinta”; já na intransitiva (ou seja, acompanhado de apenas um SN) podemos ter a acepção “cobrir de tinta”, como em *Leonardo pintava muito bem*, ou “aparecer”, como em [9].

Nossa lista de verbos e valências, portanto, se amarra a itens definidos em termos formais – em última análise, fonéticos. Podemos desprezar certas acepções, confiando em que não serão relevantes para nossos objetivos; assim, *andar* é um verbo, e não levaremos em conta seu uso para se referir a cada um dos pavimentos de um edifício (*moro no terceiro andar*). Mas note-se que isso é apenas uma conveniência para o trabalho de análise lingüística; o receptor, ao ouvir *andar*, precisa ainda apurar se se trata do verbo ou do nominal.

Em um primeiro momento, a lista consiste de verbos, sem distinções semânticas ou valenciais: *pintar* é um item apenas, para nossos objetivos. As distinções terão que ser feitas, evidentemente: *pintar*, quando é associado a dois SNs (ou seja, quando ocorre na construção transitiva), só pode significar “cobrir de tinta”. O mesmo item *pintar*, quando associado a apenas um SN, pode significar “aparecer” ou então “cobrir de tinta”. A partir daí, então, podemos construir novas taxonomias e análises de caráter mais abstrato.

Encontro a mesma solução no trabalho de Levy (1983) sobre os verbos do espanhol:

Para efeitos da construção das tabelas tomei a forma fonológica como critério básico de determinação do que é um verbo. Às vezes essa forma toma acepções distintas em cada uma das variantes de uma construção sintática, como, por exemplo, *precisar*, que tem dois sentidos distintos em:

F. precisó que tengamos el trabajo listo ‘especificou’

F. precisa dormir 10 horas para reponerse ‘necessita’

[Levy, 1983: 95]

⁸ Ou, mais exatamente, *pintou*, que se associa com um lexema que inclui *pintar* etc.

Essa me parece a única solução operacionalizável para o fenômeno, que é bem extenso e variado; basta lembrar a complexidade semântica de verbos como *andar*, *ir* (de movimento ou auxiliares); *chegar* (sem objeto em *ela chegou*, com objeto em *ela chegou a mesa para perto da janela*) etc., que é resolvida através do contexto sintático. É bom observar que essa solução se opõe à adotada por alguns linguistas. Por exemplo, Welker (2005) afirma que, ao se estabelecer a valência dos verbos

em primeiro lugar, devem ser distinguidas as diversas acepções dos verbos [...] pois cada uma tem sua valência específica.

[Welker, 2005: 81]

Como apontei acima, se fizéssemos isso (e Welker reconhece que não seria fácil), esvaziariamos boa parte do interesse da pesquisa. Não só haveria o risco de reduzir todos os verbos à monovalência, mas ficaria prejudicada a possibilidade de testar a hipótese citada acima, da relação entre o significado dos verbos e sua valência.⁹

8.5. Expressões idiomáticas

Outra questão a discutir é a das expressões idiomáticas: por exemplo, a forma *perder de vista* deve ser levada em conta como uma diátese de *perder*, ou deve ser tratada à parte? Segundo Levy (1983), as expressões devem ser tratadas à parte quando funcionam como uma unidade – ou seja, quando não podem ser consideradas simples realizações de estruturas menos elaboradas, mas se vinculam a itens léxicos específicos.

Assim, *perder de vista* não é (apenas) um exemplo da seqüência **V + Prep + N**, porque no lugar de *vista* não se pode colocar livremente outros nominais. Nesse caso o elemento que segue a preposição não é propriamente um SN, porque precisa ser composto exclusivamente do nominal, sem determinantes nem modificadores: não se diz * *perder da vista*, * *perder da minha vista* etc. Além disso, o elemento nominal não é retomável por pronome:

[10] A: Você perdeu a menina de vista?

B: Sim, perdi a menina de vista logo depois. / * perdi a menina dela logo depois.

Pelo menos em casos como o de *perder de vista*, o elemento preposicionado não pode ser topicalizado, embora SPreps complementos em geral possam:

[11] * De vista eu perdi a maioria das meninas.

[12] De graça eu trabalhei durante várias semanas.

[13] De São Paulo chegam quatro caixas por dia.

⁹ Essa solução é adotada por Anderson (1977); Starosta (1981) critica Anderson, mas seus argumentos não me parecem válidos em face da tarefa de descrever as valências verbais, e principalmente em face da necessidade de descrever o processamento da cadeia sonora por parte do receptor.

Mais uma característica que mostra o caráter unitário das expressões idiomáticas é a dificuldade que se tem em interrompê-las com interjeições ou hesitações:¹⁰

[14] Eu trabalhei... ééé... de graça.

[15] As caixas chegaram... ééé... de São Paulo.

[16] * Eu perdi ela... ééé... de vista.

Essas características se aplicam a pelo menos uma classe de expressões idiomáticas, que além de *perder de vista* inclui *prestar atenção*, *levar em conta*, *tomar / levar em consideração* e muitas outras. Elas são sintaticamente inflexíveis, e analisá-las em termos de sintagmas mais gerais (SN, SPrep etc.) é deixar escapar seu aspecto parcialmente, mas crucialmente, lexical.

Pelas razões dadas acima, as expressões idiomáticas ficam excluídas de nosso estudo, e a relação entre o verbo e seus eventuais complementos dentro dessas expressões precisa ser objeto de pesquisas específicas. Em outras palavras, não vamos considerar a ocorrência de um verbo em expressões idiomáticas como parte de sua subclassificação.¹¹

8.6. Valência do verbo ou do predicado?

Neste trabalho falo da valência dos verbos, mas naturalmente isso suscita uma pergunta: quem tem valência é o verbo ou o predicado? Quando dizemos que *comer* tem dois argumentos, estamos falando da palavra *comer* ou do esquema “comer”? Uma resposta freqüente é que a valência se aplica ao esquema, isto é, ao predicado; mas isso me parece uma simplificação indevida.

Por um lado, é verdade que o verbo enquanto entidade formal é insuficiente para determinar as valências; assim, *pintar* pode ocorrer com objeto direto (com o papel temático de Paciente), ou sem ele:

[8] Leonardo pintou a janela.

[9] Leonardo pintou lá em casa.

Essa dupla valência não pode ser atribuída ao verbo *pintar*, sem mais. Trata-se, claramente, de dois significados bem distintos: “cobrir de tinta” e “aparecer”. Como se vê, não se pode falar apenas da valência de *pintar*, temos que falar da valência de *pintar* em um ou outro significado.

¹⁰ Isso foi observado pela primeira vez por Deborah James em um artigo publicado nos anos 70 em um dos volumes da Chicago Linguistic Society.

¹¹ Estritamente falando, a ocorrência em expressões idiomáticas também subclassifica os verbos. Um estudo amplo e detalhado das expressões idiomáticas do português e de sua análise está sendo elaborado por Lúcia Fulgêncio (2007).

Uma das maneiras de dizer isso seria vincular a valência ao significado puro e simples; mas essa solução falha em muitos casos. Primeiro, temos casos em que a valência não pode ser atribuída a fatores de significado, por não haver nenhum que se possa discernir: é o caso de

[17] Leonardo gosta de marmelada.

[18] Leonardo adora marmelada.

Sem dizer que *gostar (de)* e *adorar* sejam realmente sinônimos, a relação semântica entre eles e seus complementos é certamente a mesma. Nesse caso, a que atribuir a exigência da preposição em um caso mas não no outro? Certamente aqui temos que fazer referência ao verbo enquanto item léxico, não ao predicado.

Um exemplo análogo é o do verbo *recusar*, que, com o mesmo significado, admite duas diáteses bem diferentes, exemplificadas em

[19] O deputado recusou fornecer os nomes de seus cúmplices.

[20] O deputado se recusou a fornecer os nomes de seus cúmplices.

A única maneira, a meu ver, de expressar essa dupla valência é atrelá-la ao item léxico *recusar*, já que o predicado é o mesmo nos dois contextos. Compare-se com *aceitar*, que é um antônimo de *recusar*, mas só ocorre na primeira dessas construções:

[21] O deputado aceitou fornecer os nomes de seus cúmplices.

[22] * O deputado se aceitou a fornecer os nomes de seus cúmplices.

ou com *rejeitar*, que com o mesmo significado de *recusar* também só ocorre na primeira construção:

[23] O deputado rejeitou fornecer os nomes de seus cúmplices.

[24] * O deputado se rejeitou a fornecer os nomes de seus cúmplices.

Outro caso de verbos praticamente sinônimos, mas que apresentam valências diferentes, é o de *quebrar*, no sentido de “entrar em bancarrota” e *falir*, com o mesmo sentido. *Quebrar* pode ocorrer na construção transitiva, como em

[25] A inflação quebrou meu tio.

ou na ergativa, como em

[26] Meu tio quebrou.

Mas *falir* só pode ocorrer na ergativa, embora seja exatamente sinônimo de *quebrar* nesse contexto:

[27] Meu tio faliu.

mas

[28] * A inflação faliu meu tio.¹²

Mais um exemplo é o de *roubar* e *surripiar*. *Roubar* ocorre com a Fonte como objeto direto e o Tema regido de *em*, ou então com o Tema como objeto e a Fonte regida de *de*. Mas *surripiar* (assim como seus quase-sinônimos *furtar* e *tirar*) só ocorre na segunda dessas construções:

[29] Roberto roubou mais de dez milhões da empresa.

[30] Roberto roubou a empresa em mais de dez milhões.

[31] Roberto surripiou mais de dez milhões da empresa.

[32] * Roberto surripiou a empresa em mais de dez milhões.

Se realmente o predicado é o mesmo, temos que atribuir essa diferença de valência aos verbos enquanto itens léxicos. Concluo que a valência depende tanto do verbo quanto do predicado: falamos da valência do verbo *V* na acepção *A*.

8.7. Estrutura sintática

As construções que nos interessam serão definidas em termos de certas propriedades gramaticais (formais e semânticas). Em primeiro lugar, temos a estrutura sintática da construção que, tradicionalmente, se exprime em termos de funções sintáticas como **sujeito**, **núcleo do predicado**, **objeto direto** etc.

Já vimos no capítulo 4 o sistema utilizado aqui para definir as construções. Recapitulando, o sujeito é representado não apenas como um SN em determinada posição, mas como uma variável ('X') que representa o SN que recebe o mesmo papel temático indicado pelo sufixo de pessoa-número, independentemente de sua posição da oração, ou o próprio sufixo, ou ambos. Fora isso, os demais termos são representados pela classe a que pertencem (SN, SAdj) e sua posição na oração. Assim, a construção que chamamos "transitiva" se define, em sua face formal, como **X V SN**.

Nem todos os constituintes que acompanham o verbo constam da diátese. Assim, as duas frases abaixo representam, para nós, a mesma construção (transitiva):

[33] O gato rasgou minha meia.

¹² Talvez alguns falantes aceitem essa frase. Eu não.

[34] Ontem o gato rasgou minha meia em mil pedaços.

Ambas as frases representam a realização da diátese transitiva de *rasgar*, definida formalmente como X V SN; os constituintes *ontem* e *em mil pedaços* não têm representação na diátese, sendo considerados o que na terminologia tradicional se denomina “adjuntos”. Os critérios para a inclusão de constituintes nas diáteses serão discutidos logo abaixo, na seção 8.10**.

Do ponto de vista sintático, os sintagmas que ocorrem nas diáteses podem pertencer a cinco classes: sintagma nominal (SN), verbo (V), sintagma adverbial (SAdv), sintagma adjetivo (SAdj) e finalmente SN precedido de preposição. No caso de sintagmas preposicionados, a preposição (ou preposições) vem sempre especificada; assim, o verbo *contar* tem a diátese cuja face formal é

X V *com* SN (ex.: *Jorge contava com a ajuda de Cecília*).

o que significa que esse complemento é sempre introduzido pela preposição *com*. Evito a designação geral “SPrep” (sintagma preposicionado) porque se trata de uma falsa categoria: seu potencial funcional (e semântico) é idêntico ao de um sintagma adjetivo ou ao de um sintagma adverbial, conforme o caso, e muitos verbos se constroem com uma preposição específica (*pensar em, gostar de, contar com* etc.).

Assim, temos casos de preposição seguida de SN, mas mesmo nesses casos a rigor poderíamos especificar a classe do sintagma, ou seja, *em Belo Horizonte* é, além de um exemplo de *em* SN, um exemplo de sintagma adverbial. Pode-se então perguntar se não convém incluir ambas as informações na definição, ou seja, em vez de *em* SN colocar

[*em* SN]_{SAdv}

Creio que isso seria supérfluo, porque a diátese nesses casos pode ser totalmente expressa em termos da preposição específica mais o papel temático. O constituinte ser ou não um “sintagma adverbial” não conta, desde que tenhamos condições de apurar sua forma, sua posição e seu papel temático, e tudo isso já está representado na definição da construção.

Temos dois casos que é preciso distinguir; o primeiro é o citado acima, em que uma preposição individual é exigida pelo verbo: *contar com*. O outro caso é o de verbos que, embora ocorram em geral com preposição, parecem dirigir sua exigência ao papel temático, como quer que este seja realizado. Um exemplo é o verbo *morar*, que ocorre com *em*:

[35] A professora mora em Sabará.

No entanto, outras preposições de valor locativo podem também ocorrer, satisfazendo igualmente a diátese do verbo:

[36] A professora mora ao lado da padaria.

escrever, sorrir) e os que se comportam como *engordar* (*quebrar, esquentar, desanimar, rasgar, abrir*). Ou seja, a diferença semântica entre [39] e [40] se deve à subclasse a que pertence o verbo da oração, e convém portanto levá-la em conta na definição das valências. Para isso, é necessário enriquecer o sistema de Allerton, introduzindo informações de ordem semântica na definição das construções utilizadas na subcategorização dos verbos.

8.8.2. Papel da semântica do verbo

Há, em princípio, uma maneira de manter as definições formais, sem incluir informação semântica na definição das construções. Para isso, basta que se possa derivar os traços estruturais das construções dos traços semânticos associados aos verbos individuais. Se isso for possível, poderemos identificar as duas construções exemplificadas em [39] e [40], a saber, as construções **intransitiva** (de [39]) e **ergativa** (de [40]), atribuindo a diferença de papéis temáticos e suas representações sintáticas a propriedades semânticas dos verbos. Vou discutir essa alternativa, dando as razões pelas quais, a meu ver, ela não funciona.

Antes, vou esclarecer um ponto. Há razões para crer que a estrutura de papéis temáticos associada a um verbo é derivável do significado desse verbo. Essa é a posição de C&J, a julgar por passagens como a seguinte:

Como parte de seu significado, um verbo especifica um certo número de argumentos semânticos – entidades intrinsecamente envolvidas na situação que o verbo denota.

[C&J: 173]

No entanto, ainda que os papéis temáticos sejam determinados pelo significado do verbo, sua codificação formal certamente não é. Ou seja, dado um verbo que significa *comer*, sabemos que ele deve se associar a um Agente e a um Paciente. Mas não sabemos que funções sintáticas nem que eventuais preposições são as marcas de cada um desses papéis temáticos. Essa codificação não é previsível a partir da semântica do verbo, e portanto precisa ser explicitada em algum lugar da descrição – no caso, nas diáteses verbais.

Há bastante evidência que mostra a invalidez do suposto relacionamento entre o significado do verbo e a representação sintática dos argumentos. O leitor encontrará diversos exemplos na seção 8.6**, onde se viu que há inclusive casos de verbos sinônimos (no que nos interessa) e que não obstante têm diáteses diferentes, como *roubar, furtar e lesar*, por exemplo. Dessa forma, temos que admitir que a representação formal dos argumentos precisa ser consignada na valência de cada verbo, pois não pode ser derivada de outros fatores.

Não se sabe exatamente quais informações semânticas devem ser incluídas na definição das diáteses; mas certamente o papel temático dos diversos complementos está entre elas, isso se quisermos representar a diferença entre as construções exemplificadas em [39] e em [40]:

[39] Meu cachorro latiu.

[40] Meu cachorro engordou.

Assim, substituímos a fórmula única de Allerton por duas fórmulas, definidas já não formalmente, mas em termos simbólicos:¹⁵

[41] **Construção intransitiva** (*meu cachorro latiu*)

Definição :

X V

Ag

[42] **Construção ergativa** (*meu cachorro engordou*)

Definição :

X V

Pac

Como se vê, as duas construções são formalmente idênticas, mas diferem quanto ao papel temático atribuído ao sujeito em cada caso.

A formulação das construções apresenta diversos problemas, que são discutidos nas seções seguintes.

8.9. O que incluir?

Tradicionalmente, distinguem-se os termos da oração (excluindo-se o verbo e, em geral, o sujeito) em dois tipos: complementos e adjuntos. Por exemplo, o sintagma sublinhado seria complemento em [43], mas adjunto em [44]:

[43] Helena deu um frango para sua mãe. (complemento)

[44] Helena matou um frango para sua mãe. (adjunto)

Essa diferença tem correlatos sintáticos e semânticos. Por exemplo, em [43] pode-se substituir *para* por *a*, mas em [44] essa substituição não é possível:

[45] Helena deu um frango a sua mãe.

[46] * Helena matou um frango a sua mãe.

Nessas frases, o complemento pode ser expresso por pronome oblíquo, mas o adjunto não (ou pelo menos não com a mesma facilidade):

¹⁵ Ver 4.5 para uma explicação da notação 'X' para o sujeito.

[47] Helena nos deu um frango.

[48] ?* Helena nos matou um frango.

Essas diferenças formais, como veremos, não valem para todos os casos, ou seja, às vezes o (chamado) adjunto é expresso por pronome oblíquo (*ele me quebrou o braço*), e nem sempre as preposições usadas diferem.

A oposição entre complementos e adjuntos é aceita pela maioria dos lingüistas, e figura em pontos cruciais da gramática, muito em especial nas diáteses verbais, que se formulariam em termos de complementos e não de adjuntos. No entanto, o problema de distinguir complementos de adjuntos não recebeu até hoje uma solução satisfatória. Boa parte da discussão que há se limita a notas de caráter teórico, sem preocupação de operacionalização descritiva. Como resultado, sabemos o que é um adjunto ou um complemento dentro de várias teorias; mas não temos instrumentos adequados para distinguir adjuntos de complementos ao analisarmos enunciados reais.

Para ser mais explícito: as tentativas de definição muitas vezes se limitam à simples caracterização teórica, esquecendo que a oposição entre complementos e adjuntos tem (ou deve ter) um conteúdo empírico verificável e aplicável em casos concretos de análise. Se não fosse assim, a oposição seria de importância muito reduzida. Assim, seria preciso procurar critérios rigorosos de aplicação, sem os quais a própria oposição complemento / adjunto se torna descritivamente irrelevante.

No entanto, vou argumentar neste capítulo que a distinção em pauta não é relevante para a definição das diáteses, e que a seleção do que deve ser incluído nelas precisa ser feita com base em outras considerações.

8.10. Definindo diáteses

8.10.1. Diáteses como fator de subcategorização

Uma das razões da importância presumida da oposição complemento / adjunto vem de que as diáteses verbais se formulariam em termos de complementos; isto é, os adjuntos seriam irrelevantes para a formulação das diáteses.

Em alguns casos, a oposição se manifesta de maneira clara. Assim, temos verbos que recusam ou exigem objeto direto, ou complemento de lugar, ou complemento regido pela preposição *com*,¹⁶ como:

levar (exige objeto direto)¹⁷

desmaiar (recusa objeto direto)

morar (exige complemento de lugar: *moro aqui, moro em Salvador*)

contar ‘confiar’ (exige complemento precedido de *com*: *conto com você*)

¹⁶ Já estou pressupondo uma distinção entre “complementos” e “adjuntos”, mas por ora deve-se entender que estou apenas seguindo a análise mais aceita.

¹⁷ Mais precisamente, exige um SN não sujeito. Mantenho o termo “objeto direto” para facilitar a leitura, mas, como vimos no capítulo 4, não se trata de uma função sintática legítima, à parte sua posição na oração.

e assim por diante. Mas não há nenhum verbo que exija ou recuse um vocativo, ou um adjunto adverbial de companhia. Os adjuntos se acrescentariam ou não às orações de acordo com a conveniência semântica, sem precisar adequar-se às exigências do verbo.

Desse ponto de vista, o sujeito é um adjunto, com alguma dúvida quanto a verbos como *haver* e os meteorológicos como *nevar*, *relampejar*, que aparentemente recusam sujeito. Por outro lado, contudo, o sujeito é relevante para a definição das diáteses porque o papel temático que recebe subclassifica os verbos. Ou seja, vai ser difícil estabelecer um princípio rigoroso segundo o qual as diáteses se formulam em termos de complementos, porque pelo menos um dos termos relevantes, o sujeito, é provavelmente um adjunto. Veremos mais adiante que realmente não há grandes vantagens em se adotar esse princípio.

Em vez de examinar detalhadamente a oposição complemento / adjunto, o que pretendo fazer em um texto à parte, vou me concentrar aqui na questão do que deve ser incluído na formulação das diáteses, e por quê.

Primeiro, vamos relembrar o objetivo da formulação de diáteses e da associação de cada uma a um conjunto de verbos. Isso se faz para caracterizar as propriedades gramaticais de cada verbo, ou seja, a contribuição de cada item para a sintaxe e a semântica das orações de que participa. Em outras palavras, marcamos cada verbo com os traços que o distinguem dos demais verbos: cada diátese deve, portanto, dividir o conjunto dos verbos da língua em duas subcategorias, a dos verbos que ocorrem naquela construção e a dos que não ocorrem. O efeito conjunto das muitas diáteses da língua dá como resultado o rico panorama das subcategorias dos verbos do português.¹⁸

A partir daí se coloca o problema de que elementos incluir na formulação das diáteses. Já vimos que é provavelmente verdadeiro que o significado do verbo determina os argumentos a ele associados.¹⁹ Mas isso não significa que sabendo o significado do verbo temos sua valência, por três razões:

(a) o significado do verbo não determina o modo como cada papel temático é codificado sintaticamente. Ou seja, se temos um verbo que exprime “movimento”, devemos ter uma Meta; mas essa Meta pode ser representada sintaticamente de diversas maneiras: como sujeito (*eu ganhei uma camisa nova*), como objeto direto (*os alpinistas atingiram o pico*), como *em*+SN (*chegamos em São Paulo*), como *para*+SN (*dei uma camisa nova para meu irmão*) etc. Além disso, alguns papéis temáticos são de explicitação obrigatória (como o Tema de *levar em levei o computador para cima*), e outros são de explicitação opcional.

(b) O significado do verbo, ao determinar sua configuração de papéis temáticos, não distingue casos previsíveis de casos imprevisíveis. Assim, um verbo de “ação” não apenas inclui um Agente, mas também um Lugar e um Tempo – e, no entanto, nem o Lugar nem o Tempo precisam ser incluídos na valência, porque suas possibilidades de ocorrência e sua codificação formal são previsíveis.

¹⁸ Mais precisamente, uma parte desse panorama; ver **.

¹⁹ Desconfio que seria mais correto dizer que o significado do verbo **inclui** as relações temáticas dos argumentos a ele associados.

(e) E, finalmente, a afirmação de C&J faz sentido somente se tivermos um meio de caracterizar o significado dos verbos sem utilizar, de alguma maneira, os papéis temáticos que são associados a ele.

Desse modo, e evitando discussões mais aprofundadas sobre como se caracteriza o significado de um verbo, vemos que o problema de o que incluir na formulação das diáteses permanece de pé.

Precisamos de diáteses nos casos em que a ocorrência do verbo em determinada construção não é previsível a partir de outras informações de caráter mais geral. Por exemplo, o verbo *engordar* pode ocorrer em construções negativas como

[49] Essa ração não engordou meu cachorro.

[50] Essa ração não engorda.

[51] Meu cachorro não engordou.

Mas essa ocorrência em construções negativas é inteiramente previsível a partir da ocorrência de *engordar* nas construções afirmativas correspondentes:

[52] Essa ração engordou meu cachorro. etc.

Por outro lado, isso vale para todos os verbos da língua (não há nenhum que não possa ser negado), e para todas as construções (não há nenhuma que não tenha uma negativa correspondente). Portanto, a ocorrência em construções negativas, em oposição a afirmativas, não tem relevância para a subcategorização dos verbos. Em outras palavras, o recorte que a construção transitiva afirmativa faz no conjunto dos verbos é precisamente igual ao feito pela construção transitiva negativa. Conseqüentemente, não precisamos incluir negativas e afirmativas como diáteses distintas: basta levar em conta a afirmativa.²⁰

Pela mesma razão, nunca levamos em conta a presença ou ausência do vocativo para efeitos de formulação das diáteses, porque se trata de um fenômeno totalmente independente da identidade do verbo da sentença. Em outras palavras, seja qual for o verbo da frase, podemos acrescentar um vocativo livremente; portanto, é impossível subcategorizar os verbos em verbos que admitem vocativo e verbos que não admitem. Provavelmente o mesmo ocorre com os sintagmas adverbiais de companhia: podem ser acrescentados a qualquer oração, sempre que forem semanticamente adequados, independentemente da subcategoria a que pertence o verbo. Por isso, não temos complemento vocativo, nem complemento de companhia.

Já com o objeto direto as coisas mudam. Este é representado, como sabemos, por um SN não-sujeito (definindo-se o sujeito nos termos de 4.1). E sua introdução em uma oração não é livre – ou seja, há verbos que o recusam, como por exemplo

[53] * Sônia veio a pizza.

²⁰ Ou a negativa; prefiro a afirmativa porque parece ser o membro não-marcado de cada par.

[54] * Eu gosto cerveja.

[55] * Eles desconfiam você.

Nenhum SN cabe no lugar dos termos sublinhados acima. Para o caso de um SN nessa função, portanto, podemos concluir que sua ocorrência subcategoriza os verbos entre os que a admitem (*comer, detestar, procurar...*) e os que não a admitem (*vir, gostar, desconfiar...*). Isso quer dizer que a presença ou ausência do objeto direto é relevante para a definição de diáteses – não porque o objeto direto seja “complemento”, mas porque sua ocorrência é governada pela identidade do verbo da oração.

A ocorrência desse SN não-sujeito não é tampouco consequência dos traços semânticos do verbo. Por exemplo, sabemos que *comer* se associa a um Paciente (além de um Agente), e que esse Paciente se codifica como objeto direto; mas com *morrer* o Paciente é sujeito, e o mesmo com *apanhar* em frases como

[56] Meu vizinho apanha da mulher.

Por outro lado, o objeto direto recebe papéis temáticos variados; além de Paciente, pode ser Fonte (*a multidão deixou o estádio*), Meta (*atingimos o pico de madrugada*), Causador de Experiência (*Miltinho sentiu uma pontada*) etc. Portanto, a ocorrência de um objeto direto com um verbo e o papel temático que com ele se associa são coisas que precisam ser estipuladas para cada verbo em particular.²¹

8.10.2. Graus de esquematicidade

Na seção precedente chegamos à conclusão de que o objeto direto deve figurar na formulação das diáteses verbais, porque sua ocorrência não é previsível a partir de outros fatores. No caso, é razoavelmente seguro fazer essa afirmação em termos esquemáticos, ou seja, com referência aos objetos diretos em geral, sem necessidade de distinguir ocorrências particulares. Isso parece se estender a todos os casos de SNs – em todos os casos sua ocorrência e seu papel temático parecem ser governados pelo verbo (ou seja, pela **valência** do verbo).

Mas com sintagmas de outras classes (sintagmas preposicionados e adverbiais) esse tipo de afirmação esquemática não é possível. Vamos considerar o caso de sintagmas preposicionados com o papel temático de Lugar, como em

[57] A minha prima mora em Fortaleza.

[58] A minha prima trabalha em Fortaleza.

Tradicionalmente se analisa *em Fortaleza* como “complemento” em [57] e “adjunto” em [58]. Isso quer dizer que esse sintagma deve figurar na definição da diátese de *morar* exemplificada em [57], mas não na diátese de *trabalhar* exemplificada em [58].

²¹ Isso não exclui, naturalmente, a possibilidade de haver generalizações parciais. Mas estas não são conhecidas, e por ora não temos remédio senão pressupor que a relação é idiossincrática.

Se essa análise for correta, não poderemos dizer do sintagma preposicionado com o papel de Lugar que deve aparecer nas diáteses verbais, nem que não deve aparecer. Antes, aparece na diátese de alguns verbos e não aparece na de outros verbos. Note-se o contraste com a situação do objeto direto, que sempre que ocorre é decorrência de um traço da diátese do verbo em questão.

Como reduzir essas observações a uma imagem coerente? Uma possibilidade é a de distinguir, de saída, duas situações: primeiro, aquela em que se pode decidir, **para uma categoria generalizada**, se deve ou não ocorrer nas diáteses; depois, aquela em que isso só pode ser afirmado a respeito de **uma categoria em relação a um verbo específico**. Veremos que no primeiro caso o termo “categoria” deve ser entendido de maneira bem esquemática, como “SN”, ou “objeto direto”; no segundo caso, aparecem categorias mais elaboradas, como por exemplo “preposição *em* + SN, com papel temático de Lugar”.

É provável que se chegue à conclusão de que os SNs, em qualquer caso, devem sempre figurar na definição das diáteses. Em outras palavras, qualquer que seja o verbo, os SNs que o acompanham são considerados relevantes para sua valência; isso vai se aplicar a sintagmas tradicionalmente analisados como sujeito, objeto direto, predicativo do sujeito e predicativo do objeto.²²

Por outro lado, se o termo a ser avaliado é, digamos, um sintagma preposicionado com a preposição *em* e papel temático de Lugar, essa decisão **a priori** não é possível. Aqui teremos que saber de que verbo se trata, e verificar se a ocorrência do SPrep é ou não governada pelo verbo individual que aparece na oração. Podemos então voltar aos exemplos vistos acima,

[57] A minha prima mora em Fortaleza.

[58] A minha prima trabalha em Fortaleza.

A pergunta é, para cada verbo (*morar*, *trabalhar*), se o sintagma preposicionado (aqui, *em Fortaleza*) deve ou não ser considerado parte da diátese, e portanto incluído em sua definição, ou se pode ser considerado um adjunto. Essa distinção é importante porque, se for um adjunto, o sintagma não contribui para a subcategorização do verbo; funciona, antes, como os termos que podem ser acrescentados a qualquer frase, desde que sejam semântica e discursivamente adequados, sem consideração da subclasse do verbo da oração. Ora, em [57] o SPrep de Lugar é exigido pelo verbo, ao passo que em [58] não há exigência, o que sugere uma diferença de estatuto do mesmo sintagma em frases diferentes.

A dualidade de situações que apontei acima nem sempre é percebida, a julgar pelo que se diz na literatura. Por exemplo, Allerton afirma que

Parece claro que, enquanto elaboradores como objetos e predicativos têm a ver com a valência de verbos individuais, os advérbios de tempo, lugar e ambiente²³ ficam de fora das valências.

²² Fala-se aqui apenas dos SNs de nível oracional, claro. Não contam os eventuais SNs que sejam parte de outros sintagmas (como o aposto, por exemplo).

²³ “Adverbs of environment”—não sei exatamente o que Allerton quis dizer com isso.

[Allerton, 1982: 58]

Allerton, como aliás os autores em geral, procura determinar o valor de cada sintagma para com a formulação da valência em termos generalizados, esquemáticos (SNs, SPreps de Lugar etc.). Mas, como vimos, os advérbios de lugar (sintagmas com o papel temático de Lugar) podem se relacionar com o verbo (*morar em Fortaleza*) ou não (caso de *trabalhar em Fortaleza*). Já dos objetos diretos (e mesmo dos SNs em geral) se pode afirmar que sempre se relacionam com o verbo, e portanto figuram na definição das diáteses.

Voltando aos sintagmas de Lugar em [57] e [58], note-se que não há diferença semântica entre os dois casos: ambos denotam uma localização. Encontra-se às vezes a afirmação de que um elemento como *em Fortaleza* recebe papel temático do verbo em casos como o de [57], mas não em casos como o de [58]. Ou seja, em [57] o significado “lugar” seria um papel temático atribuído pelo verbo a seu complemento, ao passo que em [58] o significado “lugar” seria devido à semântica da preposição *em*. No entanto, como a relação semântica é precisamente a mesma, e a preposição é a mesma, fica-se sem saber qual é a verdadeira razão dessa diferenciação – a não ser que seja um argumento puramente intrateórico, ou seja, não-empírico, e portanto inaceitável segundo o ponto de vista adotado aqui. Se esquecermos idéias preconcebidas e nos lembrarmos de que o que estamos tentando é justamente (entre outras coisas) **construir** uma teoria com base nos fatos, não há razão para distinguir o tipo de relação entre *em Fortaleza* e o resto da frase nos dois exemplos acima: na ausência de evidência em contrário, efeitos idênticos devem ter causas idênticas.

8.10.3. Previsibilidade

Vamos admitir que certos sintagmas têm uma fisionomia simbólica típica (ou mesmo única). Assim, um sintagma da forma *em*+SN com o papel temático de Lugar é típico, no sentido de que sintagmas com **essa** forma têm geralmente **esse** papel temático. Há exceções, como em *pensei em você*, onde *em*+SN tem o papel de Assunto, e *os hititas habitavam a Ásia Menor*, onde o papel de Lugar é expresso por um SN (*a Ásia Menor*). Mas, em algum sentido gramaticalmente relevante, esses casos são marcados, e o caso não marcado é *em*+SN com papel de Lugar. Uma indicação de que esse é realmente um caso não marcado é que quando a introdução do sintagma *em*+SN é livre (não governada pelo verbo) o papel temático é sempre o de Lugar:

[59] A Carol estudou lingüística no Canadá.
 arranjou um marido
 comia estrogonofe de ganso
 [...]

[60] A Carol tem uma tatuagem no braço.

Não parece possível introduzir um sintagma dessa forma com o papel de Assunto a não ser com certos verbos especificados – ou seja, não livremente (ver adiante a seção 8.10.5**, onde desenvolvo essa idéia).

Admitindo-se como válida essa situação, podemos dizer que qualquer caso de *em*+SN com papel temático outro que não o de Lugar, ou então do papel temático de Lugar expresso por outra forma que não a de *em*+SN é marcado,²⁴ e deve figurar nas diáteses. Além disso, mesmo casos não marcados precisam ser incluídos nas diáteses se sua realização formal for exigida pelo verbo. Assim, sintagmas da forma *em baixo de*+SN são unívocos quanto a papel temático (ou, no máximo, variam entre Lugar e Meta apenas), mas têm que ser considerados parte da diátese de verbos como *morar*, porque se diz

[61] A pobrezinha morava em baixo de uma escada.

mas não

[62] * A pobrezinha morava.

Com outros verbos, esses sintagmas não constam da formulação das diáteses, pois sua ocorrência é livre e sua semântica previsível.

Esse é o fator fundamental para a decisão: previsibilidade. As diáteses existem para fornecer informação impossível de obter em termos gerais, ou seja, que dependem das propriedades de verbos específicos. Nosso problema é portanto catalogar as situações em que a informação é imprevisível, e que portanto deve ser incluída na formulação das diáteses. Distingo então as seguintes situações:

(a) um sintagma de determinada forma tem um papel temático marcado; ou, talvez equivalentemente, um sintagma com determinado papel temático está em uma forma marcada (para aquele papel temático); ou

(b) um sintagma é de ocorrência obrigatória com determinado verbo.

Exemplos da situação (a) são

[63] Meu vizinho apanha da mulher. [*de*+SN com o papel de Agente]

[64] Eu pensei em você. [*em*+SN com o papel de Assunto]

[65] Roberto roubou a União em mais de 100 milhões. [*em*+SN com o papel de Tema]

[66] A Marília gosta de alface. [*de*+SN no papel de Causador de Experiência]

²⁴ Creio que se pode acrescentar, como não-marcados para o papel temático de Lugar, os chamados “advérbios de lugar” como *aqui, lá, ali*.

Provavelmente os sintagmas da forma *de*+SN teriam que ser sempre incluídos, porque a semântica de *de* é tão ampla que dificilmente se pode vislumbrar um uso não marcado dessa preposição.²⁵

Exemplos da situação (b) são:

[57] A minha prima mora em Fortaleza.

[67] A bagunça durou mais de uma hora.

Aqui os sintagmas *em Fortaleza* e *mais de uma hora* ocupam uma função obrigatoriamente preenchida.²⁶

[68] * Minha prima morava.

[69] * A bagunça durou.

Essa obrigatoriedade só pode ser atribuída a um traço do verbo, e por isso ela precisa constar da formulação da diátese.

8.10.4. Sintagmas preposicionados e sintagmas nominais

Essa parece ser a situação com os sintagmas introduzidos por preposição. Já os SNs devem ser todos consignados, porque um SN, sujeito ou não, não tem nenhum papel temático suficientemente típico para ser considerado o caso não marcado.

É importante enfatizar que figurar em uma diátese não significa que aquele elemento tenha que ocorrer obrigatoriamente com aquele verbo. O exemplo clássico é o do objeto direto, como em

[70] A professora está lendo Eça de Queirós.

[71] A professora está lendo.

Aqui temos, em princípio, as realizações de duas diáteses, uma da forma **X(Ag) V SN(Pac)**, e a outra da forma **X(Ag) V Ø(Pac)**.²⁷ Sabe-se que a condição de exigência não funciona para distinguir complementos e adjuntos, e tampouco como critério de inclusão na formulação de diáteses.

Com sintagmas preposicionados a situação é a mesma, embora talvez de observação mais difícil. Por exemplo, na frase

²⁵ Talvez no papel temático de Fonte; mas isso é discutível.

²⁶ Não necessariamente preenchida por um complemento de Lugar: *ele mora muito bem*. O preenchimento alternativo ocorre em outros casos: *o gato está furioso / o gato está aqui*.

²⁷ Onde o “Ø” indica um argumento semanticamente presente, com seu papel temático, mas não realizado no plano formal. Esse elemento é que distingue a construção exemplificada em [71] de casos como *a professora está chorando*, onde não é necessário subentender um Paciente.

[72] Carlinhos apanhou da mulher.

o sintagma *da mulher* realiza um elemento presente na diátese, mesmo porque seu papel temático depende do verbo – é Agente porque o verbo é *apanhar*, e com outros verbos tem outros papéis temáticos. Mas não é de ocorrência obrigatória, por causa da aceitabilidade da frase

[73] Carlinhos apanhou.

O problema é como notar a diátese representada por [28]. Resisto a colocar uma preposição “abstrata”, ou seja,

[74] X V de Ø
Pac Ag

Por outro lado, não há dúvida de que ao ouvir [28] o receptor fica sabendo que há um Agente envolvido, ou seja, Carlinhos não “apanhou” sem que alguém batesse nele. Prefiro, então, notar a diátese assim:

[75] X V Ø
Pac Ag

e fica entendido que “Ø” representa um argumento ausente, sem nenhum compromisso com sua forma caso se faça presente em outra diátese. Ou seja, “Ø” não é a representação de um SN ausente, mas de um complemento ausente, apenas. Diremos então que *apanhar* (nesse sentido) ocorre nas diáteses [75] e [76]

[76] X V de SN
Pac Ag

o que equivale a dizer que *apanhar* pode ter um complemento da forma *de+SN* com papel de Agente, ou pode não ter esse complemento, e nesse caso o Agente é indeterminado.

8.10.5. Procedimento

Acabo de formular uma proposta de critérios para distinguir, dentre os sintagmas de acompanham um verbo em uma oração, aqueles que precisam constar na diátese desse verbo. Em alguns casos, temos um critério geral, que pode ser aplicado (pelo que se sabe) sem grande perigo: trata-se dos casos de SN, seja ele sujeito, objeto direto ou predicativo do sujeito. Estes nunca ocorrem livremente, mas sempre como resultado de serem especificamente mencionados na diátese; assim, em

[77] [Daniela]_{SN} considera [Ronaldo]_{SN} [o maior jogador do mundo]_{SN}.

todos os três SNs figuram na definição da diátse, a saber,

[78] **Construção de ação opinativa**

X	V	SN	SAdj ~ SN
Opinador		Tema	Qualidade

Ao que tudo indica, isso vale para todos os complementos representados por SN.

Mas com sintagmas preposicionados a situação é outra. Sabemos que o mesmo sintagma preposicionado, com a mesma preposição e o mesmo papel temático, é analisado tradicionalmente como “complemento” em uma frase e como “adjunto” em outra:

[57] A minha prima mora em Fortaleza. (“complemento”)

[58] A minha prima trabalha em Fortaleza. (“adjunto”)

Aqui (a se aceitar a intuição tradicional) torna-se necessário um procedimento mais elaborado para identificar os constituintes a serem incluídos na diátese.

Para efeitos descritivos, propus um critério baseado na previsibilidade. Vamos agora ver como colocar esse critério em prática no caso dos sintagmas preposicionados.

8.10.5.1. Preposições unívocas

Se um SPrep for introduzido por uma preposição semanticamente unívoca, naturalmente não há necessidade de especificá-lo, a menos que sua ocorrência seja obrigatória (o que introduz um elemento de imprevisibilidade). Por exemplo, a preposição *desde* representa um papel temático que não varia conforme o verbo da oração:

[79] Ele trabalha aqui desde 1999.

[80] Eu vim cochilando desde Juiz de Fora.

Em ambos os casos trata-se de Fonte (temporal ou locativa, mas como sabemos é o mesmo papel temático). E como a presença de *desde* + SN não é obrigatória, esse SPrep não precisa constar das diáteses de *trabalhar* ou *cochilar*.

Não conheço casos de SPrep com preposição unívoca e ocorrência obrigatória.

8.10.5.2. Preposições ambíguas

Mas na maioria dos casos o papel temático introduzido pela preposição é bastante variado. Casos extremos são *de*, *em* e *com*, que são de tal modo ambíguas que talvez nem

faça sentido falar de um significado “básico” para elas. Tomando *com* como exemplo, temos:

[81] O menino sumiu com a minha agenda. (Paciente)

[82] Vou cortar essa corda com uma gilete. (Instrumento)

[83] Eu conto com vocês. (Causador de Experiência)

[84] Amélia trombou com Vânia. (Meta)

[85] O rapaz está namorando com minha sobrinha. (PRS)²⁸

e certamente outros.

Pode-se verificar facilmente que essa variedade de papéis temáticos sinalizados por *com* é governada pelo verbo na oração. Por conseguinte, SPreps introduzidos por *com* devem constar da formulação das diáteses.

Mas o leitor observará que um papel temático muito comum da preposição *com* não ocorreu na lista acima: o papel de Companhia, que aparece em

[86] A Janaína jantou com o vizinho.

Partindo de um julgamento intuitivo, não me parece que *com*+SN seja um dos complementos presentes na valência de *jantar*. E, olhando a coisa de outro lado, suspeito que o sintagma de Companhia pode ser acrescentado a orações com uma variedade muito grande de tipos verbais. Isso sugere uma hipótese, a de que Companhia seja o papel temático **não marcado** para sintagmas introduzidos pela preposição *com*. Se for isso, poderemos deixar que ele seja realizado através de uma regra de redundância, do tipo “a menos que haja indicação em contrário, *com*+SN recebe o papel temático de Companhia”. E esses casos não serão consignados na formulação das diáteses.

8.10.5.3. Marcados e não marcados

Aqui acho necessário fazer uma pequena digressão sobre a noção de **marcação** (*markedness*), à qual acabo de recorrer para caracterizar as diáteses. A proposta baseada na previsibilidade se reduz, no final das contas, a consignar nas diáteses as alternativas **marcadas**, deixando as não marcadas por conta de regras de redundância. Mas o que vem a ser exatamente essa distinção entre marcado e não marcado?

O dicionário de Trask sugere que se trata de uma noção muito importante, dizendo dela que “hoje é encarada como central em lingüística” (Trask, 2004: 187). No entanto, a leitura do verbete revela uma situação de grande incerteza quanto ao que é exatamente “ser marcado”.²⁹ Lemos que

²⁸ “PRS”= participante de relação social.

²⁹ Não que isso não seja bastante típico em nossa disciplina.

Em termos gerais, é *marcada* qualquer forma lingüística que é – sob qualquer ponto de vista – menos usual ou menos neutra do que alguma outra forma, a forma *não-marcada*. Uma forma marcada pode distinguir-se de outra também marcada pela presença de mais material, de maior quantidade de matizes de significado, por ser mais rara numa determinada língua ou nas línguas em geral, ou de vários outros modos.

[Trask, 2004: 187]

Como se vê, a definição inclui tanta coisa que acaba sendo muito pouco útil na prática. O que se faz, em vez de aplicar definições, é criar uma noção mais ou menos nova para manejar cada caso específico; e é o que vou fazer aqui, sem pedir desculpas.

Para nós, **marcada** é a atribuição de um papel temático a determinado sintagma por força da valência do verbo; e **não marcada** é a atribuição de um papel temático de maneira livre, ou seja, independentemente da valência do verbo da oração. A idéia é que certos sintagmas têm um papel temático não marcado, que aparece quando a influência da valência do verbo é neutralizada.

Para verificar isso em casos concretos, sugiro um experimento:

Digamos que o problema seja determinar o papel temático não marcado dos sintagmas da forma *com*+SN. Seleccionamos uma frase que, ao que tudo indica, está maximamente preenchida, ou seja, com todos os elementos possíveis com aquele verbo naquela acepção realizados sintaticamente. Pode ser, por exemplo

[87] O Toninho comeu a pizza.

É claro que não podemos ter **certeza** de que tudo está aí; mas é uma hipótese plausível, pelo que se sabe, que *comer* nesse sentido tem apenas o Agente e o Paciente estipulados na diátese em questão.

Isso feito, testamos o acréscimo de um SPrep, para ver com que papel temático ele aparece ali. Note-se que esse SPrep vai ser um acréscimo livre, não vinculado à valência de *comer*; o papel temático que surgir pode ser considerado o não marcado para aquele SPrep. O resultado é

[88] [O Toninho] comeu [a pizza] [com a Virgínia].
 Ag Pac Companhia

Aqui apareceu o papel de Companhia. Podemos reproduzir esse teste com outros verbos, e digamos que em todas as tentativas o papel que aparece é sempre o de Companhia.³⁰ Aí poderemos concluir que Companhia é o papel temático não marcado (ou *by default*) da preposição *com*. Quando um SPrep introduzido por *com* tem esse papel temático, não é necessário incluí-lo na diátese (a menos que sua presença seja **exigida**, ver adiante).

8.10.5.4. Ocorrência obrigatória

³⁰ Vou admitir que também em *o Toninho comeu a pizza com ketchup* o SPrep tem o papel de Companhia.

Mas o critério da previsibilidade (ou marcação) não deve ser o único, pelo menos se quisermos respeitar inteiramente as intuições tradicionais, por causa de casos como

[57] A minha prima mora em Fortaleza.

[58] A minha prima trabalha em Fortaleza.

Tradicionalmente se distingue esses dois casos porque, embora o papel temático seja o mesmo, e seja provavelmente não marcado para sintagmas da forma *em+ SN*, com *morar* o SPrep é de ocorrência obrigatória (ou, mais exatamente, alterna com um sintagma adverbial, mas não com zero), ao passo que com *morar* a ocorrência é opcional. Ou seja, em [57] temos um caso típico de ocorrência não marcada e não dependente da diátese; mas em [58] é preciso especificar na diátese que o SPrep é de ocorrência obrigatória.

Isso nos obriga a acrescentar um novo critério, a saber, que os constituintes de ocorrência obrigatória (por exigência do verbo) precisam ser marcados como tais na diátese em questão. A obrigatoriedade de ocorrência também é um caso de imprevisibilidade (já que podemos considerar a opcionalidade como o caso não marcado em geral).

A obrigatoriedade de ocorrência não é um critério sem problemas. Em muitos casos, há indicações de que se trata de efeito de fatores semânticos ou informacionais; se for assim, é possível que a obrigatoriedade não deva ser levada em conta para a formulação das diáteses, pois não seria imprevisível a partir de outros fatores. Por outro lado, levar em conta fatores semânticos do verbo para a formulação das diáteses nos leva a uma situação incômoda: inviabiliza o uso dos resultados da pesquisa em diáteses para ajudar a resolver o problema de até que ponto há correlação entre as diáteses e o significado dos verbos – relembro a citação de Levin, que repito aqui:

Se o comportamento distintivo das classes de verbos em relação às alternâncias de diáteses decorre de seu significado, qualquer classe de verbos cujos membros funcionam paralelamente quanto às alternâncias de diáteses deve ser uma classe semanticamente coerente: seus membros devem compartilhar pelo menos algum aspecto de seu significado.

[Levin, 1993: 14]

Essa me parece uma questão importante, que merece investigação. Mas para que a investigação faça sentido, é indispensável que as diáteses sejam definidas independentemente dos traços semânticos do verbo; de outra forma nos arriscaríamos em cair em circularidades do tipo “verbos com o traço semântico T têm a diátese D”, onde a diátese D se define, em parte, pela presença do traço semântico T.³¹

Em casos como o de [57] – e provavelmente em todos os casos tradicionalmente analisados como de “complementos” – há traços semânticos do verbo que o vinculam mais estreitamente com os complementos do que com os adjuntos. Mas essa informação não pode ser utilizada para decidir o que deve figurar nas diáteses e o que não deve, de modo que continuamos com o problema de frases como [57] nas mãos.

³¹ Repito aqui o argumento já visto em 8.5.1.5**.

A obrigatoriedade de ocorrência do SPrep de [57] não é uma característica geral dos elementos presentes nas diáteses (pois o objeto direto é freqüentemente de ocorrência opcional). E, de qualquer modo, obrigatoriedade é um conceito inconfiável, como já apontou Bosque (1989). Ele dá alguns exemplos que mostram claramente como o critério da supressão é falho quando se trata de distinguir complementos de adjuntos. Admitindo que o modificador (“adjunto adnominal”) é um adjunto, seria de esperar que nunca fosse indispensável à aceitabilidade da frase. Mas Bosque cita casos como

[89] As igrejas dos países escandinavos são feitas de madeira.

onde o adjunto *escandinavos*, se omitido, causa inaceitabilidade:

[90] * As igrejas dos países são feitas de madeira.

A menos que se analise *escandinavos* como complemento, esses exemplos mostram que o critério de omissibilidade não funciona, nem mesmo em sua forma mais cautelosa, como

os adjuntos são sintaticamente omissíveis: tirar um adjunto não resulta em agramaticalidade [...] Os complementos, ao contrário, podem ser obrigatórios [...] [Huddleston, 1984: 178]³²

Há provavelmente algum fator que ainda não compreendemos, responsável pela intuição que leva todos os autores a analisar o SPrep de [15] como complemento (portanto incluindo-o na diátese), mas não o de [16]. Nesse caso, vou simplesmente optar pela solução tradicional, seguindo o princípio de que é melhor fazer distinções demais (que podem ser eventualmente eliminadas) do que de menos. A justificativa deve existir, mas não está disponível no momento.³³

Ou seja, a frase [57] será considerada uma realização da diátese

[91]	X	V	SPrep ~ SAdv
	Localizando		Lugar

ao passo que [58] realiza a diátese

[92]	XV
	Ag

8.10.6. Verbos que recusam sujeito

³² “Podem ser” obrigatórios: *sic*. Exemplos como o de Bosque sugerem que a omissibilidade, pelo menos em muitos casos, depende não de fatores gramaticais, como o estatuto de complemento ou adjunto, mas de fatores de ordem informacional.

³³ Estou estudando esse problema, que deverá ser tratado em outro texto.

Um caso especial é o dos verbos tradicionalmente considerados como “sem sujeito”, a saber, *haver*, *ter* e os meteorológicos: *chover*, *nevar*, *relampejar*, *ventar* etc.

Deixando de lado *haver*, que praticamente não ocorre na língua falada, o fato de *ter* poder ocorrer com ou sem sujeito não tem nada de excepcional, dados os pressupostos deste trabalho: como definimos cada verbo em termos de sua representação fonomorfológica (ver 8.4**), o verbo será o mesmo em

[93] Tem dezenas de mangueiras no meu sítio.

[94] As flores já tinham brotado.

O problema real é a diferença de significado associada à presença ou ausência de sujeito, o que não se verifica com outros verbos. Por ora não sei como integrar esses casos na análise geral das valências. Vamos marcar *ter* como uma particularidade da língua, um verbo que pode ou não ocorrer com sujeito (como qualquer outro verbo), mas que apresenta uma alternância semântica entre as duas situações.³⁴ A valência propriamente dita de *ter* não tem nada de extraordinário; é o seu comportamento semântico que é excepcional. Tentei codificar isso, provisoriamente, na definição da construção de apresentação de existência, como se verá no capítulo 12**.

Já os verbos meteorológicos apresentam um problema diferente. *Chover*, na verdade, não tem nada de realmente notável, se considerarmos que ocorre com sujeito em frases como

[95] Choveram pedras de mais de 5 cm.³⁵

Mas *nevar*, *relampejar*, *ventar* nunca ocorrem com sujeito, e isso requer uma explicação. Vou argumentar que o fenômeno pode ser explicado em termos não-gramaticais, e que portanto esses verbos podem ser considerados normais do ponto de vista de suas valências.

A raiz do problema está na semântica de *chover*, que é mais rica do que a dos outros verbos do grupo. Ou seja, *chover* denota também a queda de outros objetos que não gotas de chuva (como em [95]), ao passo que *nevar* só se refere a neve e a nada mais. Por conseguinte, o sujeito de *nevar* (que é o seu Tema) não precisa ser expresso em virtude da Máxima de Quantidade de Grice (1975); em outras palavras, não se expressa porque é óbvio.

Assim, concluo que não há problema com a representação do sujeito dos verbos meteorológicos, e estes podem ser analisados como qualquer outro verbo: podem ocorrer com ou sem sujeito, mas alguns deles nunca ocorrem com sujeito por razões pragmáticas. Ou seja, ocorrem na construção meteorológica (sem complemento nenhum) ou então, raramente e só alguns deles, na ergativa (como em [104]).

8.10.7. Complementos especificados semanticamente?

³⁴ O mesmo ocorre com alguns outros verbos, como *estar*, *ir* etc.

³⁵ Frase aceitável em um registro bastante formal. Seria mais natural dizer *choveu pedra de mais de 5 cm*.

Vimos que a formulação das diáteses deve ser feita em termos simbólicos, incluindo portanto informação formal e semântica. No entanto, em alguns casos a informação formal pode parecer redundante. Por exemplo, o verbo *morar* ocorre em frases como

[96] Nós moramos em uma fazenda.

[97] Nós moramos aqui.

Vemos que o complemento pode ser representado, sintaticamente, por um sintagma da forma *em* + SN ou então por um advérbio, *aqui*. Essa correlação entre sintagma preposicionado e advérbio não é universal, como se vê em

[98] Ninguém confia no governo.

[99] * Ninguém confia aqui.

É claro que o que se mantém no complemento de [96] – [97] é o papel temático Lugar, comum ao sintagma *em uma fazenda* e ao advérbio *aqui*. Seria, então, o caso de definirmos o complemento nessa diátese de *morar* apenas com a indicação do papel temático Lugar?

Note-se que em todos os casos a relação de Lugar pode ser expressa alternativamente por um sintagma preposicionado (com *em*, *debaixo de* etc.) ou por um advérbio de determinado tipo (*aqui*, *lá*...). Logo, parece à primeira vista que não haveria inconveniente em formular a diátese de *morar* em [96] – [97] como

[100]	X	V	s
	Tema		Lugar

Onde ‘s’ designa qualquer tipo de sintagma que possa receber o papel temático indicado.

Essa seria uma simplificação atraente (e mesmo elegante), mas infelizmente não funciona. O problema é que com o verbo *morar* o Lugar é codificado através de um advérbio ou de um sintagma preposicionado, mas com *habitar*, verbo de semântica muito próxima, o mesmo papel temático se codifica através de um SN:

[101] Os tupinambás habitavam o litoral.

Essa é uma diferença valencial entre *morar* e *habitar*, e precisa ser expressa. Mas a notação dada em [100] escamoteia a diferença entre os dois verbos, e além do mais dá a entender que o Lugar associado a *morar* pode ser expresso por um SN (já que há SNs que têm essa potencialidade); ou seja, prevê a aceitabilidade de

[102] * Nós moramos uma fazenda.

[103] * Os tupinambás moravam o litoral.

Esses casos parecem ser comuns; para dar mais um exemplo, o papel de Meta pode ser expresso por sintagma preposicionado, mas não com SN, se o verbo é *chegar*:

[104] Os alpinistas chegaram ao pico / no pico.

[105] * Os alpinistas chegaram o pico.

Mas com *atingir* o mesmo papel temático, Meta, é expresso obrigatoriamente por SN:

[106] Os alpinistas atingiram o pico.

[107] * Os alpinistas atingiram ao pico / no pico.

Aqui se aplica o mesmo raciocínio, que nos impede de caracterizar formalmente o complemento do verbo como uma variável. Por isso, a notação de [100] não é adequada para nossos propósitos, e a construção exemplificada em [96] e [97] precisa ser formulada como

[108]	X	V	Adv _l ~ Prep	SN
	Tema		Lugar	

Concluindo, quando ao se definir uma construção um papel temático tiver várias representações sintáticas possíveis, estas deverão ser listadas uma a uma. Ou seja, não há, na definição das construções, casos de papéis temáticos de expressão formal livre. As construções que nos interessam (as diáteses verbais) devem ser totalmente definidas em termos simbólicos.

A notação exemplificada em [108], na verdade, não é totalmente explícita, porque não especifica a preposição, usando o símbolo esquemático “Prep”. No caso, isso não é inconveniente, porque com o verbo *morar* qualquer preposição serve, desde que exprima o papel temático Lugar. Mas há casos em que a identidade da preposição é relevante. Por exemplo, o verbo *confiar* exige a preposição *em*, *gostar* exige *de* etc. Certamente, a preposição precisa figurar de maneira individualizada na formulação da construção. Por exemplo, a frase

[109] Mané gosta de café.

é uma realização da construção

[110]	X	V	<i>de</i> SN
	Exp		Caus. Exp

que é uma diátese de *gostar*. Não podemos colocar “Prep” aqui, porque não basta que o complemento tenha o papel temático Causador de Experiência – é exigência do verbo que a preposição seja *de*, e nenhuma outra.

Já com o verbo *confiar* os papéis temáticos são provavelmente os mesmos,³⁶ mas a preposição tem que ser *em*:

[111] Bené confia em Zé.

Temos que analisar essa frase como realização de outra construção, de modo que *gostar* e *confiar* não têm a mesma valência, porque a diátese [110] só vale para *gostar*.

Esses casos confirmam que a informação semântica não é suficiente para caracterizar os diversos complementos; a notação simbólica é indispensável.

8.10.8. Inversões

Há ainda o problema das inversões, ou seja, de pares como

[112] Um gato morreu.

[113] Morreu um gato.

Pares como esse (já considerados em maior detalhe em 4.2.3) não são sinônimos em todos os aspectos, mas os papéis temáticos são preservados. Seguindo a orientação deste trabalho, distinguiremos duas construções se a possibilidade de inversão depender do verbo da oração, sem possibilidade de previsão a partir de traços semânticos desse verbo. Mas se essa previsão for possível poderemos consignar apenas uma das construções, relegando a outra ao campo das construções irrelevantes para efeitos de formulação de valências (como é o caso da construção topicalizada).

A tendência geral da literatura parece ser a de admitir que a posposição do sujeito é previsível, embora dependa de uma variedade bastante grande de fatores semânticos e informacionais. Ou seja, não haveria verbos individuais que, independentemente de seu significado ou do valor informacional de seus argumentos, impediriam a posposição do sujeito. Uma afirmação típica é a de que

nas línguas românicas que permitem a inversão sujeito-verbo, o sujeito não-tópico de uma sentença tética aparece obrigatoriamente em posição pós-verbal [...]

[Lambrecht, 1994: 169]³⁷

No capítulo 4 dei uma série de critérios para identificar o sujeito, inclusive o sujeito posposto. Se esses critérios funcionam, como creio que funcionam, segue-se que a posposição do sujeito é previsível, e essa é a posição que assumo aqui. Por conseguinte,

³⁶ Cf. Levin (1993: 191-2).

³⁷ Uma sentença tética é uma que “não predica uma propriedade de alguma entidade, mas simplesmente afirma [...] um fato ou um estado de coisas” [Lambrecht, 1994: 140] – o que nos interessa aqui é que se trata de uma categoria de sentenças definida semanticamente. Ver também Contreras (1983).

quando houver duas ou mais estruturas que tenham os mesmos constituintes, com a mesma distribuição de papéis temáticos, mas em ordem diferente, elas serão consideradas realizações da mesma diátese. Assim, [112] e [113] são duas realizações da diátese ergativa de *morrer*.³⁸ A construção exemplificada por ambas as frases é a ergativa, cuja estrutura sintática se exprime simplesmente como **X V**.

8.11. Problemas

8.11.1. Distinções relevantes

Há vários problemas relativos à noção de “diátese” que não estão suficientemente esclarecidos. Vou tentar formulá-los e discuti-los na medida do possível, dado o estado atual de nossos conhecimentos.

Como vimos, uma diátese é uma construção, definida em termos simbólicos, onde podem ocorrer certos verbos. Assim, quando falamos da “diátese ergativa”, estamos falando da construção ergativa, definida como se viu na seção anterior, e na qual podem ocorrer verbos como *assustar*, mas não *comer*, *castigar*, *ter*, *ser* ou *ficar*. Assim, a definição das diversas diáteses da língua se reduz à definição de um conjunto de construções consideradas relevantes para esse fim. Os verbos se subclassificam segundo a valência de cada um – ou seja, segundo o conjunto de suas diáteses.

8.11.1.1. Diáteses e não-diáteses

Uma construção não é uma diátese quando qualquer verbo pode ocorrer nela, porque nesse caso ela não define uma subclassificação dos verbos. Assim, a construção topicalizada, exemplificada em

[114] O filé eu vou comer mais tarde.

não é governada pelo verbo *comer*, porque qualquer que seja o verbo da oração sempre será possível antepor o objeto direto (isso é sujeito a certas restrições, mas estas não têm a ver com o verbo). Logo, a construção topicalizada não é relevante para a definição de diáteses: não podemos subclassificar os verbos distinguindo os que podem e os que não podem ocorrer nessa construção. O mesmo vale para a construção negativa, a interrogativa, a construção sem sujeito exemplificada em *cheguei ontem*, e, muito provavelmente, a construção de sujeito posposto como em *apareceu um policial*. Em todos esses casos, ou não há restrição (por exemplo, não há verbos que não possam ser negados), ou as restrições existentes não têm a ver com a subclasse do verbo, mas com o entorno sintático ou semântico (por exemplo, o sujeito posposto é mais comum quando é novo, e não dado).

³⁸ Ou seja, dou a esses grupos de frases o mesmo tratamento que a pares de frases que se distinguem por topicalização, por negação etc.

8.11.1.2. Restrições seletivas

É fácil ver por que a construção topicalizada não deve ser levada em conta para efeito de subclassificação dos verbos. Um problema bem mais complexo é o de distinguir as próprias construções entre elas, ou seja, saber com alguma precisão quando é que estamos lidando com duas construções e quando temos em mãos apenas duas ocorrências da mesma construção, diferentes em algum detalhe irrelevante para nossos objetivos.

Mas quando é que um detalhe é irrelevante? Se levarmos a questão a seus limites, veremos que os verbos se subclassificam de maneiras variadas e em extensões variadas. Isso se aplica muito em especial a casos de diferenciação semântica. Por exemplo, os verbos *comer* e *beber* não ocorrem, estritamente falando, nas mesmas construções, por causa de

[115] Antônio comeu a pizza / bebeu o vinho / * bebeu a pizza / * comeu o vinho.

Temos que distinguir aqui duas construções? Note-se que a diferença tem a ver com o verbo, mais precisamente com sua semântica: *beber* significa ingerir alguma coisa de certa maneira, uma maneira que exclui a possibilidade de se beber uma pizza.³⁹

Outro exemplo são os chamados “objetos cognatos” (ou “internos”) que ocorrem em frases como

[116] Machado morreu uma morte tranqüila.

Sabemos que o objeto de *morrer* só pode ser *morte* (mais uma qualificação qualquer: *tranqüila*, *heróica*...). Isso, evidentemente, se vincula à semântica do verbo – temos aqui uma construção relevante e, portanto, uma nova diátese? Como diferenciar esses casos dos exemplos clássicos de diátese? São casos chamados “shadow arguments” por Pustejovsky (1995), que, no entanto, não os conceitua devidamente, limitando-se a dizer que são “semanticamente incorporados no item léxico”.

Não se pode negar que a possibilidade de *comer* e *beber* ocorrerem com *pizza* ou *vinho* acarreta uma subclassificação desses verbos. Esse tipo de subcategorização recebe o rótulo de **restrição seletiva**. Note-se que a subclassificação do tipo exemplificado com *comer* / *beber* ocorre com quase todos os verbos, já que não se pode *ler* um bolo, nem *pintar* um fonema, e um cachorro pode *correr*, mas não pode *transcorrer*.

Para formular as restrições seletivas, é necessário lançar mão de traços semânticos inerentes aos itens léxicos: *vinho* é “líquido”, e isso não depende de onde essa palavra ocorra. Já o caráter de Paciente de um sintagma decorre da estrutura onde ocorre. Um sintagma como *os alunos* não é Paciente por si mesmo, mas recebe um papel temático que pode variar: seria Paciente em

[117] Os alunos desanimaram.

e Agente em

³⁹ *Comer*, *beber* e *tomar* dividem entre si os objetos “ingeríveis”: *tomar a sopa*, mas não **beber* /* *comer a sopa*; *tomar um café*, mas ??*beber um café*, *tomar um comprimido*, mas não **comer um comprimido* etc. Pode-se ainda acrescentar *chupar*, que vale para laranjas e jabuticabas.

[118] Os alunos bloquearam a entrada da sala.

Jackendoff (1990) é de opinião que as restrições seletivas devem ser tratadas em um nível de análise separado da subcategorização através da qual se definem as construções:

As restrições seletivas são restrições semânticas gerais sobre os argumentos, que podem entrar em detalhes muito mais finos do que simplesmente as categorias conceptuais [...] O nível lingüístico apropriado para expressar [as restrições seletivas] é a estrutura conceptual e não a sintaxe ou um presumível nível de estrutura argumental.

[Jackendoff, 1990: 51-52]

A posição dos cognitivistas não considera níveis separados de análise, mas ainda assim distingue as restrições seletivas, vinculando-as a traços semânticos específicos:

O modo clássico de tratar de fatos dessa natureza é em termos de “restrições seletivas”: *above* seleciona, como seu complemento [...], nominais de certo tipo semântico. A gramática cognitiva trata esses fatos afirmando que *above* já contém, dentro de seu perfil, uma especificação esquemática de seu trajecto e de seu pano de fundo [*landmark*] como entidades espacialmente limitadas.

[Taylor (2002: 207)]

Os traços de subcategorização que integram a definição das diáteses são também, pelo menos em parte, semânticos: um traço como a presença de objeto direto, por exemplo, é muito provavelmente motivado por fatores de significado. Ou seja, é possível que a diferença entre restrições seletivas e traços de subcategorização seja uma questão de grau (de especificidade), e não de qualidade; mas essa questão simplesmente não pode ser esclarecida no momento, dado o estágio precário de nosso conhecimento do assunto. Levin fornece um critério que entendo mais como um ideal a ser alcançado do que como ponto de partida:

[...] o verbete léxico ideal para uma palavra deve minimizar a informação fornecida sobre aquela palavra.

[Levin, 1993: 11]

Ou seja, devemos procurar generalizações sempre que possível, de maneira que a informação correspondente possa ser incluída na gramática. O problema, claro, é que para atingir essas generalizações é necessário examinar previamente os fatos em detalhe. Assim, embora sem condições de distinguir sempre com a clareza desejável esses dois tipos de fenômeno, vou deixar de lado as distinções entre os verbos que provêm daquilo que se chama, na literatura, “restrições seletivas”, considerando-as irrelevantes para efeitos de definição de diáteses. Isso significa que devemos definir construções em termos da presença, por exemplo, de elementos como um SN após o verbo (objeto direto), associado a um papel temático (por exemplo, Paciente). Mas não devemos definir

uma construção em termos de um tipo semântico específico de paciente: por exemplo, paciente representado por “líquido”, que caracteriza o verbo *beber*. Essa delimitação deve gerar uma descrição de nível suficientemente amplo para permitir a emergência de generalizações interessantes, o que se aplica não só ao caso de (*beber x comer*), mas provavelmente também ao caso do objeto cognato de.

Levin (1993), a meu ver, faz distinções desnecessárias, por serem previsíveis a partir de outros fatores, não sendo assim marcas distintivas do verbo. Por exemplo, ela distingue como duas construções a dos verbos “de dança” e a dos verbos “de corrida”, exemplificados respectivamente em frases como

[119] Eles dançaram. / Eles dançaram para dentro da sala.

[120] Eles correram. / Eles correram para dentro da sala.

Em português (como em inglês) o comportamento desses dois grupos de verbos é semelhante; quaisquer diferenças podem muito bem ser atribuídas a fatores independentes – em particular, ao significado individual de cada verbo. Daí a diferença de aceitabilidade entre

[121] Eles correram cinco quilômetros.

[122] ?? Eles dançaram cinco quilômetros.

Estritamente falando, [121] e [122] ilustram uma diferença entre *correr* e *dançar*, a saber, a de que correr é uma atividade que em geral se faz seguindo uma trajetória em direção a um destino, ao passo que dançar não tem, geralmente, uma meta espacial. Mas é preciso lembrar sempre que a classificação se faz em função de objetivos; e se nosso objetivo do momento é a descrição gramatical, essa diferença entre *correr* e *dançar* pode ser desconsiderada.

8.11.1.3. Extensão de sentido

Um exemplo da aplicação desses princípios ao problema de diferenciar ou não certas diáteses é o das construções exemplificadas em

[123] O rapaz está doente.

[124] O rapaz está em Maceió.

Devemos definir duas construções distintas, com base nesses casos? A diferença semântica associada à diferença de complementos é grande, e em muitos casos não há substituíbilidade, como em

[125] O rapaz é doente / * o rapaz é em Maceió.

No entanto, vale perguntar se as diferenças derivam de características do verbo enquanto item léxico ou se decorrem de condições de boa formação relativas a nosso conhecimento do mundo. Nosso objetivo é o de isolar (na medida do possível) os fatores léxico-gramaticais dos fatores de ordem pragmática.

Acontece que com os casos acima há indicações de que as diferenças não são léxico-gramaticais. É possível construir uma explicação para os casos de inaceitabilidade em termos do que sabemos acerca das relações no mundo real: ou seja, fica difícil vincular as aceitabilidades a itens léxicos ou a estruturas gramaticais. Assim, a impossibilidade ilustrada em [125] não tem a ver com incompatibilidade entre o verbo *ser* e o complemento de lugar, porque podemos dizer

[126] A Praia do Sobral é em Maceió.

A explicação é que sabemos que um rapaz não tem uma localização fixa, que é o que o verbo *ser* exprime nesse contexto, ao passo que uma praia tem. Outro exemplo disso é

[127] Elza está cheia de problemas / Elza é cheia de problemas.

[128] Elza está cheia de lasanha / * Elza é cheia de lasanha.

“Cheia de problemas” é algo que pode ser temporário ou permanente, ao contrário de “cheia de lasanha”, que só pode ser temporário. Se escolhermos nominais que descrevam uma relação de qualidade (ou de lugar, ou o que seja) que se toma como permanente, *estar* vai ficar excluído, e é o que acontece em

[129] Lionel é gaúcho / * Lionel está gaúcho.

Parece claro, portanto, que esse tipo de diferença não tem a ver com a valência do verbo envolvido, mas com condições que dependem de fatores globais de conhecimento do mundo, como “uma praia não pode se mover”, e “alguém pode ser permanentemente cheio de problemas”. Não há como prever as aceitabilidades em termos lingüísticos (menos ainda em termos do verbo usado), e portanto esse tipo de diferença deve ser excluído da nossa lista de construções.

Esse caso contrasta nitidamente com o de distinções relevantes como a que existe entre as construções transitiva e ergativa. A diferença aparece em

[130] Filipe rasgou o livro.

[131] O livro rasgou.

[132] Filipe ilustrou o livro.

[133] * O livro ilustrou.

Em primeiro lugar, não há razão puramente semântica para excluir [133], já que essa construção exprime a mesma mudança de estado expressa em [132], apenas com a escamoteação do agente. Apesar disso, o verbo *ilustrar* nunca pode ocorrer na construção ergativa (com sujeito paciente, relembro), sejam quais forem os outros termos da frase. É pouco provável que essa diferença possa ser reduzida a propriedades semânticas autônomas dos verbos envolvidos. É igualmente implausível que se trate de restrições de ordem pragmática, porque nada impede que um livro seja ilustrado, assim como pode ser rasgado. No entanto, para exprimir a idéia de que o livro foi rasgado temos a opção de uma construção ergativa, como em [131]; mas para dizer que o livro foi ilustrado essa opção não funciona.

Nesse caso, então, devemos atribuir a diferença aos verbos. Por isso, a diferença entre transitivas e ergativas é relevante para a classificação dos verbos, seja ela de base semântica, seja de base formal. Em outras palavras, a diferença de aceitabilidade entre [131] e [133] é de natureza léxico-gramatical, e portanto deve ser incluída em nossa lista de construções: a distinção entre as construções ergativa e transitiva continua sendo relevante para nossos objetivos – ao contrário da diferença entre [123] e [124], que pode ser atribuída a fatores independentes da valência dos verbos *ser* e *estar*.

É como se tanto em [123] quanto em [124] tivéssemos um papel temático esquemático, que pode talvez ser definido como “localização em algum tipo de espaço” – onde “espaço” pode ser entendido por extensão de sentido. Ou seja, em [123]

[123] O rapaz está doente.

estamos dizendo que o rapaz se localiza em determinado ponto em algum “espaço”. Que espaço é esse é coisa que a língua deixa por conta do usuário. No caso, como se trata de “estar doente”, e sabendo que “doente” é um estado de saúde, o receptor não tenta colocar mentalmente o rapaz em algum ponto do espaço geográfico, mas antes em um ponto de um “espaço” transferido que inclui “bem de saúde”, “doente”, “mal”, “à morte”, “bem melhor” etc.

Jackendoff (1990), seguindo uma proposta de Gruber (1965), coloca essa possibilidade da seguinte maneira:

Em um campo semântico de [EVENTOS] e [ESTADOS], as principais funções de evento, estado, trajetória e lugar são um subconjunto das que são usadas para a análise da localização e movimento espaciais.

[Jackendoff, 1990: 188]

Essa hipótese quer dizer que quando exploramos a organização de conceitos que não têm contraparte perceptual não precisamos começar do zero. Em vez disso, aplicamos uma relação transferida, adaptando mecanismos já disponíveis às novas necessidades. No caso, utilizamos as relações sintáticas e temáticas de “lugar” na função de expressar um estado de saúde.

O recurso que consiste em fornecer informação esquemática para eventual elaboração por parte do usuário é um aspecto do caráter maximamente econômico da linguagem. Repetindo o que foi dito em 7.4.2** acima, “o sinal lingüístico explícito fornece a informação mínima necessária para que o receptor, de posse de seu

conhecimento do mundo, possa construir uma paisagem mental que faça sentido. A razão principal dessa economia é a limitação dos recursos que a linguagem coloca à nossa disposição, frente à infinidade de conceitos que a experiência nos oferece.”

Os dois exemplos dados acima – o de [123]-[124] e o de [131]-[133] – são toleravelmente claros, mas essa situação favorável não se verifica em todos os casos. É de se esperar que haja freqüentemente dificuldades em decidir se o que temos em mãos é uma construção ou duas. Nesses casos, a conveniência descritiva determina que se mantenha a distinção, já que é mais fácil reunir duas classes em uma só do que dividir uma classe em duas, porque neste último caso torna-se necessário refazer todo o levantamento, reanalisando e reclassificando cada verbo. Os resultados, portanto, devem ser considerados provisórios em princípio, e todas as distinções podem ser eventualmente contestadas.

Voltando ao caso de [123] e [124],

[123] O rapaz está doente.

[124] O rapaz está em Maceió.

é necessário também perguntar qual é a semelhança semântica entre essas duas noções, que pelo menos à primeira vista parecem tão distintas: um estado de saúde (*doente*) e uma localização espacial (*em Maceió*). Já vimos que não é de se esperar que um papel temático único recubra relações totalmente sem relação uma com a outra. Isso é logicamente possível, mas vai contra o que se sabe em geral do funcionamento da linguagem; portanto, temos que procurar alguma analogia entre as diversas noções conceptuais expressas pelo papel temático proposto.

8.11.1.4. Traços léxicos inerentes: os complementos

Pode-se também discutir a conveniência de incluir na valência traços semânticos internos dos complementos, o que se faz com freqüência. Assim, o sistema de Gross (1975) utiliza categorias semânticas dos complementos:

Os *N* podem [...] ser indexados por símbolos semântico-sintáticos (N_{hum} , N_{op} , N_q , etc.) que dão indicações sobre o conteúdo do grupo nominal, e não mais sobre sua posição na frase.

[Gross, 1975: 14]⁴⁰

Esse uso de categorias semânticas mais gerais pode talvez ser entendido como um uso controlado de restrições seletivas. Por outro lado, contribui para sobrecarregar a classificação, dificultando sua utilização para efeitos teóricos.⁴¹

⁴⁰ N_{hum} pode ser traduzido como “SN com traço ‘humano’ ” e assim por diante.

⁴¹ As tabelas de Gross são tão complexas que se fez necessário publicar uma espécie de “guia” para extrair informação delas (Gardent *et al.*, 2005).

O problema, tal como o vejo, é que uma classificação não tem limites naturais, a não ser os itens individuais: se levarmos uma classificação dos verbos longe o suficiente, teremos tantas classes quantos verbos existirem na língua. É preciso, portanto, desenvolver critérios para manter a situação sob controle, sob pena de ver os traços mais relevantes submersos em uma multidão de detalhes que, para os objetivos do momento, não são necessários.

É possível, claro, seguir o exemplo de Gross e incluir traços como [+humano] nos argumentos que acompanham cada verbo. É igualmente possível deixar de fazer isso, e o que se perder em informatividade se ganhará em simplicidade e transparência das representações. Mas acredito que há razões empíricas para preferir a segunda solução, deixando de fora tais traços semânticos. Digamos que se analise a frase

[134] Marisa abriu a porta.

como tendo um sujeito marcado [+ animado]; nesse caso, a análise de [134] será diferente da de

[135] O vento abriu a porta.

porque aqui o sujeito é [- animado]. Por conseguinte, teríamos nessas frases duas diáteses diferentes do verbo *abrir*.

Mas isso contraria uma observação de fato que me parece relevante: quando o sujeito é [+ animado] e também quando é [- animado], a relação temática inclui um ingrediente comum (causação direta), e a codificação sintática é igualmente idêntica – isso nas frases acima e também em todas as outras possibilidades, como

[136] A porta foi aberta por Marisa.

[137] A porta foi aberta pelo vento.

Ou seja, há evidência de que a língua trata o sujeito de [134] e o de [135] de maneira idêntica, para efeitos de ocorrência em construções. A mesma situação parece se verificar em conexão com outras construções, como por exemplo

[138] As torradas encheram quatro caixas.

[139] A multidão encheu o estádio.

[140] Vinte soldados formam um pelotão.

[141] Vinte páginas formam um caderno.

[142] O menino esfriou.

[143] O leite esfriou.

[144] O presidente transformou o país em um caos.

[145] A nova lei transformou o país em um caos.

Ou seja, o traço [\pm animado] não é relevante para definir a possibilidade de um verbo ocorrer em determinada construção, e portanto não é relevante para a formulação das valências. Naturalmente, há casos em que as restrições seletivas causam diferenças de aceitabilidade, mas isso é outra coisa; ou seja, a diferença entre

[146] Marisa vai te entender.

[147] * O vento vai te entender.

não tem nada a ver com a valência do verbo *entender*, mas com o que sabemos da natureza do vento.

É interessante observar que as violações seletivas, como a de [147], podem ser suspensas em determinados contextos. Assim, [147] poderia ser aceitável no contexto de uma fábula, ao passo que uma violação de valência como

[148] * Marisa entendeu ao problema / com o problema / do problema.

não pode ser salva por nenhuma manipulação do contexto ou do espaço mental.

Concluo que o traço [\pm animado] não deve figurar na definição das valências verbais; e, por analogia (já que não há pesquisas abrangentes sobre o tema), estendo a mesma conclusão para os demais traços que individualizam os diversos sintagmas que funcionam como complementos verbais.

8.11.1.5. Traços léxicos inerentes: o verbo

Outra pergunta a ser feita é se devemos incluir na definição das diáteses traços inerentes ao item léxico verbal, como “verbo de ação”, “de estado”, “de mudança de estado” etc. Vou argumentar que não.

Uma das coisas que o estudo de valências deve permitir é verificar a adequação da hipótese da correlação significado-valência, expressa por Levin (1993: 14), citada várias vezes neste trabalho:

Se o comportamento distintivo das classes de verbos em relação às alternâncias de diáteses decorre de seu significado, qualquer classe de verbos cujos membros funcionam paralelamente quanto às alternâncias de diáteses deve ser uma classe semanticamente coerente: seus membros devem compartilhar pelo menos algum aspecto de seu significado.

[Levin, 1993: 14]

Trata-se de uma presumida correlação entre a valência de um verbo e traços do significado individual do verbo. Assim, chegaríamos (por hipótese) a correlações como:

- (a) todos os verbos que ocorrem na construção transitiva têm o traço “ação”;
- (b) todos os verbos que ocorrem na construção de emissão odorífera (como em *esse armário está cheirando a mofo*) têm o traço “propriedade sensorialmente perceptível (pelo olfato?)”;

e assim por diante. Mas para que essas correlações sejam válidas os traços do significado individual dos verbos precisam ficar de fora da formulação das diáteses, caso contrário viciamos a pesquisa já de saída. Por isso, é necessário ignorar os traços do significado interno dos verbos.⁴²

É muito provável – na verdade, é praticamente certo – que essa correlação seja válida até certo ponto. O problema é delimitar com rigor os limites de sua validade. Temos que adotar uma posição mais cuidadosa (como a de Langacker, 1991: 19) e investigar em que condições essa correlação se verifica.

Isso não é motivado por uma preocupação secundária, voltada para uma questão marginal. A importância da hipótese de Levin é que ela constitui uma condição básica de aplicação da Hipótese da Sintaxe Simples – que, como sabemos, é um dos pressupostos básicos da posição descritivista aqui adotada. A Hipótese da Sintaxe Simples requer a verificação de até que ponto é possível levar a descrição de base semântica. Um aspecto disso é a pesquisa da correlação entre as categorias semânticas dos verbos e suas valências: ou seja, até que ponto vai ser necessário definir valências, e até que ponto a co-ocorrência de verbos e complementos é previsível a partir da semântica dos próprios verbos? A resposta a essa pergunta nos fornecerá um importante aspecto da delimitação entre sintaxe e semântica, dentro do quadro criado pela Hipótese da Sintaxe Simples.

Ou seja, podemos fazer a seguinte pergunta:

Até que ponto verbos semanticamente semelhantes mostram valências semelhantes?

Mas para responder essa pergunta, evidentemente, é necessário definir e investigar as valências independentemente da caracterização semântica dos verbos. A pesquisa se realizará em duas frentes: a determinação das valências dos verbos e o agrupamento destes segundo seus traços semânticos inerentes (verbos de estado, de processo, de ação etc.). Uma vez realizado esse duplo levantamento, um confronto detalhado nos permitirá responder a pergunta formulada acima.⁴³

⁴² Apresjan *et al.* (1969) os incluem, mas seus objetivos vão além da formulação das valências (*apud* Vilela, 1992: 17-19). Em geral se distingue claramente a estrutura argumental da especificação do tipo de evento (cf., por exemplo, Pustejovsky, 1995: 62); pode-se dizer, portanto, que a formulação das valências inclui a estrutura argumental, mas não o tipo de evento – que no entanto é uma parte legítima do significado de um verbo.

⁴³ Já estão sendo realizadas algumas pesquisas nesse sentido, por ora ainda bastante preliminares e interessando pequenos grupos de verbos. Dois de meus alunos estão trabalhando no problema: Luciana Mazur com os verbos incoativo-terminativos, e Bruno Lima com os verbos de vitória e derrota (*ganhar*, *perder* etc.). Os resultados desses trabalhos nos fornecerão uma primeira e limitada visão da resposta que se deve dar à pergunta acima.

Uma das conseqüências da decisão de excluir traços semânticos inerentes do verbo é que podemos ter um verbo em determinada diátese apresentando mais de um significado, às vezes muito diferentes entre eles, sem relação perceptível, nem mesmo transferida. Por exemplo, o verbo *compreender* pode ocorrer na construção transitiva em duas acepções:

[149] Helena compreendeu o meu problema.

[150] A herança compreende duas fazendas e um apartamento.

Segundo o sistema aqui adotado, trata-se do mesmo verbo (já que os itens léxicos se definem em termos formais), na mesma diátese – apesar da grande diferença de significado de *compreender* nas duas frases.⁴⁴

Se decidirmos distinguir duas diáteses em [149] e [150], com base nos dois significados do verbo, teremos que enfrentar imediatamente o problema de como delimitar as acepções, sem levar em conta fatores formais: ou seja, como fazer a delimitação em termos puramente semânticos. Isso, como se sabe, é problemático, dado o caráter contínuo do tecido cognitivo; nos obrigaria a estabelecer diferenças não apenas no caso acima, mas também em casos como

[151] Helena viu o gato.

[152] Helena viu o meu problema.

onde é claro que *ver* não denota o mesmo processo em cada caso. Seria talvez necessário suspender o trabalho de investigação das valências à espera de que seja resolvido o problema da delimitação dos conceitos (ou “esquemas”; ver a respeito Perini, 1988; 1999).

Uma maneira de contornar o problema é ignorar essas diferenças semânticas na delimitação do item léxico, vinculando sua identificação a sua representação fonológica e morfológica: desse modo, *ver* é um verbo só, assim como *compreender*, e mesmo casos altamente complexos como *dar*. Isso nos fornece um ponto de partida confiável para o levantamento – e não nos impede de comentar casos notáveis de polissemia, associados ou não a diferenças de valência. Essa solução, que pode não ser a ideal, nos é imposta pelas circunstâncias em que temos que trabalhar, e é perfeitamente compatível com o caráter descritivo, portanto de certo modo preliminar, da pesquisa.

8.11.1.6. O que fica de fora

A valência de um verbo, como se vê, fica muito longe de expressar toda a riqueza de seu significado. Para repetir a velha fábula, estamos estudando apenas uma pequena

⁴⁴ As duas acepções de *compreender* partilham a diátese exemplificada em [149] e [150], mas não têm valências idênticas, porque há diáteses em que ocorre uma e não a outra. Por exemplo, só no significado de “entender” é que *compreender* ocorre na construção passiva.

parte do elefante, e apenas pelo tato. Isso é inevitável, dada a imensa complexidade do fenômeno em estudo.

Por exemplo, ficam de fora as pressuposições codificadas em muitos verbos. Assim, a valência de *aceitar* é a mesma que a de *conseguir*; mas há uma diferença importante entre os dois verbos: quando se *consegue* alguma coisa, pressupõe-se que houve um esforço para alcançar esse resultado, ao passo que se pode *aceitar* uma coisa sem nenhum esforço de nossa parte. Isso não aparece na valência, e vai ter que ser incluído em outra parte da descrição.

Igualmente, a valência de *aceitar* é a mesma de *extorquir*, mas só *extorquir* acarreta o uso de algum tipo de violência ou ameaça: outro ingrediente essencial da semântica do item que não aparece na formulação da valência.

A seleção de traços utilizada para diferenciar as construções encobrem diferenças semânticas que podem ser muito grandes. Por exemplo, a construção denominada “de vitória” se define como

[153] X V *de* SN
 Ag Pac

Um exemplo é

[154] O Vasco ganhou do Botafogo.

No entanto, a mesma construção se aplica à frase

[155] A Sílvia mudou de roupa.

Em princípio, não há realmente problema. Mas a distância semântica entre os verbos *ganhar* (no sentido de ‘vencer’) e *mudar* é um pouco desconfortável. Mas não se pode utilizar esse critério para distinguir duas construções, porque isso viciaria a possibilidade de apurar até que ponto as semelhanças valenciais coincidem com semelhanças semânticas.

O que se pode dizer é que dar a valência de um verbo é expressar aspectos selecionados de seu comportamento sintático e semântico, não exprimir sua sintaxe e sua semântica na totalidade. Mas é bom ter em mente também que o sistema aqui proposto é muito pouco delicado nas distinções que faz.

8.12. Quando um verbo não tem valência

8.12.1. Orações com dois verbos

Pode-se dizer que um verbo não tem diátese quando funciona como auxiliar, já que na oração em que ele aparece há sempre outro verbo, e as diáteses desse outro verbo prevalecem. Isso vale não apenas para os auxiliares *stricto sensu* (*ter* + particípio, *estar* + gerúndio, *ir* + infinitivo), mas também para modais como *acabar*. Exemplos com auxiliar são:

[156] Janaína tem engordado. [construção ergativa]

[157] As massas têm engordado Janaína. [construção transitiva]

E exemplos com *acabar* são:

[158] Janaína acabou engordando. [construção ergativa]

[159] As massas acabaram engordando Janaína. [construção transitiva]

Aqui as diáteses representam as do verbo *engordar*; mas em

[160] Honório acabou virando um lobisomem.

aparece a diátese de *virar*.

Acabar nessas frases não tem diátese, a não ser no sentido de que aparece com outro verbo; isso é, de certo modo, uma diátese, mas não do tipo que estamos estudando. Ou seja, certamente a possibilidade de ocorrência com outro verbo, no gerúndio, no infinitivo com *de* ou *por* etc. é um elemento que subclassifica *acabar*. Pode-se, igualmente, argumentar que esse outro verbo é um complemento de *acabar*.⁴⁵ No entanto, optei por lidar exclusivamente com complementos não-oracionais, para evitar o acúmulo de problemas neste primeiro momento do estudo das diáteses. De qualquer forma, permanece o fato de que as diáteses de *engordar* são também válidas para as seqüências *acabar por engordar*, *acabar engordando* etc.

Mas isso não significa que o item léxico *acabar* não tenha valência, porque em outras construções o mesmo verbo aparece (desta vez sozinho) com diáteses nítidas:

[161] Joaquim acabou a reunião. [constr. transitiva]

[162] A reunião acabou. [constr. ergativa]

etc.⁴⁶

Um exemplo ainda mais claro é o do verbo *deixar*. Esse verbo, quando aparece sozinho, tem sempre sujeito “agente”, isto é, controlador do evento:

[163] Joaquim deixou a mulher.

[164] Joaquim deixou a cidade.

Deixar nunca tem sujeito paciente. Mas quando vem na construção com ***de* + infinitivo**, *deixar* pode ter sujeito paciente – desde que o segundo verbo o admita:

[165] Joaquim só deixou de engordar quando adoeceu.

⁴⁵ Mas tendo a achar que não é; ver a respeito Pontes (1973).

⁴⁶ Já há um estudo da valência de *acabar* (e de outros verbos parecidos) em Mazur (2007).

Além disso, *deixar* como verbo principal se limita a sujeito animado, mas não na construção com infinitivo:

[166] Essa caneta deixou de funcionar.

[167] O leite deixou de ferver.

Ou seja, tudo se passa como se *deixar* (assim como *acabar* etc.) não tenha diátese em frases como [165], [166] e [167]: o sujeito que aparece ali é o sujeito do verbo principal, que aparece no infinitivo. Isso é confirmado por casos como

[168] * Joaquim deixou de transcorrer em paz.

O problema é a incompatibilidade de *transcorrer* com o sujeito *Joaquim*; mas a frase seguinte é aceitável:

[169] A vida deixou de transcorrer em paz.

Podemos contrastar esse comportamento de *acabar* com o de *tentar*, que apesar de ocorrer em estruturas sintaticamente semelhantes, como em

[170] Janaína tentou ficar na sala.

apresenta diátese própria, tanto é assim que o sujeito precisa ser animado:

[171] O carro ficou na garagem.

[172] * O carro tentou ficar na garagem.

Verbos como *acabar*, *deixar* acrescentam um ingrediente aspectual ao significado da oração, ao passo que *tentar* representa um predicado autônomo.

A conclusão é que cada **oração** apresenta um conjunto próprio de complementos (não-oracionais); por isso, quando uma oração tem mais de um verbo, apenas um deles contribui com suas diáteses para determinar os complementos possíveis. Nesses casos, o verbo dominante é o que aparece em forma não-finita (infinitivo, gerúndio ou particípio).

8.12.2. Os dois participios

No que diz respeito ao particípio, é preciso observar que esse termo cobre na verdade duas formas morfologicamente parecidas, mas de comportamento gramatical bem distinto, e apenas uma delas é realmente uma forma verbal associada a diáteses. Ou seja, temos que distinguir as propriedades da palavra *consertado* nas duas frases seguintes:

[173] Elisa tinha consertado o fogão.

[174] O fogão consertado é aquele ali da esquerda.

Em [173] *consertado* é invariável, ao passo que em [174] funciona como um modificador, tendo variação de gênero e número: *o fogão consertado, a geladeira consertada, os fogões consertados...* Sugiro que se denomine o primeiro **particípio verbal**, e o segundo **particípio nominal**.

Os dois particípios se distinguem segundo alguns pontos de vista importantes:

Diferença morfológica:

O particípio nominal parece fazer parte de um grupo de nominais (“adjetivos”), derivados do verbo mas não propriamente fazendo parte do lexema verbal. Embora essa derivação seja bastante regular, em certos casos há diferença formal:

[175] Manuel tinha matado / ?* morto a galinha.

[176] A galinha morta / * matada está na geladeira.

Sempre que se verifica essa dualidade (com os chamados “verbos abundantes”), a forma regular é a verbal, a irregular a nominal. Além disso, como se viu, no segundo caso há concordância nominal, o que não ocorre no primeiro.

Diferenças semânticas:

Há também diferenças semânticas, porque a relação entre o particípio nominal e o verbo correspondente não é regular como a que se verifica no caso do particípio verbal. Por exemplo, alguns particípios nominais atribuem ao núcleo de seu SN o papel temático de Paciente, e isso é em geral considerado típico desses itens: em *o fogão consertado* o fogão é Paciente do conserto. Mas há um bom número de exceções, como *um rapaz aborrecido, um palhaço divertido, uma mulher decidida, eu já estou almoçado, um curso muito puxado* etc., onde o núcleo não é Paciente. Isso sugere que nas construções exemplificadas (e que às vezes são identificadas com a passiva) o particípio não pertence ao lexema verbal.

Na voz passiva propriamente dita, esse significado anômalo não ocorre, e o particípio sempre tem seu significado “passivo”. Assim, em

[177] Ele foi aborrecido durante horas.

há uma ambigüidade: *aborrecido* pode significar “chato”, e nesse caso em geral não se analisa a frase como passiva. Mas pode significar “chateado”, e aí temos uma passiva; como se vê, nesse caso (mas não no anterior) o sujeito é Paciente.

Essas idiosincrasias nunca ocorrem dentro do paradigma verbal – por exemplo, a relação semântica entre o presente (*faço*) e o perfeito composto (*tenho feito*) é exatamente a mesma para todos os verbos. Quanto ao particípio verbal, ele atribui a seu sujeito (o

sujeito do verbo auxiliar que o acompanha) o mesmo papel temático que o verbo correspondente; o fato de haver um particípio na construção não é relevante. O particípio verbal se caracteriza por possuir as diáteses do verbo correspondente, ao passo que o particípio nominal se comporta como um adjetivo,⁴⁷ não tendo por conseguinte diáteses verbais. Assim, temos

Particípio verbal:

[178] Júlio come demais / Júlio tem comido demais. [*Júlio* = Agente]

[179] Júlio desanimou / Júlio tinha desanimado. [*Júlio* = Paciente]

[180] O rapaz gosta de Maria / confia em Maria / conta com Maria.

[181] O rapaz tinha gostado de Maria / confiado em Maria / contado com Maria.

Particípio nominal:

[182] * O rapaz gostado de Maria / * confiado em Maria / * contado com Maria.

Como se vê, o particípio nominal não tem as mesmas diáteses dos verbos correspondentes.⁴⁸ O comportamento do particípio verbal é idêntico ao das outras formas verbais que ocorrem com auxiliar, o infinitivo e o gerúndio, e ao das diversas flexões do verbo em geral. Por exemplo, passar o verbo do presente para o pretérito imperfeito ou para a forma composta *estar* + gerúndio não afeta em nada as possibilidades de diátese.

Ou, olhando as coisas de outro ângulo, a valência de um verbo afeta todas as suas formas: tanto o presente quanto o gerúndio, o particípio verbal, o infinitivo ou o pretérito perfeito de *esquentar* ocorrem na construção transitiva, na ergativa e na intransitiva. A valência se expressa de uma vez para cada verbo, sem consideração de tempo, modo, número ou pessoa. Mas o particípio nominal é especializado: se a passiva for analisada como uma diátese do verbo, então teremos a situação anômala de que uma das formas do lexema verbal só aparece em uma diátese, a saber, a passiva, e em nenhuma outra.

Na verdade, a relação entre o verbo e o particípio nominal lembra mais um caso de derivação do que de flexão: a relação do particípio nominal com o verbo é mais parecida com a que existe entre um verbo e sua nominalização (*casar* / *casamento*), ou entre um verbo e seu derivado em *-vel* (*suportar* / *suportável*). Note-se que há mesmo verbos que não possuem particípio nominal, sem serem por isso considerados defectivos (*ser*, *estar*, *ficar* e outros). Não há verbo que não tenha gerúndio, infinitivo ou particípio verbal. Mesmo os chamados defectivos apresentam ausência de certas pessoas, mas nunca de tempos inteiros.⁴⁹ Sabe-se, além do mais, que na língua escrita o particípio é a

⁴⁷ Quero dizer, como um nominal que tem potencialidade “qualificativa”.

⁴⁸ A única possível exceção é uma construção onde o particípio nominal ocorre com algo que se assemelha a um objeto direto, e às vezes mesmo com um sujeito: *a menina tomada banho, esse armário pode ser colocado porta*. Como a construção nunca foi estudada, fica aqui apenas a menção.

⁴⁹ Como as coisas na linguagem nunca são simples, há algumas suspeitas de verbos que carecem de certos tempos: o perfeito de *crer* (? *creu*) é certamente inusitado, assim como o particípio verbal de *equivaler* (? *tinha equivalido*). Mas são casos muito raros.

única forma “verbal” que recusa clíticos (**encontrado-o*, **ajudado-me*, **dado-lhe*). Essa é uma característica que não vale para nenhuma outra forma do lexema verbal. A conclusão mais provável é que o particípio nominal não é parte do lexema verbal, mas uma forma derivada.⁵⁰

Não sou o primeiro a propor uma análise nessas linhas, embora que eu saiba ninguém tenha chegado ao ponto de distinguir duas formas (em geral, mas nem sempre, homônimas) no que tradicionalmente se denomina “particípio”. Assim, Mattoso Camara já observava que

[...] o particípio foge até certo ponto, do ponto de vista mórfico, da natureza verbal. É no fundo um adjetivo com as marcas nominais de feminino e de número plural em /S/. Ou em outros termos: é um nome adjetivo, que semanticamente expressa, em vez da qualidade de um ser, um processo que nele se passa. O estudo morfológico do sistema verbal português pode deixá-lo de lado, porque morfológicamente ele pertence aos adjetivos, embora tenha valor verbal no âmbito semântico e sintático.

[Camara, 1970: 103]

Note-se que Mattoso Camara trata os particípios em bloco, sem considerar os dois tipos de comportamento gramatical que podem ser distinguidos. Não concordo com ele em identificar o particípio com os adjetivos em geral, porque o particípio ocorre em pelo menos uma construção em que os adjetivos não aparecem, a saber, na chamada construção de “particípio absoluto”, exemplificada por

[183] Feitas as contas, descobrimos que faltava mais de um milhão.

Apesar de alguns exemplos citados por Casteleiro (1981), não me parece que os adjetivos não-participais ocorram nessa construção no brasileiro atual.⁵¹ Isso mostra a necessidade de se classificar os particípios, os verbos e os adjetivos de maneira complexa, por exemplo com o uso de traços. Em termos simplificados, pode-se dizer que os particípios nominais têm traços em comum com os verbos (são derivados deles de maneira razoavelmente regular, tanto do ponto de vista morfológico quanto do semântico); têm traços em comum com os adjetivos (na verdade, como vimos, a maior parte de seu comportamento sintático); e, finalmente, têm alguns traços próprios (a ocorrência como particípio absoluto). Essa relação complexa entre as diversas formas da língua é, a meu ver, típico.

Casteleiro (1981), em uma discussão bastante ampla do problema, observa os muitos traços “adjetivais” dos particípios, aliados a alguns traços “verbais”. Ele fornece um bom número de exemplos muito ilustrativos, mas igualmente se atém à idéia de que

⁵⁰ Essa conclusão levanta questões interessantes a respeito da análise do particípio em outras línguas românicas como o francês e o italiano, onde o particípio verbal também concorda em gênero em certos contextos – embora as diáteses sejam diferentes das do verbo propriamente dito, como em português. Pode-se perguntar se a mesma conclusão a que cheguei vale para essas línguas, e levanta-se o problema de como terá surgido essa situação, diacronicamente. São perguntas importantes, mas que não nos concernem aqui, dado a abordagem sincrônica do presente estudo.

⁵¹ Os exemplos de Casteleiro são: *cheios os barris*, ... e *certas as contas*, ... Ambos me parecem inaceitáveis; devem ser exclusivos do português europeu.

se trata ou de uma forma verbal ou de uma forma adjetiva, sem levar em conta a possibilidade de uma categorização complexa. Casteleiro acaba concluindo que

“os participios passados são fundamentalmente formas verbais, que eventualmente podem funcionar como adjetivos”

[Casteleiro, 1981: 85]

Pode-se perguntar o quer dizer “fundamentalmente” nesse contexto. E “funcionar como adjetivo”, como sabemos, é problemático, pois viola a distinção entre classe e função, estabelecida no capítulo 3.

A solução de Casteleiro é transformacional, derivando (como era o padrão nos anos 80) as construções de participio a partir de estruturas passivas (*o livro aberto* < *o livro que foi aberto*), ou então de estruturas com auxiliar *ter* + participio (*as folhas caídas* < *as folhas que tinham caído*). Esta última análise é aplicada apenas aos verbos que recusam objeto direto, o que inviabiliza a derivação a partir da passiva. Ela incorre em dificuldades porque na construção fonte ocorre o participio verbal (invariável), e na construção derivada o participio nominal (variável em gênero e número). Isso resulta em termos duas fontes independentes derivando a mesma construção (passiva ou *ter* + participio > construção com participio modificador), o que é pelo menos suspeito. Finalmente, essa análise ainda requer, para alguns verbos, uma transformação do participio regular no participio regular, como em *o submarino emerso*, que teria que ser derivado de *o submarino que tinha emergido*. Esses últimos casos são raros, mas ainda assim representam uma dificuldade para a análise de Casteleiro.

8.12.3. A passiva é uma diátese verbal?

As considerações da seção precedente têm uma conseqüência importante para a análise da construção passiva – que se torna suspeita de ser uma simples construção de *ser* + adjetivo, em que a semântica do participio nominal é responsável pelo significado “passivo” (ou seja, com núcleo Paciente; cf. os derivados em *-vel*, também geralmente de “significado passivo”). Ou seja, a construção *ele foi subornado* seria estruturalmente semelhante a *ele é subornável*, que também apresenta um sujeito Paciente. Uma indicação nesse sentido é que é o participio nominal que ocorre com a passiva, e não o participio verbal – e acabamos de ver que o participio nominal muito provavelmente não faz parte do lexema do verbo. Isso se pode ver pelos casos em que os dois diferem:

[184] O assaltante foi preso. / * prendido.

Essas observações sugerem que o que temos aqui é uma construção de predicativo (verbo *ser* + nominal), e que por conseguinte esse é um caso de valência nominal e não verbal.

Outra indicação do caráter nominal do participio que aparece na passiva é que alguns de seus complementos se assemelham a complementos de adjetivos, como por exemplo em *ele foi atado de pés e mãos* (ou *amputado das duas pernas*), o que lembra a construção de adjetivo *cego do olho esquerdo*, *aleijado dos braços* etc.

Note-se, além disso, que não há certeza de que *por* + *SN* seja um complemento. Me parece mais um termo autônomo, com preposição predicadora que atribui o papel de Agente, e de ocorrência não vinculada à construção chamada passiva – tanto é assim que ocorre com a mesma semântica em construções não passivas: *a destruição da cidade pelo inimigo*. Essa estrutura de *por* + *SN*, com o significado de Agente, nunca ocorre como complemento de uma forma inequivocamente verbal; ela é típica de estruturas de núcleo nominal.

Um terceiro exemplo, talvez mais próximo do caso da passiva, é o de derivados em *-nte*, como em

[185] O espetáculo foi chocante.

[186] Esse livro é desmotivante.

Ao que parece, o derivado em *-nte* causa a atribuição de uma qualidade, e não de uma ação, ao sujeito de *ser*. Isso dificulta a ocorrência de derivados de verbos que exprimem preferivelmente ações e eventos, e por isso temos *chocante*, *(des)motivante*, *impressionante*, mas não **comente*, **pintante*, **adormecente*.

Há vários paralelismos entre essa construção e a passiva: a estrutura sintática (*SN* + *ser* + nominal derivado); a existência de derivados em *-nte* para muitos radicais verbais; e a atribuição de um papel temático típico ao sujeito de *ser*. No entanto, nunca ninguém analisou [185] como uma diátese, ou uma “voz”, do verbo *chocar*. Faz pensar se não seria adequado encarar a passiva da mesma maneira.

O particípio que ocorre na passiva é praticamente sinônimo de seu homônimo que ocorre em construção predicativa:

[187] O livro que foi lido em classe.

[188] O livro lido em classe.

Em [188] não se vê, geralmente, uma passiva, mas apenas um modificador de *livro*. Mas o caso de [187] – [188] é exatamente paralelo ao de

[189] O livro que é amarelo.

[190] O livro amarelo.

o que sugere uma análise idêntica para as duas duplas: se em [188], [189] e [190] não há passiva, não há razão para que haja passiva em [187] – ou seja, analisa-se [187] como representando a mesma construção que aparece em [189].

Resta mencionar casos de passivas sem ativas correspondentes, como em

[191] D. Marlene é antipatizada por todos os moradores do prédio.

[192] * Todos os moradores do prédio antipatizam D. Marlene.

[193] D. Marlene é muito falada.

[194] * Todos falam muito D. Marlene.

Em ambos esses casos a forma “ativa” exige preposição: ... *antipatizam com D. Marlene*, ... *falam de D. Marlene*. Mas não se conhece nenhum caso de formas do paradigma verbal que exijam preposição, enquanto outras não exigem – por exemplo, um verbo que exija preposição com seu complemento, mas só em determinados tempos. Outro exemplo semelhante é *fadado*: esse particípio é usado em estruturas que podemos chamar “passivas”, como

[195] Ele é fadado a perder namoradas para os outros.

mas o verbo correspondente, *fadar*, embora ainda exista nos dicionários, já desapareceu da linguagem corrente, falada ou escrita.⁵²

Há indicações, portanto, de que a passiva, ainda que seja uma diátese, não é uma diátese **verbal**. Ficaria no grupo das diátesses nominais, e o particípio nominal que ocorre nela seria semelhante a um adjetivo em *-vel*: há radicais verbais que aceitam essa formação, outros não: *lavável*, *legível*, mas **sível*, **estável*, **permanecível*, assim como há radicais, em grande parte os mesmos, que não aceitam passiva: *ter*, *permanecer*, *estar*, *ser*...⁵³ As diátesses nominais ficam fora do nosso estudo.

A construção passiva é um verdadeiro cavalo de batalha da sintaxe moderna, e negar sua existência como construção autônoma acarreta certas responsabilidades. É verdade que a passiva sempre foi considerada até certo ponto uma anomalia. Como transformação (sua análise normal na teoria gerativa padrão de Chomsky, 1965), era singular porque envolvia não apenas a movimentação de constituintes, mas a mudança de funções sintáticas e o acréscimo de elementos em geral significativos como o verbo *ser* e a preposição *por*.

Semanticamente, sempre houve um certo desconforto porque o modelo previa que não haveria diferenças de significado proposicional entre as diversas versões superficiais de uma mesma estrutura profunda, mas a passiva era exceção. Em certos casos que envolvem quantificadores, havia diferenças proposicionais entre passivas e ativas, como em

[196] Todo mundo nesta sala fala três línguas.

[197] Três línguas são faladas por todo mundo nesta sala.

⁵² Também em inglês há alguns verbos tradicionalmente analisados como só podendo ocorrer na passiva (Huddleston, 1984: 440; Levin, 1993: 107). Não me parece que haja tais casos em português; em inglês, considerá-los verbos leva à esdrúxula situação da existência de verbos que não têm nenhuma forma “verbal”: não têm presente, nem passado, nem gerúndio, nem variam em pessoa etc. Na verdade, só têm o particípio (nominal) – é o caso de se perguntar por que são chamados verbos.

⁵³ Em Cunha (1981: 95), certamente por lapso, um dos exemplos de agente da passiva (e, portanto, de oração passiva) é justamente uma construção com adjetivo em *-vel*: *Todo abismo é navegável a barquinhos de papel* (frase de Guimarães Rosa).

Embora as duas frases sejam (segundo o modelo padrão) derivadas da mesma estrutura profunda, e portanto deveriam ter o mesmo significado proposicional, é bastante claro que as duas não têm acarretamentos idênticos, e [196] pode ser verdadeira em situações em que [197] não é. Esse caso é excepcional entre as transformações definidas na teoria padrão, e foi uma das motivações para a modificação do modelo.

Mas como estamos apenas estudando as diáteses verbais, basta mostrar que estas não estão ligadas aos participios nominais – ou seja, que, o participio nominal não é uma forma verbal ao lado do presente do indicativo ou do progressivo com *estar* + gerúndio. Isso fica bastante claro pelo que se mostrou acima. A conclusão é que parece mais adequado não contar a construção passiva como uma diátese verbal (contrariando a prática de Levin, 1993 e Goldberg, 1995).

Igualmente, a argumentação acima nos leva a excluir a construção chamada “de participio absoluto” do elenco das diáteses verbais. Essa construção se exemplifica em

[198] Terminadas as apresentações, Marquinhos subiu em cima da mesa.

Note-se que é o participio nominal que aparece aqui:

[199] Suspensa / ?? suspendida a reunião, voltamos para o hotel.

Se o participio nominal for realmente estranho ao lexema verbal, teremos que excluir o participio absoluto do elenco das diáteses verbais.

8.13. Concluindo

Para efeitos da descrição das valências verbais, a decisão do que incluir ou não na formulação das diáteses depende de um eixo de previsibilidade que comporta diversos graus, até o momento ainda mal conhecidos. Diante da urgência de realizar o levantamento das valências dos verbos, o mais que se pode fazer no momento é criar procedimentos que nos forneçam indicações quanto ao grau de previsibilidade da ocorrência e papel temático dos diversos constituintes que acompanham o verbo nas orações. A partir daí, obtemos um conjunto de critérios suficiente para nos orientar na formulação das diáteses verbais.

Evidentemente, nem tudo fica resolvido, e em particular persiste o problema de determinar, para cada associação sintagma / papel temático, qual é o caso não marcado. Em um primeiro momento, acredito que será indispensável recorrer a uma avaliação um tanto impressionística a respeito da marcação desta ou daquela associação simbólica (forma + papel temático). Mas os procedimentos estão aí, e podem ser aplicados sempre que houver dúvida – eles funcionam como procedimentos de laboratório, que não decidem questões, mas fornecem indicações de qual é a solução mais plausível para cada caso particular.

8.14. Notas para a pesquisa futura

Algumas observações podem ser úteis para um futuro aperfeiçoamento do sistema aqui adotado. Acabamos de ver que, semanticamente, a ausência de certos complementos (em especial o objeto direto) pode ter efeitos diferenciados. Assim, em

[200] Sônia já comeu.

a representação semântica inclui, obrigatoriamente, um Paciente esquemático (“indeterminado”), ao passo que em

[201] A menina sorriu.

isso não ocorre.

Observa-se que a situação de [200] (Paciente esquemático) ocorre de preferência com verbos que têm objeto especificado com frequência, ao passo que a situação de [201] é típica de verbos que nunca, ou só em casos especiais, ocorrem com objeto – *sorrir*, por exemplo, só ocorre, e raramente, com objeto cognato (*sorriu um sorriso amarelo*). Note-se que o objeto cognato não acrescenta um elemento referencial, servindo apenas como suporte para um qualificativo; assim pode-se dizer que, semanticamente falando, *sorrir* nunca tem Paciente.⁵⁴ Como isso, afinal, acaba subclassificando os verbos, optei por distinguir os dois casos: considero [200] e [201] como realizações de construções distintas. Elas se diferenciam pela presença, em [200], de um Paciente não representado formalmente e indeterminado, ao passo que [201] não inclui esse elemento semântico. Mas é provável que isso se relacione com traços semânticos do próprio verbo, o que pode nos levar, eventualmente, a identificar os dois casos (Paciente esquemático e Paciente inexistente).

Em certos casos a omissão do objeto direto remete a um esquema bastante elaborado. Assim, embora o Paciente (subentendido) de [200] seja algo comível, e só isso, em

[202] O marido dela bebe.

entende-se que se trata de bebida alcoólica. Isso depende de um traço do verbo, e não me parece que seja um traço semântico: é idiossincrático, e portanto deve em princípio ser incluído na valência. No entanto, o fenômeno é tão mal conhecido (nem sequer se sabe se se limita a SNs) que não foi possível incluí-lo nos critérios de diferenciação das diáteses.

Em outros casos, as condições de opcionalidade diferem mesmo com verbos semanticamente semelhantes; um caso bem claro é o de *falar* e *dizer*, que diferem entre si no que diz respeito à expressão de seu objeto direto. Assim, temos:

[203] Meu primo falou uma bobagem. / Meu primo falou.

[204] Meu primo disse uma bobagem. / * Meu primo disse.

⁵⁴ O mesmo vale para *dormir* (mas não *cochilar*) e *morrer* (mas não *falecer*).

Esses exemplos sugerem que pelo menos em alguns casos pode ser inevitável marcar a possibilidade de omissão do objeto direto individualmente no item léxico do verbo. Uma decisão a respeito vai depender de mapeamento extenso do léxico.

A opcionalidade de alguns complementos constitui um ponto potencialmente delicado, porque as coisas parecem ser mais flexíveis do que o esperado. Assim, dizemos que o objeto direto de *fazer*, ao contrário do de *comer*, é obrigatório (mais exatamente, *fazer* ocorre em construções com objeto direto e não nas sem objeto direto). Mas há circunstâncias em que o objeto direto de *fazer* pode ser omitido. A principal, claro, é o contexto anafórico, que fica excluído do estudo das diáteses (esse parece ser um princípio universalmente aceito, embora nem sempre explicitado; ver 3.4.5). Por isso, a possibilidade de ocorrência sem objeto direto em

[205] Ela prometeu fazer o rosbife mas não fez.

não nos interessa no momento de estabelecer e definir as diáteses de *fazer*.

Mas existem outras situações, que estão sendo igualmente excluídas, e que ainda não foram delimitadas de maneira tão precisa quanto a situação anafórica. Pode-se usar frases como

[206] Eu só apóio o Waldir porque ele prometeu e fez.

[207] É mais fácil prometer do que fazer.

Aqui não temos a situação anafórica clássica de [205], onde *fazer* ocorre duas vezes, e o objeto direto da primeira ocorrência fornece a pista para recuperar o paciente da segunda. Por outro lado, pode-se vislumbrar uma espécie de preparação no contexto precedente, que facilita a ocorrência de *fazer* sem objeto, porque em outras circunstâncias isso não é possível:

[208] * Esse político não fez.

Outro fator que parece ter alguma relevância é o aspecto verbal. Embora [96] seja claramente inaceitável, me parece que [97] é bem melhor:

[209] ? Esse é um político que faz.

Há indicações de que a possibilidade de omitir um complemento se relaciona com o grau de especificidade (ou de elaboração) da semântica do verbo. Assim, o verbo *fazer*, que como vimos dificilmente ocorre sem objeto direto, tem um significado muito esquemático, exprimindo uma gama muito ampla de relações entre o sujeito e o objeto: basta comparar o que quer dizer *fazer* em *ele fez um bolo*, *ele fez uma bobagem*, *ele fez sucesso na festa*, *ele fez uma careta*, *ele fez falta* etc., onde o conteúdo do verbo varia amplamente em função do complemento. Já *comer* exprime uma ação muito menos esquemática, de modo que se pode facilmente descrever o que é que esse verbo significa em geral: algo como “colocar algo na boca e ingeri-lo para alimentar-se”. E,

correspondentemente, *comer* ocorre sem objeto com muita freqüência, já que permite que o receptor preencha o lugar do Paciente com um conceito razoavelmente elaborado.⁵⁵

Outro exemplo pode ser o dos verbos *durar* e *demorar*. Podemos dizer

[210] A festa durou quatro horas.

[211] A festa demorou quatro horas.

Mas a omissão do SN pós-verbal só dá bons resultados com *demorar*:

[212] ?? A festa durou.

[213] A festa demorou.

Atribuo isso à semântica mais rica de *demorar*, que inclui uma referência a um período longo de tempo, ao passo que *durar* é muito mais neutro a respeito.

Poderíamos ver aí um caminho para determinar os fatores da opcionalidade. Mas a linguagem não costuma colaborar com o lingüista, de modo que há outros casos em que exatamente o verbo de semântica mais elaborada é o que exige o complemento. Assim, é fácil usar *comer* sem objeto, mas o mesmo não ocorre com *devorar*, embora *devorar* inclua a informação de que algo foi comido com violência, *comer* sendo neutro a respeito.

E, finalmente, há casos em que não se consegue vislumbrar diferença de grau de elaboração, e ainda assim os verbos diferem quanto à possibilidade de omitir o objeto: compare-se *vencer* (com ou sem objeto) e *derrotar* (só com objeto).

Dada a complexidade do fenômeno, e nosso pouco conhecimento dele, todos esses fatores vão ser desconsiderados no presente estudo; no atual estado do conhecimento, simplesmente não há alternativa. Direi portanto que *fazer* é um verbo que só ocorre em construções com objeto direto, nunca sem, no que se distingue de *comer*. Essa é, como já apontei, a solução mais prudente, porque é muito mais fácil reunir classes separadas do que separar uma classe em duas. Mas as observações acima indicam um tema muito interessante de pesquisa. Seria importante levantar muitos casos e avaliar cada um para ver se há possibilidade de explicá-lo em termos semânticos. Os fatos, naturalmente, podem ser devidamente pesquisados a partir das listas elaboradas utilizando construções separadas; estas vão permitir apurar se os verbos que podem ocorrer com ou sem complemento possuem alguma característica semântica própria, que os distinga dos verbos que só ocorrem com complemento e dos que só ocorrem sem complemento.

Para resumir o problema, então, podemos dizer que os tipos de complementos que ocorrem com cada verbo (SN; preposição + SN etc.) podem ser estabelecidos com bastante segurança. Já quanto à obrigatoriedade ou não desses complementos, temos que adotar uma solução de compromisso, à espera de coisa melhor.

⁵⁵ Alguns verbos de significado esquemático, entretanto, admitem a ocorrência sem complemento, mas nesses casos não se subentende um argumento elíptico, e o verbo assume um significado idiossincrático. É o caso de *dar*, que ocorre sem complemento em *ela dava para todo mundo*.

Capítulo 9: Classes de verbos: problemas e planos

9.1. Classificando verbos

Estabelecemos no capítulo anterior algumas diretrizes para a definição das construções em que ocorrem os verbos do português. Como vimos, os verbos se classificam de acordo com as construções em que podem ocorrer, isto é, de acordo com suas valências. Essa classificação, como qualquer outra, é apenas uma das possibilidades que existem para a classificação dos verbos, como vimos em algum detalhe no capítulo 3.

Como as valências são indubitavelmente um aspecto importante da descrição gramatical, a classificação dos verbos segundo as valências é igualmente importante. Ao investigar como se classificam os verbos segundo as valências, estamos pesquisando um importante aspecto da interface entre a gramática e o léxico: um componente essencial do nosso conhecimento da língua. Vou iniciar este capítulo examinando alguns aspectos dessa classificação, assim como alguns problemas que ela suscita, sendo que nem todos estão satisfatoriamente resolvidos atualmente.

É claro que não se pode dizer muita coisa de definitivo a respeito da classificação dos verbos, pois esta depende de levantamentos ainda não realizados. Mas já é possível estabelecer trilhas mais ou menos promissoras para a pesquisa futura, mesmo com base em nosso conhecimento fragmentário dos agrupamentos verbais da língua. Vou começar por examinar um desses agrupamentos, que apresenta alguns aspectos interessantes que poderão (ou não) se confirmar posteriormente.

Ao se falar de “classes”, é bom lembrar que não se trata, a não ser ocasionalmente, de itens de comportamento gramaticalmente idêntico. A diversidade gramatical é muito grande, bem maior do que dão a entender os estudos tradicionais. Termos como “substantivo”, “preposição” e “verbo transitivo” não passam de abreviaturas para feixes de traços distintivos; e a única maneira exaustiva de tratar da classificação de um item é considerar todos os seus traços, pelo menos os relevantes de determinado ponto de vista. No entanto, as abreviaturas são úteis porque representam uma “calibragem” dos traços em vista de um objetivo descritivo particular. Isso foi exposto na seção 3.5, e será revisto adiante em conexão com as classes verbais estudadas neste capítulo.

9.2. Verbos “transitivos”, “ergativos” e “transitivo-ergativos”

Sabemos que “transitiva” e “ergativa” são duas das construções que figuram na lista de diáteses verbais. A transitiva é exemplificada por

[1] O frentista encheu o tanque.
 Agente Paciente

e a ergativa ocorre em

[2] O tanque encheu.
Paciente

Agora vamos considerar os verbos da língua, limitando-nos a sua propriedade de ocorrer em uma ou outra dessas construções. Desse ponto de vista (muito limitado), podemos distinguir quatro classes de verbos: primeiro, temos os que ocorrem na transitiva, mas não na ergativa -- um exemplo é *comer*. Em segundo lugar, temos verbos que ocorrem na ergativa, mas não na transitiva, como *desmaiar*. Depois, temos verbos que ocorrem nas duas, como *encher*. Restam os verbos que não ocorrem em nenhuma das duas construções, como *ser* ou *ter*; estes vamos deixar de lado por ora, para nos concentrarmos no exame das três primeiras classes. Para facilitar a referência, chamarei os verbos como *comer* **verbos transitivos**; os como *desmaiar*, **verbos ergativos**; e os como *encher*, **verbos transitivo-ergativos**.¹

A identificação dos transitivo-ergativos como grupo não é novidade; algo muito semelhante foi estudado para o inglês por Fillmore (1970), e é um dos ingredientes da tipologia proposta por Dixon e Aikhenvald (2000). Levin (1993) também os inclui em sua classificação dos verbos do inglês, mas dividindo-os em diversos subgrupos. O que ainda não se fez é, primeiro, um estudo detalhado, necessário porque os transitivo-ergativos não são uma classe homogênea; e, depois, o mapeamento da ocorrência de tais verbos e sua extensão dentro do léxico.

Eu disse acima que classes assim definidas são heterogêneas. Vamos ver como isso se aplica às nossas três classes. Entre os verbos transitivos temos *comer*, e também *enviar*, que igualmente ocorre na construção transitiva, mas não na ergativa:

[3] O Sérgio enviou mais de cem cartas.

[4] * As cartas enviaram.

Mas isso não quer dizer que esses dois verbos sejam idênticos. Por exemplo, *enviar* ocorre também com um complemento de Meta (tradicionalmente, “objeto indireto”), que não pode ocorrer com *comer*:

[5] O Sérgio enviou um pacote ao seu tio.

[6] * O Sérgio comeu um sanduíche ao seu tio.

¹ O termo “transitivo” é usado aqui em sentido diferente do que normalmente se acha nas gramáticas tradicionais. Neste livro, um “verbo transitivo” é um verbo que pode ocorrer na construção transitiva, e não pode ocorrer na ergativa. Assim, *comer* é transitivo, mas *encher* não (é transitivo-ergativo). Os termos “transitivo”, “ergativo” e “transitivo-ergativo” se referem a **classes**, nunca a **funções** (esclareço a distinção entre classe e função na seção 3.4.4). Talvez fosse melhor usar outro termo. Mas eu fico sempre na dúvida entre usar um novo termo, complicando a nomenclatura a ser adquirida pelo leitor, e usar um termo tradicional em sentido parcialmente diferente. No caso, tenho poucos escrúpulos em usar o termo “transitivo” porque este é utilizado de maneira tão inconsistente na gramática tradicional que na verdade não corresponde a nenhum conceito coerente.

Ou seja, *comer* e *enviar* pertencem a uma mesma classe se considerarmos apenas os dois traços “ocorrência na construção transitiva” e “ocorrência na construção ergativa”. Veremos que mesmo assim essa classe tem interesse gramatical; mas é bom ter sempre em mente como sua delimitação é parcial.

Passando aos verbos ergativos, temos *desmaiar* e também *desaparecer*; ambos ocorrem na ergativa, mas não na transitiva:

[7] A menina desmaiou / a menina desapareceu.

[8] * O sujeito desmaiou a menina / * o sujeito desapareceu a menina.

Mas ainda aqui há diferenças: *desaparecer* pode ocorrer em uma construção com Agente e Paciente, só que nesse caso exige uma preposição:

[9] O sujeito desapareceu com a menina.

Já *desmaiar* não ocorre nessa construção:

[10] * O sujeito desmaiou com a menina.²

Ou seja, aqui também a classe dos verbos ergativos mostra sua heterogeneidade.

Finalmente, temos os verbos transitivo-ergativos, como *encher*. Essa classe é muito grande no português do Brasil, e tende a aumentar. Ela é também bastante heterogênea: por exemplo, *encher* e *mudar* são verbos semelhantes no sentido de que ocorrem nas construções ergativa e transitiva:

[1] O frentista encheu o tanque.

[2] O tanque encheu.

[11] O governo mudou as regras.

[12] As regras mudaram.

No entanto, *mudar* pode ocorrer em uma construção em que o Paciente vem regido de preposição, ao passo que *encher* não pode:

[13] Elas mudaram de roupa.

[14] * Elas encheram do tanque / de tanque.

Essas observações nos alertam para o que se pode esperar de classes como as três que estamos considerando: têm interesse descritivo, mas não são homogêneas de todos os

² Claro que essa frase pode ocorrer, mas com sujeito Paciente, significando “o sujeito desmaiou juntamente com a menina”. Mas não é a mesma construção exemplificada em [9], que tem sujeito Agente.

pontos de vista. Essa situação é típica em gramática, de maneira que classes definidas em termos limitados não podem ser evitadas. As classes só fazem sentido, em geral, quando vinculadas a um objetivo descritivo; e é com classes assim limitadas que temos que lidar quase sempre. Para voltar a um ponto já visto, que interesse tem para nós, estudando sintaxe e semântica, a conjugação a que pertence um verbo? No entanto, isso é fundamental para o estudo da morfologia.

Vou agora expor algumas afirmações gramaticais que podem ser feitas a respeito das três classes que acabamos de estabelecer. Essas afirmações mostram como classes limitadas são, ainda assim, gramaticalmente interessantes.

9.3. Casos de complementaridade

Em certos casos as duas construções em estudo requerem verbos especializados; assim, *curar* e *matar* só aparecem na construção transitiva; na construção ergativa, são substituídos, respectivamente, por *sarar* e *morrer*:

[15] O doutor Sérgio curou minha filha. / ?? ... sarou minha filha.

[16] Minha filha sarou. / * Minha filha curou. [onde *minha filha* é Paciente]

[17] O vizinho matou o gato. / * O vizinho morreu o gato.

[18] O gato morreu. / * O gato matou. [onde *o gato* é Paciente]

As relações simbólicas entre as frases [15]-[16] e [17]-[18] acima são exatamente as que se verificam entre as construções transitiva e ergativa: na primeira o objeto direto é Paciente e o sujeito é Agente; na segunda não há Agente expresso, e o sujeito é Paciente. *Matar* e *curar* podem aparecer sem objeto direto, mas nesse caso o sujeito é interpretado como Agente, não se tratando, portanto, de construção ergativa:

[19] As plantas curam.

[20] Catapora também mata.

Ver também o par *cair* / *derrubar*, e possivelmente outros. Tais casos mostram que alguns verbos aparecem em pares, ligados pelo significado. Um dos membros do par é transitivo, e o outro é ergativo; e o par preenche exatamente o lugar de um verbo transitivo-ergativo, verbo esse que não existe na língua.

9.4. Expansão da classe dos transitivo-ergativos na língua falada

A classe dos transitivo-ergativos tende a se expandir na língua falada, à custa das outras duas classes. Vou examinar brevemente essa tendência – deixando claro que aqui se trata apenas de chamar a atenção para um fenômeno diacrônico, fazendo sugestões e

perguntas, sem tentar um estudo em profundidade, que não caberia dentro dos objetivos do presente trabalho.

Nos exemplos abaixo, tomo a forma padrão como cronologicamente anterior à forma coloquial, o que fornece uma direção à presumível mudança. Isso é plausível, mas não é necessariamente verdadeiro: as formas coloquiais podem vir coexistindo com as formas do padrão há séculos. Isso é outro ponto a ser esclarecido através de pesquisa devidamente documentada.³

9.4.1. Ergativos em português padrão

Há muitos verbos transitivo-ergativos em português padrão, como por exemplo *ligar, acabar, começar, ferver, encher, explodir*, etc.:

[21] Papai ligou o carro.

[22] O carro não ligou.

[23] O professor começou a conferência às três em ponto.

[24] A conferência começou às três em ponto.

[25] Fervi um pouco de água para o chá.

[26] A água ferveu até secar.

[27] Minha loja vende aquecedores solares.

[28] Aquecedores solares vendem muito no inverno.

Com certos verbos, a construção ergativa exige um pronome reflexivo na língua padrão, mas esse pronome não é usado no coloquial, o que torna esses verbos idênticos, no coloquial, aos transitivo-ergativos. Por exemplo, no padrão temos

[29] O professor interessou a turma pela matéria.

[30] A turma se interessou pela matéria.

mas no coloquial esta última frase seria

[31] A turma interessou pela matéria.

³ O fenômeno já foi observado anteriormente; por exemplo, Whitaker-Franchi (1989: 27) afirma que “as construções ergativas constituem um processo sintático em expansão no Português do Brasil.” Eu diria, antes, que a expansão se verifica nas **classes** de verbos, não propriamente nas construções.

ou seja, sem o *-se*.⁴

O mesmo fenômeno se dá com *aposentar, casar, calar, confundir, cumprimentar, desiludir, divorciar, entusiasmar, impressionar, separar, suicidar* e outros verbos. O resultado é que na língua coloquial um verbo como *interessar* acaba ocorrendo nas duas construções em pauta, a transitiva e a ergativa, e portanto deve ser classificado como transitivo-ergativo.

O fenômeno do desaparecimento do reflexivo na construção ergativa não é universal; com alguns verbos o *se* nunca é usado: *engordar, cozinhar, explodir*; com outros, parece haver variação regional: *quebrar, divorciar, calar*; finalmente, com alguns verbos o reflexivo se mantém em todas as variantes: *dar-se, gabar-se, retirar-se, abster-se, tratar-se, dar-se conta, encontrar-se*.⁵

9.4.2. Ergativos se tornam transitivo-ergativos

Um dos aspectos da expansão da classe dos transitivo-ergativos é ilustrado por verbos utilizados na construção transitiva na língua coloquial, embora esse uso não seja encontrado no padrão, onde apenas o uso ergativo é atestado. Por exemplo, ouvem-se sentenças como as seguintes:

[32] Esse sapato dói o meu pé.

[33] Você precisa funcionar o motor durante uns minutos.

onde os verbos *doer* e *funcionar* aparecem na construção transitiva, o que faz deles transitivo-ergativos, pois também ocorrem (no coloquial como no padrão) em frases ergativas como

[34] Meu pé está doendo.

[35] O motor funcionou durante uns minutos.

Três exemplos recentes de uso transitivo de verbos que são ergativos no padrão são:

[36] Temporal desaba casa.

[37] A enxurrada transbordou a represa.

[38] As fortes chuvas transbordaram os córregos.⁶

⁴ Há bastante variação aqui; observei que muitos falantes do Sul e do Rio de Janeiro utilizam obrigatoriamente o *se* em frases como essa. Nos exemplos, considero o dialeto mineiro.

⁵ No entanto, suspeito que estes últimos verbos são em geral limitados à língua escrita.

⁶ [36] e [37] acima foram manchetes do jornal *Hoje em Dia*, de Belo Horizonte, em janeiro e fevereiro de 2002; [38] ocorreu no rádio em dezembro de 2005.

9.4.3. Transitivos se tornam transitivo-ergativos

Também se observa o processo oposto, isto é a transformação de transitivos em transitivo-ergativos. Temos exemplos com *varrer* e *lavar*, normalmente considerados transitivos (ocorrem com sujeito Agente, e objeto, quando há, Paciente), em frases como a seguinte:

[39] Parece que esse carro nunca lavou.

[40] Esse quarto nunca varre.

Aqui temos um sujeito Paciente e nenhum objeto direto, o que faria dessas frases casos normais de construção ergativa.

É preciso distinguir esses casos dos de transitivos que ocorrem sem objeto, mas com sentido recíproco ou reflexivo (o sujeito é ao mesmo tempo Paciente e Agente). Isso os coloca fora da classe dos transitivo-ergativos tais como definidos acima. Um caso de verbo usado sem objeto com sentido recíproco é

[41] Carlos e Selma nunca beijam em público.

Com sentido reflexivo encontramos

[42] Ela vestiu de branco para a festa de Iemanjá.

[43] O menino enxugou com a minha toalha.

Essas frases não podem ser consideradas exemplos de construções ergativas por causa do papel temático do sujeito, que é Agente (e igualmente Paciente) em todos os casos.

9.5. Efeitos sintáticos da ergatividade

Mencionam-se na literatura diversos efeitos sintáticos que se consideram associados à ergatividade. Embora vários desses efeitos sejam aparentemente nítidos em outras línguas românicas, veremos que em português a evidência é menos clara.

Segundo Salvi e Vanelli (1992), os ergativos apresentam muita facilidade de ocorrência com sujeito posposto, ao contrário de verbos em construção transitiva. Assim, em italiano a posição normal do sujeito de um ergativo é depois do verbo, principalmente quando há um elemento adverbial no início da sentença. Em português, encontramos sujeitos pospostos nas mesmas condições:

[44] Ontem caiu uma aranha na cabeça do Gabriel.

[45] De repente bateu a porta.

Mas há muitos casos em que o sujeito pode ocorrer posposto mesmo quando não se trata de construção ergativa:

[46] De repente pulou um cabrito no colo da menina.

[47] No show de ontem cantaram três tenores famosos.

O que realmente bloqueia a posposição do sujeito é a presença de um objeto direto:

[48] * De repente pulou um cabrito o muro do jardim.

[49] * No show de ontem cantaram três tenores famosos o Hino Nacional.

Como nem *cantar* nem *pular* são ergativos, vê-se que essa propriedade não é diretamente ligada à ergatividade do verbo ou da construção.

Além disso, mesmo com verbos ergativos a posição posposta é inaceitável em muitos casos:

[50] A turma entusiasmou.

[51] * Entusiasmou a turma.

As condições exatas de ocorrência do sujeito posposto não são bem conhecidas. Existe uma literatura a respeito (um exemplo recente é Silva, 2001), mas não me parece que a questão tenha sido suficientemente esclarecida, e ainda demanda pesquisa. Mas de qualquer forma não creio que se possa dizer que a facilidade de posposição do sujeito seja uma característica da construção ergativa.

9.6. Correlatos semânticos

9.6.1. *Relação complexa entre valência e significado*

Levin (1993) afirma que

o comportamento de um verbo, em particular no que se refere à expressão e interpretação de seus argumentos, é em grande parte determinado por seu significado.

[Levin, 1993: 1]

Essa posição pode ser útil como hipótese inicial. Mas, evidentemente, ela requer comprovação a cada passo. É perfeitamente plausível que haja também interferência de fatores não-semânticos, isto é, muitos verbos podem ser marcados idiossincraticamente quanto a traços como a aceitação de certos complementos, ou quanto à ocorrência ou não em certas construções. Esses casos existem, e já vimos alguns exemplos (dei alguns em 9.5.4). Até que ponto tais casos são típicos ou anômalos é coisa que só a investigação detalhada poderá responder – e só a partir daí é que se poderá avaliar o valor heurístico de afirmações como a de Levin. Alguns fatos relativos às três classes que estamos estudando podem dar indicações sobre a correlação entre a forma e o significado de que fala a autora.

Vamos considerar o caráter transitivo ou transitivo-ergativo de certos pares de verbos relacionados pela ocorrência ou não do prefixo *des-*. Por exemplo, *descosturar* é transitivo-ergativo:

[52] A menina descosturou a manga do vestido.

[53] A manga do vestido descosturou.

Mas *costurar* se comporta como um transitivo:

[54] A menina costurou a manga do vestido.

[55] * A manga do vestido costurou.

Talvez o fator em jogo aqui seja que *costurar* geralmente se entende como uma ação, envolvendo um Agente humano, ao passo de *descosturar* é algo que pode ocorrer espontaneamente, por acidente. *Costurar* é diferente de *colar* porque uma coisa pode colar por acidente:

[56] As páginas colaram e eu não consegui abrir o livro.

Mas a ação de costurar é muito específica: envolve necessariamente o uso de linha, agulha, amarração das peças através de pontos etc. – algo que dificilmente ocorreria sem o intermédio de um agente. Em outras palavras, o evento de costurar envolve necessariamente “controle sobre o desencadeamento do processo”, no dizer de Cançado (2003: 100), ao passo o evento de descosturar não pressupõe tal controle. Em tais casos, a construção ergativa não é adequada, o que com certeza deriva do fato de que a ergativa é tipicamente usada para omitir o Agente (ver a respeito Moreira, 2005).

Outros pares que funcionam como *costurar* / *descosturar* são:

montar / *desmontar*
colorir / *descolorir*
pentear / *despentear*

O primeiro membro desses pares talvez possa, com a devida preparação, ocorrer em construções ergativas:

[57] O pelo da gata era muito rebelde, mas com esse creme ele penteou facilmente.

[58] O motor era muito difícil de montar; aí nós lubrificamos bastante, e (??) ele finalmente montou.

Esses verbos, basicamente transitivos, parecem ser susceptíveis ao contexto, a ponto de ocorrerem na construção ergativa se houver uma preparação suficiente. Outros verbos transitivos, como *entender*, *comer* ou *matar*, parecem ser totalmente irreduzíveis: não há manipulação de contexto que nos leve a entender *onça mata* ergativamente (a onça como Paciente).

Um caso parecido é o de *fazer*: esse verbo não pode ocorrer na construção ergativa.⁷ Mas *desfazer*, que tem traços semânticos semelhantes aos dos verbos com *des-* vistos acima, é transitivo-ergativo:

[59] Lili desfez a costura.

[60] A costura desfez.

Novamente se trata de verbos que exprimem algo que pode, com grande probabilidade, ocorrer espontaneamente – muito embora a construção ergativa não exclua necessariamente um Agente subentendido.

Outro fato que ilustra a complexidade da correlação entre semântica e valência é apontado por Levy (1983), e tem a ver com certos verbos pertencentes ao mesmo campo semântico, mas que aceitam complementos oracionais diferentes. Assim, Levy observa que *escribir* ‘escrever’ aceita sem problemas completivas oracionais no indicativo:

[61] F. escribió que los había visto. ‘F. escreveu que os tinha visto’
[Levy, 1983: 31]

Mas outros verbos que exprimem maneiras específicas de escrever não aceitam complemento oracional; pegando exemplos do português, temos:

[62] * Fulano caligrafou que vinha.

[63] * Fulano datilografou que vinha.

[64] * Fulano rabiscou que vinha.

Note-se que em outras circunstâncias esses verbos têm diáteses idênticas: *ele escreveu / caligrafou / datilografou / rabiscou / digitou uma carta* etc. É possível que essas restrições tenham algo a ver com a complexidade (isto é, maior restrição extensional) de verbos como *caligrafar* etc., apontada por Farber e Usón, que sustentam que

⁷ No entanto, já ouvi a frase *espera um pouco que o café tá fazendo*. Para o português brasileiro, nada é sagrado.

[...] verbos semanticamente mais complexos são, via de regra, sintaticamente mais restritos do que os verbos mais básicos com os quais se relacionam em uma hierarquia semântica. [...] Em geral, os verbos mais específicos perdem ambientes sintáticos (privilégios de ocorrência).

[Farber e Usón, 1999; *apud* Wald, 2001: 859]

Ou seja, *rabiscar* é semanticamente mais complexo (pois exprime uma maneira específica de escrever), e portanto tem privilégios de ocorrência mais restritos (não pode ocorrer com oração completiva).

Casos como esses nos levam a imaginar que a afirmação de Levin pode ser correta em linhas gerais, mas com muitas restrições. A extensão do fenômeno, assim como os fatores que o influenciam, é algo a verificar.

9.6.2. Definindo semanticamente os transitivo-ergativos

O verbo *perder* parece, à primeira vista, um transitivo-ergativo, por causa de frases como

[65] Felipe perdeu o queijo.

[66] O queijo perdeu.

No entanto, temos aqui duas acepções radicalmente diferentes de *perder*: em [65] trata-se de ‘não saber onde colocou’, mas em [66] é ‘apodrecer’.⁸ Veremos que há razões que nos levam a distinguir esse caso dos casos de transitivo-ergativos estudados até o momento.

Vimos acima que um verbo é transitivo-ergativo quando cabe nas duas construções, a transitiva e a ergativa – como é certamente o caso de *perder*. Por outro lado, há certamente uma diferença importante entre *perder* e os verbos transitivo-ergativos. Com os transitivo-ergativos que estudamos, a construção transitiva inclui, como um de seus ingredientes, o significado da construção ergativa correspondente; ou seja, o significado de

[1] O frentista encheu o tanque.

seria analisável mais ou menos como “o frentista causou o evento de {o tanque encher}”. É claro que essa análise não se aplica ao par [65]-[66], porque [65] não significa “Felipe causou o evento de {o queijo perder (= ‘apodrecer’)}”. Podemos exprimir essa relação semântica da seguinte maneira:

[67] Relação semântica entre as construções transitiva e ergativa:

Construção transitiva: X V SN₂ = {CAUSAR (X) [P (SN₂)] }

⁸ Essa acepção de *perder* ‘apodrecer’ é normal para muitos falantes; outros a aceitam, embora não a usem ativamente.

Construção ergativa: $X V = [P(X)]$

Um verbo será transitivo-ergativo somente se essa relação semântica valer para sua ocorrência na construção transitiva e na ergativa. Se incluirmos essa condição em nossa definição, *perder* não será um transitivo-ergativo, muito embora ocorra em ambas as construções.

9.6.3. Relações entre diáteses

O caso de verbos como *perder* levanta um problema bastante difícil para o sistema aqui desenvolvido de notação das valências. A se seguir estritamente esse sistema, a valência de *perder* vai ser a mesma que a de *encher*, ou seja, ambos são transitivo-ergativos. Mas isso oculta o comportamento semântico muito diferente dos dois verbos, pois *encher* tem o mesmo significado nas duas construções (transitiva e ergativa), e o significado destas é relacionado de maneira sistemática como está em [67], ao passo que *perder* é idiossincrático desse ponto de vista. É evidente que esses fatos vão ter que figurar em algum lugar na descrição da língua; a pergunta é se devem ser incluídos na descrição das valências.

Note-se que a idiossincrasia de *perder* é decorrente não de alguma de suas diáteses, mas da **relação** existente entre duas delas. Ou seja, aqui entra um novo fator, que não consideramos até o momento, a relação entre construções. A se incluir propriedades como essa do verbo *perder* no sistema de valências, teremos que levar a cabo um levantamento geral da lista de diáteses, definindo a relação semântica entre elas sempre que partilharem algum verbo, e procurando eventuais idiossincrasias (já sabemos que *pintar* é outro caso, que afeta as construções transitiva e intransitiva). Ora, no momento essa tarefa é impossível, em vista do estágio inicial em que se encontra o estudo das valências verbais. Por isso, vou ignorar o comportamento de *perder* e *pintar* para efeitos da formulação das valências, embora reconhecendo sua importância e reservando-o para pesquisas futuras. Será também necessário apurar a extensão desse tipo de fenômeno, para ver quantos verbos ele afeta, e se o faz de mais de uma maneira. Na próxima seção faço algumas especulações preliminares sobre o assunto.

A relação entre diáteses envolve mais do que isso: existem também relações de acarretamento de existência, da forma “se um verbo ocorre na construção A, então necessariamente (ou: muito provavelmente) ocorre também na construção B”. Por exemplo: se um verbo ocorre na construção ergativa, com certa probabilidade ocorre também na transitiva. Esse tipo de acarretamento tem relevância gramatical, pois tudo indica que faz parte do nosso conhecimento da língua (ver as “regras de correlação” na seção 12.2). Uma criança, ao aprender a falar, faz hipóteses baseadas neles, e essas hipóteses, quando incorretas, precisam ser refutadas pelos dados a que ela tem acesso. No contexto desta proposta, porém, as relações entre diáteses, no que pese sua importância, terão que ficar na sombra. Temos que limitar o campo da pesquisa para estabelecer linhas diretivas: isso nos é ditado pelas limitações dos recursos disponíveis e do nosso equipamento cognitivo.⁹

⁹ Um meteorologista que estuda os furacões de Bangla Desh deveria, em princípio, levar em conta o bater das asas das borboletas no Chile. Mas como fazer isso, no atual estágio do conhecimento?

9.6.4. Polissemia induzida

Acabamos de ver que a ocorrência de um verbo em diferentes diáteses pode acarretar diferença de significado – relembro que estamos tomando como base o verbo enquanto forma fonológica e morfológica, ou seja, *pintar* é o mesmo verbo em *pintei o quarto* e em *o rapaz pintou lá em casa*.

Outro exemplo é o de *colocar*; esse verbo, em uma construção como

[68] Alda colocou o peixe na geladeira.

denota um deslocamento provocado por um Agente (*Alda*) sobre um Tema (*o peixe*) até uma Meta (*na geladeira*). Mas quando *colocar* aparece na construção transitiva, como em

[69] Alda colocou todos os pastéis.

já não se trata de deslocamento de um lugar para outro, mas de venda dos pastéis. Como se vê, temos aqui basicamente o mesmo fenômeno observado anteriormente para *perder*.

No entanto, essa alteração semântica observada no verbo (e conseqüentemente no papel temático dos complementos) não depende necessariamente de uma mudança de diátese. Em muitos casos a alteração semântica se processa independentemente de qualquer mudança de diátese, como efeito da interação do verbo com o significado de seus complementos. Esse é um fenômeno que chamarei *polissemia induzida*, e que se observa em casos como

[70] Alda colocou todos os pastéis. (‘vendeu’)

[71] Alda colocou uma camisa branca. (‘vestiu’)

[72] Alda pôs uma camisa branca. (‘vestiu’)

[73] ?? Alda pôs todos os pastéis. (*certamente não* ‘vendeu’)

Ou seja, quando o objeto é *todos os pastéis*, o verbo *colocar* assume o significado de ‘vender’, mas o mesmo não acontece com seu quase-sinônimo *pôr*. Essa diferença tem algo a ver com o significado do objeto direto – no caso, o fato de que entendemos *pastéis* como alguma coisa a ser vendida. Mas, naturalmente, tem também algo a ver com as propriedades do verbo *colocar*, em oposição a *pôr*.

O fenômeno da polissemia induzida é muito comum; limitando-nos ao verbo *colocar*, temos ainda:

[74] Alda colocou algumas questões muito pertinentes. [‘formulou’]

[75] Alda colocou uma pastelaria. [‘abriu (negócio)’]

[76] Alda colocou todos os sobrinhos. [‘deu emprego’]

A vinculação da polissemia induzida com a valência é que a acepção selecionada às vezes (mas só às vezes) depende da diátese em que o verbo ocorre. Assim, o significado de ‘abrir (negócio)’ é próprio da diátese transitiva (sujeito-verbo-objeto), ao passo que o de ‘posicionar’ ocorre quando a diátese é a **de Objeto Tema e SPrep Meta**, como por exemplo em

[77] Alda colocou o peixe na geladeira.

Agora, pode-se acrescentar um sintagma de significado locativo a [75], e o aspecto formal da frase vai ficar idêntico ao de [77]:

[78] Alda colocou uma pastelaria no Mercado Central.

No entanto, me parece que o sintagma locativo é complemento em [75] e adjunto em [78], pelo menos a julgar pelos critérios disponíveis. Assim, a anteposição do sintagma locativo se faz livremente em [78], sem que surja necessariamente o valor contrastivo:

[79] No Mercado Central, Alda colocou uma pastelaria. [não contrastivo]

mas o efeito da mesma operação em [75] causa o aparecimento necessário da acepção contrastiva:

[80] Na geladeira, Alda colocou o peixe.

Se isso for verdade, então poderemos dizer que há vinculação entre a polissemia induzida e a valência de *colocar*: esse verbo pode assumir o significado de ‘vender’ só quando ocorre na construção transitiva.

No entanto, a vinculação com a valência só se verifica em determinados casos. Em outros, o verbo na mesma diátese varia de significado conforme a relação seletional que o une ao sujeito:

[81] Alda pôs o ovo na geladeira.

[82] A galinha pôs um ovo debaixo do sofá.

Estamos lidando aqui com um fenômeno mais próximo da restrição seletional do que da valência, e a eventual conexão da polissemia induzida com a valência é acidental. Tudo se passa como se a seleção deste ou daquele significado para o verbo dependesse de fatores ligados ao conhecimento do mundo, associados à polissemia do item léxico. A única coisa que me faz vacilar nessa interpretação é a impossibilidade de entender *no Mercado Central* como complemento em [78] – se é que isso é um fato. De qualquer modo, proponho reduzir a polissemia induzida a um processo ligado às restrições

selecionais, e não às valências do verbo – o que coloca a polissemia induzida fora do campo de interesses do presente estudo.

Voltando ao caso de *colocar* e *pôr*: digamos que *colocar* tem, entre outras, a acepção de ‘vender’, e *pôr* não. Isso é freqüente – já se observou que a ocorrência de sinônimos perfeitos é excepcional nas línguas. Por alguma razão, a acepção ‘vender’ só ocorre quando o verbo está na construção transitiva, e é identificada pelo receptor com base em seu conhecimento do mundo: pastéis são itens vendáveis, peixe também, mas a geladeira não é um lugar onde se vendem coisas, ao passo que o Mercado Central é. Isso explica a interpretação dada a [77] e [78].

A interpretação dada a frases com *colocar* é probabilística, porque é possível entender a frase

[71] Alda colocou uma camisa branca. (‘vestiu’)

no sentido de que Alda vendeu uma camisa branca, e mesmo

[70] Alda colocou todos os pastéis.

no sentido de que ela vestiu os pastéis (fantasia de carnaval?). Isso, a meu ver, reforça a teoria de que estamos lidando aqui com um fenômeno puramente de interpretação de base pragmática; não é o caso das valências, cuja interpretação semântica é firmemente ancorada na sintaxe.

9.6.5. Concluindo

Voltando ao caso de *perder*, a pragmática não pode ser responsabilizada por sua variação semântica. Temos portanto mais de um tipo de variação, e um emaranhado de casos e de fatores determinantes: pragmáticos no caso de [70] - [71], semânticos (e codificados no item léxico) no caso de *perder* e de *pintar*. É uma área muito vasta de investigação, pelo que sei ainda totalmente virgem. Para efeitos deste estudo, terei que deixar essas interessantes questões de lado; para nós, portanto, a valência de *encher* e a de *perder* é a mesma.

Diremos que *perder* ocorre na construção transitiva e também na ergativa, mas não que é um transitivo-ergativo, já que um verbo é transitivo-ergativo quando ocorre nas duas diáteses de tal forma que a relação [67] se aplique a elas.¹⁰ Para efeitos de uma simples lista, entretanto, *pintar* e *perder* são idênticos a *encher* e aos demais transitivo-ergativos.

É óbvio que a diferença entre esses verbos precisa ser representada em algum lugar, pois é um componente do conhecimento da língua. Assim, há, como vimos, uma tendência a converter verbos transitivos ou ergativos em transitivo-ergativos, mas não se verifica nenhuma tendência a ampliar o pequeno grupo representado pelos verbos *pintar* e *perder*. Pode-se argumentar, com certa plausibilidade, que o falante ao aprender novos verbos está de certo modo predisposto a encontrar transitivo-ergativos relacionados

¹⁰ Não se trata de uma relação transformacional, mas de uma espécie de *via-rule* semântica que é, com toda probabilidade, parte do conhecimento lingüístico do usuário da língua.

através da condição [67], mas certamente não há nenhuma tendência sistemática a ampliar o elenco de verbos que se comportam como *perder* ou *pintar*. Até onde podemos ver, são casos excepcionais e idiossincráticos, ao passo que os transitivo-ergativos como *encher* representam uma grande parte dos verbos da língua: uma lista muito preliminar elaborada a partir de 120 verbos muito freqüentes na língua mostra que 99 deles são transitivo-ergativos. Ou seja, podemos admitir sem perigo que o número dos transitivo-ergativos em português é, pelo menos, da ordem das centenas, e provavelmente bem maior.

Isso tem importância descritiva e também importância teórica. A importância descritiva provém do fato de que as características dos verbos transitivo-ergativos terão que ser formuladas para centenas de itens, e portanto seria conveniente ter um rótulo para abreviá-las: em vez de listar as características uma a uma, diremos que qualquer verbo que as possua se chama “transitivo-ergativo”.

Teoricamente, esse rótulo tem importância porque a aprendizagem da língua se faz em grande parte por generalizações (formulação de hipóteses). Portanto, é plausível que uma criança ou adulto que esteja aprendendo o português acabe criando em sua memória uma categoria correspondente aos verbos transitivo-ergativos; e, ao escutar e entender uma frase como

[83] A mesa molhou toda.

parta para a hipótese de que também deve ser possível dizer

[84] O cachorro molhou a mesa.

Sabemos que essa hipótese nem sempre funciona: com *sair*, só ocorre a construção ergativa, e com *derrubar* só a transitiva. Mas a probabilidade de sucesso é suficientemente grande para que valha a pena formular uma regra de que um verbo que cabe em uma das construções também cabe na outra. Em outras palavras, pode-se dizer que a relação descrita em [67] é, em grande medida, **regular**. E se numa gramática compensa distinguir regularidades de irregularidades, então compensa manter o rótulo de “transitivo-ergativos” para os verbos que satisfazem a definição. Por outro lado, *perder* é uma irregularidade.¹¹

9.7. Construção pseudo-ergativa

Em certos casos, observa-se uma diátese semelhante à ergativa, mas que não deve ser identificada com ela por causa de algumas diferenças importantes. Assim, *fazer* ocorre normalmente na construção transitiva, como em [85], mas também aceita uma construção semelhante em certos pontos à ergativa em [86]:

[85] O Dr. Roberto faz três transplantes por semana. [transitiva]

¹¹ Na discussão desta seção, não considere uma solução baseada na separação de *perder* em dois itens léxicos independentes. Em outro trabalho (Perini, 1999) mostrei que isso leva a circularidades insustentáveis.

[86] Ele fez transplante, e ainda está em tratamento.

Não considero [86] um caso de construção ergativa porque o papel temático do sujeito nesses casos é Possuidor (ou Lugar), ao contrário da verdadeira ergativa, que tem sujeito Paciente:

[87] Ele fez transplante ~ Δ fez transplante **nele**. [Lugar]

[88] Léo pôs ponte de safena ~ Δ pôs ponte de safena **em Léo**. [Lugar]

[89] Ele cortou o cabelo ~ Δ cortou o cabelo **dele**. [Possuidor]

[90] Marina pintou o cabelo ~ Δ pintou o cabelo **de Marina**. [Possuidor]

Mas essa não é a única diferença. Nessas construções (que chamarei **pseudo-ergativas**), o sujeito também tem um papel temático próprio, que tem a ver com “controle”. Assim, podemos dizer

[91] Tito fez uma cirurgia ~ Δ fez uma cirurgia **em Tito**

mas

[92] * O barraco fez uma reforma *~ Δ fez uma reforma **no barraco**

Aparentemente, [92] é inaceitável porque o barraco não poderia ter controle sobre o evento da realização da reforma. Note-se que [91] é apropriada se Tito fez a cirurgia voluntariamente, mas não se ele chegou inconsciente ao hospital e a cirurgia se fez à sua revelia.

Parece que a pseudo-ergativa só ocorre em casos especiais, que envolvem posse inalienável, tipicamente de partes do corpo. As pseudo-ergativas ainda estão por ser devidamente estudadas.

9.8. A chamada “construção média”¹²

9.8.1. Colocando o problema

Vimos que um verbo é ergativo ou transitivo-ergativo quando pode ocorrer na chamada “construção ergativa”, exemplificada por

[93] A lata encheu rapidamente.

¹² Uma versão anterior desta seção foi publicada na revista *Scripta* (Perini, 2005).

Estou classificando como “ergativa” toda e qualquer construção que tenha as características listadas na seção 4.5. Segundo essa definição, temos outro exemplo de construção ergativa em

[94] Esse artigo só vende no verão.

No entanto, para alguns autores [94] seria exemplo de outra construção, denominada **média**. Por exemplo:

A alternância média não deve ser confundida com a alternância transitiva/[ergativa] [...] Primeiro, a construção média difere da [ergativa] [...] porque não denota um evento; ou seja, não precisa ter uma referência temporal específica. Depois, a construção média sempre pressupõe um Agente (*Crystal vases shatter easily*), ao passo que a [ergativa] não precisa (*The crystal vase shattered*).

[Levin, 1993:4]¹³

Eu gostaria de questionar essa distinção mencionada por Levin, que é aparentemente seguida por muitos autores atuais. O problema, evidentemente, é o de distinguir ou não a construção ergativa da média. Vou argumentar que essa distinção não é necessária. Mas antes de apresentar meus argumentos será preciso discutir os objetivos de fazer esse tipo de distinção; ou seja, vamos nos perguntar para que é preciso estabelecer distinções entre as diversas construções gramaticais.

Existe, em princípio, uma infinidade de distinções possíveis em gramática. No entanto, nem todas são consideradas importantes, e há uma razão por trás disso. A razão, acredito, é que toda classificação tem um objetivo. Assim, os verbos *cair* e *desaparecer* serão classificados juntos ou não, segundo os objetivos do momento. Se nosso objetivo for descrever a possibilidade de ocorrência em construções transitivas ou ergativas, *cair* e *desaparecer* estarão juntos, pois ambos são verbos ergativos – só podem ocorrer na construção ergativa. Mas se o objetivo for a descrição da flexão verbal, então esses verbos terão que ser separados, pois pertencem a conjugações diferentes. Em outras palavras, nenhuma classificação é válida ou inválida em si; será válida ou inválida em função de um objetivo descritivo particular (ver o capítulo 3 para uma discussão do problema da categorização).

O que temos que perguntar agora é qual o objetivo de se distinguir, digamos, a construção ergativa da construção transitiva ou da média. Vamos considerar o objetivo de descrever as diáteses verbais. Nesse caso, não se trata pura e simplesmente de exprimir diferenças sintáticas, nem semânticas, mas de procurar as diferenças que se correlacionam com a seleção de verbos específicos em significados específicos.

¹³ Para facilitar a leitura, substituo o termo **incoativo**, que é o usado por Levin, por **ergativo**. O texto original é o seguinte: “The middle alternation should not be confused with the transitive/inchoative alternation [...] First, the middle construction differs from the inchoative construction [...] in not denoting an event; that is, it need not have a specific time reference. Second, the middle construction always implies an agent (*Crystal vases shatter easily*), while the inchoative construction need not (*The crystal vase shattered*).” [Levin, 1993: 4, nota 2]

É por isso que não falamos de uma diátese topicalizada, baseando-nos na diferença entre

[95] A menina comprou uma bicicleta.

[96] Uma bicicleta, a menina comprou.

Acontece que não há verbos que exijam ou recusem objeto topicalizado, e portanto a relação que chamamos “topicalização” é irrelevante para a definição de diáteses. Pela mesma razão não distinguimos diáteses baseadas no tempo verbal, como em

[97] A menina comprou uma bicicleta.

[98] A menina vai comprar uma bicicleta.

e assim por diante. Já a ergativa pode ser considerada uma diátese, no sentido que nos interessa, porque há verbos que não podem ocorrer nela; e, naturalmente, também precisamos distinguir a construção transitiva da ergativa, em virtude dos muitos casos examinados neste livro.

Portanto, se quisermos discutir a conveniência ou não de se distinguir a construção “média” de

[94] Esse artigo só vende no verão.

da construção “ergativa” de

[99] O leite ferveu.

teremos que investigar se essa distinção é relevante para a definição de diáteses verbais. É minha opinião que não é relevante, e que portanto a distinção entre construção média e construção ergativa é supérflua.

Não há, que eu saiba, um verbo que possa ocorrer na construção ergativa e não na média, ou vice-versa. Os traços semânticos apontados por Levin como características da construção média são derivados automaticamente de fatores tais como a semântica do verbo e o aspecto verbal da sentença. Por exemplo, se em [94] não temos um evento, mas uma propriedade (“é propriedade desse artigo só vender no verão”), isso pode ser atribuído às condições sob as quais se diz de um artigo que “vende”: *vender*, nesse sentido, exprime uma propriedade, mas poderia exprimir um evento em

[100] Sinto muito, mas o chocolate já vendeu todo.

E se aí há um Agente subentendido, isso se deve ao significado de *vender*, que exprime um fato que não pode ocorrer espontaneamente. Isso não tem nada a ver com a construção, mas antes com a maneira como concebemos o evento que chamamos *vender*.

Um caso oposto é *esfarinhar*, que é algo que pode acontecer a um pão sem que nenhum Agente interfira:

[101] O pão esfarinhou todo.

Se isso for correto, teremos que concluir que a oposição entre as construções média e ergativa não tem relevância para a classificação dos verbos, e portanto deixa de ter utilidade, e deve ser abandonada. Vamos explorar essa possibilidade.

Para começar, note-se que uma das propriedades apontadas por Levin não é exclusiva da (presumível) construção média, já que frases ergativas também podem pressupor um Agente: nada impede que se entenda [93] como sendo o resultado da ação de uma pessoa, por exemplo.

[93] A lata encheu rapidamente.

A diferença real, deste ponto de vista, é que o fato de uma lata ficar cheia pode ou não ser resultado da ação de um Agente, ao passo que o fato de um artigo ser vendido só pode ser resultado da ação de um vendedor (e também de um comprador). Essa diferença está codificada na semântica dos próprios verbos, assim como no nosso conhecimento dos processos de “encher” e de “vender”. Se o contexto deixar claro que há um Agente envolvido, entenderemos [93] como resultado de sua ação – mas para descrever isso não é necessário definir uma construção específica, já que há fontes adequadas, independentemente motivadas, que nos fornecem essa informação. Aliás, a própria frase citada por Levin (*crystal vases shatter easily*), ou pelo menos sua tradução portuguesa, *vasos de cristal quebram facilmente*, não me parece em absoluto requerer um Agente em sua interpretação semântica. Acho sintomático que Levin, em seus exemplos, tenha utilizado o mesmo verbo (*shatter* ‘quebrar’) para exemplificar tanto a construção ergativa quanto a média.

Agora voltemos à propriedade da construção média de não denotar um evento, e portanto não ter referência temporal específica. Digamos que estivéssemos estudando construções transitivas como

[102] A chuva derrubou as flores.

[103] A chuva atrapalha as atividades esportivas.

Em [102] temos um evento, com referência temporal específica (“passado”), e em [103] temos uma afirmação de caráter universal, sem referência específica. Mas isso não nos autoriza a definir aqui duas “construções” diferentes. A diferença vem, no caso, do aspecto do verbo, perfectivo em [102] e imperfectivo em [103]. Sabe-se que o perfectivo tem a propriedade de fixar referências temporais; e a diferença entre um evento e uma afirmação universal está justamente na presença ou ausência de referência temporal. Tanto é assim que podemos permutar os tempos e aspectos dos verbos, e o resultado será que a referência temporal específica de [102] desaparece, e aparece uma em [103]:

[104] A chuva derruba as flores.

[105] A chuva atrapalhou as atividades esportivas.

Se essa permuta de tempo/aspecto é suficiente para explicar as diferenças semânticas em questão, não há nenhuma necessidade de se definir duas “construções” diferentes.

Voltando às médias e ergativas, me parece que o caso aqui é perfeitamente análogo ao que illustrei com [102] – [105]. A frase [93] tem referência temporal específica porque o verbo está no perfeito; se o mudarmos para o presente, a referência desaparece:

[106] A lata enche rapidamente.

Já com [94],

[94] Esse artigo só vende no verão.

fica mais difícil introduzir referência específica; isso não se deve à construção, mas à semântica do verbo *vender*, no sentido de “ter aceitação entre os compradores”. Não se pode dizer que um artigo *vende* com base em um único evento, ou em alguns poucos. Um artigo só “vende” quando os clientes o compram com frequência, habitualmente – daí a dificuldade de se introduzir referência temporal puntual. Fica mais fácil se o aspecto for não-puntual: *esse artigo só vendia no verão*. Por essas razões, não distinguirei aqui a construção ergativa da construção média proposta por alguns autores; e direi que a construção é ergativa nos exemplos acima.

Vale acrescentar, entretanto, que não me parece que todas as oposições diatéticas sejam necessariamente deriváveis da semântica do verbo e de seus afixos. Por exemplo, Fellbaum (1992) observa que *sell* ‘vender’ ocorre na construção ergativa (para ela, “média”), mas *buy* ‘comprar’ não, e o mesmo pode ser dito a respeito de *vender* e *comprar*, respectivamente:

[107] Sombrinhas vendem mais no verão.

[108] * Sombrinhas compram mais no verão.

Por enquanto, temos que atribuir isso a marcas idiossincráticas dos verbos envolvidos: *vender* é transitivo-ergativo, ao passo que *comprar* é transitivo.

Verbos como *vender* constituem contra-exemplo à condição proposta por Cançado (2002), seguindo Whitaker-Franchi (1989), que estipula que a construção ergativa não é aceita por verbos que acarretam para seu sujeito, necessariamente, a propriedade de ser desencadeador do processo com controle. Um dos exemplos de Cançado é *modelar*:

[109] José modelou um vaso de barro.

[110] * O vaso de barro modelou.

Acredito que Cançado se refere à exigência de um Agente desencadeador e controlador para o evento de “modelar”. Para *modelar*, a condição funciona; mas não funciona para *vender*, porque apesar de o evento de “vender” acarretar necessariamente um Agente (desencadeador com controle), a construção ergativa é possível.

9.8.2. *Testando uma hipótese*

Resumindo o que foi dito até o momento, a construção média não deve ser considerada uma diátese independente porque não é instrumento de subclassificação dos verbos – seus efeitos podem ser explicados através de outros fatores, como o aspecto verbal, e não como propriedade autônoma do verbo da sentença.

Levin (1993) afirma o contrário:

Tem havido algum debate [...] sobre se realmente existe uma alternância média que seja distinta da alternância [transitiva/ergativa], ou se se trata de uma alternância apenas. Verbos que mostram a alternância [transitiva/ergativa] são encontrados na construção média, mas há um bom número de verbos que ocorrem na construção média que não mostram a alternância [transitiva/ergativa].

[Levin, 1993: 26]¹⁴

E, com efeito, Levin tem, em sua lista de tipos de verbos, diversos verbos que ocorrem em uma das diáteses, mas não na outra; sua afirmação parece, portanto, ser válida para o inglês. A questão que nos interessa, entretanto, é se ela vale para o português, e suspeito que a resposta é negativa. Para verificar a validade da afirmação de Levin, examinei sua lista de subclasses de verbos (Levin, 1993: 111-176). Em todos os casos em que a diátese transitiva era possível, mas não a média, ou vice-versa, testei as alternativas com verbos do português, com os resultados dados a seguir.¹⁵ Naturalmente, quando uma frase abaixo é marcada como caso de construção “média”, trata-se de uma simples referência à análise de Levin, adaptada ao exemplo português, e por isso a designação fica entre aspas. Para mim, são todas **ergativas**.

9.8.2.1. Casos em que a ergativa é possível, mas não a média

¹⁴ Como fiz algumas adaptações nessa passagem, dou aqui o original: “[...] there has been some debate [...] about whether there really is a middle alternation that is distinct from the transitive/inchoative alternation or whether there is only a single alternation. Verbs that display the transitive/inchoative alternation are found in the middle construction, but there is a number of verbs found in the middle construction that do not display the transitive/inchoative alternation.” [Levin, 1993: 26]

¹⁵ Não estou incorrendo no erro de utilizar resultados de uma língua para validar uma análise de outra. Usei a lista de Levin apenas como uma fonte de idéias para construir uma lista de verbos portugueses, testados em termos de seu uso em português. Parti da idéia de que seria de esperar que pelo menos alguns dos verbos que aceitam a ergativa e não a média, ou vice-versa, em inglês teriam comportamento semelhante em português. Essa hipótese, como vimos não foi confirmada.

Stand verbs (item 9.2 de Levin, 1993): não há verbos correspondentes em português – isto é, as diáteses dos verbos que traduzem os *stand verbs* em português são diferentes. Por exemplo, com *stand* temos:¹⁶

[111] The books stood on the table. [ergativa]

[112] * Tall books stand on tables easily. [“média”]

A tradução portuguesa de [111] precisa utilizar verbos como *ficar* ou *estar*, que ocorrem em ambas as construções:

[113] Os livros ficaram / estavam na mesa [ergativa]

[114] Livros altos ficam em estantes estreitas. [“média”]

Pour verbs (9.5): em português, esses verbos aceitam as duas construções. Assim, para os exemplos de Levin

[115] Water pours onto the plants. [ergativa]

[116] * Water pours easily onto the plants. [“média”]

temos em português frases aceitáveis:

[117] A água derramou nas plantas. [ergativa]

[118] Água derrama fácil. [“média”]

9.8.2.2. Casos em que a média é possível, mas não a ergativa

Esses são os casos mencionados por Levin na passagem citada acima, e com efeito parecem ser numerosos em inglês. Em português, porém, tudo indica que não existem. Em todos os casos examinados, o português aceita livremente a ergativa ou a “média”.

Cut verbs (21.1):

Exemplos de Levin:

[119] * The bread cut. [ergativa]

[120] Whole wheat bread cuts easily. [“média”]

¹⁶ Os exemplos em inglês são os de Levin, com numeração minha. Estou usando “ergativa” onde Levin usa *inchoative*.

Em português, temos

[121] Meu braço cortou todo. [ergativa]

[122] Banana corta facilmente. [“média”]

[123] Quando serrei as pernas, o tampo da mesa serrou também. [ergativa]

[124] Essa madeira serra fácil. [“média”]

Carve verbs (21.2):

Exemplos de Levin:

[125] * The stone carved. [ergativa]

[126] Marble carves easily. [“média”]

A situação é a mesma que para os verbos *cut*, semanticamente muito semelhantes. Em português, com alguns ambas as construções são possíveis, a julgar por

[127] O Zé não soube usar a máquina, e o filé moeu todo. [ergativa]

[128] Essa carne é dura, não mói. [“média”]

Com outros verbos, como *esculpir*, nem a ergativa nem a (presumível) “média” é aceitável.

Shake verbs (22.3):

Exemplos de Levin:

[129] * The sugar whipped into the cream. [ergativa]

[130] Sugar whips into cream easily. [“média”]

Em português, temos:

[131] ? O açúcar amassou na margarina. [ergativa]

[132] ? O açúcar amassa em margarina ou em manteiga. [“média”]

Eu diria que as frases [131] e [132] são ambas ligeiramente estranhas, não muito naturais. Mas, primeiro, elas o são em grau idêntico; e, depois, os demais verbos do grupo de *amassar* (seguindo-se a listagem de Levin), ou se comportam como *amassar*

(*embolar, bater, amarrar, ajuntar*) ou então não aceitam nenhuma das duas construções (*sacudir*). O que interessa é que não há casos claros em que uma das construções é aceitável e a outra inaceitável.

Tape verbs (22.4):

Exemplos de Levin:

[133] * The label taped to the cover. [ergativa]

[134] Labels tape easily to that kind of cover. [“média”]

Em português:

[135] A tinta grudou no banco. [ergativa]

[136] Tinta gruda. [“média”]

Disassemble verbs (23.3):

Exemplos de Levin:

[137] * The handle unscrewed from the box. [ergativa]

[138] That new handle unscrews easily. [“média”]

Em português:

[139] O cabo desparafusou de repente. [ergativa]

[140] Esse cabo desparafusa. [“média”]

Amuse verbs (31.1):

Exemplos de Levin:

[141] *The children amused. [ergativa]

[142] Little children amuse easily. [“média”]

Em português:

[143] O menino assustou com o palhaço. [ergativa]

[144] Esse menino assusta fácil. [“média”]

9.8.2.3. Discussão dos exemplos

Creio que os exemplos acima são suficientes para mostrar que os fatos do português são significativamente diferentes dos do inglês (tais como apresentados por Levin). Ou seja, não encontrei em português um único caso claro de verbo que aceite a construção ergativa mas não a “média”, ou vice-versa. Tenho que concluir, até melhor juízo, que a situação mencionada por Levin não se verifica em português.

Como se vê nos exemplos do português, é freqüentemente necessário criar um contexto para que a construção em teste possa ser aceita. Isso não me parece um problema, já que, primeiro, frases normalmente aparecem em contexto; e, depois, sabemos que o que se chama *construção média* não passa da ocorrência da construção ergativa em determinadas situações. A própria Levin reconhece que na maioria das vezes a construção média é acompanhada por “um elemento adverbial ou modal” (Levin, 1993: 26). Esse elemento está, a meu ver, justamente fornecendo o contexto aspectual que caracteriza essas frases, e que, pelas razões vistas, não acarreta o aparecimento de uma nova diátese.

9.8.3. A proposta de Camacho (2003)

O artigo de Camacho (2003) merece comentário porque trata especificamente do português, e reflete uma concepção a meu ver insuficientemente clara que às vezes se tem do problema. Devo salientar que Camacho só leva em conta dados do português padrão escrito, o que significa que ele não considera a possibilidade de frases como *a cadeira quebrou* (ao lado de *a cadeira se quebrou*) que, se incluídas, o levariam possivelmente a introduzir modificações em sua proposta. Assim, a presença do pronome *se* é, para Camacho, uma marca necessária da construção que ele chama “média” – e que eu prefiro considerar apenas um caso especial da construção ergativa.

O trabalho de Camacho não aborda o problema discutido nas seções 10.8.1 e 10.8.2 acima, ou seja, o da distinção entre as construções ergativas e as chamadas “médias”. Aliás, até onde posso ver, Camacho chama de “média” ambas as construções, isto é, não faz distinção entre médias e ergativas. Isso se depreende, por exemplo, da passagem seguinte:

As sentenças formalmente ativas cujo sujeito é não-afetado serão consideradas sentenças básicas de diátese ativa, enquanto as sentenças formalmente ativas cujo sujeito é afetado serão consideradas sentenças básicas de diátese média.

[Camacho, 2003: 92]

É claro que segundo essa definição frases tipicamente ergativas como

[145] Os alunos assustaram.

precisam ser analisadas como exemplos da construção “média”.¹⁷

Isso quer dizer que o que Camacho pretende é mostrar a necessidade de definir uma construção distinta da transitiva; essa diátese se identifica com a que chamo “ergativa”. Eu não considero o reflexivo uma marca necessária da construção ergativa, ponto em que me distancio da posição de Camacho. Mas no essencial concordo com ele quanto à necessidade de definir essa construção (para mim “ergativa”, para ele “média”).

No entanto, não concordo com a argumentação de Camacho, que a meu ver é insuficiente para demonstrar o ponto. O artigo de Camacho exemplifica um problema bastante comum na literatura lingüística, o de tentar demonstrar um ponto de vista sem ter o cuidado prévio de definir com a maior clareza possível os termos fundamentais a serem usados na discussão. Não se trata de deficiência desse autor, mas um verdadeiro vício, bastante encontrado entre os lingüistas.

A argumentação de Camacho se baseia na proposta de Kemmer (1994) que

enumera uma lista de dez tipos de situações altamente relevantes, denominadas tipos de situação medial [...]: cuidados corporais (Latim *lavo-r*, português *lavar-se*), movimento não-translacional (Latim: *revertor*, português: *virar-se*) [...] eventos espontâneos (francês *s'évanouir*; sem tradução no português; alternativa: *originar-se*) [...]

[Camacho, 2003: 95-96]

E Camacho prossegue (aparentemente, ainda seguindo Kemmer), afirmando que

línguas que dispõem de um marcador medial para representar gramaticalmente esses estados de coisas devem ser consideradas sistemas mediais.

[*id.*, p. 96]

Ele conclui daí que o português é um sistema medial, ou seja, comporta uma “voz média”.

Vejo problemas nesse raciocínio. A lista de Kemmer presumivelmente foi estabelecida sem referência aos fatos do português; ou seja, por mais adequada que seja para outras línguas (os exemplos dados são do latim e do francês), sua adequação ao português é algo a investigar, não algo que se possa pressupor de início.

Nesse ponto, a argumentação de Camacho revela a tendência comum de analisar uma língua em termos de outra, em nome de um universalismo que acaba se reduzindo a circularidade. Assim, Camacho defende que

¹⁷ Não fica totalmente claro no artigo o que vêm a ser sentenças “básicas”, que são as únicas estudadas por Camacho, a julgar por sua afirmação de que “este trabalho interessa-se [...] pela relação do sujeito com situações expressas em sentenças básicas, não em sentenças derivadas.” [Camacho, 2003: 92] Trata-se, presumivelmente, de uma noção vinculada a um modelo transformacional de análise. No entanto, os exemplos dados sugerem que as construções aqui chamadas “ergativas” não seriam derivadas, mas básicas; na verdade, não se distinguiriam das “médias” de Camacho. Por exemplo, Camacho procura distinguir “médias” de “reflexivo-recíprocas”, mas nunca menciona a necessidade de diferenciar “médias” de “ergativas” (ou “inacusativas”; ou ainda “incoativas”, na nomenclatura inadequada de Levin.

A associação da medialidade com intransitividade no inglês pode constituir evidência indireta e comparativamente reveladora para postular que os clíticos de verbos como *barbear-se*, *vestir-se*, *lavar-se* sejam considerados marcadores mediais e não pronomes reflexivos. Para essa classe de verbos, o inglês não admite o reflexivo.

[*id.*, p. 96]

Um dos exemplos é *John washed* ‘John se lavou’, sem o reflexivo *himself*.

A argumentação aqui encerra uma falácia fundamental: a de que dados de uma língua são relevantes, automaticamente, para a análise de outra. Esse tipo de raciocínio é utilizado com a justificativa de que é preciso comparar línguas para efeito do estabelecimento de universais. Mas os dados, tratados dessa maneira, não são confiáveis: se o português tem uma voz média **porque** o inglês (nas frases que traduzem os exemplos portugueses) não emprega o reflexivo, que valor comparativo tem essa observação? Se a análise do português foi feita, mesmo parcialmente, com base em dados do inglês, a comparação das duas línguas, nesse particular, está viciada de início.¹⁸

O “marcador medial”, para Camacho, é o reflexivo: teríamos uma média em

[146] Ela se lavou.

porque esse verbo expressa “cuidado corporal”, e ocorre com *se*.

Isso faria sentido se pudéssemos provar que todo verbo que se encaixa nas categorias semânticas citadas por Kemmer é sempre acompanhado (em português escrito) por um reflexivo. Mas há contra-exemplos: Camacho cita o francês *s'évanouir*, e comenta: “sem tradução no português; alternativa: *originar-se*”. Mas *s'évanouir* tem tradução em português, a saber, *desmaiar* – que, mesmo na língua escrita, não ocorre com reflexivo. Substituir *desmaiar* por *originar-se*, a meu ver, vicia a argumentação, cujo objetivo, nesse momento, seria mostrar que todos os verbos que exprimem “evento espontâneo” são reflexivos. É fácil fazer uma lista de verbos que, como *desmaiar*, dispensam *se* e estão em campos semânticos próximos: *morrer*, *falecer*, *desfalecer*, *adoecer* etc.

Ou seja, não é verdadeira a afirmação de que todo verbo que se encaixa nas categorias semânticas citadas por Kemmer é reflexivo em português. Mas, ainda que fosse, teríamos que demonstrar que essa construção é uma diátese verbal legítima. Na verdade é: mostrei acima que a construção ergativa (que, ao que parece, se identifica com a “média” de Camacho) é base para uma diátese verbal. Mas Camacho não mostra isso, perdendo-se em uma argumentação baseada na aceitação, aparentemente sem crítica, da proposta de Kemmer (1994).

Um dos ingredientes semânticos propostos por Kemmer é o de “cuidados corporais”, presente em verbos como *lavar-se*. A proposta de Kemmer parece ser uma tentativa de caracterizar interlingüisticamente a semântica da voz média. Mas isso só faz

¹⁸ Usar a análise de uma língua para inspirar a de outra é justificável, com as devidas precauções, no chamado “contexto de descoberta” – ou seja, a análise justificada para uma língua é utilizada como fonte de idéias para a análise de outra. Mas não se admite o mesmo no “contexto de justificação” – ou seja, nunca é válido afirmar que uma análise é boa para a língua X **porque** é boa para a língua Y.

sentido na presença de uma definição **independente** de “voz média”; e, além do mais, depende de confirmação empírica para cada nova língua que se examine. No entanto, Camacho parece tomá-la como um princípio a ser seguido na identificação de construções médias:

um verbo pronominal, como *vestir-se*, por exemplo, não será considerado uma instância de reflexividade, mesmo que haja um predicado transitivo correspondente. A razão disso é que *vestir-se* representa uma situação de cuidado corporal, o que é por definição uma atividade própria da entidade iniciadora e/ou controladora do evento.

[*id.*, p. 97]

O fato de que *vestir-se* expressa cuidado corporal não pode ser a razão para que essa construção não seja considerada reflexiva. Falta aqui, pelo menos, uma definição de “reflexiva” que seja independente do fato, em si inegável, de que *vestir* denota um tipo de cuidado corporal.

Na falta de definições aceitáveis para a construção média (ou para a reflexiva), ficamos com essas afirmações vagas ou circulares. Não é de espantar, então, que o próprio Camacho se confunda na prática, ao tentar aplicar a classificação a casos concretos. Vimos que para ele *vestir-se* não é reflexivo (e é, presumivelmente, médio). Logo adiante, ele afirma que

nas construções médias, o clítico não permite, por um lado, comutações com outros termos possíveis do mesmo paradigma e, por outro, não estabelece com o sujeito uma relação semântica de correferência e sintática de coindexação.

[p. 98]

Mas isso é falso, considerando os próprios exemplos de Camacho, porque, primeiro, o clítico pode certamente ser comutado com outros termos do paradigma: *eu te vesti, ela vestiu todos os atores* etc. Depois, em *vestir-se* o clítico está justamente estabelecendo uma relação de correferência com o sujeito: *a atriz se vestiu* significa que a vestidora e a vestida são a mesma pessoa.

Camacho tenta escapar dessa contradição afirmando que

[com *vestir* o] uso transitivo é uma situação incomum, marcada, e só é possível de ser enunciada nas situações em que o iniciador do estado de coisas não está em condições de controlá-lo, como se observa em (6a-b) abaixo.

(6) a. *Após o banho, a tia vestiu o bebê com cuidado.*

b. *Depois que Antônio fraturou o braço, sua esposa o veste todos os dias.*

[p. 97]

Não concordo com as intuições de Camacho; e arrisco dizer que a maioria dos brasileiros aceita essas frases como perfeitamente normais, e não “incomuns”. Nesses casos não é verdadeiro que o iniciador não pode estar em condições de controlar o evento; poderíamos dizer, por exemplo,

[147] A esposa de Antônio o veste todos os dias.

sem subentender nenhuma incapacidade de Antônio de vestir-se por si próprio.

Sumariando esta resenha de Camacho (2003), direi que (a) em princípio, concordo com Camacho em que há uma diátese, que ele chama “média” e eu “ergativa”, válida para a subclassificação dos verbos do português; (b) no entanto, não me parece que Camacho tenha mostrado isso; tentei sanar essa lacuna acima, nas seções 10.8.1 e 10.8.2; (c) acredito que pelo menos em parte o fato de Camacho não conseguir solucionar o problema convenientemente provém da falta de uma conceituação clara e operacionalizável de “diátese”, o que procurei fornecer no capítulo 9; e (d) finalmente, não me parece que as categorias semânticas de Kemmer (1994) sejam adequadas para descrever os verbos que aceitam a diátese ergativa em português.

Quanto a este último ponto, gostaria de acrescentar que ele é de grande interesse, em princípio, pois se liga à idéia de que o comportamento sintático dos verbos é (em grande parte) determinado por seu significado. Levin (1993) coloca assim essa hipótese:

[...] o comportamento de um verbo, em particular no que diz respeito à expressão e interpretação de seus argumentos, é em larga medida determinado por seu significado.

[Levin, 1993: 1]

No entanto, a lista de Kemmer não parece ser ainda a demonstração, para o português, da validade dessa hipótese para o português. O problema continua de pé, e merece consideração.

9.8.4. *Sumário e conclusão*

Procurei mostrar que, tendo em vista os dados disponíveis, não se justifica a distinção entre construção ergativa e construção média em português. A razão principal é que essa distinção não é instrumental na subclassificação dos verbos – ou seja, não existem verbos que exijam ou recusem uma delas, mas não a outra; conseqüentemente, a distinção entre ergativa e média não tem utilidade para o estabelecimento de diáteses verbais. Esses verbos parecem existir em outras línguas (a julgar pelos dados de Levin, 1993, para o inglês), mas em português não se conhecem exemplos. As diferenças apontadas para distinguir “ergativas” de “médias” são decorrentes de fatores independentes, não relacionados com uma presumível propriedade da construção propriamente dita: o aspecto verbal e fatores ligados ao nosso conhecimento do mundo.

O português tem uma construção **ergativa** que se opõe a outras construções tais como a transitiva. Cada construção se define através de suas propriedades formais (sua estrutura), mais a distribuição dos papéis temáticos pelos diferentes termos da oração. Isso quer dizer que uma construção, tal como aqui se define, é uma entidade **simbólica** (seguindo-se a terminologia de Langacker, 1999), ou seja, uma estrutura formal específica associada a certos traços semânticos específicos.

9.9. Terminologia

A terminologia aqui adotada requer algum comentário, para evitar confusões. Não há consistência na literatura quanto à maneira de designar as construções que chamo “transitiva” e “ergativa”, nem quanto ao nome dado às classes de verbos definidos em função de ocorrência nessas construções.

Os verbos transitivo-ergativos são às vezes chamados simplesmente “ergativos” – o que, como observa Trask (1997: 79) “é incorreto e questionável”. A objeção de Trask é motivada, creio, pela idéia de que somente verbos que ocorrem exclusivamente na construção ergativa (como *morrer*) deveriam ser chamados “ergativos”. A terminologia que preferi usar evita, espero, a crítica de Trask: chamo *morrer* de **ergativo**, e *matar*, que somente ocorre na construção transitiva, **transitivo**. Já *começar*, que ocorre em ambas as construções, será chamado **transitivo-ergativo**. Alguns autores, como Renzi *et al.* (1988) chamam os transitivo-ergativos de “inacusativos”.¹⁹

Um verbo é **ergativo** se só ocorre na construção ergativa; é **transitivo** se só ocorre na construção transitiva; e é **transitivo-ergativo** se pode ocorrer nas duas. A nomenclatura que adoto evita a confusão freqüentemente encontrada na literatura. Trask define um “verbo ergativo” como

Um verbo que pode ser usado intransitivamente, com sujeito paciente, ou transitivamente, com objeto paciente e sujeito agente.

[Trask, 1997: 79]

E define um “verbo inacusativo” como

[u]m verbo intransitivo cujo sujeito não é agente, como *die, melt, sink* ou *explode*. Muitos desses verbos podem também ser usados transitivamente: *The ice melted; The sun melted the ice*.

[p. 227]

Não é difícil verificar que as duas definições cobrem exatamente os mesmos verbos (é bom observar que as definições não se devem a Trask; ele apenas as coligiu da literatura, e inclusive, como vimos, considera incorreta a denominação “verbo ergativo” para *melt* etc.).

Alguns autores definem a construção “causativa” (correspondente à minha transitiva) em termos transderivacionais, ou seja, uma construção com verbo e objeto direto seria “causativa” apenas quando existe na língua uma construção semanticamente próxima com um verbo dos tradicionalmente chamados “causativos” (*fazer, mandar, deixar*) seguidos de uma oração infinitiva ou não. Assim, [148]

[148] O orador emudeceu o auditório.

seria causativa porque existe também

¹⁹ Sou também culpado de ter rotulado esses verbos como inacusativos (Perini, 2002). Definitivamente não pretendo entrar em discussões puramente terminológicas, e meu uso desses termos segue simplesmente o uso de certos autores. Mas há problemas etimológicos, como o de chamar “ergativa” (de ἔργον ‘trabalho’) uma construção caracterizada justamente pela ausência de ação.

[149] O orador fez o auditório emudecer.

Mas a frase

[150] Marquinhos amarrou o cachorro.

não seria exemplo de construção causativa porque não há paralelo com *fazer, deixar, mandar*, etc. Essa é a concepção que deduzo, por exemplo, de Bittencourt (1984).

No modelo descritivo que adoto, que é estritamente superficial, e portanto não-transformacional, tais definições não têm lugar. Cada construção é gerada independentemente, e se define por suas características próprias, sem referência a eventuais construções semântica ou formalmente semelhantes. Assim, defino a construção transitiva apenas pelos termos sintáticos que inclui, mais o papel temático que cada um deles desempenha. No caso da construção transitiva, o sujeito é Agente.

Capítulo 10: A valência na interface entre semântica e estrutura informacional

10.1. Introdução

O mecanismo estudado nos capítulos precedentes procura explicitar, entre outras coisas, a atribuição de papéis temáticos aos diferentes complementos de um verbo. Como vimos, esse processo é controlado por fatores estruturais (funções sintáticas) e léxicos (valência do verbo). É nesses termos que a atribuição de papéis temáticos é estudada na literatura. No entanto, o brasileiro falado, como outras línguas, inclui estruturas de tópico, onde as funções sintáticas são diferentes das tradicionais. Nesses casos, o mecanismo de atribuição dos papéis temáticos é parcialmente diferente do que geralmente se admite, em especial no que se refere às funções sintáticas envolvidas. Neste capítulo apresento algumas notas preliminares sobre esse fenômeno, ainda não estudado até hoje (no que pesem observações de interesse de certos autores, citados no decorrer do capítulo).

Devo enfatizar que os fatos estudados neste capítulo não estão integrados na análise que proponho para as valências verbais. Aqui só posso oferecer sugestões preliminares, que poderão orientar pesquisas futuras.

Em geral se admite que os papéis temáticos são atribuídos pelo verbo com base nas funções sintáticas dos diversos complementos. Assim, o verbo *derreter* atribui o papel de Agente a seu sujeito e o de Paciente a seu objeto direto:

[1] O calor derreteu o gelo.

Isso é condicionado à construção em que o verbo ocorre; *derreter* pode atribuir o papel de Paciente ao sujeito se a construção for a ergativa:

[2] Gelo derrete a zero grau.

[3] O gelo derreteu.

Mas, em qualquer caso, a função sintática é tomada como crucial para a atribuição correta dos papéis aos diversos complementos. Sem dúvida, nesses casos o fator determinante na atribuição de papéis temáticos é a valência do verbo. Isso explica a diferença do papel temático de *o gelo* em [3] e de *o vizinho* em

[4] O vizinho sorriu. (construção intransitiva)

A estrutura sintática é idêntica (sujeito + verbo), mas em [3] o sujeito é Paciente, e em [4] é Agente. Isso se deve às propriedades diferentes dos verbos *derreter* e *sorrir*, ou seja, ao fato de que eles têm valências diferentes: o verbo *sorrir* não ocorre na construção ergativa, ao passo que *derreter* ocorre. E em geral se define uma construção através de

determinados papéis temáticos associados a determinadas funções sintáticas; assim, a construção ergativa se caracteriza como tendo “sujeito Paciente” e assim por diante. Essa é a posição gerativista usual:

Dizemos que o verbo *kill* toma dois argumentos aos quais atribui um papel [temático]: atribui o papel de AGENTE ao argumento sujeito da sentença, e o papel de PACIENTE ao argumento objeto. O verbo **marca tematicamente** [*theta-marks*] seus argumentos.

[Haegeman, 1991: 41]

A vinculação dos papéis temáticos às funções sintáticas parece ser também pressuposta por lingüistas que trabalham segundo outros modelos. Por exemplo, Langacker diz, ao comentar a codificação de uma ação que envolve agente e paciente, que

Prototipicamente, uma oração que codifica essa situação realça [*profiles*] a cadeia de ações inteira, escolhendo o agente para ser seu sujeito e o paciente para ser seu objeto direto.

[Langacker, 1991: 216]¹

Ou seja, pode-se afirmar que a posição canônica encontrada na literatura é a de que a atribuição dos papéis temáticos se faz com base nas funções sintáticas. Neste capítulo vou mostrar que nem sempre é assim.

10.2. Casos problemáticos

Há casos que levam a imaginar que nem sempre a função sintática, entendida em termos tradicionais, é relevante para a atribuição dos papéis temáticos. Me refiro a estruturas de tópico, como por exemplo

[5] Esse cano sai fumaça.

Frases desse tipo foram analisadas por Pontes (1986; 1987) como contendo um tópico sentencial (no caso, *esse cano*) externo à estrutura da oração. Aqui vou examinar as condições que governam a atribuição do papel temático ao sintagma inicial, que é o tópico sentencial. Os dados provêm, em grande parte, do estudo pioneiro de Pontes (1987), colhidos na observação informal da fala espontânea: de gravações, conversas etc. Os exemplos de Pontes, que são muitos e muito variados, são identificados pela sigla [EP].² Os outros exemplos foram colhidos por mim.

Aqui vou mostrar que o sistema habitualmente aceito de atribuição de papéis temáticos falha em muitos casos semelhantes ao de [5]. Para ser mais explícito: vou

¹ Não encontrei uma discussão explícita da atribuição dos papéis temáticos em Langacker. Mas essa passagem deixa claro que ele considera a atribuição com base no verbo pelo menos o caso normal (“prototípico”).

² Em alguns casos modifiquei ligeiramente a grafia usada por Pontes.

analisar casos em que o papel temático não pode ser totalmente derivado da estrutura sintática explícita, e onde portanto não se pode falar de “atribuição” desses papéis a partir da estrutura sentencial. Em tais sentenças, o papel temático do sintagma que funciona como tópico depende de uma colaboração entre a estrutura (o “tópico”, definido como uma função sintática) e restrições de ordem pragmática. Como Pontes mostra nos trabalhos citados, tais sentenças são extremamente comuns no brasileiro falado, e não podem ser marginalizadas como casos excepcionais.

No caso de [5], *esse cano* tem o papel temático de Fonte (a origem de um movimento); no entanto, a estrutura é de um sintagma nominal, sem a preposição que tipicamente marca a Fonte com verbos da classe de *sair*, como em

[6] Sai fumaça desse cano.

O problema a ser abordado aqui é o de que fator governa a atribuição do papel temático em [5], e em frases de tópico³ em geral.

10.3. O ponto de partida da interpretação

Antes de abordar o problema propriamente dito, vou considerar algumas questões prévias que nos permitirão colocar o problema em seus devidos termos. A principal dessas questões é a determinação do ponto a partir do qual o receptor interpreta uma seqüência gramatical.⁴ Já tratei disso anteriormente, e agora volto ao tema em conexão com a interpretação das sentenças de tópico.

Há muitas teorias sobre o processamento de material lingüístico, mas todas elas reconhecem que o receptor parte das seguintes informações:

- (a) a seqüência material (fonética ou gráfica) fornecida pelo emissor;
- (b) o conhecimento que o receptor tem da gramática de sua língua; e
- (c) o conhecimento que o receptor tem do mundo em geral.

Isso exclui, notadamente, estruturas enriquecidas com elementos não realizados foneticamente, como categorias vazias, marcas de coindexação, limites de constituintes e sua hierarquia. Pode ser, em princípio, que se tenha que lançar mão desses recursos (isso é uma questão a discutir), mas certamente não é possível **partir** deles, já que não fazem parte nem da materialidade do enunciado, nem do conhecimento prévio que o receptor tem da língua e do mundo.

³ “Frases de tópico” segundo a definição de Pontes, ou seja, justamente aquelas que fogem à possibilidade de análise em termos tradicionais. No caso, ficariam excluídos casos de simples ocorrência anteposta de algum termo da oração, como os de topicalização: *dessa toca não sai coelho* etc.

⁴ E igualmente o ponto ao qual o emissor chega no final de sua tarefa codificadora. Aqui vou considerar apenas o ponto de vista do receptor, mas o argumento funciona igualmente para o emissor, *mutatis mutandis*.

Por exemplo, uma análise possível (embora eu não esteja disposto a adotá-la) é a de que a frase

[5] Esse cano sai fumaça.

inclui, em uma estrutura subjacente, a preposição *de* antes de *esse cano*. Essa afirmação pode (talvez) trazer vantagens para a análise em geral, mas não resolve o problema de por que o sintagma *esse cano* é interpretado como Fonte, embora não seja marcado por preposição na seqüência fonética que constitui a base material da transmissão da mensagem do emissor ao receptor. Se há uma preposição subjacente aí, ela precisa ser restaurada pelo receptor, com base em alguma informação previamente disponível. A questão que coloco é justamente que tipo de informação prévia é utilizada para isso – a meu ver, aliás, não a restauração de uma preposição, mas diretamente a atribuição do papel de Fonte ao sintagma *esse cano*. Essa pergunta precisa ser respondida, independentemente da opção teórica que se possa adotar.

A restauração de uma preposição abstrata em estruturas como [5] é altamente implausível, inclusive pelo seguinte: para saber que preposição restaurar (por exemplo, *de*, mas não *para* ou *por*) o receptor precisa identificar o papel temático de *esse cano*. Se for Fonte, a preposição será *de*, se for Meta será *para* ou *por*. Mas, uma vez identificado o papel temático, a restauração da preposição não faz mais sentido, pois ela só foi “inventada” para marcar o papel temático.

10.4. O papel temático do tópico

Como vimos, o tópico pode ser conectado semanticamente à oração por um papel temático. É preciso reconhecer que em certos casos é difícil vislumbrar essa conexão. Um exemplo é

[7] Eu, graças a Deus, é só café. [EP]

Aqui o tópico é *eu*, mas não há um verbo explícito e portanto não se pode atribuir ao tópico um papel temático. Outro exemplo é

[8] Eu agora, 'cabou desculpa de concurso, né? [EP]

Aqui temos um verbo (*'cabou*), mas não parece possível encaixar *eu* em sua grade temática. Esses casos não serão estudados aqui, embora constituam um tema interessante de pesquisa.

Em geral, porém, o tópico pode ser colocado com certa facilidade em um dos encaixes do esquema valencial do verbo. O problema, como vimos acima, é que podem faltar as marcas formais que caracterizam o papel temático: em [5], falta uma preposição. Vale perguntar, então, como é que o receptor resolve o problema da atribuição do papel temático.

Uma pista bastante óbvia é o conhecimento que o usuário da língua tem das valências de cada verbo. Isso deve ser o que permite a atribuição do papel de Fonte a *esse cano* em [5],

[5] Esse cano sai fumaça.

Ou seja, *sair* pode ocorrer em uma construção que admite um Tema (o elemento que se desloca) e uma Fonte (o lugar de onde o movimento se origina). A Fonte é marcada por preposição, normalmente *de*, mas não quando se trata de um tópico. Este, seguindo ao que parece uma regra geral, aparece sem preposição pelo menos na maioria dos casos. Assim, podemos ter

[6] Sai fumaça desse cano.

com a Fonte em posição de não-tópico, e portanto marcada obrigatoriamente por preposição; ou então

[5] Esse cano sai fumaça.

Mas só raramente ocorre a estrutura com Fonte topicalizada e com preposição:

[9] Desse cano sai fumaça.

Esta última alternativa é possível, mas não parece ser a maneira usual de codificar a mensagem na língua falada.

A interpretação de [5] pode seguir as seguintes linhas: primeiro, o receptor, ao ouvir a seqüência fonética, lança mão de seu conhecimento da valência de *sair*, que inclui a seguinte diátese:

[10] <i>sair</i> /	X	V	<i>de</i> SN
	Tema		Fonte

Ora, essa diátese, tal como está formulada, não pode ser aplicada a [5], cuja estrutura não confere com a dada em [10]. Em [5] o SN Tema, *fumaça*, está depois do verbo (e isso vamos deixar de lado),⁵ e além disso não há nenhum sintagma preposicionado para assumir o papel de Fonte. Temos aqui um problema, pelo menos, de formulação de diáteses. Devemos criar uma nova diátese para enriquecer a grade valencial de *sair*? Essa seria a solução mais imediata. Mas a meu ver deixa de lado algumas observações que são bastante óbvias, razoavelmente seguras e de certo interesse.

Primeiro, o fenômeno não se limita ao verbo *sair*, nem a nenhum grupo específico de verbos, mas parece ser de natureza geral, valendo para uma grande variedade de categorias sintáticas e semânticas, quando topicalizadas (ver os exemplos da seção 10.8). Em segundo lugar, o papel temático recuperado nem sempre se vincula a um verbo, mas

⁵ Trata-se do tradicional problema da posposição do sujeito; ver a respeito a seção 4.2.3.

é às vezes interno a um SN. Esse caso se verifica em particular com o papel de Possuidor,⁶ como em

[11] Meu carro furou o pneu.

onde é necessário vincular semanticamente *meu carro* com o Possuidor do pneu (“o pneu do meu carro”). Aqui, evidentemente, não se pode resolver o problema através de uma nova diátese para *furar*. Concluo que precisamos não de um conjunto de novas diáteses para os verbos envolvidos, mas de um mecanismo de atribuição de papéis temáticos a sintagmas em posição identificada como de tópico.

O primeiro problema, pois, é a identificação de um sintagma como tópico. Voltando ao caso de [5] e [6], sabemos que a valência de *sair* inclui um Tema, codificado como SN, e uma Fonte, codificada como SPrep. Note-se que essa maneira de codificar a Fonte não pode ser alterada no caso geral, ou seja, em posição não-topicalizada a preposição é obrigatória:

[12] * Sai fumaça esse cano.⁷

Essa frase não se beneficia do mecanismo, qualquer que seja, que atribui a um SN (sem preposição) o papel de Fonte; em vez, ela é inaceitável, ao que tudo indica justamente por falta da preposição – o que mostra que o mecanismo em questão é exclusivo de frases com tópico sentencial.

10.5. Identificando o tópico

O primeiro passo no mecanismo de atribuição do papel temático a um tópico sentencial é a identificação desse tópico. Sabemos que ele ocorre em primeiro lugar na oração; pelo menos como hipótese inicial, vamos admitir que o primeiro sintagma da oração (não sendo o próprio verbo) é sempre identificado como tópico. A discussão dos exemplos [5], [6] e [12] mostra que o tópico, assim definido, é processado de maneira diferente dos sintagmas não-tópicos: por exemplo, o sintagma *esse cano* pode ser interpretado como Fonte se for tópico, mas não se estiver em outra posição na oração.

Em alguns casos, o problema da atribuição do papel temático ao tópico é trivial; por exemplo, quando o tópico é sujeito:

[13] O leite esfriou.

Aqui o SN *o leite* recebe o papel de Paciente, e mesmo se esse papel ocorrer em alguma outra posição será representado por um SN, aliás o único da oração:

⁶ Falando em termos bem amplos. É possível que esse papel temático seja diferente do que encontramos em *a casa de mamãe*, mas não podemos entrar nesse problema aqui.

⁷ Pode ocorrer uma frase seqüencialmente semelhante a [12], mas com entonação diferente: *sai fumaça, esse cano* – com uma cesura clara depois de *fumaça* e entonação baixa em *esse cano*, o que marca o sintagma final como tópico à direita. Apesar da posição diferente do tópico, ele continua marcado como tal, e entra no sistema geral de atribuição de papel temático exclusivo de tópicos.

[14] Já esfriou o leite?

Nesses casos pode-se dizer que ambos os sistemas de atribuição de papel temático – o sistema de base sintática e o sistema exclusivo para tópicos – dão o mesmo resultado, e na verdade não vejo como apurar qual deles está funcionando em [13] e [14].

As coisas mudam quando um tópico recebe um papel temático que é normalmente atribuído a um sintagma marcado diferentemente. É o caso de

[5] Esse cano sai fumaça.

onde a Fonte é representada por um SN, o que não acontece normalmente, pelo menos com o verbo *sair*, pois a preposição é obrigatória em

[6] Sai fumaça desse cano.

A pergunta a ser formulada é: como é que o receptor identifica o papel temático de *esse cano* em [5], se falta a marca sintática de Fonte?

10.6. O mecanismo de atribuição do papel temático

Já vimos que não pode ser simplesmente uma nova diátese de *sair*. E, de qualquer maneira, não poderíamos marcar como Fonte todo SN inicial de frases com *sair*, por causa de frases como

[15] A pracinha, sai pela porta da frente.⁸

onde o tópico é um SN, como em [5], mas tem o papel de Meta, não de Fonte. Isso inviabiliza a solução do problema através da formulação de uma nova diátese.

Acredito que o que entra em jogo aqui são fatores de ordem pragmática (conhecimento do mundo). Voltando a

[5] Esse cano sai fumaça.

o receptor atribui a *esse cano* o papel de Fonte porque é simplesmente o papel mais plausível, dado o significado básico da sentença. O processo deve ser algo baseado em conhecimento do mundo, inclusive as probabilidades estimadas de ocorrência desta ou daquela situação.

É evidente que o processo de base pragmática coexiste com outro, que interpreta necessariamente *desse cano* (com preposição!) como Meta, independentemente de considerações de plausibilidade – ou seja, o sistema tradicional de atribuição de papel temático com base na sintaxe. Creio que vale a pena especular um pouco sobre as características desse sistema complexo, mesmo porque em geral sua existência não é

⁸ Frase inventada, mas testada com sucesso: soa normal para os falantes.

reconhecida. Digamos, então, que o receptor aplica um (quase-)algoritmo da seguinte forma:

1º passo:

Para cada sintagma:

é Tópico?

Se a resposta for *não*, aplique-se o sistema léxico-sintático, a saber:

(c) a valência do verbo *sair* exige um Tema e um outro argumento, que pode ser Fonte ou Meta;⁹

(d) a Fonte é marcada com *de* e a Meta é marcada com *para* ou *por*;

(e) o Tema é um SN (sujeito).

Decisão: Atribuir o papel temático de cada sintagma de acordo com sua marca formal.

Se a resposta for *sim*, passar ao →

2º passo:

o sintagma está marcado por preposição?

Se a resposta for *sim*, aplique-se o 1º passo.

Se a resposta for *não*, aplique-se o sistema misto, a saber:

Decisão: Atribuir ao tópico o papel temático presente na valência do verbo, levando em conta as considerações pragmáticas de plausibilidade, dado o contexto e a probabilidade estimada de cada situação. Algo como

(a) *esse cano* é Fonte ou Meta? É mais comum a fumaça sair de um cano (Fonte) do que sair para um cano (Meta);

(b) logo, *esse cano* é Fonte.

Como se vê, o sistema, mesmo em sua fase parcialmente pragmática (2º passo) é crucialmente subordinado à valência do verbo, no que esta tem de semântico. No entanto, a parte formal da valência – no exemplo visto, a presença da preposição – é relaxada, de modo que o sintagma *esse cano* pode receber o papel de Fonte, o que nunca aconteceria se ele não fosse tópico.

No caso de

[15] A pracinha, sai pela porta da frente.

o processo seria semelhante, mas a decisão do 2º passo teria que ser em favor de Meta (a gente sai para a pracinha pela porta da frente), porque em geral não se sai de uma pracinha através de uma porta. Mesmo quando isso acontece (por exemplo, sair da pracinha para entrar em casa) o ponto de vista adotado é outro: fala-se de entrar em casa, não de sair da pracinha.

⁹ Ou ambos, ou ainda talvez outros papéis, que vou deixar de lado.

No entanto, mesmo o 2º passo não pode ser definido em termos exclusivamente semântico-pragmáticos, porque depende da definição do sintagma como tópico – e o tópico é uma função sintática, definida pela posição linear na seqüência. Ou seja, o sistema é inevitavelmente complexo, incluindo fatores sintáticos e semânticos (no 1º passo) e sintáticos, semânticos e pragmáticos (no 2º passo).

10.7. Sumário: um sistema complexo

Há algumas lições a serem deduzidas da situação acima descrita. Em primeiro lugar, chamo a atenção para o caráter heterogêneo do fenômeno da atribuição dos papéis temáticos. Em vez de um sistema único e coerente, como se encontra, por exemplo, nas propostas gerativistas atuais, temos um conjunto de recursos que são utilizados conforme as exigências do momento: podemos atribuir papéis temáticos com base na valência do verbo, mais a estrutura sintática; ou então com base na valência, mais considerações de cunho pragmático; ou ainda independentemente da valência verbal (caso de *meu carro furou o pneu*).¹⁰

A pesquisa recente tem revelado que essa situação ocorre em muitas áreas da estrutura da língua. Culicover e Jackendoff (2005) tentaram captar essa propriedade das línguas com sua hipótese da Caixa de Ferramentas (*Toolkit Hypothesis*), que diz essencialmente que as línguas dispõem de uma bateria bastante variada de recursos para realizar sua tarefa de relacionar sons e significados, e que os utiliza, em grande medida, de acordo com as circunstâncias do momento.

Por outro lado, o funcionamento desse mecanismo heterogêneo parece obedecer a certas restrições que são, elas próprias, interessantes. Primeiro, a atribuição de base pragmática não é livre, mas se subordina à valência verbal – por exemplo, em [5] o tópico não poderia ser Paciente, porque *sair* não admite Paciente em sua grade valencial. Assim, a informação pragmática funciona (nesses casos pelo menos) para selecionar, das possibilidades definidas pelo léxico, a mais plausível. Além disso, essa parte do sistema só funciona em frases onde se identifica um tópico.

Outra observação é que a atribuição pragmática parece ser menos taxativa do que a atribuição de base puramente léxica (semântica). Assim, o papel temático de Agente do sujeito de

[16] Manuel cortou a maçã.

não admite alternativa: nessa construção, com esse verbo, o sujeito é Agente, a pronto. Agora, em casos como

[5] Esse cano sai fumaça.

vimos que o tópico é geralmente Fonte, mas seria possível, dado um contexto apropriado, interpretá-lo como Meta – por exemplo, se estivéssemos descrevendo uma lareira, onde o

¹⁰ Naturalmente, vai ser preciso acrescentar algo ao sistema para dar conta de *meu carro furou o pneu*. Creio que um dos fatores em jogo é a valência (nominal) do núcleo do SN, no caso *pneu*.

cano em questão tem a função de levar a fumaça para fora, mas está sendo visto do lado de dentro.

10.8. Exemplos: papéis temáticos atribuídos pragmaticamente

Agora vou dar alguns exemplos de construções de tópico, com os papéis temáticos atribuídos, ao que tudo indica, através da aplicação do 2º passo no sistema descrito. A julgar pela pequena amostra disponível, os seguintes papéis temáticos são particularmente freqüentes:

Lugar

[17] O seu regime entra muito laticínio? [EP]

[18] Aquela escola deles, rouba tudo! [EP]

[19] Minha casa deu ladrão. [EP]

Causador de experiência

[20] Meu cabelo dessa vez eu não gostei nem um pouco. [EP]

[21] A Joana não se deve confiar. [EP]

Nos casos acima se verifica a omissão da preposição, de maneira que o tópico acaba tendo a estrutura de um SN. Nesses casos, como vimos, é indispensável lançar mão de informação pragmática para atribuir o papel temático correto ao tópico. Já em outros casos a estrutura contém marcas formais suficientes para que seja possível processá-la com base em informação léxico-gramatical:

Assunto

[22] Tudo isso tem estudo a respeito. [EP]

Aqui parte da locução prepositiva *a respeito de* é conservada, de modo que não há dúvida quanto ao papel temático a ser preenchido.

Outros casos envolvem sujeito ou objeto direto, como

[23] Esse negócio de tópico eu tou examinando desde o semestre passado. [EP]

[24] Esse sapato dói o pé da gente.

Aqui é possível atribuir o papel temático sem utilizar informação pragmática.¹¹

Finalmente, há muitos exemplos em que o sistema esboçado neste capítulo não é suficiente para dar conta da atribuição do papel temático ao tópico. O mais freqüente, de longe, é o caso já ilustrado na frase [11], de Possuidor:

Possuidor

[25] A lanterna, já comprou as pilhas? [EP]

[26] Aquela casa ali, é engraçado o jardim. [EP]

[27] Meus óculos, você apanhou a capa? [EP]

[28] Esse negócio o prazo acaba. [EP]

[29] Esse rádio estragou o ponteiro. [EP]

[30] Essa primeira parte é interessante a crítica. [EP]

[31] A Belina deita o banco, sabe? [EP]

Quanto a esses exemplos, é interessante notar que eles envolvem o desmembramento de um sintagma nominal, sendo apenas uma parte dele topicalizada. Isso se verifica também em casos de topicalização “normal”, isto é, sem omissão da preposição que é marca do Possuidor:

[32] Desse menino eu só conheço a mãe, o pai não.

Para casos como [11] e [25 - 32], naturalmente, vai ser necessário acrescentar novas partes ao mecanismo de atribuição descrito na seção 10.6.

Um comentário final: essas estruturas provavelmente não receberam ainda a atenção ampla e detalhada que merecem porque são pouco freqüentes em muitas línguas (por exemplo, em inglês). Nesse particular, o português brasileiro, com sua grande liberdade de topicalização, apresenta uma grande riqueza e revela recursos até agora pouco suspeitados de sinalização simbólica.

10.9. Tópicos sem papel temático?

Exemplos como os que estudei acima são os mais freqüentes. Deles pode-se dizer que a sentença inteira se enquadra dentro do esquema simbólico de uma oração, onde figura um verbo e sua valência devidamente preenchida. Mas também ocorrem casos em que não se pode ver com clareza um papel temático vinculando o tópico ao verbo ou ao

¹¹ O que não quer dizer que não seja utilizada às vezes; mas não é possível apurar quando isso ocorre e quando não. Note-se em [32] o uso transitivo de *doer*, bastante comum na fala de Minas.

restante da sentença. Aqui o tópico comparece em sua função discursiva, sem muito respeito pela sintaxe ou pela semântica sentencial; por exemplo,

[33] Essa minha barriga, só jejum. [EP]

[34] A grande maioria desse pessoal, gente, as mulheres são verdadeiras heroínas. [EP]

[35] Já o Jornal do Brasil, você viu a crônica do Drummond? [EP]

Esses casos, certamente interessantes, ficam fora do âmbito do presente estudo. Eles se aproximam dos casos de tópico discursivo, e talvez se identifiquem com eles.

10.10. Examinando (e rejeitando) uma alternativa

Uma análise que se poderia propor para [5]

[5] Esse cano sai fumaça.

é a de que *esse cano* é o sujeito, e *fumaça* o objeto direto – reduzindo [5] a uma estrutura do tipo usual, não de tópico. Vejo dois problemas com essa hipótese. Primeiro, seria necessário criar, para o verbo *sair*, uma nova diátese, porque sabemos que em outros casos o sujeito desse verbo é Tema:

[36] Os gases tóxicos saíram desse cano.

Agora teremos que definir uma nova diátese, na qual o sujeito de *sair* é a Fonte e o objeto direto é o Tema. Isso não é impossível, em princípio (basta comparar a expressão do Paciente em frases transitivas e ergativas), mas pelo menos requer mais comprovação.

Mas há outro argumento que milita contra a análise em questão, a saber, a possibilidade de interpretar *esse cano* como Meta. Vai ser necessário complicar a alternância, acrescentando a possibilidade de uma nova diátese, na qual o sujeito (presumido) é Meta, e o objeto direto Tema. E para cada caso de tópico que tenha um papel temático diferente do que teria se não fosse tópico será preciso criar uma nova diátese. Isso, a meu ver, perde de vista o caráter essencialmente único do fenômeno, e reduz a uma coincidência o fato de que todos os casos envolvem sintagmas nominais em posição inicial na sentença.

10.11. Breve resenha da literatura

Até onde pude verificar, as perguntas acima não foram tratadas em detalhe na literatura. Em geral só se pode encontrar algumas menções rápidas, como relato a seguir.

Na clássica antologia de Li (1976), nenhum artigo ataca a questão da atribuição de papéis temáticos aos tópicos.

Os trabalhos de Pontes (1986, 1987), fonte da maioria dos exemplos aqui estudados, não consideram a questão do papel temático do tópico. Pontes se limita a fazer uma breve observação, utilizando a nomenclatura tradicional para as funções sintáticas (objeto indireto, objeto direto, complemento nominal, adjunto circunstancial, adjunto predicativo, sujeito), embora a meu ver esteja na verdade considerando papéis temáticos. Nessa passagem, ela apenas dá uma lista de frases, agrupadas segundo a função do tópico (1987: 18-19).

Lambrecht (1994) é o único autor que assume uma posição explícita sobre o estatuto sintático e semântico do tópico na sentença. Afirma ele que o tópico sentencial

[...] não pode ser um constituinte – nem argumento nem adjunto – da oração à qual é pragmaticamente associado. Antes, ele deve ser analisado como um elemento extra-oracional sintaticamente autônomo, cuja relação com a oração não é a relação gramatical de sujeito ou objeto mas a relação pragmática de “a respeito de” e relevância.

[Lambrecht, 1994: 192-193]

Mas, primeiro, não concordo que um tópico como *esse cano* em [5] não tem função sintática na oração. Muito embora não seja sujeito nem objeto, e provavelmente tampouco adjunto adverbial, prefiro dizer que *esse cano* em [5], apesar de não ter análise tradicional,¹² tem uma função sintática, porque se coloca em uma posição típica que lhe permite ser identificado como tópico e portanto sofrer o processo de atribuição de papel temático que é exclusivo dos tópicos. Não me parece que isso seja outra coisa que não sintaxe.

Mas, além das relações pragmáticas citadas por Lambrecht, esse sintagma se vincula à oração por uma relação semântica, a saber, o papel temático de Fonte – que é, em outras frases, expressa por um complemento marcado pela preposição *de* e que, com certeza, faz parte da estrutura da oração, preenchendo uma vaga na valência do verbo.

Em contraste, em casos como os de [33-35] é realmente difícil identificar alguma relação semântica dessa natureza, de modo que para essas frases a afirmação de Lambrecht provavelmente se sustenta. Mas esses não parecem nem sequer ser o caso típico; frases do tipo de [5] são freqüentes e absolutamente normais da fala. Essa é a situação no brasileiro falado, e fica em aberto a possibilidade de que outras línguas mostrem comportamento diferente.

Lambrecht assume uma posição curiosa: ele dá a entender que a ausência (segundo ele) de vinculação sintática acarreta a impossibilidade de relação semântica:

[Referindo-se a casos como *The typical family today, the husband and wife both work*.:] o SN tópico não é anaforicamente ligado a um argumento, explícito ou nulo, em nenhuma das proposições que se referem ao tópico. Segue-se daí que o sintagma tópico não pode ser um argumento na oração à qual se associa.

[Lambrecht, 1994: 193]

¹² Seria, no máximo, um anacoluto, considerado uma espécie de deformidade sintática.

Naturalmente, a partir dessa posição Lambrecht se desinteressa de investigar o mecanismo de atribuição de papel temático ao tópico: para ele essa atribuição não ocorre (é o que entendo da passagem citada). No entanto, os fatos desmentem essa conclusão, pois é claro que o tópico, em muitos casos, preenche claramente um dos encaixes semânticos da valência do verbo. Para Lambrecht, isso só acontece quando o tópico é retomado por um elemento anafórico no restante da frase, algo como

[37] Esse sorvete, eu vou guardar ele no freezer.

Essas estruturas existem, mas a retomada pronominal não é obrigatória. E, de qualquer maneira, não vejo por que uma relação semântica só possa ser considerada como presente sob condição de ser sintaticamente representada à maneira tradicional. Uma relação semântica é parte dos dados brutos iniciais com que o lingüista deve lidar: essa relação é acessível ao usuário da língua, e está presente ou não, conforme seu julgamento. Ora, em

[5] Esse cano sai fumaça.

qualquer falante concorda que *esse cano* tem uma relação semântica evidente com o verbo *sair*, a saber, a de Fonte (e às vezes Meta, como vimos). Isso é um fato, e não tem que ser discutido nem condicionado: basta ser observado.

Suspeito que a posição de Lambrecht vem de certa confusão, que é muito freqüente, entre função sintática e função semântica. O tópico não teria função sintática tradicional¹³ – não é sujeito nem objeto etc. – logo não teria função semântica a ligá-lo ao verbo. Mas são duas noções diferentes, que têm que ser mantidas cuidadosamente distintas; pode-se dizer que em muitos casos a função semântica (o papel temático, por exemplo) é deduzida pelo receptor a partir da função sintática e de marcas como as preposições. Por outro lado, há sempre uma função sintática (ou seja, a de tópico), ainda que não conste do elenco tradicional; e a partir dela, e da informação pragmática, é possível inferir funções semânticas como, no caso, o papel temático.

10.12. Conclusão

Concluo que a atribuição do papel temático aos diferentes complementos de uma sentença é um processo complexo. Embora seja em geral dependente dos traços de valência do verbo principal, difere da visão tradicional por duas características importantes, quando se aplica a sintagmas em função de tópico sentencial:

- (a) a codificação sintática que é parte da valência verbal (em particular, a presença de preposições) pode ser relaxada. Nesses casos, a falta de marca explícita causa, em princípio, o surgimento de ambigüidade;

¹³ Não tem função sintática concebida à maneira tradicional. Para mim, como vimos, o tópico tem uma função sintática, que se define por sua posição inicial na sentença.

(b) a ambigüidade é geralmente resolvida através do recurso a informação de ordem pragmática.

Um efeito desse mecanismo de atribuição de papel temático aos sintagmas tópicos é que a ambigüidade, embora resolvida no contexto, permanece latente, de modo que a mesma estrutura pode ser interpretada, em contextos diferentes, com papéis temáticos diferentes para o mesmo sintagma – situação que não ocorre em casos de atribuição a sintagmas não-tópicos, que são semanticamente unívocos desse ponto de vista.

Capítulo 11: Um catálogo de construções

11.1. O Catálogo das Estruturas do Português

Uma das etapas iniciais da pesquisa das diáteses da língua é a elaboração de uma lista de construções relevantes para sua formulação. Nessa lista, procura-se sistematizar a caracterização das diferentes construções; cada construção recebe uma descrição formal, uma descrição dos traços semânticos que a distinguem e, quando necessário, notas sobre eventuais problemas de análise. E, como vimos, algumas construções constituem **diáteses (verbais)**: um verbo pode aparecer em uma ou mais diáteses, e esse conjunto de diáteses constitui sua **valência**. Essa lista de diáteses é um fragmento do **Catálogo das Estruturas do Português**, estágio fundamental na descrição da língua. O Catálogo inclui, além das construções relevantes para a determinação das valências verbais, grande número de outras construções que ou são relevantes para a definição de valências não-verbais (de adjetivos, substantivos, advérbios), ou são construções oracionais não relevantes para a definição de valências (como a topicalizada). Além disso, claro, há a especificação das construções sub-oracionais, como o sintagma nominal, o sintagma adjetivo etc. Acrescente-se a tudo isso a distribuição e interpretação dos diversos adjuntos.

Goldberg (1995) formula assim a importância da lista de construções:

As construções são entendidas como as unidades básicas da língua. Os padrões sintagmáticos são considerados construções se algo de sua forma ou significado não é estritamente previsível a partir das propriedades de suas partes componentes ou de outras construções. Ou seja, postula-se uma construção na gramática se for possível mostrar que seu significado e/ou sua forma não deriva composicionalmente de outras construções existentes na língua. [...] As construções [são] as entidades da gramática que precisam ser listadas.

[Goldberg, 1995: 4]

Quando Goldberg especifica que uma construção não deve ser derivada composicionalmente de outras construções, deve estar se referindo apenas a construções de nível oracional – ou seja, um período composto não seria uma construção independente, mas seria derivado de seus componentes. Prefiro assumir uma posição mais “básica”, considerando construções inclusive os períodos compostos (= orações complexas), que aliás não entram no estudo relatado neste livro. Por outro lado, se levamos em conta construções sub-oracionais, é claro que muitas são composicionalmente derivadas, no sentido de que são montadas a partir de construções menores (SNs, SAdjs etc.).

A lista de construções deve ter como objetivo ser tão exaustiva quanto possível, para possibilitar a avaliação da extensão dos fenômenos estudados. Gross (1979), comentando a falta de preocupação com esse aspecto da análise na lingüística atual, observa:

[...] a determinação da extensão léxica das formas lingüísticas não tem “status” na G[ramática] G[erativa]. Quando essa determinação foi feita em francês [para casos de “raising” / MAP], encontramos três verbos do tipo de *to seem*, e mais

estatística não tem nada a ver com o problema formal. Mas ela mostra que o tipo *to seem* é muito limitado – de certo modo, estereotipado – ao passo que o outro tipo é produtivo, no sentido de que pode afetar novos verbos e novas construções de verbos.

[Gross, 1979: 867]

Pode-se acrescentar que, para efeitos práticos de descrição de uma língua, é certamente importante saber se uma construção vale para três verbos ou para seiscentos. Isso é parte da organização da língua, e tem repercussões práticas e teóricas evidentes.

11.2. Listas e regras

A lista, entretanto, não esgota a descrição das construções possíveis. No caso, mencionado por Goldberg, de construções previsíveis a partir de outras construções, é necessário formular essa previsão, o que pode ser feito, por exemplo, através de regras. Essas regras se entendem como relações sistemáticas entre estruturas reais, e não como transformações que as derivam a partir de formas abstratas. Alguns autores as chamam *via-rules*; vou chamá-las “regras de correlação”. Um exemplo são as frases com elemento topicalizado, que sempre se correlacionam com versões não topicalizadas:

[3] Meu tio comprou essa casa em 1974.

[4] Essa casa, meu tio comprou em 1974.

Aqui se verifica que tanto a forma quanto o significado de cada uma dessas frases é previsível a partir da outra, de modo que se sabemos que [3] é aceitável, automaticamente sabemos que existe uma outra sentença da forma [4] que é também aceitável. Esses casos são irrelevantes para a formulação de diáteses: a correlação entre [3] e [4] é propriedade da própria estrutura, e não de seus itens léxicos, em particular não do verbo *comprar*. Qualquer estrutura da forma [SN₁ V SN₂ ...] estabelece a existência na língua de uma estrutura da forma [SN₂, SN₁ V ...] – e vice-versa, porque não vejo razão estrutural para uma delas ser considerada mais “básica” do que a outra. O que nos leva a dizer que [4] é a topicalização de [3], e não que [3] é a destopicalização de [4], é apenas a tradição.¹

O papel das regras em conexão com o Catálogo merece comentário. Distingo basicamente dois tipos de regras, a que podemos chamar, informalmente, regras de correlação e regras de estrutura sintagmática.

11.2.1. Regras de correlação

As **regras de correlação** são as que descrevi acima, que explicitam acarretamentos de existência entre construções; quando categóricas, têm a forma seguinte:

Se um verbo ocorre na construção A, também ocorre na construção B.

Quando essa correlação não admite exceções, uma das construções pode ser omitida da subcategorização do verbo, pois é prevista pela própria regra. É o caso, visto acima, da topicalização: para qualquer verbo, dada a presença, digamos, de um SN objeto direto, prevê-se a existência de outra oração, onde esse SN ocorre no início da construção, com o mesmo papel temático; exemplos são [3] e [4]. Outros exemplos são orações negativas e interrogativas fechadas (*yes-no questions*), já que não há orações que não possam ser negadas ou interrogadas. Como vimos, essas regras são irrelevantes para nossos objetivos porque não têm papel na subcategorização dos verbos – embora, obviamente, precisem ser incluídas na gramática de algum modo.

Mas em muitos casos a correlação é parcial, e tem a forma

Os verbos *a,b, ... n* ocorrem na construção C e também na D.

sendo que outros verbos podem ocorrer apenas na construção C ou apenas na D. Aqui, naturalmente, a ocorrência em C ou em D deve ser incluída na valência dos verbos, já que não é possível prever uma ocorrência a partir da outra. Mas isso não quer dizer que não valha a pena formular a regra, porque ela pode representar um conhecimento probabilístico. Digamos que a maioria dos verbos em questão ocorram em C e D, e alguns poucos verbos só em C ou só em D. Nesse caso, pode-se dizer que a regra é parte do conhecimento da língua dominado pelos falantes.

Um exemplo é a correlação entre a construção ergativa e a transitiva; observa-se que a maioria dos verbos que ocorrem na ergativa também ocorrem na transitiva, mas há exceções, como *morrer, desmaiar e doer*. E é significativo que exista uma tendência a utilizar essas exceções também na construção transitiva, ou seja, transformar verbos ergativos em verbos transitivo-ergativos. Dei vários exemplos no capítulo 10, com os verbos *doer, funcionar, desabar, transbordar*. Uma interpretação natural desse fato seria que a regra de correlação que liga (para muitos verbos) a ocorrência na construção ergativa à ocorrência na transitiva seria psicologicamente real, e teria tendência a expandir sua aplicação – um exemplo do conhecido fenômeno da “pressão estrutural” (ou “analogia”), que resulta na eliminação de irregularidades e exceções.

Outro exemplo de correlação é ilustrado nas frases [5] e [6] abaixo:

[5] Paula e Armando brigaram.

[6] Paula brigou com Armando.

Note-se que as duas frases não são sinônimas (em [6] fica claro que a iniciativa foi de Paula, em [5] não há essa afirmação); mas há uma forte expectativa de que se um verbo ocorre na construção [5], deve também ocorrer na construção [6]. Isso vale para *brigar, encontrar, trombar, namorar, casar* etc. – mas não com *esbarrar* (que requer *em* em vez de *com*, e não acarreta que Paula tomou a iniciativa):

[7] Paula e Armando esbarraram.

[8] Paula esbarrou em Armando.

As regras de correlação formam uma rede de extrema complexidade na língua. Parte dessa rede foi estudada pelos lingüistas da linha gerativista (sob o rótulo de

“transformações”), mas a maior parte está por descrever.² O levantamento realizado por Gross (1975) para o francês fornece muitos exemplos, e mostra como é vasto o número de idiosincrasias léxicas no controle das correlações.

11.2.2. Regras de estrutura sintagmática

Além das regras de correlação, pode-se formular **regras de estrutura sintagmática**, que teriam a função de estipular as construções possíveis na língua, e que podem assumir a forma das “regras de base” (ou **PS rules**) da gramática gerativa.³

É evidente que a apreensão dessas regras tem interesse tanto prático quanto teórico. Prático, porque nos permite expressar a lista de maneira mais compacta, evitando a redundância inevitável em um rol de estruturas parcialmente idênticas. E teórico, porque sabemos que as pessoas aprendem e utilizam a língua evitando, quando possível, a memorização pura e simples de listas, e procurando generalizações.⁴

A elaboração das regras de estrutura sintagmática do português depende do levantamento das diáteses e outras construções, e deve progredir paralelamente a ele.

11.3. Pesquisando as regras por indução

No que pese o interesse das regras de correlação e de estrutura sintagmática, elas não devem ser formuladas desde o primeiro momento – antes, devem ser apreendidas indutivamente a partir de levantamentos muito amplos de verbos da língua. Começamos pelo estabelecimento de listas puras e simples de verbos devidamente associados a suas valências, da maneira mais completa possível. Em um segundo momento, e de posse de listas abrangentes, pode-se passar à tarefa de formular as regras de estrutura sintagmática e de correlação. Trata-se de examinar a lista completa das construções e construir um sistema de regras capaz de enumerá-las com exatidão. Como se vê, a tarefa pressupõe a elaboração da lista; e a lista, uma vez elaborada, funcionará como controle de qualidade do sistema de regras, permitindo verificar a que ponto ele é completo e exato, gerando e relacionando todas as construções existentes e apenas elas.

Em princípio, minha posição é semelhante à de Levin (1993), mas discordo dela no que diz respeito à condução da pesquisa. Levin organiza suas construções utilizando “alternâncias”, que são relações entre construções. Assim, ela apresenta a alternância “causativo-incoativa” (que para nós seria “transitivo-ergativa”), que inclui as frases

[9] Janet broke the cup. ‘Janet quebrou a xícara’

[10] The cup broke. ‘a xícara quebrou’

² No entanto, as regras de correlação se distinguem das transformações porque relacionam estruturas superficiais observáveis, ao passo que as transformações (em sua versão gerativista) se entendem como operações de derivação a partir de estruturas subjacentes abstratas. As regras de correlação se aproximam mais das transformações originais, propostas por Harris (1957).

³ Mas ficando claro que elas definem estruturas superficiais observáveis, não estruturas subjacentes como é o caso das PS rules.

⁴ É bom sempre repetir que a memorização de listas e de casos individuais é parte integrante e importante do conhecimento da língua. Lúcia Fulgêncio (2007) mostra um aspecto dessas listas, que, no caso são surpreendentemente amplas: limitando-se a expressões idiomáticas de uso corrente no português

Haveria, ao que parece, uma ligação especial entre essas duas estruturas, ou seja, em nossos termos, elas seriam relacionadas através de uma regra de correlação. Embora eu reconheça a eventual importância dessas correlações (como já explicitiei acima), insisto em que a pesquisa deve, de início, desprezá-las, justamente porque é importante que as correlações sejam estabelecidas com base em observações empíricas confiáveis.

Quanto às inegáveis semelhanças entre frases como [9] e [10], elas podem ser descritas sem referência a uma relação de correlação entre elas. Primeiro, é inevitável que se tenha que expressar na gramática a semelhança semântica entre essas duas frases: para citar as formas portuguesas, em [9] *a xícara é paciente*, assim como em [10]. Isso decorre inevitavelmente da interpretação que temos que atribuir a essas construções. Depois, teremos de qualquer maneira que dizer que o verbo *quebrar*, assim como uma longa lista de outros verbos (*assustar, desanimar, rasgar...*), pode ocorrer tanto na construção exemplificada em [9] (transitiva) quanto na que aparece em [10] (ergativa). É assim que se expressa o relacionamento entre [9] e [10]: quanto à semântica, são parcialmente idênticas, e há muitos verbos que ocorrem em ambas – mas não todos os verbos.

Temos, portanto, dois tipos de relação entre construções. Primeiro, temos o acarretamento de existência, expresso pelas regras de correlação: se um verbo ocorre em A, também ocorre em B. E, depois, temos semelhanças formais ou semânticas, que a meu ver não requerem expressão independente, além da fornecida pela descrição propriamente dita. Note-se que, pelo que se sabe, as semelhanças formais e semânticas não acarretam correlação, nem vice-versa: podemos, em princípio, ter correlação entre frases muito diferentes formal ou semanticamente, ou frases quase sinônimas mas não correlacionadas (por exemplo: *eu sou sobrinho do Emílio / o Emílio é meu tio*). O grau de coincidência entre correlação e semelhança formal e semântica é algo a ser investigado, não pressuposto. Se verificarmos que há uma forte tendência à semelhança semântica entre pares de frases correlatas, teremos um forte argumento em favor da hipótese de que as valências verbais são determinadas pelo significado dos verbos.

11.4. Relevância da lista para a descrição

A pesquisa da lista de construções gera uma série de subprodutos úteis para o equacionamento e solução de outros problemas de análise. Dentre esses subprodutos, saliento os seguintes:

11.4.1. Definição dos papéis temáticos

A investigação de diáteses pode dar subsídios à determinação dos papéis temáticos: ou seja, uma resposta à pergunta: “Quais são as relações semânticas (dentre as quase infinitas possíveis) que são gramaticalmente relevantes, devendo portanto receber o status de papéis temáticos?”

Para dar um exemplo: digamos que a dúvida seja entre três candidatos a papel temático: “agente volitivo”, definido como o causador voluntário de um evento; “agente não-volitivo”, definido como o “agente”, mas não voluntário; e “paciente”, definido como o elemento que muda de estado em virtude de um evento. Quantos papéis temáticos temos aqui: um, dois ou três? Em termos de conhecimento do mundo, é claro que são três, porque nunca confundimos algo que ocorre por ação de uma pessoa com algo que acontece por causa do vento. E tentamos não confundir, em geral, causas e

efeitos. A questão é se a língua respeita essa distinção através de uma distinção estrutural.

Podemos começar examinando a distinção entre “agente” e “paciente”, que é universalmente admitida, e por isso mesmo nunca justificada. Seria interessante ver por que ela nos parece tão inquestionavelmente relevante para efeitos gramaticais – talvez isso nos forneça alguns critérios úteis para casos mais problemáticos.

O que nos autoriza a separar esses dois papéis temáticos na descrição do português são fatos como os seguintes:

(a) “agente” e “paciente” se associam, de modo sistemático, a funções sintáticas próprias. Por exemplo, com um grande número de verbos o sujeito exprime o “agente” e o objeto direto exprime o “paciente”; note-se, aliás, que não há casos de objeto direto exprimindo o “agente”;

(b) existem pares de construções que são formalmente idênticas, e cuja diferença reside apenas em que uma tem o papel de “agente” atribuído a determinado termo, enquanto a outra tem o papel de “paciente” atribuído a esse mesmo termo. E cada uma dessas construções define uma subclasse diferente (e razoavelmente ampla) de verbos. Por exemplo, sabemos que a seqüência [SN V] pode ter o SN (sujeito) “agente” ou então “paciente”; o primeiro caso ocorre com os verbos *matar*, *comer*, *beliscar* etc. O segundo, com os verbos *encher*, *desanimar*, *esquentar*, e também *desmaiar*, *morrer* etc. Não há praticamente nenhuma sobreposição das duas subclasses.

Para descrever esses fatos precisamos, crucialmente, distinguir “agentes” de “pacientes”, o que nos fornece um argumento decisivo para considerá-los papéis temáticos distintos em português.

Podemos agora utilizar os mesmos testes para verificar se deve haver distinção entre “agente” volitivo e não-volitivo – uma questão que, ao contrário da anterior, é controversa. Ora, pelo que posso ver – e só o levantamento extenso poderá confirmar ou não essa impressão – não há associação sistemática de “agente” voluntário e “agente” não-voluntário a funções sintáticas diferenciadas. Temos, por exemplo, sujeitos voluntários e sujeitos não-voluntários, mas isso se condiciona diretamente à semântica do verbo ou dos sintagmas nominais envolvidos, e não à função sintática:

[11] Carlinhos bateu a janela. (voluntário)

[12] A ventania bateu a janela. (involuntário)

ou então

[13] Carlinhos reafirmou sua posição. (voluntário)

[14] * A ventania reafirmou sua posição.

O que está acontecendo aqui pode ser descrito em termos de adequação conceptual: sabemos que a ventania não pode tomar decisões voluntárias, nem afirmar coisa nenhuma. Não se trata de papéis temáticos distintos, porque estes têm uma semântica autônoma, não dependente do “bom senso” representado pela adequação conceptual.

A conclusão é que Agente é um papel temático único, cujo valor esquemático se elabora, segundo o contexto, como um elemento dotado de volição própria ou como um causador não-voluntário, inanimado. Em outras palavras, a oposição volitivo/não-volitivo não parece ser relevante para a descrição dos papéis temáticos da língua.

O exame da lista (uma vez que esta seja ampliada até se tornar razoavelmente abrangente) pode dar subsídios para responder certas perguntas fundamentais para a compreensão da estrutura da língua; em particular, as duas perguntas complementares:

- Que recursos formais possui o português para exprimir cada um dos papéis temáticos?
- Que papéis temáticos podem ser expressos por cada uma das funções sintáticas da língua?

11.4.2. Outras contribuições

Os traços semânticos utilizados para a definição das construções também fornecem indicações para a descrição semântica dos verbos envolvidos. Por exemplo, a construção denominada **de sentido recíproco obrigatório**, exemplificada por

[15] Helena está namorando Paulo.

denuncia um tipo particular de verbo (*namorar*, *esposar*) que impõe uma relação simétrica entre os dois argumentos, que conseqüentemente não podem ser rotulados como Agente e Paciente. Notar a diferença entre *namorar* e *paquerar*: o segundo envolve um Agente, o primeiro não, de modo que Helena pode paquerar Paulo sem que Paulo paquere Helena. Logo, *paquerar* não contém o ingrediente “sentido recíproco obrigatório”.

Outra possível contribuição do estudo das diáteses vem de que, para formular as construções, é preciso utilizar algum tipo de relação sintática. Digamos que se consiga fazer muitas formulações utilizando relações concretas como a ordem dos termos. Isso será um bom argumento em favor da hipótese de que as funções abstratas são desnecessárias.⁵

Tudo isso, naturalmente, depende da elaboração da lista das valências verbais: uma lista muito ampla de verbos, cada um deles relacionado com suas diáteses.⁶ E a elaboração da lista depende, por sua vez, do esclarecimento de bom número de questões teóricas e metodológicas preliminares, que são em grande parte objeto de discussão no presente livro.

⁵ Ver os capítulos 4 e 5 para uma colocação do problema das funções sintáticas

Capítulo 12: Notação e lista de diáteses

Neste capítulo sumario as convenções adotadas para notar as construções e apresento uma lista de construções que são consideradas diáteses, por subcategorizarem os verbos que ocorrem nelas. A lista, evidentemente, é parcial e provisória, pouco mais que uma amostra. Apesar de ter envidado grandes esforços no sentido de estabelecer critérios objetivos para a seleção dos constituintes a serem consignados nas diáteses, fui obrigado em diversos pontos a fazer decisões seguindo a tradição. Ou seja, as formulações têm um componente intuitivo (leia-se “arbitrário”). Essa situação, por mais lamentável que seja, é a meu ver inevitável no atual estágio da pesquisa. E não me parece aconselhável adiar o levantamento das valências verbais até que se disponha de critérios objetivos para todos os casos – mesmo porque o estabelecimento desses critérios depende em parte de uma visão geral do sistema, que vai se revelando à medida que se amplia o estudo a novas classes de verbos. Ou seja, a tarefa de definir as diáteses verbais do português deve progredir à medida que se forem estudando os verbos individuais; temos aqui material para anos de pesquisa.

Procurei organizar a lista de diáteses pela forma sintática, acrescentando exemplos e, quando for o caso, comentários. As construções são numeradas para facilitar a referência na lista de verbos a ser eventualmente elaborada. Assim, quando se fala da construção C8, trata-se da construção **de objeto de medida**, exemplificada por *Alceu pesava cem quilos*. A numeração não obedece a nenhuma ordem, e alguns números estão faltando – isso porque a lista é um instrumento de trabalho para mim e meus colaboradores, e achamos conveniente nunca mudar o número de uma diátese, mesmo quando alguma é eliminada da lista no decurso da pesquisa.

12.1. Convenções de notação

Na formulação das construções utilizei algumas convenções, arroladas e explicadas a seguir; como se verá, a maior parte são convenções comuns e bem conhecidas:

(a) Símbolos categoriais

SN = sintagma nominal

V = verbo

SAdj = sintagma adjetivo

SAdv = sintagma adverbial

SPrep = sintagma preposicionado (usado apenas nos nomes das construções)

Prep = preposição (quando possível, dou a preposição individual, por exemplo, *de SN*, *por SN* etc.)

X = sufixo de pessoa-número e / ou SN identificado como “sujeito” pela regra de identificação dada na seção 4.5.

Os símbolos categoriais são colocados, na formulação das diáteses, na ordem em que aparecem. A exceção é o símbolo ‘X’, que representa uma unidade que pode aparecer em diversas posições. O sufixo de pessoa-número ocorre sempre após o verbo

mas o SN sujeito pode ocorrer imediatamente antes do verbo ou depois dele. No entanto, como vimos no capítulo 4, esse SN é sempre identificado (na língua falada, que é a que nos interessa) sem margem de dúvida. Por outro lado, ‘X’ pode representar o sufixo, o sujeito ou ambos. Representa apenas o sufixo em *cheguei ontem*; apenas o sujeito em *Antônio roncar* [...], e representa ambos em *Antônio ronca*. ‘X’ se coloca, por convenção, antes do verbo.

(b) Papéis temáticos

αRef (usado para indicar a correferência de dois ou mais elementos)

Ag = Agente

Apresentando

Beneficiário

Caus.Exp. = Causador de experiência

Causa

Designação (um nome atribuído a algo)

Designando (entidade que recebe um nome)

Exp = Experienciador

Fonte

Instrumento

Localizando

Lugar

Medida

Meta

Opinador

Pac = Paciente

Ponto de vista

Possuído

Possuidor

PRS = Participante de relação social

Qualificando (entidade à qual é atribuída uma qualidade)

Qualidade

Tema (o elemento cuja mudança de lugar é expressa)

Zero = representa ausência de papel temático, com verbos avalentes como *ventar*.

(c) Outras convenções

∅ = associado a um papel temático, indica um argumento indeterminado, ou seja, um papel temático subentendido, mas não preenchido lexicamente; por exemplo,

X	V	∅
Ag		Pac

onde se entende que há um paciente indeterminado, como em *Pedro já comeu*.

Note-se que não utilizei, nas definições, as funções sintáticas tradicionais de objeto direto, predicativo etc. – embora esses termos apareçam nas denominações, para

na medida do possível, a ordenação dos sintagmas. A exceção, discutida em 4.2 – 4.5, é o sujeito, que deve ser reconhecido como um elemento excepcional. As diferentes posições do sujeito são totalmente previsíveis, e portanto irrelevantes para a definição das diáteses.

(d) Construções complexas

Finalmente, como se viu na seção 8.7.2, existem construções complexas, que denotam dois eventos simultâneos, paralelos e interdependentes. Cada um desses eventos determina uma relação entre complementos e papéis temáticos. O exemplo visto (que é um dos dois conhecidos) é a construção Transitiva de troca, que define a subclasse dos verbos de troca: *trocar, vender, comprar, barganhar, permutar*. Nesse caso, a definição especifica separadamente os dois eventos e suas respectivas estruturas temáticas, assim:

Construção Transitiva de troca

Definição:

	X	V	SN	<i>com/de</i> SN	<i>por</i> SN
Ev.1:	Fonte		Tema	Meta	
Ev.2:	Meta			Fonte	Tema

Um exemplo é

[3] Xavier vendeu um lote a Azevedo por uma boa quantia.

De posse dessas convenções, espero que o leitor não tenha dificuldade em interpretar as fórmulas utilizadas para notar as diáteses na lista abaixo.

12.2. A lista das diáteses do português

A lista que se segue compreende pouco mais de 70 diáteses, contando subtipos. Existem outras mais, talvez algumas centenas; e certamente muitas das que estão aqui precisarão ser reformuladas eventualmente.

Apesar de todas as incertezas que tive que expressar no decorrer deste livro, há também algumas certezas; e entre elas a principal é a crença no papel central do estudo de valências para a descrição da estrutura da oração em português. Por isso, gostaria de enfatizar uma vez mais a importância da descrição das valências verbais como um dos problemas principais que se colocam à frente dos estudiosos de gramática portuguesa nos próximos anos; e espero que as discussões e a lista de diáteses que apresento neste livro possam servir de orientação inicial para os que se dedicarem a essa tarefa descritiva, que é árdua e longa, mas crucial para o progresso do conhecimento em nossa área.

DIÁTESES DO PORTUGUÊS BRASILEIRO
(lista parcial)

Construções X V SN

C1. Transitiva

Zezé comeu a pizza

Definição :

X	V	SN
Ag		Pac

Com mudança de estado do Pac.

comer, fazer ... (muitos verbos)

Subtipo:

C1.2. de Contato

Zezé tocou a parede
A mãe acariciou o bebê

Sem mudança de estado do Pac.

Inclui: de objeto cognato (ele dormiu um sono tranqüilo); de contato (ele tocou a parede); de movimentação (Sílvia empurrou o armário) – neste caso, a movimentação é inferida pragmaticamente, mas pode não ter ocorrido; portanto, deve ser outro caso de contato.

lamber, coçar ...

C64. Transitiva de Sujeito Tema

A multidão abandonou o estádio.

Definição:

X	V	SN
Ag		Pac
Tema		Fonte

abandonar, deixar

C5. de Sujeito locado

As torradas encheram quatro caixas.
A multidão lotava o estádio.

Definição :

X	V	SN
Ag		Pac
Localizando		Lugar

Há variação do papel temático, talvez dependente do aspecto verbal: em a

Estou mantendo a mesma análise, supondo que essa diferença seja derivável do contexto.

encher, lotar, preencher, ocupar, habitar (*mas não morar*)

C39. de Apreciação

Todos aproveitaram a festa.
Carolina adora chocolate.
O quindim está me tentando.

Definição:

X	V	SN
Exp		CausExp

ver, ouvir, enxergar (*mas não olhar, cheirar, que têm sujeito Agente; cf. C56*)

C56. de Percepção

O cachorro cheirou a carne.
A menina olhou os bombons.

X	V	SN
Exp		CausExp
Ag		

Difere de C39 porque o sujeito é também Agente.

olhar, cheirar, escutar

C32. de Objeto transferido

Yeda pegou sarna.

Definição:

X	V	SN
Meta		Tema

pegar, contrair, adquirir

C33. de Objeto possuído

Minha prima tem dois carros.
O velho possuía mais da metade da cidade.

Definição:

X	V	SN
Possuidor		Possuído

ter, possuir

C6. Estativa de identificação

Vinte soldados compunham o pelotão
Beth é a rainha da Inglaterra.

Definição:

X	V	SN
α Ref		α Ref

α Ref é " termo de igualdade" (ou melhor, de co-referência).

ser, compor (*mas não estar*).

C7. de Transformação

Honório virou um lobisomem

Definição:

X	V	SN
Tema		Meta
α Ref		α Ref

C7 difere de C6 porque C7 exprime uma correferência incoativa.

Tradicionalmente, SN₂ é "predicativo"; mas é sempre um SN, nunca um SAdj, e não tem relação de concordância nominal com o sujeito.

Ficar às vezes ocorre nessa construção: ela ficou um monstro depois da plástica, mas geralmente exige SAdj: Honório ficou doido.

Caso muito próximo, mas com se obrigatório: tornar-se (+ SN ou SAdj).

O sujeito é também Paciente.

C8. de Objeto de medida

Ele pesava mais de cem quilos
Esse cachorrinho custa seiscentos reais.

Definição:

X	V	?SN
?Localizando		Medida

Às vezes o complemento do verbo não parece ser um SN: Ele pesava cada vez mais (notar que o complemento é exigido; ver também C49). O sujeito é "Theme" segundo Jackendoff (1972: 44).

Nesses casos, Allerton (1982) fala de "objoid", como uma maneira de dizer que não é propriamente um objeto direto (ver discussão em 5.1.6)

Andrews (1985, apud Dowty, 1991: 554) chama a Medida de extent, segundo Dowty "a little-studied thematic role". Segundo Dowty, teríamos adjunto em I walked a mile (v. C49), e complemento em this weighs five pounds; a razão é a possibilidade de omitir, que não se sustenta (ver cap. 6).

SNs de tempo parecem se comportar como adjuntos; cf. eu espirrei a prova inteira; o ano que vem vou a Curitiba.

pesar, medir, custar

Subtipo:

C8.2. de Duração

A conferência durou uma hora.

durar

C49. de Agente e medida

Eles andaram mais de dez quilômetros

A menina correu a escola toda

Definição:

X	V	?SN
Ag		Medida

Sobre o “?SN”, ver C8.

C10. de Relação social

Helena está namorando Paulo

Definição:

X	V	SN
PRS		PRS

“PRS”: “participante de relação social”.

namorar, esposar

Existe um sinônimo com SN composto: Helena e Paulo estão namorando, mas só funciona com namorar, não com esposar.

Subtipo:

C10.2. de Relação social assimétrica

Helena está paquerando Paulo.

Definição:

X	V	SN
PRS		PRS
Ag		

paquerar, cortejar (arcaico)

C11. de Sujeito tema

O ministro elegeram o vice

Definição:

X	V	SN
Tema		Meta
(Ag)		

O sujeito parece ser, em geral, Agente.

alcançar, atingir

C9. de Sujeito instrumental

Essa faca não vai cortar o filé.

O martelo amassou a lata.

Essa teoria vai resolver nossos problemas.

Definição:

X	V	SN
Instr		Pac

O sujeito não é Agente porque a faca não “causa” o evento, e é necessário subentender um Agente, como em Zé amassou a lata com o martelo (C28).

C43. Pseudo-ergativa

Ernesto cortou o cabelo.

Os cachorrinhos já nasceram dente.

Ela cresceu o cabelo até a cintura.

Definição:

SN	V	SN
Possuidor		Pac
		Possuído

O Possuidor pode também ser Agente, mas não obrigatoriamente. Aqui temos uma relação que não se relaciona (totalmente) com a valência do verbo, mas tem a ver também com a relação valencial entre dois Ns (Ernesto, possuidor de cabelo).

C59: de Sujeito Agente-Meta

O policial aceitou o dinheiro.

Definição:

X	V	SN
Meta		Tema
Ag		

aceitar, recusar, extorquir, recuperar, conseguir

C62: Transitiva de sujeito Meta

A menina ganhou um sorvete.

Definição:

X	V	SN
Meta		Tema

Difere de C59 por não ter sujeito Agente.

ganhar, receber, herdar, obter

Construções de X V

C2. Transitiva de objeto elíptico

Zezé comeu.

Essa menina lê até de noite.

Definição :

X	V	Ø
Ag		Pac

A notação “Ø Pac” é provisória. Este caso deve ser incluído nos casos gerais de indeterminação de argumento estudados pelo Moreira (2005). Os subtipos, entretanto, dependem da subclasse do verbo e devem ser incluídos em sua valência.

comer, ler ...

Subtipos (provavelmente dependentes da semântica do verbo ou de fatores pragmáticos):

C2.2. Pac indeterminado (“típico”)

Zezé comeu

C2.3. Pac privilegiado

Zezé cuspiu. (*isto é, “cuspe”*)

Zezé piscou. (*isto é, “o olho”*)

acenar, abanar, beber

C2.4. Pac correferente

Ela vestiu em cinco minutos.

A menina cortou toda com a gilete. [?]

C2.5. Pac recíproco

Eles cumprimentaram de longe.

Parece que fica melhor com interpretação recíproca. Aqui temos um sujeito Agente, ao contrário de C4, onde o Agente é indeterminado.

C3. Intransitiva

Zezé sorriu.

Zezé está viajando.

Definição :

X V

Ag

Distingue-se de C2 por não incluir um Paciente.

sorrir, viajar, pular ...

C29. Intransitiva de sujeito instrumental

Essa caneta não escreve.

Essa faca não corta.

Essa teoria não ajuda.

Definição:

X V Ø

Instr Pac

Possivelmente subtipo de C2 (ver também C9).

C4. Ergativa

Zezé engordou.

Definição :

X V

Pac

engordar, encher, assustar, decepcionar ...

C61: de Sujeito Meta

Ela recebe de quinze em quinze dias.

Definição:

X V

Meta

receber (*não ganhar, que aparentemente exige SN₂ ou um SAdv*)

<i>Construções de X V SN SPrep</i>

C30. Dativa

Carminha deu 200 reais a sua neta.

Carminha vendeu o carro para sua neta.

Definição:

X V SN

Ag

Prep SN

Beneficiário

*O objeto é Tema; provavelmente não é Paciente, porque não muda de estado. A preposição é **a** ou então **para**, que tende a substituir **a** no coloquial. Alguns analisam a Meta como Beneficiário (Wenceslau, 2003). É possível que os verdadeiros casos de Beneficiário tenham o SPrep como adjunto: Carminha fez um bolo de aniversário para sua neta. Note-se que é possível fazer um bolo sem ser “para” alguém, mas “vender” ou “dar” acarreta um destinatário (“meta”). Já se observou também (Pereira et al., 1993) que **a** só ocorre na dativa; com beneficiário só ocorre para. Mas isso pode ser tratado em termos de regência do verbo (dar + a/para, fazer + para). Os verbos que ocorrem com **a** parecem ser minoritários. Pelo menos, ?? Carminha vendeu um carro a sua neta me parece pouco natural.*

dar, vender, entregar, fornecer, emprestar, alugar, passar ...

C12. de Objeto fonte

Roberto roubou o Estado em mais de 100 milhões.

Definição:

X	V	SN	em SN
Ag		Fonte	Tema

roubar, lesar (*mas não furtar, tirar, surripiar*)

C52. de SPrep Tema

A faxineira limpou o teto de teias de aranha.

Vou esvaziar essa gaveta dos documentos.

O governo privou Zenóbio de sua pensão.

Definição:

X	V	SN	de SN
Ag		Pac	Tema

limpar, esvaziar, privar, espoliar

C13. de Objeto tema

Roberto roubou mais de 100 milhões do Estado

A faxineira limpou as teias de aranha do teto

Vou tirar os documentos dessa gaveta

Definição:

X	V	SN	de SN
Ag		Tema	Fonte

O complemento da preposição é opcional com alguns verbos (roubar, limpar), mas parece ser obrigatório com outros (tirar) – o que quer dizer que pode ser necessário postular outra construção

roubar, limpar, tirar, surripiar (*mas não lesar*)

C17. de SPrep meta

O governo transformou o país em um chiqueiro
 Meu pai converteu o quintal em um jardim

Definição:

X	V	SN	em SN
Ag		Pac	Meta

Nesta construção (ou em outra parecida) há a possibilidade de ocorrência separada de Fonte: Meu pai converteu o barracão de depósito em residência. Com reduzir, ocorre uma construção semelhante, mas com a preposição a: o governo reduziu o país a um chiqueiro.

transformar, converter, metamorfosear

C18. de SPrep paciente com objeto direto

O governo fez um chiqueiro do país
 A escola fez de mim um desiludido.

Definição:

X	V	SN	de SN
Ag		Meta	Pac

O compl. da preposição é também Fonte. A ordem inversa é muito comum: o governo fez do país um chiqueiro.

fazer (*aparentemente só esse verbo*)

C21. de SPrep recíproco

A cozinheira separou as claras das gemas.
 A cozinheira misturou as claras com as gemas.
 Ele confundiu Beth com um cachorro.
 Nós vamos ligar o motor ao gerador / no gerador.

Definição:

X	V	SN	PrepSN
Ag		Pac ₁	Pac ₂

Os dois pacientes em geral têm exatamente o mesmo papel. Além disso, nota-se que: (i) é possível trocá-los de lugar sem alteração de significado: a cozinheira separou as gemas das claras; (ii) a construção é sinônima de uma com objeto coordenado: a cozinheira separou as claras e as gemas. Isso nem sempre funciona perfeitamente: ele confundiu Beto com um cachorro não é provavelmente um sinônimo perfeito de ele confundiu um cachorro com

Beto, *o que sugere que há alguma diferença entre os dois papéis, ou então que as frases com confundir representam outra construção.*

Subtipos:

Quatro subtipos, segundo a preposição exigida pelo verbo: *de, com, a, em*

separar, misturar, juntar, confundir, ligar

C23. de Objeto tema e SPrep meta

Jair jogou o chapéu na cama.

Maria empurrou a porta contra o marido.

Sílvia chutou o chinelo para debaixo da cama.

Definição:

X	V	SN	Prep SN (~ Adv)
Ag		Tema (Pac?)	Meta

A preposição pode ser qualquer uma que exprima lugar aonde (= Meta): em, contra, (para) debaixo de, em cima de...

jogar, empurrar, chutar, atirar, arremessar ...

C28. de Agente, Paciente e Instrumento

Selma cortou o bilhete com a faca.

Definição:

X	V	SN	comSN
Ag	Pac	Instr	

cortar, rasgar, dobrar ...

C31. Anti-dativa

A escola premiou minha filha com uma bolsa.

Vou presentear Camila com um carro novo.

Descrição:

X	V	SN	com SN
Ag		Meta	Tema (Instr?)
Fonte			

premiar, presentear, recompensar, pagar, (?indenizar)

C37: de SPrep causa

O povo culpou o ministro pela crise.

X	V	SN	<i>por</i> SN
Ag		Pac	Causa

Caso muito raro de complemento com o papel de Causa.

aplaudir, maldizer, elogiar, perdoar, condenar

C67: de Posse inalienável com SPrep

O assaltante feriu Ernesto na perna.

Definição:

X	V	SN	<i>em</i> SN
Ag		Possuidor	Pac
			Possuído
			Lugar

Ver nota a C43.

C44: de Ação opinativa com *como*

Carvalho me vê como um especialista.

Todos te reconhecem como o melhor da turma.

Definição:

X	V	SN	<i>como</i> SN
Opinador		Caus.Exp	Qualidade

Como, apesar das gramáticas tradicionais, funciona como preposição.

ver, reconhecer, ter (literário)

C70: de Nomeação com SPrep

O diretor te designou como assessor direto.

Essa quadrilha me marcou como a próxima vítima.

Definição:

X	V	SN	<i>como</i> SN
Ag		Pac	Qualidade

designar, nomear, eleger, marcar

C71: de Designação

O professor vingou a menina de burra

A gente chamava o estilingue de “bodoque”.

Definição:

X	V	SN	<i>de</i> SN
Ag		Designando	Designação

Designando é a entidade que recebe uma Designação. É possível que o Designando seja um Paciente, em nível esquemático.

chamar, xingar, batizar, rotular, apelidar

<i>Construções de X V SPrep</i>

C14. de Objeto indireto

Marília gostava de alface.
 José desconfiava de Marília.
 José confiava em Marília.
 O menino (se) interessa por Química.
 Eu conto com você. [?]

Definição:

X	V	Prep SN
Exp		Caus.Exp.

Subtipos:

Quatro subtipos, segundo a preposição exigida pelo verbo: *de, por, com, em*

gostar, desconfiar, confiar, interessar, contar

C58: de Vitória [Bruno Lima 2006]

O Vasco ganhou do Botafogo.
 O menino mudou de roupa.
 Ela troca de marido todo ano.

Definição:

X	V	<i>de</i> SN
Ag		Pac

Fora ganhar, os verbos não são de vitória e derrota.

mudar, trocar, variar, ganhar

C15. de Tema e fonte

A gerente saiu da firma
 O cachorro fugiu de casa

Definição:

X V *de* SN
Tema Fonte

O Tema nem sempre é Agente: a fumaça sai desse cano.

sair, fugir

C16. de Tema e meta

O menino entrou no carro
Ricardo voltou para BH
Ênio foi para a Itália.
O cachorro correu para o quintal.
Sara foi na / à casa do namorado.

Definição:

X V *Prep* SN
Tema Meta

O Tema pode também ser Agente, mas nem sempre. A Meta pode ser expressa por Prep + um elemento adverbial: o cachorro correu para debaixo da cama/ para lá. em certos casos a preposição não ocorre: o menino entrou debaixo da cama / lá.

Uso metafórico evidente: essa estrada vai até Salvador; Mário chegou aos 60 anos.

Subtipos

Segundo a preposição exigida pelo verbo: *em, para, a*.
entrar, voltar, ir, correr

C19. Recíproca de SPrep

Amélia trombou com Vânia.
O asteróide chocou com Júpiter.

Definição:

X V *com* SN
Tema Meta
Meta Tema

Forma homônima: C20 (sujeito “agente”).

trombar, chocar

C20. de Movimento e choque

Amélia trombou com Vânia (propositalmente).

X V com SN
Ag Meta

Forma homônima: C19 (sujeito não é Agente).

trombar, ?esbarrar

C24. de Emissão odorífera

O armário está cheirando a naftalina.
Essa fruta sabe a mofo. [*arcaico*]

Definição:

X V a SN
??Tema ?Causador Exp.

A determinação dos papéis temáticos é um problema: a relação natural é pouco comum.

cheirar, feder, recender [*aparentemente só esses; arcaicos: saber, trescalar, soar*]

C34. de SPrep de ponto de vista

A fábrica caiu de produção.
Depois de 1880, Machado cresceu de reputação.
A gasolina aumentou de preço.

Definição:

X V de SN
Tema Ponto de vista

Talvez o sujeito seja (também?) Paciente, já que é afetado e muda de estado.

cair, crescer, diminuir, aumentar

C41. de Modo de proceder

Filipe procedeu como um cavalheiro.
A menina comportou de maneira inadequada / mal.

Definição:

X V como SN ~ SAdv
Ag Modo

O Modo precisa ser expresso: não há construção parecida sem como SN ~ SAdv.

proceder, comportar (*no padrão, e para muitos falantes, comportar-se*)

C45. Partitiva

Comi daquele bolo.

Definição:

X V *de* SN
 Ag Pac (partitivo)

O Paciente se entende como parte de um todo maior, expresso este pelo SN. Não é claro se deve ser um novo papel temático.

comer, tomar, beber (*mas não* mastigar, devorar, engolir)

C54. de Residência

Marcelo morou em Recife.

Definição:

X V *em* SN ~ SAdv
 Localizando Lugar

*O complemento pode ser representado por advérbios dêiticos: aqui, lá...
 Distingue-se de C25 porque não admite SAdj. Mas às vezes ocorre SAdv: ela mora muito bem.*

residir, ficar, habitar, viver

C55. de Aparecimento e desaparecimento

O menino sumiu com a minha agenda.

Definição:

X V *com* SN
 Ag Pac

Os mesmos verbos ocorrem em C4 (ergativa); C55 pode ser considerada uma variante supletiva da transitiva C1.

sumir, consumir, desaparecer; ?aparecer

C57. de Derrota [Bruno Lima 2007]

O Botafogo apanhou do Vasco.

O Botafogo perdeu do / para o Vasco.

Definição:

X V *de* SN
 Pac Ag

Com perder a preposição pode ser para: perdeu pro Vasco

apanhar, perder (*apenas esses*)

C63: de Relação social com SPrep

Helena casou com Paulo.

Definição:

X	V	com SN
PRS		PRS

casar, namorar, noivar

Subtipo:

C63.2:

Essa teoria não casa com os dados
Suas declarações não batem / não combinam / chocam com a realidade.

Será outro PT, ou uma extensão do PRS? Trata-se de uma relação recíproca, social ou não (?).

C47. de SPrep paciente sem objeto direto

Toninho acabou com a festa.

Definição:

X	V	com SN
Ag		Pac

Só é atestada com verbos incoativos: terminar, acabar, começar.

C48. de como SN meta

Marli acabou como presidente do clube.

Definição:

X	V	como SN
Tema		Lugar
		Meta

Relembro que como é preposição.

acabou, terminou

<i>Outras construções</i>

C25. Estativa

Clara está doente.

Clara é chata.

Raquel está aqui

O hotel fica na praça.
 A praia do Sobral é em Maceió.
 Ela está / é cheia de espinhas.
 Ela está / * é cheia de lasanha.

Definição:
 X V SAdj ~ SAdv
 Localizando (*)

(*) *O papel temático depende da semântica do constituinte: Qualidade; “situação” ou “lugar” (ou seja, provavelmente Lugar); mas em todos os casos se trata de uma característica atribuída ao sujeito*

ser, estar, continuar (*exprimem qualidade real*); parecer (*exprime qualidade aparente*)

A rigor, *ser* (qualidade inerente) é diferente de *estar*, *continuar* (qualidade accidental); mas talvez isso não deva aparecer na diátese (ver 9.5.1.3)

C26. Estativa de lugar

Raquel está aqui.
 O hotel fica na praça.

Definição:
 X V SAdv ~ Prep SN
 Localizando Lugar

Estou admitindo que o complemento de lugar é sempre representado por SAdv (aqui) ou Prep SN (em Niterói), e por isso considero as duas possibilidades em uma única construção. É possível que eu esteja errado nesse particular.

Possivelmente, subtipo de C25. Ver 9.5.

estar, continuar (*mas não ser*).

C35: de Ação opinativa

Todos acham Camila bonita.
 Considero Guga o maior tenista do mundo.

Definição:
 X V SN SAdj ~ SN
 Opinador Qualificando Qualidade

Opinador é o agente de uma ação mental. Talvez possa ser identificado com o Agente – cf. decidir, que teria sujeito agente; pensar etc.

Qualificando é a entidade à qual se atribui uma qualidade. Talvez, em nível mais esquemático, se identifique com o Localizando.

achar considerar incluir

C53: de Nomeação

O presidente nomeou Sara ministra da Educação.

Definição:

X	V	SN	SAdj ~ SN
Agente		Pac	Qualidade

nomear, eleger

C40. de Apresentação de existência

Tem quatro carros na garagem.

Definição:

V	SN
	Apresentando

ter (haver *só na escrita*)

Subtipo:

C40.2. de Apresentação de existência com concordância

Existem duas soluções para o problema.

Difere de C40 porque, em certos registros, há concordância do verbo com o SN que, ainda assim, normalmente aparece depois do V. A rigor, teria que ser definida independentemente, para caracterizar o sujeito:

X	V
Apres.	

Mas existir com concordância é bastante raro no coloquial.

C46. Locativa de meta

Ela acabou debaixo da mesa.

Ele ficou doido.

Definição:

X	V	Prep SN ~ SAdj
Qualificando		Qualidade ~ Lugar
Tema		Meta

*A alternância **Qualidade ~ Lugar** recorre em outros casos.*

*O sujeito é também **Paciente**, pelo menos em alguns casos.*

acabar, ficar, terminar

C66. Meteorológica

Ventou muito hoje

Vai chover.

Definição:

V

Zero

chover, ventar, relampejar, nevar

Chover *também ocorre na ergativa, raramente*: choveram pedras de 5 cm.

Construções complexas

C65. Transitiva de troca

Esaú trocou seu direito de primogenitura com Jacó por um prato de lentilhas.

Manuel comprou um carro de Helena por R\$ 10.000,00.

Definição:

	X	V	SN	PrepSN	porSN
Ev.1:	Fonte		Tema	Meta	
Ev.2:	Meta			Fonte	Tema

Entende-se que os dois eventos são simultâneos e interdependentes. A preposição do primeiro SPrep pode ser com, de ou a / para, dependendo do verbo.

Subtipos:

Três subtipos, segundo a preposição exigida pelo verbo: *de, com, a/para*.

comprar, vender, trocar (*este, raramente; mais comum em C22*)

C22. de Permutação

Eu troquei meu carro por um cavalo.

Definição:

	X	V	SN	por SN
	Ag			
Ev. 1:	Fonte		Tema	
Ev. 2:	Meta			Tema

trocar, permutar, substituir, barganhar (?)

A preposição normal é por. Permutar ocorre com por e com (segº Borba et al., 1990); Fernandes (1942) só consigna com.

Para facilitar a consulta, acrescento o índice das construções da lista:

13.2. A lista das diáteses do português

Construções X V SN

C1. Transitiva

Zezé comeu a pizza

C64. Transitiva de Sujeito Tema

A multidão abandonou o estádio.

C5. de Sujeito locado

As torradas encheram quatro caixas.

C39. de Apreciação

Todos aproveitaram a festa.

C56. de Percepção

O cachorro cheirou a carne.

C32. de Objeto transferido

Yeda pegou sarna.

C33. de Objeto possuído

Minha prima tem dois carros.

C6. Estativa de identificação

Vinte soldados compunham o pelotão

C7. de Transformação

Honório virou um lobisomem

C8. de Objeto de medida

Ele pesava mais de cem quilos

C49. de Agente e medida

Eles andaram mais de dez quilômetros

C10. de Relação social

Helena está namorando Paulo

C11. de Sujeito tema

O alpinista alcançou o pico

C9. de Sujeito instrumental

Essa faca não vai cortar o filé.

C43. Pseudo-ergativa

Ernesto cortou o cabelo.

C59: de Sujeito Agente-Meta

O policial aceitou o dinheiro.

C62: Transitiva de sujeito Meta

A menina ganhou um sorvete.

Construções de X V

C2. Transitiva de objeto elíptico

Zezé comeu.

C3. Intransitiva

Zezé sorriu / Zezé está viajando.

C29. Intransitiva de sujeito instrumental

Essa caneta não escreve.

C4. Ergativa

Zezé engordou.

C61: de Sujeito Meta

Ela recebe de quinze em quinze dias.

Construções de X V SN SPrep

C30. Dativa

C12. de Objeto fonte

Roberto roubou o Estado em mais de 100 milhões.

C52. de SPrep Tema

A faxineira limpou o teto de teias de aranha.

C13. de Objeto tema

Roberto roubou mais de 100 milhões do Estado

C17. de SPrep meta

O governo transformou o país em um chiqueiro

C18. de SPrep paciente

O governo fez um chiqueiro do país

C21. de SPrep recíproco

A cozinheira separou as claras das gemas.

C23. de Objeto tema e SPrep meta

Jair jogou o chapéu na cama.

C28. de Agente, Paciente e Instrumento)

Selma dobrou o bilhete com a régua.

C31. Anti-dativa

A escola premiou minha filha com uma bolsa.

C37: de SPrep causa

O povo culpou o ministro pela crise.

C67: de Posse inalienável com SPrep

O assaltante feriu Ernesto na perna.

C44: de Ação opinativa com como

Carvalho me vê como um especialista.

C70: de Nomeação com SPrep

O diretor te designou como assessor direto.

C71: de Designação

A gente chamava o estilingue de “bodoque”.

<i>Construções de X V SPrep</i>

C14. de Objeto indireto

Marília gostava de alface.

C58: de Vitória

O Vasco ganhou do Botafogo.

C15. de Tema e fonte

A gerente saiu da firma

C16. de Tema e meta

O menino entrou no carro

C19. Recíproca de SPrep

Amélia trombou com Vânia.

C20. de Movimento e choque

Amélia trombou com Vânia (propositalmente).

C24. de Emissão odorífera

O armário está cheirando a naftalina.

C34. de SPrep de ponto de vista

A fábrica caiu de produção.

C41. de Modo de proceder

Filipe procedeu como um cavalheiro.

C45. Partitiva [= Transitiva, com partitivo no OD?]

Comi daquele bolo.

C54. de Residência

Marcelo morou em Recife.

C55. de Aparecimento e desaparecimento

O menino sumiu com a minha agenda.

C57. de Derrota

O Botafogo apanhou do Vasco.

C63: de Relação social com SPrep

Helena casou com Paulo.

Outras construções**C25. Estativa**

Clara está doente.

C26. Estativa de lugar

Raquel está aqui.

C35: de Ação opinativa

Todos acham Camila bonita.

C53: de Nomeação

O presidente nomeou Sara ministra da Educação.

C44: de Ação opinativa com como

Carvalho me vê como um especialista.

C40. de Apresentação de existência

Tem quatro carros na garagem.

C46. Locativa de meta

Ela acabou debaixo da mesa.

C47. de SPrep paciente (sem objeto direto)

Toninho acabou com a festa.

C48. de comoSN meta

Marli acabou como presidente do clube.

Construções complexas**C65. Transitiva de troca**

Manuel comprou um carro de Helena por R\$ 10.000,00.

C22. de Permutação

Eu troquei meu carro por um cavalo.

Referências

- AGUILAR, Rafael Cano (1981) *Estructuras sintácticas transitivas en el español actual*. Madrid: Gredos.
- ALLERTON, D. J. (1982) *Valency and the English verb*. London: Academic Press.
- ALTMAN, C., HACKEROTT, M. e VIOTTI, E. (orgs) (2002) *Revista do GEL. Número Especial em Memória de Carlos Franchi*. São Paulo: Contexto.
- ANDERSON, John M. (1971) *The grammar of case: towards a localistic theory*. Cambridge: Cambridge University Press.
- ANDERSON, John M. (1977) *On case grammar: prolegomena to a theory of grammatical relations*. London: Croom Helm.
- ANDOR, József (2004) “The master and his performance: an interview with Noam Chomsky” *Intercultural Pragmatics*, 1, 1.
- APRESJAN, Y.D.; MEL’CUK, I.A. e ZOLKOVSKY, A.K. (1969) “Semantics and lexicography: towards a new type of unilingual dictionary”. In Kiefer, F. (org) *Studies in syntax and semantics*. Dordrecht: Reidel.
- BARLOW, Michael e FERGUSON, Charles A., orgs. (1988) *Agreement in Natural Language: Approaches, Theories, Description*. Stanford, CSLI.
- BASILIO, Margarida (2004) *Formação de palavras no português do Brasil*. São Paulo: Contexto.
- BERLIN, B. e KAY, Paul (1969) *Basic color terms: their universality and evolution*. Berkeley: University of California Press.
- BHAT, D. N. S. (1991) *Gramatical relations: The evidence against their necessity and universality*. London e New York: Routledge.
- BITTENCOURT, Vanda O. (1984) “Sobre as estruturas causativas sintéticas no português”. *Ensaio de Lingüística 11*, UFMG.
- BLOOMFIELD, Leonard (1933) *Language*. London: Geroge Allen & Unwin.
- BORBA, Francisco da Silva (2003) *Organização de dicionários: uma introdução à lexicografia*. São Paulo: Ed. UNESP.
- BORBA, Francisco da Silva et al. (1990) *Dicionário Gramatical de Verbos do Português Contemporâneo do Brasil*. São Paulo: Ed. UNESP.
- BOSQUE, Ignacio (1989) *Las categorías gramaticales: relaciones y diferencias*. Madrid: Síntesis.
- BYBEE, Joan L. e MODER, Carol Lynn (1983) “Morphological classes as natural categories” *Language*, 59, 2.
- BRENTARI, D., LARSON, G. e MACLEOD, L. orgs. (1988) *Parasession on Agreement in Grammatical Theory*. Chicago Linguistic Society.
- C&J – ver Culicover e Jackendoff (2005).
- CALVIN, William H. (2007) “A evolução do pensamento” *Scientific American Brasil*, 17.
- CAMACHO, Roberto G. (2003) “Em defesa da categoria de voz média no português”. *DELTA*, 19:1, p. 91-122.
- CAMARA Jr., J. Mattoso (1969) *Problemas de lingüística descritiva*. Petrópolis: Vozes.
- CAMARA Jr., J. Mattoso (1970) *Estrutura da língua portuguesa*. Petrópolis: Vozes.
- CANÇADO, Márcia (2002) “Uma aplicação da teoria generalizada dos papéis temáticos: Verbos psicológicos” *Revista do GEL*, número especial.
- CANÇADO, Márcia (2003) “Um estatuto teórico para os papéis temáticos”, in Müller et al. orgs. (2003)

- CASTELEIRO, João Malaca (1981) *Sintaxe transformacional do adjetivo*. Lisboa: Instituto Nacional de Investigação Científica.
- CHAFE, Wallace L. (1970) *Meaning and the structure of language*. Chicago: University of Chicago Press.
- CHAFE, Wallace L. (1976) “Givenness, contrastiveness, definiteness, subjects, topics and point of view” in Li, Charles (org.) *Subject and Topic*. New York: Academic Press.
- CHOMSKY, Noam (1957) *Syntactic Structures*. The Hague: Mouton.
- CHOMSKY, Noam (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- CHOMSKY, Noam (1977) *Dialogues avec Mitsou Ronat*. Paris: Flammarion.
- CHOMSKY, Noam A. (1982) *Lectures on Government and Binding*, Dordrecht, Holland: Foris.
- COLE, Peter e MORGAN, Jerry (1975) *Syntax and Semantics 3 - Speech Acts*, New York: Academic Press.
- CONTRERAS, Heles (1983) *El orden de palabras en español*. Madrid: Ed. Cátedra.
- CRAIK, Kenneth (1943) *The nature of explanation*. Cambridge: Cambridge University Press.
- CULICOVER, Peter W. e JACKENDOFF, Ray S. (2005) *Simpler syntax*. Oxford: Oxford University Press.
- CUNHA, Celso (1981) *Gramática de base*. Rio de Janeiro: FENAME.
- CUNHA, Celso e CINTRA, Lindley (1985) *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira.
- DIK, Simon C. (1980) *Studies in functional grammar*. London: Academic Press.
- DIXON, R. M.W. e AIKHENVALD, Alexandra Y. (2000) *Changing valency: Case studies in transitivity*. Cambridge University Press.
- DONNELLAN, K. (1966) “Reference and definite descriptions”. *Philosophical review*, 75.
- DOWTY, David (1989) “On the semantic content of the notion ‘thematic role’”. in Partee, B.; Chiechia, G.; e Turner, R. (orgs) *Properties, types and meanings*. Dordrecht: Kluwer.
- DOWTY, David (1991) “Thematic proto-roles and argument selection” *Language*, 67, 3.
- DUHEM, Pierre (1908) *Σωζειν τα φαινόμενα : Essai sur la notion de théorie physique de Platon à Galilée* [tradução inglesa: *To Save the Phenomena – an essay on the idea of physical theory from Plato to Galileo*. Chicago: The University of Chicago Press, 1971]
- FARBER, Pamela B. e USÓN, Ricardo M. (1999) *Constructing a Lexicon of English Verbs*. Berlin e New York: Mouton de Gruyter.
- FELLBAUM, Christiane (1992) resenha de Dixon, R.M.W., *A new approach to English grammar, on semantic principles*. *Language* 68, 3.
- FERNANDES, Francisco (1940) *Dicionário de verbos e regimes*. Porto Alegre: Globo.
- FERNANDES, Francisco (1950) *Dicionário de regimes de substantivos e adjetivos*. Porto Alegre: Globo.
- FILLMORE, Charles J. (1970) “The grammar of *hitting* and *breaking*” in Jacobs and Rosenbaum, orgs. *Readings in English Transformational Grammar*. Waltham, Mass: Ginn (1970).
- FILLMORE, Charles J. e KAY, Paul (1993) *Construction grammar*. Inédito. Berkeley: University of California

- FOX, B. e HOPPER, P. (orgs.) (1994) *Voice: Form and Function*. Amsterdam: John Benjamins.
- FRANCHI, Carlos (2003) “Teoria da adjunção: predicação e relações ‘temáticas’”. *Revista de estudos da linguagem*, v. 11, nº 2. UFMG.
- FRANCHI, Carlos e CANÇADO, Márcia (2003) “Teoria generalizada dos papéis temáticos”. *Revista de estudos da linguagem*, v. 11, nº 2. UFMG.
- FRIES, Charles C. (1952) *The structure of English*. New York: Harcour Brace.
- FULGÊNCIO, Lúcia (2007) *Expressões fixas e idiomatismos no português do Brasil*. Belo Horizonte: Tese de doutorado, PUC-Minas.
- GARDENT, Claire *et al.* (2005) “Extracting subcategorisation information from Maurice Gross’ grammar lexicon”. Texto de internet.
- GIORGI, Alessandra e LONGOBARDI, Giuseppe (1990) *The Syntax of Noun Phrases*, Cambridge: Cambridge University Press.
- GIVÓN, Talmy (1990) *Syntax: a functional-typological introduction*. Amsterdam: John Benjamins.
- GOLDBERG, Adele E. (1995) *Constructions: a construction grammar approach to argument structure*. Chicago: The University of Chicago Press.
- GOULD, Stephen J. (2000) *The lying stones of Marrakech*. New York: Three Rivers Press.
- GREENBAUM, Sydney (1969) *Studies in English adverbial usage*. London: Longmans.
- GRICE, H. Paul (1975) “Logic and conversation”, in Cole e Morgan, eds., 1975.
- GROSS, Maurice (1969) “Remarques sur la notion d’objet direct en français”, *Langue Française*, 1.
- GROSS, Maurice (1975) *Méthodes en syntaxe*. Paris: Hermann.
- GROSS, Maurice (1979) “On the failure of generative grammar”, *Language*, 55, 4.
- GRUBER, Jeffrey S. (1965) *Studies in Lexical Relations*. Tese de doutorado, MIT.
- HAEGEMAN, Liliane (1991) *Introduction to Government and Binding Theory*. Oxford: Blackwell.
- HARRIS, Zellig S. (1951) *Structural Linguistics*. Chicago: The University of Chicago Press.
- HARRIS, Zellig S. (1957) “Co-occurrence and transformation in linguistic structure”, *Language*, 33, 3.
- HAUY, Amini B. (1983) *Da necessidade de uma gramática-padrão da língua portuguesa*. São Paulo: Ática.
- HOCKETT, Charles F. (1968) *The State of the Art*. The Hague: Mouton.
- HUDDLESTON, Rodney (1984) *Introduction to the Grammar of English*. Cambridge University Press.
- HUDSON, Richard (1992) “So-called ‘double objects’ and grammatical relations”. *Language*, 68, 2.
- ICHIHASHI-NAKAYAMA, Kumiko (1994) Resenha de Bhat (1991). *Language*, 70, 1.
- ITKONEN, Esa (1974) *Linguistics and Metascience*. Kokemäki: Societas Philosophica et Phaenomenologica Finlandiae.
- JACKENDOFF, Ray S. (1972) *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- JACKENDOFF, Ray (1983) *Semantics and Cognition*. Cambridge, MA: MIT Press.
- JACKENDOFF, Ray S. (1990) *Semantic structures*. Cambridge, MA: MIT Press.
- JOHNSON-LAIRD, P. N. (1981) “Mental models in cognitive science”, in

- Norman, Donald A., org., *Perspectives in cognitive science*. Norwood, NJ: Ablex.
- KAC, Michael B. (1992) *Grammars and grammaticality*. Amsterdam: John Benjamins.
- KAY, Paul e MCDANIEL, C.K. (1978) "The linguistic significance of the meanings of basic color terms" *Language*, 54.
- KEENAN, Edward L (1976) "Toward a universal definition of 'subject'". In Li, (org) (1976).
- KEMMER, S. (1994) "Middle voice, transitivity and the elaboration of events" in Fox, B. e Hopper, P. (1994).
- KUHN, Thomas (1962) *The Structure of Scientific Revolutions*. Chicago: Univ. of Chicago Press [trad. port. *A Estrutura das Revoluções Científicas*. São Paulo: Objetiva, 1975].
- KURY, Adriano da Gama (1985) *Novas lições de análise sintática*. São Paulo: Ática.
- LAKOFF, George (1970) *Irregularity in Syntax*, New York: Holt, Rhinehart & Winston.
- LAKOFF, George (2004) *Don't think of an elephant! Know your values and frame the debate*. White River Junction, VT: Chelsea Green.
- LAKOFF, George e JOHNSON, Mark (2002) *Metáforas da vida cotidiana*. Campinas: EDUC/Mercado Letras.
- LAMBRECHT, Knud (1994) *Information structure and sentence form*. Cambridge: Cambridge University Press.
- LANGACKER, Ronald W. (1987) *Foundations of cognitive grammar. Vol I: Theoretical prerequisites*. Stanford, Stanford University Press.
- LANGACKER, Ronald W. (1988) "Autonomy, agreement, and cognitive grammar". In Brentari, Larson and MacLeod, orgs., 1988.
- LANGACKER, Ronald W. (1991a) *Concept, image, and symbol*. Berlin, Mouton de Gruyter.
- LANGACKER, Ronald W. (1991b) *Foundations of Cognitive Grammar - vol. II, Descriptive Application*. Stanford: Stanford University Press.
- LANGACKER, Ronald W. (1997) "Constituency, dependency, and conceptual grouping". *Cognitive Linguistics*, vol. 8-1.
- LANGACKER, Ronald W. (1999) *Grammar and Conceptualization*. Berlin and New York: Mouton de Gruyter.
- LAPOINTE, Steven G. (1980) *A Theory of Grammatical Agreement*. Amherst: Tese de doutorado, Univ. of Massachusetts.
- LAPOINTE, Steven G. (1988) "Toward a unified theory of agreement", in Barlow & Ferguson (orgs) (1988).
- LAUNEY, Michel (1992) *Introducción a la Lengua y a la Literatura Náhuatl*. México : UNAM.
- LEVIN, Beth (1993) *English Verb Classes and Alternations – a Preliminary Investigation*. Chicago: The University of Chicago Press.
- LEVINE, Robert D. (2002) Resenha de Uriagereka, J., *Rhyme and Reason: An Introduction to Minimalist Syntax*. *Language*, 78, 2.
- LEVY, Paulette (1983) *Las completivas objeto en español*. México: El Colegio de México.
- LI, Charles N., org. (1976) *Subject and topic*. New York: Academic Press.
- LIBERATO, Yara G. (1980) *Sobre a oposição dado/novo*. Belo Horizonte : dissertação de mestrado, UFMG.
- LIBERATO, Yara G. (1997) *A Estrutura do SN em Português: uma Abordagem Cognitiva*. Belo Horizonte: tese de doutorado. IIFMG

- LIMA, Bruno A. F. (2007) *Verbos de vitória e derrota*. Belo Horizonte: dissertação de mestrado, PUC-Minas.
- LONGACRE, Robert E. (1968) *Grammar Discovery Procedures*. The Hague: Mouton.
- LUFT, Celso Pedro (1999) *Dicionário prático de regência verbal*. São Paulo: Ática.
- LYONS, J. (1977) *Semantics*. Cambridge: Cambridge University Press.
- MARTIN, John W. (1975) “Concordância”. *Revista Brasileira de Linguística*, vol. 2, nº 2.
- MARTINET, André (1979) *Grammaire Fonctionnelle du Français*. Paris : CRÉDIF-Didier.]
- MAZUR, Luciana C. S. (2007) *Verbos incoativos no português falado: um estudo de suas valências*. Belo Horizonte: dissertação de mestrado, PUC-Minas.
- MATEUS, Maria Helena M. et al. (1983) *Gramática da Língua Portuguesa*, Coimbra: Almedina
- MINSKY, Marvin (1975) “A framework for representing knowledge”, in Winston (org.) (1975).
- MORAIS, Vinícius de (1956) “Se todos fossem iguais a você”. Jobim e V.M. Music Productions.
- MOREIRA, Élcio Costa (2005) *Indeterminação de agente no português do Brasil*. Diss. de Mestrado, PUC-Minas.
- MÜLLER, Ana Lúcia, NEGRÃO, Esmeralda e FOLTRAN, Maria José (orgs.) (2003) *Semântica Formal*. São Paulo: Contexto.
- NAPOLI, Donna Jo (1993) *Syntax: Theory and Problems*. Oxford University Press.
- OUHALLA, Jamal (1999) *Introducing Transformational Grammar: From Principles and Parameters to Minimalism*. New York: Oxford University Press.
- PAUL, Hermann (s/d [1884]) *Princípios fundamentais de história da língua*. Lisboa: Fundação C. Gulbenkian.
- PAWLEY, A. e SYDER, F. H. (1983) “Two puzzles for linguistic theory: Native-like selection and native-like fluency”. In Richards, J. e Schmidt, R. orgs., *Language and Communication*. London: Longman.
- PEREIRA, Maria Angela B., LEMLE, Miriam e MENEZES, Humberto P. (1993) “A noção de adjunto rediscutida”. *Anais do VII encontro nacional da ANPOLL*. Goiânia: ANPOLL.
- PERINI, Mário A. (1977) *Gramática do Infinitivo Português*. Petrópolis: Vozes.
- PERINI, Mário A. (1988) “Expressões Linguísticas e a noção de ‘esquema’”, *Cadernos de Estudos Linguísticos*, UNICAMP, vol. 14.
- PERINI, Mário A. (1989) *Sintaxe portuguesa: metodologia e funções*. São Paulo: Ática.
- PERINI, Mário A. (1995) *Gramática Descritiva do Português*. São Paulo: Ática.
- PERINI, Mário A. (1999) “Sobre o conceito de ‘item léxico’: uma proposta radical”. *Palavra* (PUC-RJ), pp. 140-163 [republicado in Albano, Eleonora et al., orgs., *Saudades da língua*. Campinas: Mercado Letras, p. 285-310 (2003)].
- PERINI, Mário A. (2004) “Categorização linguística: caminhos e becos na análise de um fenômeno complexo”. Conferência na PUC-Minas, 2004.
- PERINI, Mário A. (2005) “Ergativas e médias em português”. *Scripta*, 8, 16.
- PERINI, Mário A. (2007) “Por uma descrição gramatical mais concreta: sujeito e concordância no português do Brasil” Conferência de abertura do V Congresso da Associação Brasileira de Linguística, Belo Horizonte.
- PERINI, Mário A.; Fulgêncio, Lúcia; e Rehfeld, Maria Bernadete (1984) “Percepção linguística e fatiamento” *Letras de Hoje*, nº 58, pp. 45-84.
- PERINI, Mário A. et al. (1996) *O Sintagma Nominal em Português*:

- Estrutura, Significado e Função*. Número especial da *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG.
- PLANK, Frans (org.) (1983) *Objects – Towards a theory of grammatical relations*. London: Academic Press.
- PONTES, Eunice (1972) *Estrutura do verbo no português coloquial*. Petrópolis: Vozes.
- PONTES, Eunice (1973) *Verbos auxiliares em português*. Petrópolis: Vozes.
- PONTES, Eunice (1986) *Sujeito: da sintaxe ao discurso*. São Paulo: Ática.
- PONTES, Eunice (1987) *O tópico no português do Brasil*. Campinas: Pontes.
- POPPER, Karl (1959) *The Logic of Scientific Discovery*. Londres: Hutchinson [trad. port. *A Lógica da Pesquisa Científica*. São Paulo: Cultrix, 1975].
- PUSTEJOVSKY, James (1995) *The generative lexicon*. Cambridge, MA: MIT Press.
- RAVIN, Yael (1990) *Lexical semantics without thematic roles*. Oxford: Clarendon Press.
- RENZI, Lorenzo (org.) (1988) *Grande Grammatica Italiana di Consultazione*. Bologna: Il Mulino.
- ROBINSON, Orrin W. (1992) *Old English and its Closest Relatives – a Survey of the Earliest Germanic Languages*. Stanford University Press.
- ROJAS NIETO, Cecilia (1988) *Verbos locativos en español*. México: UNAM.
- ROSCH, Eleanor (1973) “Natural categories”. *Cognitive Psychology* 4:328-350.
- ROSCH, Eleanor (1975) “Cognitive representation of semantic categories” *Journal of Experimental Psychology*, 104.
- ROSCH, Eleanor (1978) “Principles of categorization”. In E.rosch e B.B. Lloyd (orgs), *Cognition and categorization*. Hillsdale: Lawrence Erlbaum.
- ROSS, John R. (1967) “On the cyclic nature of English pronominalization”. In *To Honor Roman Jakobson*. The Hague: Mouton.
- ROSS, John R. (1972) “Endstation Hauptwort: the category squish”, *Proceedings of the Chicago Linguistic Society*, 8.
- RUMELHART, David E. e ORTONY, Andrew (1976) “The representation of knowledge in memory”. La Jolla, CA: Center for Human Information Processing, University of California, San Diego.
- RŮŽIČKA, Rudolf (1977) “Three aspects of valence”. In Abraham, W. (org.) *Valence, semantic case and grammatical relations*. Amsterdam: Benjamins.
- SALVI, Giampaolo e VANELLI, Laura (1992) *Grammatica Essenziale di Riferimento della Lingua Italiana*. Istituto Geografico De Agostini.
- SANDERS, Gerald (1983) “Adverbials and objects”, in Plank (1983).
- SARAIVA, Maria Elizabeth F. (1997) *Buscar menino no colégio : a questão do objeto incorporado em português*. Campinas : Pontes.
- SAUSSURE, Ferdinand de (1916) *Cours de Linguistique Générale*. Paris : Payot. [Reedição de 1969]
- SCHANK, Roger e ABELSON, R. (1977) *Scripts, plans, goals, and understanding: an inquiry into human knowledge*. Hillsdale, NJ: Erlbaum.
- SCHŪTZE, Carson T. (1996) *The empirical base of linguistics: grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- SELLS, Peter (1985) *Lectures on contemporary syntactic theories*. Stanford, CA: CSLI.
- SILVA, Gláucia V. (2001) *Word Order in Brazilian Portuguese*. Berlin and New York: Mouton de Gruyter.
- STAROSTA, Stanley (1981) Resenha de Anderson (1977). *Language*, 57, 3.
- TESNIÈRE, Lucien (1959) *Eléments de syntaxe structurale*. Paris: Klincksieck.
- TAYLOR, John R. (1989) *Linguistic categorization: prototypes in linguistic theory*. Oxford: Clarendon Press

- TAYLOR, John R. (2002) *Cognitive grammar*. Oxford University Press.
- TOZZER, Alfred M. (1977) *A Maya grammar*. New York: Dover.
- TRASK, R. L. (1992) *A dictionary of grammatical terms in linguistics*. London: Routledge.
- TRASK, R. L. (1997) *A Student's Dictionary of Language and Linguistics*. London: Arnold.
- TROUBETZKOY, N.S. (1939) *Principes de phonologie*. Paris: Klincksieck. [Reedição de 1970]
- VARNHAGEN, Francisco A. de (1854) *História geral do Brasil*. São Paulo: Melhoramentos. [Reedição de 1956]
- VILELA, Mário (1992) *Gramática de valências: teoria e aplicação*. Coimbra: Almedina.
- VILELA, Mário e KOCH, Ingedore Villaça (2001) *Gramática da língua portuguesa*. Coimbra: Almedina.
- WALD, Benji (2001) Resenha de Farber e Usón (1999). *Language*, 77, 3.
- WELKER, Herbert (2005) “A valência verbal em três dicionários brasileiros”. *Linguagem e ensino*, Vol. 8, N° 1.
- WENCESLAU, Fábio de Lima (2003) *Verbos beneficiários: um estudo na interface entre semântica e sintaxe*. UFMG: Dissertação de Mestrado.
- WHITAKER-FRANCHI, Regina C. M. (1989) *As construções ergativas: um estudo sintático e semântico*. UNICAMP: Dissertação de Mestrado.
- WIERZBICKA, Anna (1996) *Semantics: primes and universals*. Oxford: Oxford University Press.
- WILLEMS, D. (1981) *Syntaxe, lexique et sémantique: les constructions verbales*. Gent: Uitgeverij Universa.
- WINSTON, P. (org) (1975) *The psychology of computer vision*. New York: McGraw-Hill.

Índice remissivo

- adjetivo 1.5.1
- adjunção 6.3.4
- adjunto
 - adnominal 6.6
 - definição tradicional 6.3
 - e complemento **cap. 6**
 - adverbial 5.4
- anafórico, contexto 3.4.5
- anomalia 2.3
- catálogo das estruturas do português 12.1
- categoria vazia 1.8.2
- categorização **cap. 3**
 - cruzada 3.3.1
 - dos verbos 10.1
 - e funções 3.4.4
 - múltipla 3.3.1
 - por objetivos 3.4.3
 - tradicional 3.3; 3.4.2
- classes, classificação – *ver* categorização
- clichês 3.3.1
- clítico 4.2.2
- cognitivistas
 - posição quanto à categorização 3.3.3
- complemento
 - conceituação provisória 6.8
 - definição tradicional 6.3
 - de medida 5.1.6
 - e adjunto **cap. 6**
 - e papel temático 6.7
 - especificado semanticamente? 9.5.7
 - nominal 6.6
 - opcionalidade 9.5.6
 - papel na formulação das diáteses 6.2; 6.3.5
 - reflexos na sintaxe 6.5
- concordância verbal 4.2.1
- constituente 1.1.
 - abstrato 1.8.2
- construção 4.5; **cap. 9**
 - catálogo **cap. 12**
 - complexa 9.7.2
 - de alçamento 5.1.6
 - definição 4.5; 5.1.1; 9.3
 - e o significado do verbo 9.3.2
 - ergativa 4.1.7; **cap. 10**
 - e pronome pessoal 10.5.1.2
 - e sujeito posposto 10.5

- intransitiva 9.3.2
- média 10.8
- não-diatéticas 9.5.1
- notação **cap. 9**
- notação de Goldberg 7.5
- papel na descrição 9.2
- passiva 4.2.4; 5.1.6; 9.6.3
- pseudo-ergativa 10.7
- transitiva 4.1.7; **cap. 10**
- conteúdo semântico
 - dos complementos 9.5.1.4.
 - do verbo 9.5.1.5.
- contexto extra-lingüístico
 - uso na interpretação 1.2
- córpus 1.5.1
- critério-teta 2.4
- descrição
 - e realidade psicológica 1.8.1
 - e taxonomia 3.6
 - e teoria 1.5; 1.6; 2.1; 2.2
 - guias 1.8
 - necessidade 2.2.1
 - postulados 2.4
 - superficial 2.4
- detalhamento 7.4
- diátese **cap. 9**
 - e construções 4.1.7; 4.2.4
 - lista **cap. 13**
 - passiva? 9.6.3
 - relações entre diáteses 10.6.3
- espanhol 1.3
- esquemas 5.8.; 8.6.3
- estrutura superficial 2.4
- estruturalistas
 - posição quanto à categorização 3.3.2
- expressões idiomáticas 9.5.2
- “fechamento” 1.8.2.2; 4.1.8
- forma
 - e conceito 1.1; 1.4
 - traço de 2.4.2
 - como decorrência do significado 3.2
- francês 1.4.2
- funções **cap. 4; cap. 5**
 - e classes 3.4.4
 - e definição das construções 4.1
 - identificação 4.1
 - papel na gramática 4.1.3; 5.7
- gerativistas
 - posição quanto à categorização 3.3.2
- gramática

- e léxico 2.3.2
- inglês 1.4.2
- inversão (sujeito-verbo) 9.5.8
- latim 1.4.2
- lexema 1.4
 - e palavra 1.4.2
 - classificação 3.4.1
- léxico
 - e gramática 2.3.2.
- lingüística
 - como “história natural” 1.5.1; 1.6
 - de cópula 1.6
 - objeto 1.3
 - “terreno conquistado” 1.7
- listas
 - e regras 12.2
- memória
 - semântica 1.2
 - episódica 1.2
- metáforas
 - e a definição das construções 9.5.1.
- metodologia 1.6
 - bases 1.7
 - posição de Gross 7.2
- morfema
 - classificação 3.4.1
- náhuatl 5.1.1
- objeto direto 5.1
 - e papéis temáticos 5.1.2
 - elíptico 10.1.2
 - segundo a gramática gerativa 5.1
 - posições na oração 5.1.3
- objeto indireto 5.5
- objóide 5.1.6
- palavra
 - e lexema 1.4.2; 3.4.1
 - classificação 3.4.1
- papéis temáticos **cap. 8**
 - Agente 1.8.2; 5.1.5.1; 8.6.2
 - atribuição pragmática **cap. 11**
 - Beneficiário (=Meta) 8.6.2.5
 - caráter esquemático 8.3.4; 8.6.3
 - conceituação 8.1; 8.3
 - critérios de distinção 8.3.2; 8.6
 - critério de necessidade 8.6.1; 8.6.2
 - critério de semelhança semântica 8.6.3
 - cumulativos 8.7.1
 - definidos transformacionalmente 8.6.5
 - diferentes das RCTs 8.3.1
 - dos adjuntos 8.1 (n)

- Lugar 8.3.1
- Meta 8.6.2.5.
- na definição das construções 8.2; 13.1
- Paciente 5.1.5.1; 8.6.2
- representação sintática 8.5.2
- utilidade descritiva 8.2
- vinculados a nominais 8.2
- paradigma (em ciência) 1.6
- particípio
 - verbal 9.6.2
 - nominal 9.6.2
- polissemia induzida 10.6.4
- potencial funcional 3.4.1
- potencial referencial 3.4.1
- predicativo 5.6
- pronome 1.5.1
- protótipos 3.7; 8.4
 - como contínuos 8.4
 - e a descrição gramatical 8.4.2
- regras 2.3.2
 - calibragem 3.5
 - de correlação 12.2.1
 - de estrutura sintagmática 12.2.2.
 - disponibilidade simultânea 5.8
 - e listas 12.2
 - funcionamento e interação 5.8
 - ordenação 5.8
 - tipologia 2.5
 - uso no desempenho 5.8
- regularidade 2.3.1
- relações conceptuais temáticas (RCT) 8.1
 - como elaboração dos papéis temáticos 8.3.3
- relatório descritivo 1.5.1
- restrições seletivas
 - na definição das construções 9.5.1
- russo 1.3
- seleção de evidência 1.5.1; 1.6
- significado
 - traço de 2.4.2.
- simbolização
 - traço de 2.4.2
- símbolos categoriais
 - usados na formulação das construções 13.1
- sintaxe
 - simples (hipótese da) 2.4.3
- SN (sintagma nominal)
 - semântica 4.1.8
- SPrep (sintagma preposicionado) 9.5.9
- sujeito **cap. 4**
 - ambigüidade 10.1.2

- ausência 4.2.4
- definição 4.2.4
- do gerúndio 4.1.4
- identificação 4.1.5
- na definição das construções 4.5
- oculto 1.8.2
- posposto 4.2.3; 9.5.8
- taxonomia 1.5.1; 3.1
 - e descrição gramatical 3.6
 - e protótipos 3.7
- terminologia
 - das classes verbais 10.9
 - variação excessiva 1.8.1
- topicalização 1.8.2
- tópico 10.2
 - codificação sintática 11.5
- transformação 2.4
- valência **cap. 9**
 - abordagens tradicionais 7.1
 - e classificação 10.1
 - e previsibilidade 8.5
 - e significado 10.6
 - e tópicos **cap. 11**
 - do verbo ou do predicado? 9.5.4
 - proposta de Allerton 7.3
 - proposta de Levin 7.4
- verbo
 - classificação 10.1
 - como classe 1.5.1
 - critérios de contagem 9.5.2
 - ergativo 10.2; 10.5
 - monovalente 5.1.4; 9.5.2
 - sem sujeito 9.5.6.2
 - sem valência 9.6
 - subcategorização 9.4.1
 - transitivo 10.2
 - transitivo-ergativo 10.2; 10.6.2