

# **Um “erro” demasiadamente humano? A seleção dos adjetivos na predicação secundária orientada para o sujeito**

An “all-too-human” error? The selection of adjectives  
in subject-oriented secondary predicate

Ronaldo Teixeira Martins  
UNIVÁS

## **Resumo**

O emprego adjetivo de advérbios e o emprego adverbial de adjetivos constituem dois dos principais problemas de concordância nominal ainda não resolvidos pelo ReGra, principal sistema de revisão gramatical automática disponível para o português do Brasil. O problema, aparentemente simples, se complica porque não se pôde perceber, nas ocorrências estudadas, uma regularidade subjacente que permitisse à ferramenta verificar quando as formas adjetivas podem funcionar como advérbio ou quando as formas adverbiais podem funcionar como adjetivo. A procura dessa regularidade foi o motor deste texto, e seu objetivo esteve vinculado à formalização de um conjunto de princípios que viesse a impedir a realização de transformações dessa natureza sempre que estivessem em desacordo com o que prescreve a disciplina gramatical da língua portuguesa.

## **Palavras-Chave**

Revisão gramatical e de estilo automática, Predicado secundário, Adjetivo, Advérbio.

## Abstract

The use of adjectives as adverbs and of adverbs as adjectives constitute two of the main unresolved problems of noun agreement in the scope of ReGra, the most important grammar and style checker for Brazilian Portuguese. The problem, apparently simple, turns out to be intricate because it was not possible to find an underlying regularity that would enable the tool to check when adjectives may function as adverbs, and adverbs as adjectives. This text is a result of this quest, and our goal was to formalize a set of principles that would prevent such changes whenever they were not in accordance with the standard Portuguese grammar.

## Keywords

Grammar and style checking, Secondary predicate, Adjective, Adverb.

## Introdução

O emprego adjetivo de advérbios e o emprego adverbial de adjetivos constituem dois dos principais problemas de concordância nominal ainda não resolvidos pelo ReGra, ferramenta computacional de revisão gramatical automática que vem sendo desenvolvida, desde 1993, pelo Núcleo Interinstitucional de Linguística Computacional (NILC). O problema, aparentemente simples, se complica porque não se pôde perceber, nas ocorrências estudadas, uma regularidade subjacente que permitisse à ferramenta verificar quando as formas adjetivas podem funcionar como advérbio ou quando as formas adverbiais podem funcionar como adjetivo. A procura dessa regularidade foi o motor deste texto, e seu objetivo esteve vinculado à formalização de um conjunto de princípios que viesse a impedir a realização de transformações dessa natureza sempre que estivessem em desacordo com o que prescreve a disciplina gramatical da língua portuguesa. Este texto, relatório do percurso para a consecução deste objetivo, está assim subdividido: na primeira seção exponho o cenário do projeto ReGra, para que se possa entender por que o problema constitui efetivamente um problema; na segunda seção, o problema é tratado de forma mais detalhada, para que se possa compreender por que as intervenções até agora bem-sucedidas se devem antes a regras pontuais, de alcance muito limitado, especializadas no contexto de ocorrência de alguns homônimos muito comuns (como as formas *muito*, *pouco* e *bastante*, entre outras); a terceira seção explora alguns dos pressupostos metodológicos que, por força do contexto, orientaram a procura de alternativas teóricas para a solução do problema; a quarta seção apresenta a hipótese teórica que se revelou mais pertinente para o tratamento das fronteiras entre adjetivos e advérbios; e a quinta e última seção discute algumas das implicações resultantes da adoção da hipótese delineada.

## 1. Do contexto

O objetivo principal do Projeto ReGra<sup>1</sup> é o desenvolvimento de uma ferramenta computacional para revisão gramatical automática de textos referenciais<sup>2</sup> produzidos, em português do Brasil, por usuários de nível de escolaridade médio. O Projeto vem sendo desenvolvido, desde 1993, no Núcleo Interinstitucional de Linguística Computacional (NILC), que vincula pesquisadores ligados à USP / São Carlos, à UNESP / Araraquara e à UFSCar. Existem várias versões comerciais disponíveis da ferramenta. A versão mais recente pode ser encontrada no conjunto de ferramentas de auxílio à escrita que integra o Microsoft Office 2010.

Do ponto de vista das estratégias de processamento, o Projeto ReGra conheceu pelo menos duas fases muito distintas. Na primeira fase, predominava uma abordagem direcionada ao erro, em que se procuravam sequências predeterminadas que constituiriam desvio em relação à norma padrão da língua portuguesa. Nessa etapa, expressões fixas (como “a nível de”, “cujo o”, “apesar que”, e muitas outras) ou padrões sintáticos determinados (como a ocorrência de crase diante de palavras masculinas) eram procurados no texto do usuário e, quando encontrados, ativavam as mensagens de erro e as sugestões incorporadas à ferramenta. Não havia, então, processamento propriamente linguístico. A sentença, definida como cadeia de caracteres, era percorrida linearmente e processada como uma rede de transição simples. Como consequência, o escopo de atuação da ferramenta era muito restrito: problemas envolvendo itens lexicais não contíguos não eram abordados.

Para ampliar o espectro de problemas gramaticais tratados pela ferramenta, a segunda fase do Projeto fez uma opção diversa, introduzindo uma análise linguística mais refinada. Preservou-se a abordagem pontual no que ela tinha de mais eficiente: a localização e a correção (rápida) de erros padronizados; e acrescentou-se a preocupação com o desenvolvimento de estratégias formalizáveis que permitissem proceder ao diagnóstico correto de outros tipos de problemas de expressão. No entanto, sem a possibilidade de replicar os processos naturais de atribuição de sentido a um texto, acabou-se por privilegiar o cotexto linguístico e dele retirar todas as instruções necessárias à revisão das sentenças do usuário. Sobrelevou-se, assim, a análise sintática automática, voltada para a identificação das funções sintáticas de cada uma das palavras da oração. A disponibilidade dessas informações tornou possível a elaboração de regras genéricas, não comprometidas com o tamanho ou com o formato da sentença.

A implementação linguística e computacional dessa nova abordagem tem oferecido, porém, problemas incontáveis. Saliente-se, em primeiro lugar, que não se pode contar aqui com muitos dos dispositivos normalmente empregados por outras ferramentas de processamento automático para recuperar a estrutura sintática da sentença. Não podemos confiar, por exemplo, nas relações de concordância e regência estabelecidas pelo usuário, porque elas constituem, em última análise, o objeto mesmo de nossa revisão. Pressupor que, em uma sentença do usuário, o sujeito esteja concordando com o verbo é esvaziar de sentido a tarefa da correção, porque a concordância verbal representa exatamente uma das questões que nos propomos a verificar e, se necessário, corrigir. Os mecanismos sintáticos disponíveis para revisão gramatical são, portanto, extremamente restritos. O único critério efetivamente inventariável é a ordem das palavras na sentença, que não consubstancia um procedimento infalível para a definição de estruturas de constituintes.

Outra dificuldade concerne à estrutura postulada para o dicionário que constitui a base de dados a partir da qual opera o analisador sintático. Como estamos vinculados à análise e à correção de textos reais, efetivamente produzidos por usuários da língua, desde o início se impôs o tratamento (robusto) de todo o repertório das palavras do português, em lugar da elaboração de léxicos restritos a domínios específicos ou a universos de discurso muito particulares. Vetada, portanto, a simplificação necessária ao tratamento exaustivo de apenas algumas das lexias da língua, os consequentes problemas operacionais fizeram com que o dicionário utilizado não contivesse senão as classes gramaticais dos verbetes e algumas informações de natureza flexional, tradicionalmente repertoriadas pela tradição lexicográfica do português.<sup>3</sup> A única informação sintática disponível concerne à transitividade das formas verbais, frequentemente sobrecategorizada em função das extensões normalmente registradas pelos dicionaristas da língua, que contemplam mesmo formas raras e arcaicas da estrutura predicamental dos verbos.<sup>4</sup>

Com a ausência de outras categorias linguísticas e, principalmente, de informação semântica no léxico, sobrevalorizou-se o papel da sintaxe. Embora a centralidade da componente sintática já tenha sido suficientemente relativizada em formalismos gramaticais recentes, que já não ousam contar apenas com a hierarquização da informação morfossintática para estabelecer as relações intrassentenciais, não nos restou saída senão utilizar um modelo bastante simples, que pouco aposta na riqueza da informação lexical. Foi formulada, então,

uma gramática livre de contexto, com aproximadamente 600 regras de reescrita categorial, cujos símbolos não terminais correspondem às categorias previstas pela Nomenclatura Gramatical Brasileira (NGB), e cujos símbolos terminais correspondem exatamente às (poucas) categorias presentes no dicionário. Sobre essa gramática opera um *parser* descendente, recursivo, que retorna como resultado a primeira análise tornada possível pela aplicação das regras gramaticais.<sup>5</sup> A incorporação desses resultados ao revisor gramatical promoveu considerável redução do número de não-intervenções indevidas (falsos negativos), com a identificação de problemas de concordância e regência antes não detectados. No entanto, passou a ser também observado um número perturbador de intervenções indevidas (falsos positivos) no texto do usuário, fruto de casamentos malsucedidos durante o processamento sintático automático.

Para reduzir o problema, decidiu-se pela atribuição de prioridades às regras gramaticais e pela implementação de um algoritmo de *parsing* limitado a frases nucleares. Para efeito de processamento, e em que pese a simplificação aí empreendida, frases não nucleares foram reduzidas a frases nucleares pela transformação de conjunções e de algumas fronteiras lexicais (como a contiguidade de formas finitas dos verbos) em fronteiras oracionais. A ordenação do conjunto de possibilidades sintáticas da língua portuguesa acompanhou a frequência de uso – determinada em análise manual de um fragmento do *corpus* –, para que se pudesse contrapor, à pouca confiabilidade da concordância e da regência, um critério estatístico capaz de comprovar as relações de dependência que se observam dentro da sentença.<sup>6</sup>

A adoção dessas estratégias, embora tenha efetivamente reduzido o número de erros da análise sintática automática, não foi capaz de evitar o problema da ambiguidade lexical, principal fonte de ruído durante o processamento automático da sentença. Para tentar contornar esse problema, decidiu-se que também o dicionário que serve de apoio ao ReGra deveria conter uma ordenação da frequência de ocorrência das partes do discurso para cada forma ambígua. Embora a iniciativa tenha tido o mérito de isolar os arcaísmos, os regionalismos e outros casos de uso localizado (e bastante incomum) das palavras, é necessário reconhecer que, na maior parte dos casos, a frequência de uso não funciona como critério razoável de desambiguação lexical. Outros mecanismos de desambiguação lexical revelaram-se necessários, e essa necessidade conduziu à elaboração de um conjunto de regras que permitisse calcular a classe da palavra a partir do seu

contexto de ocorrência. Trata-se, pois, de uma tentativa de retirar do dicionário parte da responsabilidade pela decisão, deixando-a também a cargo da distribuição da palavra na sentença, com a consideração dos contextos mínimos, à esquerda e à direita, junto aos quais ela aparece.

No entanto, a combinação das duas estratégias de desambiguação [a frequência do uso e o contexto de ocorrência], ainda que tenha permitido eliminar alguma parte da ambiguidade lexical, mostrou-se também insatisfatória. De um lado, porque as duas estratégias supõem correto o texto que se deseja corrigir, e nesse sentido viciam a análise da ferramenta. De outro, porque a aplicação em textos reais fez perceber que, não apenas os itens lexicais mas também os contextos (formados, obviamente, por outros itens lexicais) são predominantemente ambíguos. E a análise combinatória que nos permitiria selecionar uma entre as várias possibilidades categoriais de uma dada palavra provou ser capaz apenas de estender a ambiguidade de itens lexicais vizinhos para o sintagma por eles formado. A ambiguidade, portanto, persistiria.

A resolução da ambiguidade categorial provou, pois, exigir estratégias de ação bem mais refinadas do que a ordenação pela frequência de uso ou a determinação dos contextos de ocorrência. Dessa necessidade de uma representação mais rica e estruturada do conteúdo lexical emergiu a proposta desta pesquisa, e o seu objetivo geral foi exatamente a busca de alternativas disponíveis, teoricamente consistentes, formalizáveis, para uma organização minimamente ambivalente do repertório de itens lexicais.

## 2. Do problema

O problema da ambiguidade categorial se revela de forma particularmente grave porque, envolvidos na construção de uma ferramenta de revisão gramatical, não podemos nos fiar (completamente) no texto que serve como dado de entrada. Nosso ponto de partida é um texto normalmente precário, com a incidência de lapsos de língua, erros de grafia e problemas de digitação, em que não apenas as relações de concordância e regência são pouco confiáveis, mas a própria marcação de fronteiras sentenciais deve estar sujeita à revisão. Nesse contexto, inóspito ao processamento automático, a extração das categorias da sentença deve ser feita de forma extremamente cautelosa, porque, no desenvolvimento do texto, o usuário humano explicita – principalmente por meio da sintaxe de concordância,

de regência e de colocação – hipóteses a respeito da categoria das palavras da língua que nem sempre coincidem com o que prescrevem os gramáticos e os lexicógrafos. Frequentemente encontraremos, em textos reais, verbos transitivos indiretos sendo tratados como verbos transitivos diretos (*A menina obedeceu o pai*), pronomes pessoais do caso reto funcionando como oblíquos (*Ele viu ela ontem*), ou pronomes relativos empregados em discordância com a norma-padrão da língua (*A ideia onde todo mundo estava de acordo com ela estava errada*). Como esses papéis não podem estar previstos na base lexical que nos serve de apoio, é forçoso reconhecer e tratar essas reclassificações lexicais, sob o risco de, a manter uma rigidez excessiva, inviabilizarmos a análise sintática e, por consequência, a identificação exata dos problemas do usuário. Neste texto nos debruçaremos sobre um desses problemas específicos, bastante comum para usuários não tão familiarizados com a norma padrão da língua portuguesa: trata-se da aparente instabilidade entre adjetivos e advérbios homônimos, ou entre o uso adjetivo e o uso adverbial de formas da língua. Dessa instabilidade nascem muitos dos chamados erros de concordância nominal, cuja identificação e resolução integram um dos objetivos da ferramenta que ora desenvolvemos.

Vejam alguns exemplos de troca:

a) do advérbio pelo adjetivo:

- (1) *Tinha menas gente na festa* (em vez de *Tinha menos gente na festa*)
- (2) *Os soldados estavam alertas* (em vez de *Os soldados estavam alerta*)
- (3) *Os rios deságuam longes* (em vez de *Os rios deságuam longe*)
- (4) *Ela estava meia cansada* (em vez de *Ela estava meio cansada*)
- (5) *Comprei muitos poucos bois* (em vez de *Comprei muito poucos bois*)
- (6) *Ele correu mau* (em vez de *Ele correu mal*)
- (7) *A roupa custava cara* (em vez de *A roupa custava caro*)
- (8) *A espada feriu funda* (em vez de *A espada feriu fundo*)
- (9) *A menina chegou rápida* (em vez de *A menina chegou rápido*)
- (10) *Elas brincavam próximas da casa* (em vez de *Elas brincavam próximo da casa*)

b) do adjetivo pelo advérbio:

- (11) *A roupa está caro* (em vez de *A roupa estava cara*)
- (12) *Há bastante alunos aqui* (em vez de *Há bastantes alunos aqui*)
- (13) *Elas mesmo fizeram isso* (em vez de *Elas mesmas fizeram isso*)

- (14) *Comprei muito pouco bois* (em vez de *Comprei muito poucos bois*)
- (15) *Os times jogaram desfalcado* (em vez de *Os times jogaram desfalcados*)
- (16) *Estaremos aberto no domingo* (em vez de *Estaremos abertos no domingo*)
- (17) *Eles fizeram horas extra* (em vez de *Eles fizeram horas extras*)
- (18) *Elas sempre chegam junto* (em vez de *Elas sempre chegam juntas*)
- (19) *Segue anexo a foto* (em vez de *Segue anexa a foto*)
- (20) *Julguei desnecessário as recomendações* (em vez de *Julguei desnecessárias as recomendações*)

Uma das tarefas do ReGra deveria ser alertar o usuário para o fato de que todas as construções acima listadas não obedecem ao que prescrevem as gramáticas normativas da língua portuguesa, caracterizando casos de solecismo de concordância.<sup>7</sup> Seriam, pois, sentenças agramaticais, sabendo-se que a ideia de gramaticalidade aqui pertinente não pode ser a postulada pela tradição da linguística descritiva, particularmente o gerativismo, em que se confunde com a noção de aceitabilidade, por qualquer dos falantes nativos da língua, de determinadas estruturas sintáticas. Em função da própria natureza da ferramenta – conservadora por definição e essencialmente normativa – o conceito de gramaticalidade aqui esposado se aproxima da ideia de adequação a uma variedade sócio-histórica específica da língua portuguesa, a chamada “norma culta”, cujo juízo de aceitabilidade é tomado como privilégio de um grupo muito restrito de falantes, os dicionaristas e os gramáticos, de suposta envergadura para licenciar ou reprovar o uso das formas linguísticas. Não pertence ao objetivo deste trabalho a discussão acerca da pertinência e da validade desse tipo de juízo de valor para as diferentes variedades linguísticas; importa aqui que, em situações mais formais de produção da escrita, os redatores que fizerem uso das expressões acima listadas poderão ser punidos ou discriminados por as terem utilizado em detrimento das formas licenciadas pelas autoridades gramaticais. À ferramenta caberia justamente alertar o usuário para essa possibilidade.

No caso das sentenças assinaladas, o alerta já se verifica em relação às do número (1) ao (6) e do número (12) ao (20), mas quase sempre de forma indireta. No caso da sentença (1), a correção ocorre porque a forma *menas* não consta no léxico. Trata-se, portanto, de um erro interpretado como ortográfico. Em (2), (3), (4), (5), (6), (12), (13) e (14), a correção se deve a regras pontuais, especializadas nos contextos de ocorrência das formas *alerta*, *longe*, *meio*, *muito*,

*mau, bastante, mesmo e pouco*, que representam casos de desvio extremamente recorrentes na língua portuguesa. A correção das sentenças que se estendem do número (15) ao (20) acompanha o resultado da análise sintática empreendida. Às palavras sublinhadas nessas sentenças não corresponde nenhum tipo de ambiguidade problemática no léxico (isto é, a classificação simultânea como adjetivo e advérbio): todas estão categorizadas (prioritariamente) como adjetivo (ou como adjetivo e substantivo), o que conduz ao preenchimento da posição de predicativo do sujeito (caso de (15), (16), (18) e (19)), da posição de predicativo do objeto (caso de (20)) e da posição de adjunto adnominal (caso de (17)), que estabelecem necessariamente uma relação de concordância nominal com o núcleo por elas modificado. Restam, porém, as sentenças (7), (8), (9), (10) e (11), nas quais a ferramenta, indevidamente, não intervém, e que oferecem maior dificuldade porque, diferentemente dos casos anteriores, envolveriam fenômenos cuja regularidade não se pôde ainda perceber.

Essa dificuldade de regularização será aqui tomada como evidência de que a má-formação das sentenças de (1) a (20) deriva de trocas que ocorrem em contextos bastante diferentes e que, portanto, não poderão ser equacionadas por um mesmo conjunto de regras de revisão. Embora constituam comportamentos de mesma natureza – a troca de adjetivo por advérbio ou de advérbio por adjetivo –, esses problemas não possuem uma etiologia comum e podem e devem ser reagrupados em dois subconjuntos razoavelmente distintos: o derivado das trocas provocadas pela interferência de uma forma verbal, e o derivado das trocas que ocorrem à revelia do verbo, no interior do sintagma nominal. Assim, seriam dois os conjuntos de problemas:

a) sem interferência do verbo:

- (1) Tinha menas gente na festa (em vez de Tinha menos gente na festa)
- (4) Ela estava meia cansada (em vez de Ela estava meio cansada)
- (5) Comprei muitos poucos bois (em vez de Comprei muito poucos bois)
- (12) Há bastante alunos aqui (em vez de Há bastantes alunos aqui)
- (14) Comprei muito pouco bois (em vez de Comprei muito poucos bois)
- (17) Eles fizeram horas extra (em vez de Eles fizeram horas extras)

b) com interferência do verbo:

- (2) *Os soldados estavam alertas* (em vez de *Os soldados estavam alerta*)
- (3) *Os rios deságuam longes* (em vez de *Os rios deságuam longe*)

- (6) *Ele correu mau* (em vez de *Ele correu mal*)
- (7) *A roupa custava cara* (em vez de *A roupa custava caro*)
- (8) *A espada feriu funda* (em vez de *A espada feriu fundo*)
- (9) *A menina chegou rápida* (em vez de *A menina chegou rápido*)
- (10) *Elas brincavam próximas da casa* (em vez de *Elas brincavam próximo da casa*)
- (11) *A roupa está caro* (em vez de *A roupa estava cara*)
- (13) *Elas mesmo fizeram isso* (em vez de *Elas mesmas fizeram isso*)
- (15) *Os times jogaram desfalcado* (em vez de *Os times jogaram desfalcados*)
- (16) *Estaremos aberto no domingo* (em vez de *Estaremos abertos no domingo*)
- (18) *Elas sempre chegam junto* (em vez de *Elas sempre chegam juntas*)
- (19) *Segue anexo a foto* (em vez de *Segue anexa a foto*)
- (20) *Julguei desnecessário as recomendações* (em vez de *Julguei desnecessárias as recomendações*)

No primeiro caso, os solecismos de concordância envolveriam relações de dependência sintática que se refeririam à própria estrutura do sintagma nominal: a quantificação (1, 4, 5, 12, 14) e a delimitação de fronteiras lexicais (17). No caso de (17), por exemplo, o problema parece ser de natureza principalmente morfológica, relativa à estrutura de “hora extra”, que pode ser definida como uma lexia complexa (ou mesmo composta: “hora-extra”), no que o problema se equipararia, em larga medida, ao verificado na formação de plural de outras tantas expressões multipalavras da língua portuguesa: “salários-família”, “caixa de fósforos”, “banca de jornal” etc. A suposta adverbialização do adjetivo “extra”, neste caso, corresponderia a um processo de formação de um novo composto, ainda não integrado à variedade oficial da língua. O fenômeno seria, portanto, menos sintático do que lexical. Nos casos de quantificação, observa-se mais explicitamente a referência à estrutura interna do sintagma nominal, mas os fenômenos observados independem, novamente, da forma verbal. Os exemplos remetem ao problema da definição da classe dos quantificadores do português, hoje distribuídos pela classe dos adjetivos e dos advérbios, sem que venham a integrar um todo articulado. Ainda que ambos os casos – de delimitação de fronteiras lexicais e de quantificação – venham a merecer tratamento mais geral do que o estabelecido atualmente, o fato é que as regras de revisão já implementadas vêm produzindo resultados razoavelmente satisfatórios para essas situações, de que

não nos ocuparemos aqui. Privilegiaremos, neste texto, o tratamento dos demais casos, mais relacionados à estrutura da própria sentença.

No entanto, será preciso também observar que os membros desse segundo grupo, em que intervém a forma verbal, também derivam de situações de troca diferentes, que merecem ser assinaladas. Em (13), por exemplo, o uso da pró-forma “mesmo” cumpriria um papel enfático de marcada expletividade, que também poderia ser considerado no quadro de outros fenômenos do português do Brasil, como o indicado, por exemplo, na clivagem de (13') *Elas é que fizeram isso*. Embora pareça estar novamente envolvida, nesse caso, a estrutura do sintagma nominal – a forma “elas mesmas” constituiria uma locução pronominal –, a relação que se estabelece entre “elas” e “mesmo” parece ser, em algum sentido, tributária da estrutura da própria sentença, na medida em que (13'') *Elas mesmas fizeram isso para elas mesmas* parece não ser uma sentença bem-formada da língua portuguesa.<sup>8</sup>

Outro caso que merece tratamento à parte consiste em (7) *A roupa custava cara*, em que parece ser violada a estrutura de uma colocação da língua portuguesa, porque o verbo “custar”, em sentido intransitivo, selecionaria um conjunto extremamente reduzido de formas com as quais poderia coocorrer: “caro”, “barato”, “muito”, “pouco”, “nada”. Haveria aqui, portanto, como no caso de “hora extra”, uma nova componente locucional, agora envolvendo um verbo. Evidência dessa seletividade do verbo poderia ser derivada da agramaticalidade de (7) *A roupa custava dispendiosamente*, a indicar que mesmo os advérbios de parentesco semântico com “caro” seriam vetados neste contexto.

Deveremos, portanto, isolar também essas duas formas, e concentrarmos nos demais casos em que as ocorrências parecem convergir para um fenômeno de natureza mais geral, menos associado às singularidades lexicais que afetam cada uma das formas repertoriadas, e que não poderiam ser resolvidos, portanto, por meio de regras pontuais, especializadas no contexto de ocorrência de um número muito limitado de entradas.

Analisemos, a título de exemplo, a sentença (9), emblemática dos problemas que se reproduzem nas demais. Para o usuário, que redigiu *A menina chegou rápida*, trata-se de um verbo intransitivo (*chegou*) seguido de um predicativo do sujeito (*rápida*). A hipótese do usuário é bastante razoável se considerarmos que a coordenação entre verbo intransitivo e predicativo do sujeito é uma estrutura sintática extremamente produtiva na língua portuguesa. Na NGB, esse tipo de construção recebe o nome de predicado verbo-nominal e,

como (9), ilustram essa estrutura sentenças como *A menina chegou atrasada*, *A lua ia grande e bela*, *Virgília entrou risonha e sossegada*, *Fabiano marchava teso*.<sup>9</sup> O problema se instala quando percebemos que, para a gramática normativa, a estrutura correta para a sentença (9) não seria a de um predicado verbo-nominal, mas a do predicado verbal, caracterizado pela adjunção adverbial ao verbo.<sup>10</sup> Para a tradição gramatical, a ideia de rapidez, materializada no adjetivo “rápida”, refere-se antes ao processo descrito pelo verbo (“chegar rapidamente”) do que a uma qualidade do sujeito (“menina rápida”). Trata-se, portanto, de uma ideia que desempenha a função de modificadora do verbo, e que por isso deveria vir materializada como expressão de natureza adverbial, e não como expressão de natureza adjetiva (modificadora do nome), como de fato ocorreu. Assim, a hipótese do usuário sobre a estrutura de (9) estaria errada, e como reflexo desse erro se observaria a lexicalização de “rápida” (adjetivo) em vez de “rápido” ou “rapidamente” (advérbios).<sup>11</sup>

O problema que se coloca aqui é como ensinar à máquina isso que acabamos de perceber intuitivamente. Como mostrar para a máquina que (9) *A menina chegou rápida* não é uma sentença bem-formada da língua portuguesa? Se proibíssemos a coordenação entre quaisquer adjetivos e verbos intransitivos, corrigiríamos também (21) *A menina chegou atrasada*, que não apresenta nenhum problema gramatical. Se proibíssemos a coordenação entre o verbo intransitivo e todas as formas marcadas pela ambiguidade categorial entre adjetivo e advérbio (caso de *rápido*), corrigiríamos, com sucesso, (7), (8), (9) e (10), mas vetaríamos (22) *A lua ia alta*, que é perfeitamente bem-formada. O problema é, pois, recusar (9) sem deixar de aceitar (21) e (22). Em outras palavras: trata-se de reduzir o número de omissões (não intervenções indevidas: caso atual de (9)) sem incrementar o número de hipercorreções (intervenções indevidas) da ferramenta. Haverá algum modelo teórico que explique por que, embora as três sentenças sejam (aparentemente) idênticas do ponto de vista morfológico e do ponto de vista sintático, apenas a primeira delas não é aceita pelos gramáticos?

Antes de passar adiante, é preciso dizer que dois pontos incomodam na proposição deste problema. O primeiro é o fato de a má-formação de (9) *A menina chegou rápida* ser bem mais controversa do que a que se observa em *A menina chegou amanhã* e *A menina chegou durante dez horas*, por exemplo. No primeiro caso, está patente uma incompatibilidade temporal entre o verbo e o adjunto adverbial; no segundo caso, a incompatibilidade será aspectual [isolada a possibilidade de interpretação iterativa da sentença]. Nos dois casos, a anomalia

parece evidente, e desconfiamos que sentenças como essas, por seu caráter exageradamente esdrúxulo, raramente serão produzidas pelos falantes, salvo erro de digitação e excluídas algumas situações de resto muito idiossincráticas, como a de produção deste texto. No entanto, não é o que se observa em relação a (9), cuja previsão nos manuais de correção gramatical denuncia a frequência da ocorrência.<sup>12</sup> Assim, é lícito interpretar que, para um conjunto razoavelmente numeroso de falantes da língua, (9) não é malformada. Tratar-se-ia menos de uma anomalia semântica do que de uma inadequação às regras da gramática normativa [e, nesse sentido, seria menos um problema para a linguística teórica do que para a sociologia da língua].

O segundo ponto incômodo no problema diz respeito à natureza da ocorrência de (9). Tem sido observado, com alguma frequência, o desvirtuamento do comportamento linguístico do falante em virtude de expectativas falsas (ou imagens inadequadas) do interlocutor e do que o interlocutor julga ser um bom texto. Cláudia Lemos (1977) observou, relativamente à produção de redações nos exames vestibulares, a interferência de “estratégias de preenchimento”, segundo as quais o redator incorreria em uma preocupação excessiva com a forma do texto, em detrimento de seu conteúdo ou de sua função comunicativa, resvalando em procedimentos de hipercorreção que estariam em desacordo com a sua competência linguística real. Na produção de textos escolares, importaria mais preencher estruturas pré-moldadas do que efetivamente criar estruturas. Britto (1983) expande as conclusões da autora para textos produzidos em outras instâncias, e é de crer que o procedimento seja mais usual do que se imagina. A existência de inúmeras evidências de hipercorreção na passagem da oralidade para a escrita apenas confirmaria a possibilidade de (9) revelar não mais do que um conflito entre conhecimento linguístico real e insegurança do uso, o que desautorizaria, para a sentença, a ideia de que esteja efetivamente envolvida ali uma legítima ambiguidade entre o uso adjetivo e o uso adverbial das palavras da língua.

Não pode pertencer ao objetivo deste texto o tratamento minucioso desses dois pontos. Cabe apenas dizer que ambos conformam atalhos disponíveis para evitar o problema [na medida exata em que o transferem para outro domínio], e que não serão trilhados no desenvolvimento deste percurso, porque escapam ao seu objetivo último: prover mecanismos de representação da adequação de (21) *A menina chegou atrasada* e (22) *A lua ia alta* e da pressuposta inadequação de (9) *A menina chegou rápida*. Desautorizar a gramática normativa, que teria

em (9) mais uma evidência de sua precariedade teórica; desautorizar o falante da língua, que estaria ainda em processo de aquisição da escrita; desautorizar a própria ocorrência, que manifestaria antes um problema de desempenho (ou uso) do que o conhecimento efetivo que o falante tem das estruturas linguísticas, são possibilidades teóricas que, embora legítimas, não nos levarão à consecução dos objetivos propostos. De fato, mesmo que venhamos, contra a gramática, aceitar (9) como ocorrência possível, pertencente à língua portuguesa [dada a sua enunciação por falantes nativos], como explicar a capacidade que têm determinados falantes, supostamente mais autorizados, de perceber na sentença a má-formação que nela percebem? Apenas preconceito linguístico? Essa é, em última instância, a grande questão deste trabalho, já que não estamos preocupados aqui em justificar ou não a correção de sentenças da língua portuguesa, mas em representar, para a máquina, um conhecimento algoritmizável que a leve a corrigir as sentenças que o gramático corrigiria e a não corrigir aquelas que ele não corrigiria. Seremos obrigados, portanto, a admitir, com a gramática normativa, que (9) não é bem formada, e que nosso objetivo é antes a formalização da competência metalinguística dos gramáticos, e não de todo o grupo de falantes da língua.

Da delimitação do problema emergem ainda duas dúvidas, fundamentais ao desdobramento deste texto. A primeira diz respeito ao nível de análise linguística em que o problema se situa; a segunda concerne à natureza do objeto de investigação. No primeiro caso, parece estarmos investigando um fenômeno sintático, já que em última análise estamos dedicados à resolução de casos de concordância nominal ou, no limite, da representação da estrutura sintagmática da sentença. O problema é basicamente o de descobrir se um determinado predicado é verbal ou verbo-nominal, ou seja, se admite ou não o predicativo do sujeito. No entanto, o problema sintático parece derivar de uma constatação semântica: a de que determinados itens lexicais, por suas propriedades de sentido, podem ou não funcionar como predicativo do sujeito ou adjunto adverbial. Como já foi visto, em (9) *A menina chegou rápida.*, a má-formação deriva da percepção de que o sentido de *rápida* vincula-se ao sentido de *chegou* e não ao sentido de *a menina*. Não houvesse o problema semântico (caso de (21) *A menina chegou atrasada*, por exemplo), não haveria problema sintático.<sup>13</sup>

A segunda dúvida diz respeito ao objeto de investigação. A consideração das sentenças (9) *A menina chegou rápida* e (21) *A menina chegou atrasada* faz crer que o problema é de ordem lexical. Como a única diferença entre (9) e (21)

reside na palavra que ocupa o núcleo do predicativo do sujeito, é de supor que o tratamento das formas *rápida* e *atrasada* conduza a uma diferenciação que explique a má-formação de (9) e a boa-formação de (21). Aparentemente, há inúmeras pistas, na própria língua, a respeito das diferenças que as caracterizam: a existência do advérbio *rápido* e a não existência do advérbio *atrasado*; a existência do verbo *atrasar*, de que a forma *atrasado* seria, em última análise, o particípio passado; a existência dos substantivos *rapidez* e *atraso*, e os processos de derivação (sufixal, em *-ez*; e regressiva, em *-o*, respectivamente) que os constituem. Mas a essas diferenças de inspiração morfológica corresponde um comportamento sintático muito semelhante, na medida em que (22) a (27), abaixo, são sentenças bem formadas da língua portuguesa:

(22) *A menina rápida chegou.*

(23) *A menina atrasada chegou.*

(24) *A menina é rápida.*

(25) *A menina é atrasada.*

(26) *A menina está rápida.*

(27) *A menina está atrasada.*

A única diferença que, como falante, consigo perceber entre o uso sintático de *rápida* e *atrasada* está no preenchimento da posição de predicativo do sujeito dos predicados verbo-nominais. Qual seria, portanto, a diferença entre *rápida* e *atrasada* que poderia explicar a boa-formação de (21) e a má-formação de (9)? Fosse a má-formação comum a todo uso predicativo de *rápida*, o problema estaria resolvido. Mas a consideração dos exemplos anteriores prova que a) ou devemos tratar diferentemente o predicativo do sujeito dos predicados nominais e o predicativo do sujeito dos predicados verbo-nominais, ou b) devemos buscar em *rápida* e *atrasada* propriedades emergentes (e não imanentes), que, com base no contexto, reorganizem suas possibilidades combinatórias. Nos dois casos, já não estaríamos operando (apenas) no nível lexical.

Por fim, é sempre interessante lembrar que o problema não se restringe à sentença (9), aqui tomada como ilustração. A consideração das sentenças (8), (10) e (11) faz perceber que o problema é menos simples do que aparentemente pode sugerir, extrapolando a classe dos verbos de movimento ou os adjetivos relacionados à ideia de tempo. Em todos estes casos, ainda não tratados pela ferramenta de revisão gramatical, observa-se exatamente o mesmo fenômeno: a

troca do adjunto adverbial pelo predicativo do sujeito ou o inverso, com resultados agramaticais. Se considerarmos que várias das sentenças já tratadas o foram de forma *ad hoc*, perceberemos que todo o conjunto de sentenças do segundo grupo, à exceção de (7) e (13), envolve o mesmo tipo de problema, apesar da diversidade dos adjetivos, advérbios e verbos ali presentes. Enfim, o problema que se coloca aqui, ilustrado pela sentença (9), não é exatamente incomum.<sup>14</sup> E é tanto mais intrigante quanto percebemos que a própria gramática normativa prevê o emprego adverbial de adjetivos ou o emprego adjetivo dos advérbios (cf. nota 7), sem que por isso venham a estar corretas as sentenças (1) a (20). Essas são exatamente as questões que vêm permitir a enunciação da pergunta-problema deflagrada da pesquisa que aqui se pretende realizar: como determinar se um predicado é ou não verbo-nominal? Ou, dito de outra forma: por que algumas formas adjetivas não podem ocupar a posição de predicativo do sujeito em predicados que contêm núcleos verbais?

### 3. Dos adjetivos

A resposta à pergunta-problema enunciada na seção anterior parece, à primeira vista, trivial. Algumas formas adjetivas não podem ocupar a posição de predicativo do sujeito dos predicados verbais exatamente porque os predicados envolvidos são verbais, ou seja, não admitem ou não preveem posição para o predicativo do sujeito. O conteúdo semântico representado (equivocadamente) pelo adjetivo não estaria modificando o sujeito, mas o processo expresso pelo verbo. A resposta trivial esconde, porém, pelo menos duas proposições bem mais obscuras: 1) por que esse conteúdo tem o poder de indicar antes as circunstâncias que caracterizam o processo verbal do que as qualidades do sujeito? e 2) por que isso não ocorre com todos os conteúdos exprimíveis por adjetivos?

A consideração dessas duas interrogações permite já perceber que o problema tem duas faces: a primeira, a face do adjetivo; a segunda, a face da relação entre o adjetivo e o verbo. A exploração do primeiro caso nos conduz ao conceito de adjetivo e suas possíveis subclassificações. Apenas a proposição de uma subclasse dos adjetivos permitiria explicar por que, na pergunta-problema, o pronome indefinido utilizado foi “algumas” e não “todas”. A exploração do segundo caso conduz à descrição dos processos pelos quais os sintagmas, numa sentença, subcategorizam outros sintagmas, impedindo que eles tenham um caráter adjetivo. Para responder à pergunta-problema, será fundamental, no

desenvolvimento do texto, responder a essas duas questões subsidiárias. E essas respostas não podem perder de vista o objetivo geral da pesquisa, que restringe de início as possíveis vias de análise, sob o risco de se produzir um conhecimento que, embora válido, não se preste ao pretendido aprimoramento da ferramenta.

Aceitas as restrições indicadas na seção anterior, há pelo menos um roteiro teórico disponível para a solução do problema. Trata-se da suposição – a que referirei, de forma bastante genérica, como “analítica” – de que o significado da sentença é motivado pelo significado das palavras ou dos sintagmas que a compõem. Reconhece-se, pois, a autonomia do significado no nível subsentencial da língua, e acredita-se que esse significado pré-sentencial desempenha efetivamente algum papel na construção do significado da sentença. Dessa hipótese deriva logicamente a tese do isomorfismo sintático-semântico: haveria equivalência entre a estrutura de constituintes sintáticos e a estrutura de constituintes semânticos da sentença. A versão mais conhecida dessa tese é aquela que postula a precedência da estrutura sintática sobre a estrutura semântica, ou seja, em que a organização dos constituintes do nível semântico representaria uma função (não necessariamente linear) da organização dos constituintes do nível sintático. Nesse caso, postula-se a existência de um mecanismo de projeção, a partir do qual seria feita a “interpretação” da estrutura sintática da sentença.

Levado às últimas consequências, este roteiro analítico sugere que, assim como o significado das sentenças estaria diretamente relacionado ao significado de suas partes, o significado das partes mínimas da sentença (os morfemas lexicais) poderia ser decomposto em unidades menores, constituindo um feixe de traços de significação (ou traços semânticos) que estabeleceria as condições necessárias e suficientes de pertencimento do morfema a uma determinada classe e, no mesmo movimento, o individualizaria, tornando-o singular no conjunto dos outros morfemas pertencentes à mesma classe. Trazida para o âmbito da discussão do problema, essa possibilidade ofereceria uma solução para a primeira subquestão anunciada no início da última seção. A análise componencial disponibilizaria um poderoso instrumento para representar as semelhanças e as diferenças entre os elementos (como *rápida* e *atrasada*) que compõem a classe dos adjetivos, e viria, portanto, explicar por que alguns deles podem e outros não podem ocupar a posição de predicativo do sujeito dos predicados dotados de núcleo verbal.

Na medida em que nos interessa a apreensão de regras genéricas, válidas para um conjunto a princípio indeterminado de situações em que o fenômeno

se realiza, a nossa investigação não pode pretender fazer referência a todo o conteúdo semântico dos adjetivos envolvidos, devendo ser orientada para a busca de características gerais, comuns a todos e apenas aos membros da subclasse procurada. Torna-se útil, então, a decomposição do significado lexical em traços semânticos subordenados, capazes de diferenciar os membros de uma mesma categoria, agrupando-os em subconjuntos que compartilhariam particularidades semânticas às quais poderia ser reputada a propriedade de poder exercer esta ou aquela função sintática.

Nesse sentido, uma primeira alternativa teórica é o modelo, já clássico, de Katz e Fodor (1963). Para os autores, o significado dos itens lexicais é composto de uma parte sistêmica, que reflete “quaisquer relações semânticas sistemáticas existentes entre aquele item [lexical] e o resto do vocabulário da língua”, e de uma parte assistêmica, específica de cada item lexical, que reflete “o que é idiossincrático quanto ao significado daquele item”. No primeiro caso, fala-se em “marcadores semânticos” (*semantic markers*), e na possibilidade de eles conduzirem a regras genéricas para a descrição e a explicação dos fenômenos semânticos (como a anomalia ou a redundância, por exemplo). A parte idiossincrática do significado lexical, a que os autores referem como “distinguidores” (*distinguishers*), não desempenharia papel relevante na teoria semântica na medida em que sua especificidade impediria a formalização de procedimentos generalizáveis.

Dessa forma, o problema, no modelo de Katz e Fodor, seria encontrar o marcador semântico ou o conjunto de marcadores semânticos que pudesse diferenciar um e outro item lexical. Nessa direção, uma forma de representar a oposição entre *rápida* e *atrasada* consistiria em encontrar traços semânticos que garantissem apenas ao segundo a possibilidade de ocupar a posição de predicativo do sujeito dos predicados compostos com verbos nocionais. No entanto, devo dizer que não me foi possível encontrar entre as classificações estudadas alguma que diferenciasse os dois itens lexicais sob suspeição. A separação entre adjetivos qualificativos e determinativos (como faz a gramática tradicional), entre adjetivos atributivos e predicativos (cf. BOLINGER, 1967) ou entre adjetivos categoremáticos e sincategoremáticos (cf. BORGES, 1979) não resolve o problema, visto que as formas *rápida* e *atrasada* pertencem, em todas essas classificações, ao mesmo subconjunto: são ambas adjetivos qualificativos, predicativos, categoremáticos. Se não, vejamos.

Se admitirmos que adjetivo qualificativo é “a palavra que se une ao substantivo para exprimir a qualidade do objeto ou do ser, ou da noção designada por esse substantivo”, em oposição a adjetivo determinativo, que serve “para fazer com que esse substantivo seja atualizado numa frase” (cf. DUBOIS *et alii*, 1973, p. 21), será fácil concluir que tanto *rápida* quanto *atrasada* podem ser qualificativos ou determinativos, a depender do contexto. No contexto *A menina chegou* –, são ambos adjetivos qualificativos, porque ambas as formas são utilizadas, não para indicar uma instância específica (entre as várias instâncias disponíveis) da classe das “meninas”, mas para consignar um valor (não necessariamente diferencial) a um membro já específico dessa classe. O problema é que apenas *atrasada* poderia, segundo a gramática, desempenhar esse papel.

Nesse aspecto, a classificação de Bolinger (1967) parece ser mais adequada, uma vez que o objetivo do autor é provar que a estrutura atributiva tem estatuto teórico tão primitivo quanto a predicativa. Na Teoria Padrão da gramática gerativo-transformacional, o uso atributivo dos adjetivos era tratado como resultado de transformações de relativas. O fato de que muitos adjetivos atributos não possuem uso predicativo (*a total stranger*, mas *\*the stranger is total*) enfraquece seguramente essa posição, conduzindo o autor a pesquisar os critérios segundo os quais nem todos os adjetivos podem desempenhar funções atributivas e predicativas. No caso aqui endereçado, no entanto, o problema concerne apenas em parte ao uso atributivo ou predicativo dos adjetivos, pois tanto *A menina é rápida* ou *A menina está rápida*, quanto *A menina é atrasada* ou *A menina está atrasada* são sentenças possíveis na língua portuguesa.

Por fim, se admitirmos, com Borges (1979) (com base em Zuber), que categoremáticos são os adjetivos para os quais é verdadeira a implicação  $NP_1$  é  $NP_2$  A  $\bar{P}$   $NP_1$  é  $NP_2$  e  $NP_1$  é A, devemos também admitir que *rápida* e *atrasada* compartilham o mesmo traço, já que *Maria é uma menina atrasada*  $\bar{P}$  *Maria é uma menina* e *Maria é atrasada*, da mesma forma que *Maria é uma menina rápida*  $\bar{P}$  *Maria é uma menina* e *Maria é rápida*

Este roteiro teórico parece nos conduzir, inevitavelmente, para a percepção de que há algo mais envolvido no fenômeno do que a simples subclassificação dos adjetivos. Em outras palavras: somos levados a admitir que a inserção de itens lexicais na sentença não respeita apenas informação categorial, mas envolve também informação sobre a combinação de categorias gramaticais, especialmente adjetivos e verbos. Já que nos afastamos da classe dos adjetivos para considerar a forma pela qual as relações entre adjetivos e verbos são atualizadas em uma

sentença, torna-se imperioso aderir a um novo modelo descritivo, menos solipsista do que o modelo anterior, restrito à investigação da referência de palavras isoladas. Por seus irrecusáveis méritos formais, embora sua computabilidade seja antes uma premissa metodológica do que um objetivo para fins de aplicação, decidiu-se investigar o problema da relação entre o adjetivo e o verbo sob o espectro da teoria da gramática gerativa ou, mais especificamente, no escopo da Teoria da Regência e da Ligação (ou Government and Binding Theory, GB), tal como exposta, principalmente, por Raposo (1992). Embora seja essa já uma apropriação secundária, porque interpretação da interpretação dos textos originais que conformam o aparelho teórico da GB, o objetivo deste trabalho, a descoberta de princípios que permitam a formalização de regras para o controle do uso adjetivo dos advérbios e do uso adverbial dos adjetivos, autoriza, pelo menos neste primeiro momento, uma abordagem mais esquemática dos conceitos propostos pelo modelo adotado.

A ideia de que a inserção dos itens lexicais na sentença não respeita apenas informação sobre o item lexical mas também sobre suas possibilidades combinatórias, por seu “quadro de subcategorização” ou por sua “estrutura argumental”, vem sendo proposta pela teoria gerativa desde Chomsky (1965). Estabelece-se ali que as regras de inserção lexical não seriam exatamente da mesma natureza das regras de reescrita categorial, mas constituiriam antes um mecanismo (de inserção lexical) autônomo, conformando mais uma parte do modelo gramatical (já constituído, a esta época, pela componente categorial e pelo léxico). Na GB, esse mecanismo passaria a estar associado ao Princípio da Projeção, segundo o qual “as representações em cada nível sintático (estrutura-D, estrutura-S e LF) são projetadas do léxico, isto é, observam as propriedades temáticas e de subcategorização dos itens lexicais” (RAPOSO, 1992, p. 294). A exploração desse princípio constituirá, em larga medida, a hipótese teórica encontrada para a representação da pressuposta má-formação de (9) *A menina chegou rápida* e da boa-formação de (21) *A menina chegou atrasada*, mas será necessário realizar antes um movimento terminológico: aproximar a ideia de “má-formação” ao conceito de “agramaticalidade”.

Consideremos novamente, a título de exemplo, a sentença (26) *A menina está rápida*, até aqui tratada como incontroversamente bem formada, e comparemo-la a duas outras sentenças, também já apresentadas, (24) *A menina é rápida* e (27) *A menina está atrasada*. Acrescentemos a este pequeno corpus de análise algumas outras construções: (28) *?A menina está tola*, (29) *A menina é*

*tola*, (30) ??*A menina é desmaiada* e (31) *A menina está desmaiada*. Embora todas essas sentenças possam ser consideradas gramaticais, quero crer que sejam gramaticais com maior ou menor grau de certeza. Para pelo menos um falante nativo do português [o autor deste texto], as sentenças (24), (27), (29) e (31) parecem ser mais bem formadas (mais naturais) do que (26), (28) e (30). Em função disso, devo admitir, neste texto, que o juízo de gramaticalidade comporta um gradiente de certeza ao longo do qual as sentenças poderiam ser consideradas mais ou menos gramaticais, como se houvesse, entre o estatuto de gramaticalidade e de agramaticalidade, uma semigramaticalidade, zona cinzenta, de delimitação incerta, governada por princípios de categorização mais ambíguos do que os propostos pelos modelos clássicos, baseados em um conjunto discreto e finito de condições necessárias e suficientes.

Parece-me que os adjetivos *rápida* e *inteligente*, mercê de sua carga semântica e das associações culturais por ela deflagradas, não poderiam, pelo menos em princípio, participar de enunciados marcados pela transitoriedade da asserção. Esses adjetivos, como muitos outros, parecem denotar antes uma propriedade intrínseca (imaneente) dos nomes por eles modificados, e não um estado que pudesse ser a eles circunstancialmente adjudicado. Corresponderiam, em larga medida, à ideia de “qualidade” (de natureza essencial), em oposição à “situação” (de caráter acidental), no sentido que esses termos recebem em Aristóteles (*Cat.* I.16 a 15 e ss.). Evidentemente que, por força dos mecanismos de ampliação do âmbito de significação de uma palavra (particularmente da possibilidade metonímica ou, mais precisamente, sinedótica de substituição do permanente pelo provisório), poderiam ser ambos os adjetivos utilizados em outros contextos mais temporais, mas estariam condicionados, então, à possibilidade de consignaçoão de um sentido menos literal (i.e., menos composicional) às expressões de que participariam. À medida que essa possibilidade estivesse vetada, seja pela forte contradição das expectativas constituídas em relação ao conjunto de referências do falante, seja pela declarada proeminência da interpretação literal, que se revelaria muito mais provável e factível, a ampliação da extensão dos adjetivos seria frustrada, dando origem aos problemas aqui reportados.

A ser aceita essa hipótese, observar-se-ia, então, a possibilidade de uma nova classificação dos adjetivos, subdivididos agora em dois grupos: o dos adjetivos que indicam propriedades imanentes (ou permanentes), e o dos adjetivos que manifestam propriedades transcendententes (ou transitórias) dos nomes a eles relacionados. A primeira classe corresponderia ao que Bolinger

(1967) chamou de adjetivos “não temporários” (ou “costumeiros”); a segunda classe, aos adjetivos “temporários” (ou “não costumeiros”).<sup>15</sup> Segundo o autor, apenas os adjetivos não temporários poderiam ser usados atributivamente na língua inglesa. Por isso:

(32a) *The girl is foolish.*

(32b) *The foolish girl.*

(*foolish* = adjetivo não temporário)

(33a) *The girl is faint.*

(33b) \**The faint girl.*

(*faint* = adjetivo temporário).

No entanto, é preciso convir que, no português, o comportamento dos adjetivos temporários e não temporários seria diferente do observado para o inglês. A depender do contexto, ambas as classes de adjetivos podem exercer tanto funções atributivas quanto predicativas:

(32a') *A garota é tola.*

(32b') *A garota tola.*

(*tola* = adjetivo não temporário)

(33a') *A garota está desmaiada.*

(33b') *A garota desmaiada.*

(*desmaiada* = adjetivo temporário<sup>16</sup>)

Como se pode observar, contrariamente a (33b), (33b') é uma expressão bem formada da língua portuguesa. No entanto, e diferentemente do que ocorre no inglês, em que não há marcação formal entre as duas possibilidades do verbo “to be” (temporária e não temporária), (32a”) e (33a”) abaixo teriam estatuto de boa formação controverso, sendo particularmente digno de nota o fato de que o uso predicativo do adjetivo temporário com uma forma verbal não temporária (como a indicada pelo verbo “ser”) parece ser menos autorizado do que o uso predicativo do adjetivo não temporário com uma forma verbal temporária (indicada pelo verbo “estar”):

(32a”) ?A garota está tola.

(33a”) ??A garota é desmaiada.

Deve ser observado, principalmente, que a restrição do uso predicativo de algumas formas adjetivas não se confunde, nesse caso, com o fenômeno, também reportado por Bolinger (1967), segundo o qual há adjetivos que não podem exercer (nunca) função predicativa. A sentença do inglês “*The stranger is total*” será malformada independentemente do valor temporário ou não temporário do verbo “to be”. O que se percebe a partir dos exemplos acima é algo diferente: embora “tola” e “desmaiada” possam exercer função predicativa, não o podem fazer (ou não o podem fazer com o mesmo grau de boa-formação) no mesmo contexto sintático. Estabelece-se uma incompatibilidade entre a cópula e o adjetivo que, em larga medida, denuncia a pertinência, em ambos os casos, da categoria de “temporiedade”, como quer Bolinger (1967), ou “imanência”, como foi referida pouco acima. Qualquer que seja o caso, a má-formação dessas duas sentenças se deve ao não atendimento de critérios de seleção de natureza eminentemente semântica: o predicativo do sujeito deveria concordar em temporiedade ou imanência com o verbo de ligação, sob o risco de anomalia semântica. Nesse sentido, a má-formação tanto de (32a”) quanto de (33a”) muito se aproximaria da presente em (34) *Incolores ideias verdes dormem furiosamente* (Chomsky), (35) *Quadruplicidade bebe procrastinação* (Russell) e (36) *Quinta-feira está na cama com sexta-feira* (Ryle). Ou seja, as sentenças seriam sintaticamente bem formadas (visto que atualizariam uma combinatória sintática prevista pela gramática da língua), mas padeceriam de má-formação semântica tendo em vista que não seriam interpretáveis. A possibilidade de (32a”) e de (33a”) parecerem menos semanticamente anômalas do que (34), (35) e (36) derivaria exclusivamente do número e do nível das incongruências semânticas presentes em cada caso.

No entanto, é forçoso observar que o problema de (9) *A menina chegou rápida* é aparentemente menos de natureza semântica do que sintática, já que a interpretabilidade do enunciado resta preservada. O problema tem evidentemente uma matriz semântica – o sentido de “rápida” –, mas a sentença (e sua frequência de ocorrência o prova) não deixa de ser interpretável. A incompatibilidade que se observa, então, é de outra ordem: a sentença atualizaria uma estrutura de predicação que não seria possível na língua portuguesa. Trata-se, pois, de um problema de natureza sintática. A restrição a (9) seria, principalmente, consequência do estabelecimento de um impedimento sintático para o preenchimento, por um adjetivo, de uma posição que estaria sintaticamente reservada para o

advérbio. Essa hipótese – a de que a má-formação de (9) é de caráter predominantemente sintático –, que está diretamente associada ao estatuto do predicado secundário no português brasileiro, será explorada na próxima seção.

#### **4. Dos predicados secundários orientados para o sujeito**

Segundo Foltran (1999; 2003), o conceito de “predicado” – e, por extensão, o conceito de “predicativo” – tem estatuto controvertido no quadro teórico da gramática gerativo-transformacional. A autora reporta duas posições teóricas bem definidas, opostas pelo estatuto que conferem ao predicado: para a vertente configuracional, o predicado seria uma projeção máxima; para a vertente não configuracional, o predicado participaria de um outro tipo de estrutura (semântica), de natureza proposicional. Na vertente configuracional, a autora identifica duas diferentes interpretações do predicado: a derivada de Williams (1980; 1983), que trata o predicado em termos de papéis temáticos; e a derivada de Rothstein (1983; 1997), para quem nem toda predicação envolve atribuição de papel temático (a autora cita, explicitamente, o caso de estruturas com sujeitos expletivos, verbos na passiva e verbos de alçamento). No primeiro caso, a relação de predicação seria principalmente semântica; no segundo, é marcadamente sintática (“a relação entre sujeito e predicado é codificada na sintaxe”).

O principal argumento a favor da posição de Rothstein parece ser o de que seu poder explanatório seria superior às abordagens “semânticas” (configuracionais ou não configuracionais) do fenômeno da predicação. À medida que a autora dissocia a predicação da atribuição de papel temático, é capaz de estabelecer uma distinção entre dois tipos de constituintes: os que são tematicamente marcados (isto é, que recebem um papel temático), e que viriam a constituir argumentos, elementos fechados, já saturados, que não dependeriam de outros constituintes para alcançar sua integridade sintática; e os que não são tematicamente marcados, elementos abertos, parasitas de outros constituintes, de que seriam dependentes para saturação sintática. Essa saturação – convém observar – estaria primariamente associada às diáteses sintáticas desses elementos mais do que à sua valência semântica, que seria antes contingencial. A vantagem competitiva dessa hipótese está associada à reorganização (sintática) do fenômeno da predicação, que passaria a ser concebido em dois diferentes níveis: a predicação primária e a predicação secundária, assim definidas:

- (I) A é um predicado primário de B sse A predica de B, e A e B estão numa relação de c-comando mútuo e B não é q-marcado fora da relação de predicação com A;
- (II) A é um predicado secundário de B sse A é predicado de B, e A e B estão numa relação de c-comando mútuo e B é q-marcado por um núcleo não contido em A. (ROTHSTEIN, 1997, *apud* FOLTRAN, 2003, p. 56-57)

A partir dessa definição, se pode perceber que os problemas que vêm sendo aqui analisados conformam exatamente um dos casos de predicação secundária referidos por Rothstein. A distinção é particularmente importante porque é menos vaga que o conceito de “predicativo”. De fato, para a gramática tradicional, o “rápida” de (9) *A menina chegou rápida* e o “rápida” de (24) *A menina é rápida* exerceriam, rigorosamente, a mesma função sintática, de predicativo do sujeito, o que impediria que se explicasse porque (24) é bem formada e (9) não. Na abordagem de Rothstein, constituiriam, respectivamente, um predicado secundário e um predicado primário,<sup>17</sup> distinção que permitiria reendereçar o problema definido na seção anterior.

Outra distinção que se torna a esta altura pertinente diz respeito à dupla possibilidade de predicação secundária, reiterada pela oposição entre predicativo do sujeito e predicativo do objeto, extraída da gramática tradicional, e ilustrada por (37) e (38) abaixo, retirados de Bisol (1975):

- (37) O menino chegou cansado.  
 (38) Pedro encontrou Maria doente.

Para efeito deste texto, será importante restringirmos desde já o escopo desta investigação aos predicados secundários do primeiro tipo, normalmente referidos por “predicado secundário do sujeito”, “predicado secundário orientado para o sujeito” ou “predicado secundário circunstancial”. Evita-se, dessa forma, o debate, aqui desnecessário, sobre a unicidade dos sintagmas nominais envolvidos no segundo caso, frequentemente interpretados como “small clauses”, o que desautorizaria, por sua vez, a própria postulação de predicação secundária em muitos desses casos.

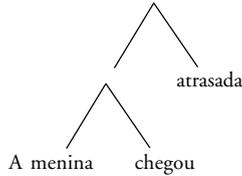
Como vimos no capítulo anterior, a restrição a (9) *A menina chegou rápida* não seria propriamente um problema de seleção semântica (de onde o resíduo de interpretabilidade), mas do estabelecimento de um impedimento sintático

para o preenchimento, por um adjetivo, de uma posição que não poderia ser por ele ocupada. Visto que esta restrição somente se verificaria para parte dos adjetivos, fomos levados a admitir que existiriam duas classes de adjetivos, uma das quais não poderia exercer a função de predicado secundário orientado para o sujeito. Para que possam ser explicadas as razões desse impedimento, torna-se imperioso considerar, ainda que muito esquematicamente, e de forma apenas provisória, a estrutura do predicado secundário orientado para o sujeito no português brasileiro. No entanto, e para que não se perca o foco desta discussão, será necessário fazer a abstração dos outros muitos fatores envolvidos na estruturação do predicado secundário, e responsáveis pela agramaticalidade de um sem-número de outras construções, como (39) *Maria ganhou na sena esperançosa*, (40) *Pedro sabe a lição seguro* e (41) *Maria tem um carro feliz*, referidas por Foltran (1999). Como essas construções não envolvem, a princípio, a oposição aqui discutida entre adjetivos e advérbios, não faremos referência aos problemas nelas envolvidos que, como demonstrou a autora, estariam associados ao valor aspectual do predicado principal.

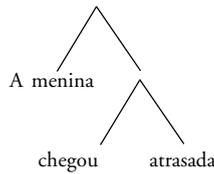
Em primeiro lugar, e para que se possa propor uma configuração sintática para (9) e outras sentenças assemelhadas, caberia discutir o estatuto da predicação secundária orientada ao sujeito, particularmente em sua relação com a predicação primária. Em relação a esse ponto, parecem ser duas as possibilidades: estabelecer que entre a predicação primária e a secundária há uma relação de coordenação ou afirmar que a relação entre elas é de subordinação. No primeiro caso, (9) *A menina chegou rápida* deveria ser interpretada como (9') *A menina chegou e a menina estava rápida*; no segundo caso, (9) corresponderia a (9'') *A menina estava rápida ao chegar*. Admitamos, inicialmente, e de acordo com Foltran (1999), que a predicação secundária orientada ao sujeito envolva relações desta segunda espécie, ou seja, de subordinação. Isso implica dizer que, do ponto de vista semântico, (9) seria equivalente a (9''), ou seja, que a predicação “rápida” de “a menina” somente faz sentido no contexto de “chegar”, relação que não fica tão evidente a partir de (9'), em que o intervalo temporal de “estar” não precisa, necessariamente, coincidir com o expresso pelo verbo matriz.

O dado curioso em relação à adoção dessa hipótese consiste em perceber que, de acordo com (9''), a predicação primária estaria subordinada à predicação secundária. Ou seja, do ponto de vista da configuração sintática, a predicação primária (“chegar”), e não a predicação secundária (“atrasada”), exerceria o papel de adjunto. Isso implicaria dizer que a estrutura de (9) seria a indicada em (9'')

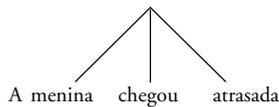
abaixo, contra as duas possibilidades registradas para (9'), mais frequentemente adotadas na literatura:



(9'') A menina estava atrasada ao chegar



(9') A menina chegou e estava atrasada



(9') A menina chegou e estava atrasada

Embora Foltran (1999) rejeite a leitura (9') *A menina chegou e estava cansada*, opera, em todos os seus exemplos, com configurações que afirmam o caráter adjuntivo da predicação secundária, resguardando para o predicado “chegou” a posição de verbo matriz. Para a autora, é o predicado secundário que tem estatuto de adjunto:

Esse entendimento é devido ao fato de esse constituinte não ser um argumento semântico da projeção verbal com a qual se combina. O predicado adjunto não é selecionado pelo verbo e nem contribui para a descrição da ação verbal (FOLTRAN, 1999, p. 83-84).

Uma evidência empírica a favor do estatuto de predicado secundário como adjunção seriam, segundo a autora, que reporta McNally (1997), as dificuldades de extração: “dos adjuntos, quando é possível a extração, só podemos extrair material T-regido” (p. 84). Do contrário, seria violado o princípio das categorias vazias (ECP). No entanto, é forçoso observar que, pelo menos no caso assinalado, essa seria antes uma evidência a favor do estatuto de adjunção da predicação primária, como se pode perceber abaixo:

- (42) A menina tomou o sorvete de morango doente de febre.
- a. O que a menina tomou doente de febre?
  - b. Como a menina tomou o sorvete de morango?
  - c. ?De que a menina tomou o sorvete doente de febre?
  - d. ?De que a menina tomou o sorvete de morango doente?

Se admitirmos que, em (42), nem “de morango” nem “de febre” são T-regidos, quero crer que não se estabelecem diferenças nas possibilidades de extração desses dois adjuntos, estejam eles adjungidos ao predicado primário ou ao predicado secundário. O estatuto de adjunção do predicado secundário, pelo menos em relação ao teste proposto, revela-se, portanto, matéria de disputa.

Um segundo problema concerne à integridade do predicado secundário na estrutura sintática da sentença. Para Foltran, que acompanha a literatura, o predicado secundário orientado ao sujeito não deve ser representado como uma *small clause*. A autora se vale de várias evidências empíricas e toma a resposta a vários dos testes propostos como determinantes: ainda que muitas das situações normalmente referidas como predicação secundária possam configurar casos de *small clause*, observa-se que: a) os casos reportados (principalmente por Stowell) estão limitados à predicação secundária orientada para o objeto e b) há *small clause* apenas nos complementos dos verbos de marcação excepcional de caso. A posição da autora, que segue os argumentos de Rothstein, é a de que os predicados secundários orientados ao sujeito e o sujeito não formam um único constituinte.

A hipótese esposada pela autora, no entanto, parece mais aplicável à interpretação (9’), por ela mesma rejeitada. Minha intuição de falante parece indicar que, no caso de (9), o verbo “chegar” atribui papel temático ao constituinte oracional “A menina [estava] rápida” (a uma *small clause*, portanto, no sentido proposto por Stowell), e não apenas ao sintagma nominal sujeito “A menina”. Isso implica dizer que, pelo menos nos casos analisados, o predicativo do sujeito e o sujeito formam, sim, um único constituinte sintático, como indicado abaixo:

- (9) \*A menina chegou [ PRO [rápida]].  
 (21) A menina chegou [ PRO [atrasada]].

Parece haver pelo menos três evidências favoráveis a essa posição. A primeira delas diz respeito à interpretação do advérbio. Segundo Stowell, a impossibilidade de o verbo ter alcance sobre o verbo matriz seria uma prova de

que ele estaria inserido em uma *small clause*. Foltran oferece, como contra-exemplo, a sentença (43) abaixo:

(43) Ele cantou a música repetidamente fora do tom.

Para a autora, o “repetidamente” teria alcance tanto sobre o “fora do tom” quanto sobre o “cantou”, o que indicaria que, no caso, não teríamos uma *small clause*. Embora a questão possa ser motivo de controvérsia, não reconheço, como falante nativo do português, o alcance de “repetidamente” sobre o predicativo do sujeito em (43). O alcance do advérbio parece estar limitado ao verbo matriz. A conclusão deriva, particularmente, da implausibilidade de *?Ele estava repetidamente fora do tom*. Parece-me, nesse caso, que não se satisfaz a pré-condição de dupla combinatória semântica, ou seja, de que o advérbio possa modificar as duas estruturas, do verbo matriz e do predicativo do sujeito. O argumento é, pois, elusivo. O “repetidamente”, nesse caso, não estaria no interior da *small clause*, mas fora dela, pelo que nada provaria. Como o próprio Stowell indica, os testes para sustentar a posição de *small clause* para os predicados secundários orientados para o sujeito são pouco expostos à comprovação empírica, porque a posição de especificador seria ocupada por uma categoria vazia. Nos predicados secundários orientados para o objeto, em que a posição de especificador estaria preenchida, a identificação do escopo do advérbio seria mais transparente:

For small clauses, constituency tests may be applied with definitive results only when the subject of small clause is overt, as in [I consider John foolish], since the location of a null category is difficult to establish conclusively – especially when its existence is in dispute, as in [John is considered foolish]. Hence I will only consider constituency tests applied to small clause with overt subject (STOWELL, 1987, p.11).

De qualquer forma, se utilizássemos, por exemplo, o advérbio “razoavelmente”, que poderia modificar tanto o verbo matriz (“Ele cantou a música razoavelmente”) como o predicativo do sujeito (“Ele estava razoavelmente fora do tom”), o resultado da interpretação seria outro: a única leitura admissível para (44) abaixo – assim entendo – seria a de que “Ele estava razoavelmente fora do tom ao cantar a música”.

(44) Ele cantou a música razoavelmente fora do tom.

Nesse caso, a conclusão seria exatamente a inversa: o fato de advérbio não poder alcançar o verbo matriz provaria que, no caso, estaríamos diante de uma *small clause*. O exemplo, no entanto, não é muito feliz, porque “razoavelmente” pode também ser interpretado como advérbio de intensidade, o que conturbaria a igualdade de condições das estruturas candidatas à modificação. Em todo o caso, é importante aqui assinalar que a interpretação do advérbio, pelo menos nos casos de predicação secundária orientada para o sujeito, resta como objeto de disputa, e não se oferece, pelo menos até o momento, como evidência definitiva de que não estamos diante de um constituinte único nos casos assinalados.

Uma outra evidência diz respeito à (im)possibilidade de coocorrência de dois advérbios na mesma sentença, um dos quais modificando o verbo matriz, e outro modificando o predicativo do sujeito. Foltran apresenta, como evidência da não unicidade do predicado secundário orientado ao sujeito, o exemplo de (45):

(45) \*A Maria frequentemente chega raramente cansada.

Nesse caso, e diferentemente do que ocorreria em situações de *small clause* (como em *Eu frequentemente considero a Maria raramente inconveniente*), a impossibilidade de coocorrência dos dois advérbios indicaria que, novamente, estaríamos diante de uma estrutura de adjunção. Contra esse argumento, além da já referida dificuldade de proceder a testes empíricos com a predicação secundária orientada ao sujeito, pode ser oferecido o exemplo de (46) abaixo, que me parece gramatical:

(46) A Maria frequentemente chega realmente cansada.

Foltran analisa também os testes relativos aos efeitos das condições de sujeito, mas ela mesma reconhece que os testes, nesse caso, são inconclusivos. Os contra-argumentos restantes são retirados de McNulty, para quem a estrutura com categoria vazia proposta por Stowell não poderia ser licenciada, dado que PRO ocorreria em posição regida, contra condição estabelecida pela Teoria da Ligação. Adicionalmente, a autora invoca o alto custo teórico e a fragilidade empírica da adoção de uma categoria vazia no caso indicado. Os argumentos, nesses casos, são, como a própria autora diz, de natureza intrateórica, pelo que a opção da autora contra a representação da predicação como *small clause* assume um caráter predominantemente arbitrário, derivado das escolhas metodológicas que lhe seriam pertinentes para o desenvolvimento da tese.

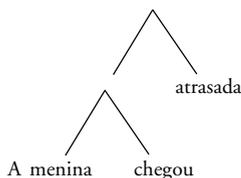
A recusa da possibilidade de representação da predicação secundária como *small clause* poderia ser mantida se, ao fim e ao cabo, atingíssemos alguma configuração sintática que permitisse explicar os fenômenos envolvidos nesse tipo de construção. Não é, porém, o que ocorre. Embora faça um longo recenseamento da literatura, à procura da categoria à qual o predicado secundário estaria adjungido, a autora chega a um beco sem saída: as várias categorias candidatas – VP e IP, principalmente – são, por fim, rejeitadas:

todas as tentativas de se elaborar uma configuração sintática que desse conta das diferentes construções de predicação secundária, inclusive da checagem dos traços *f* do adjetivo, que são [-interpretáveis], foram frustradas. Essa questão fica, portanto, para trabalhos futuros (FOLTRAN, 1999, p. 102).

Embora este texto, muito mais modesto, e de alcance muito mais restrito, não possa pretender estabelecer uma configuração sintática para a predicação secundária, cabe aqui apontar que o beco sem saída a que chegou a autora pode estar de alguma forma relacionado às premissas adotadas. Embora compartilhe, com a autora, a posição de que a relação entre a predicação primária e a predicação secundária é de subordinação (e não de coordenação), quero crer que, como indica a paráfrase (9”) *A menina estava rápida ao chegar*, a única solução para a sentença é a de que não está efetivamente envolvida, no caso, uma estrutura de predicação secundária, mas duas estruturas de predicação primária, que ocorrem em diferentes níveis de constituição da sentença. Assim, do ponto de vista da estruturação semântica, a representação para (9) seria a primeira, e não as duas outras indicadas abaixo:

- (9) atrasada (chegar (menina))  
 \*chegar (atrasada (menina))  
 \*chegar (menina) · atrasada (menina)

A representação sintática dessa estrutura, como indicado acima, deveria ser (9”) abaixo:



(9'') A menina estava atrasada ao chegar

Nesse caso, o predicado secundário não desenvolveria, propriamente, uma relação de adjunção, mas ocuparia uma posição nuclear, de que a sentença constituiria projeção máxima. Trata-se, evidentemente, de hipótese heterodoxa, contraintuitiva, particularmente porque inverte o estatuto normalmente atribuído aos verbos nocionais em predicados verbo-nominais. Mas respeita-se, nesse caso, a estrutura semântica indicada acima, segundo a qual o predicado “atrasada” predica de “A menina chegou”, o que parece estar mais de acordo com a leitura proposta: “A menina estava atrasada ao chegar”.

No entanto, a principal razão pela qual a configuração (9'') merece ser adotada, em oposição às configurações adversárias, diz respeito à agramaticalidade (9). A se admitir que a sentença, em construções como a indicada, constitui uma projeção máxima do predicado secundário, e não do predicado primário, a predicação secundária poderia ser reduzida a uma estrutura de predicação primária com verbo de ligação marcado pela temporariedade, caso de “estar”. Bastaria, para tanto, que admitíssemos que a estrutura da predicação secundária constitui, de fato, um IP, como propõe, por exemplo, McNulty.

McNulty, contra Stowell, não admite que em (9) e (21), PRO possa ser regido por AP. A autora parte de (46) para afirmar que os mesmos princípios que impedem PRO de ser regido por VP impediriam o AP de reger PRO em (9) e (21).

(46) [<sub>C'</sub> [<sub>I'</sub> I tried [ PRO [<sub>V'</sub> to go]]].

(9) \*A menina chegou [<sub>AP</sub> PRO [rápida]].

(21) A menina chegou [<sub>AP</sub> PRO [atrasada]].

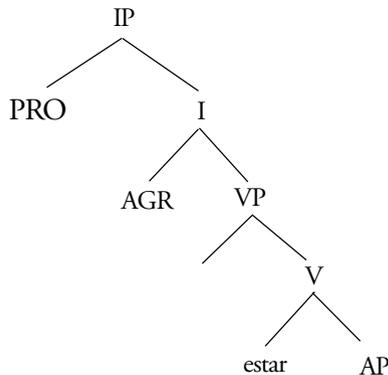
PRO precisaria ser regido, portanto, por um elemento externo à *small clause*. A autora supõe, então, que a *small clause* projeta uma categoria funcional I', sem flexão, que seria adjungida a V', no caso dos predicados secundários orientados para o sujeito. Assim, (9) e (21) seriam representadas como:

(9) \*A menina [<sub>V</sub>, chegou [<sub>I</sub>, PRO [<sub>AP</sub> rápida]]].

(21) A menina [<sub>V</sub>, chegou [<sub>I</sub>, PRO [<sub>AP</sub> atrasada]]].

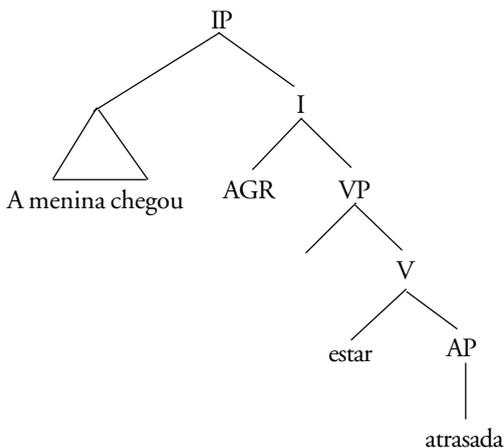
A análise da autora conduz, no entanto, à regência de PRO, o que estaria em oposição à Teoria da Ligação, produzindo, portanto, uma redução ao absurdo. Na medida em que nem I” nem V” são barreiras para PRO, PRO seria regido por AGR, que o m-comanda, e (9) e (21) não seriam, portanto, licenciadas. Foltran indica que esta análise é tendenciosa, porque, no âmbito da própria teoria gerativista, postula-se que “se a categoria IP domina uma barreira, ela própria se transforma numa barreira por herança”. Como em (9) e (21) AP não é complemento e, portanto, não é uma categoria L-marcada, constituiria uma barreira, propriedade esta que seria transferida para IP, permitindo que PRO fosse licenciado como sujeito da *small clause*.

Para efeito da representação da agramaticalidade de (9) *A menina chegou rápida*, no entanto, as representações indicadas acima são de bem pouca utilidade. Qualquer que seja a projeção máxima da *small clause* (se AP, no caso de Stowell, se IP, no caso de McNulty), a má-formação resta inexplicada. Será preciso introduzir, na estrutura, algum mecanismo que impeça que adjetivos como “rápida” possam ocorrer na posição de núcleo do sintagma adjetivo. A única alternativa que parece razoável, embora seja também muito precária, consiste na proposição de que a projeção máxima da *small clause* compreenderia um sintagma verbal de ligação, como indicado abaixo:



Não nos interessa, neste momento, indicar em que nível dessa estrutura estaria PRO, que McNulty representa na posição [Spec, IP], aqui mantida por constituir

posição não-regida. O fato é que, no limite, a própria categoria PRO poderia ser dispensada, se admitirmos, por exemplo, que todo o constituinte formado por “A menina chegou” poderia ocupar uma das posições previstas nessa estrutura.



De qualquer forma, o preenchimento da posição de núcleo do AP deveria satisfazer o critério de seleção semântica reivindicado pelo verbo “estar”, de que contivesse o traço [+costumeiro] ou [-inerente], responsável, por exemplo, pela já referida (e pressuposta) semiagramaticalidade de “A menina está tola” ou de “A menina está inteligente”. A ser aceita essa estrutura, entendo que se encerraria aqui o papel da sintaxe, e as demais restrições passariam a ser de natureza eminentemente semântica, como as já indicadas por Foltran para explicar outros casos de agramaticalidade na predicação secundária. No entanto, será preciso observar que existe pelo menos um problema para a abordagem proposta: a realização fonológica de “estar”, que não se verifica na estrutura superficial da sentença. Será particularmente difícil sustentar, sem a presença do verbo, a pertinência de uma projeção máxima de natureza verbal dentro da estrutura da *small clause*. Esse problema, no entanto, não concerne apenas à configuração proposta, mas à configuração da própria predicação primária, que vem sendo frequentemente interpretada, também ela, e com o sujeito, como *small clause*, a quem o verbo de ligação atribuiria um único caso sintático (nominativo). É forçoso observar, por exemplo, que, embora não seja esse o caso do português, o verbo de ligação nem sempre se realiza fonologicamente nas línguas naturais.

Em outras palavras: a abordagem proposta transfere, para a estrutura da predicação primária, os problemas relativos à seleção dos adjetivos na posição de núcleo do sintagma adjetivo da predicação secundária, o que, em última instância, fragiliza a pertinência da solução proposta.

A estrutura proposta, embora ofereça uma alternativa descritiva para o problema representado em (9) – e, por extensão, para os problemas também indicados por (10) e (11) – deve estar sujeita a um conjunto de testes de confirmação que, por força do tempo, escapam ao objetivo deste trabalho. As principais conclusões extraídas desta análise são, evidentemente, controvertidas, porque amparadas em um conjunto muito restrito de dados, e são sujeitas à validação em outros contextos. A verdade ou a falsidade das hipóteses consideradas é, no entanto, menos relevante para este texto do que o seu significado: a possibilidade de consideração da agramaticalidade (9) de uma perspectiva estritamente sintática, despregada dos critérios de seleção semântica, inadequados pela observância da interpretabilidade das construções analisadas. Por ora, registre-se aqui apenas que, fora da definição do predicado secundário como uma *small clause*, ou como predicação primária de uma predicação primária, não se concebem estratégias para a representação da suposta agramaticalidade de (9) e da gramaticalidade de (21), objeto último desta descrição.

## Conclusões

São principalmente duas as conclusões (práticas) extraídas deste trabalho: no espectro teórico aqui explorado, 1) é possível estabelecer um conjunto de regras formais que proibam o uso adjetivo dos advérbios e o uso adverbial dos adjetivos nos predicados secundários; 2) não é possível estabelecer um conjunto de regras formais que permitam o uso adjetivo dos advérbios e o uso adverbial dos adjetivos nos predicados secundários. Ambas as conclusões se tornam (in)operacionalizáveis pelo mesmo conjunto de procedimentos:

- 1) o léxico deve subclassificar os adjetivos em temporários (ou não costumeiros) ou em não temporários (ou costumeiros). Essa subclassificação pode ser feita pela consignação de traços [+temp] e [-temp] aos itens lexicais. O critério de atribuição dos traços é de ordem inapelavelmente subjetiva, e deve estar idealmente associado à possibilidade de ocorrência no contexto de predicativo

do sujeito de verbos temporários (como “estar”) e não temporários (como “ser”). As opções são exclusivas, ou seja, a mesma forma adjetiva não pode possuir ambos os traços, a menos que estejam envolvidas acepções diferentes. No entanto, o uso não temporário de formas temporárias, assim como o uso temporário de formas não temporárias, devem ser previstos pelo sistema, sob a forma de recurso figurativo (como metonímia ou sinédoque);

- 2) devem estar igualmente previstas, no léxico, na estrutura argumental dos verbos, suas valências semânticas: [+temp] e [-temp], de forma que serão licenciados, como núcleos de predicados secundários, apenas os adjetivos que respeitem as possibilidades combinatórias previstas pelo verbo. A imensa maioria dos verbos será classificada como [+temp]; apenas alguns verbos de ligação serão classificados como [-temp]. Ou seja, a temporalidade do verbo estará em larga medida relacionada à sua capacidade de portar conteúdo. Essa relação impede uma associação direta da categoria de temporalidade aqui referida a outras categorias verbais (como aspecto, tempo ou modo), características de todos os verbos;
- 3) deve ser reforçado o Princípio de Projeção, segundo o qual não pode haver predicado secundário com valor [-temp] associado a sintagma dominante (verbal) de valor [+temp]. O nó dominante (VP, projeção do valor de temporalidade do verbo) constituirá, neste caso, barreira para a ligação do predicado secundário à posição (spec) de sujeito do VP. Em sendo observado esse caso, a ferramenta deve acusar erro de sintaxe, indicando que o verbo esperaria um advérbio, e não um adjetivo na posição do predicado secundário;
- 4) não será possível, no entanto, à ferramenta ponderar o valor das relações dentro da sentença, de forma a prever, por algum mecanismo estocástico, a possibilidade de outras relações virem a preponderar, dando origem a diferentes graus de gramaticalidade. Nesse sentido, a ferramenta persistirá recusando a possibilidade (metonímica) de o adjetivo temporário ser tratado como não temporário, ou vice-versa. Essa limitação é diretamente decorrente da polaridade do aparelho teórico, para o qual as sentenças ou são absolutamente gramaticais ou agramaticais, sem a possibilidade de um estágio ou posição intermediária.

Os procedimentos aqui abordados merecem, obviamente, uma exploração mais detalhada, e devem se submeter ainda à prova da experiência. No entanto, não pode ser objetivo deste trabalho a verificação empírica das hipóteses teóricas ora construídas, particularmente porque a atribuição de traços a verbos e adjetivos, no caso do Projeto ReGra, deve ser feita para o conjunto de todas as

formas que, na base lexical, estejam assim categorizadas. Mais importante se revela, para efeito deste texto, a contradição em que incorre: a proibição do uso adverbial dos adjetivos nas situações desautorizadas pela gramática implica, por ora, a proibição de todo e qualquer uso adverbial dos adjetivos, em função do determinismo sintático que rege, não apenas o aparelho teórico aqui utilizado, mas os pressupostos metodológicos que inspiram o lugar de sua aplicação, uma ferramenta de revisão gramatical automática. A superação dessa limitação, associada à representação do desconhecimento ou da dúvida que o falante possa registrar em relação à sua própria língua materna, depende ainda da apropriação de estratégias teórico-metodológicas que venham a permitir que se represente, de forma consistente, a diferença entre adjetivos que podem ser usados adverbialmente e adjetivos que não podem ser usados adverbialmente. Sem que essa distinção esteja disponível em modelos teóricos formais de descrição da linguagem, o “erro” humano, porque absolutamente inexplicável como “erro”, continuará fora do escopo de atuação da máquina.

## Notas

<sup>1</sup> Sobre o Projeto ReGra, consulte-se Martins *et alii*. (1998).

<sup>2</sup> Por textos referenciais, entendam-se aqueles assim chamados por Vanoye (1988). São textos que privilegiam, entre as funções da linguagem, a função referencial, ou seja, aquela centrada no referente da mensagem, segundo a tipologia proposta por Jakobson. Com esta restrição, pretende-se retirar do escopo da ferramenta o tratamento de textos predominantemente emotivos, poéticos, fáticos, conativos e metalinguísticos, que envolvem complicadores que não foram, por ora, considerados.

<sup>3</sup> Para os substantivos, estão disponíveis o gênero e o número; para os adjetivos, o gênero, o número e o grau; para os verbos, o número, a pessoa, o tempo e o modo; para os pronomes, o tipo, o gênero, o número, a pessoa e, para os pronomes pessoais, também o caso; para os artigos, o tipo, o gênero e o número. Sobre a estrutura do dicionário consulte-se Nunes *et alii* (1996b).

<sup>4</sup> A título de exemplo, considere-se a transitividade proposta por Ferreira (1986) para o verbo *esquecer* (p. 710): transitivo direto (*A humanidade jamais esquecerá os crimes do nazismo*), transitivo indireto (*Esqueceram-lhe os dias felizes*), intransitivo (*Tudo esquece neste mundo*) e pronominal (*Se eu de ti me esquecer, nem uma lágrima caía sobre o sepulcro em que eu jazer*).

<sup>5</sup> Sobre o modelo de análise sintática adotado no Projeto ReGra veja-se Martins (1997) e Nunes *et alii.* (1996a).

<sup>6</sup> A postulação dessa ordem preferencial para a execução das regras de produção – o que implica graus diferenciados de precedência para as possibilidades sintáticas da língua – encontra respaldo em teorias do desempenho que postulam a existência de um viés de estruturação *default* (cf. WINGFIELD, 1993). A imprescindibilidade de *backtrackings* em sentenças com inversões sintáticas radicais ou de desdobramento inesperado (como as *sentenças-labirinto*) ilustra de forma razoavelmente convincente a existência dessa tendência para se realizar a análise sintática preferencialmente em certa direção (considerada aqui como canônica), em prejuízo de outras mais raras ou de relações sintagmáticas mais obscuras.

<sup>7</sup> É interessante observar que a gramática normativa admite o emprego adverbial dos adjetivos (“ela fala alto”, “ela canta baixo”) e o emprego adjetivo dos advérbios (“ela está toda molhada”, “tenho comprado muito caras as minhas loucuras”), mas em situações específicas. Os dois casos dizem respeito ao fenômeno de atração, reportado por Câmara Jr. (1977, p. 63), que se verifica “a) para maior harmonização morfológica dentro da frase, ou b) para uma harmonização com outro padrão sintático significativamente equivalente (cruzamento).” No entanto, embora possam envolver algumas das características descritas por Câmara Jr. (caso explícito de 4, citada pelo próprio autor), todas as sentenças assinaladas são condenadas, sem exceção, por pelo menos três gramáticas brasileiras, aqui tomadas como referência: Cunha e Cintra (1985), Rocha Lima (1992) e Bechara (1992).

<sup>8</sup> A má-formação dessa sentença parece ser motivo de controvérsia já que um dos avaliadores deste texto entende que a construção é bem formada. Como falante nativo do português, no entanto, resisto a aceitar construções que marquem a ênfase em mais de um constituinte, porque entendo que a ênfase, como a topicalização, é um fenômeno exclusivista: em uma mesma construção, não se pode topicalizar ou enfatizar mais de um constituinte, sob o risco da agramaticalidade. No exemplo em tela, soa-me estranha a ênfase simultaneamente atribuída ao sujeito e ao destinatário da ação, como o comprovaria a variante “Elas mesmas fizeram isso para eles mesmos” ou “Eles mesmos fizeram isso para elas mesmas”.

<sup>9</sup> Exemplos retirados de Kury, 1993, p. 27.

<sup>10</sup> Nesse sentido, a sentença (9) estaria mais relacionada às sentenças *Ela chegou cedo, A lua ia tranquilamente, Virgília entrou rápido, Fabiano marchava muito*, do que às anteriormente referidas.

<sup>11</sup> É interessante antecipar que a sentença (9) *A menina chegou rápida* não constitui, em princípio, um caso de má-formação sintática, na medida em que é possível a

interpretação “a menina chegou e estava rápida”, como se observa, por exemplo, no contexto (9a) *Quando Maria saiu do país, há uns três anos, ela era muito lenta, levava horas para fazer o que quer que fosse. Agora a menina chegou rápida, esperta, ágil: executa a mais complicada das tarefas em poucos minutos*. A possibilidade dessa interpretação, que será discutida mais adiante, seguramente enfraquece uma das principais premissas que servem de base a esta pesquisa. No entanto, quero supor que, em situações mais comuns, e desconsiderado o contexto, é extremamente mais plausível que o falante, durante a enunciação de (9), tenha tido por intenção dizer “a menina chegou rapidamente”, e que o tenha feito utilizando uma estrutura sintática inadequada (de predicado verbo-nominal), segundo a perspectiva da gramática tradicional. De qualquer forma, as sentenças (7) e (8), que compartilham com (9) o mesmo problema, dificilmente poderiam admitir dupla interpretação, porque \**“a roupa custava e estava cara”* e ?*“a espada feriu e estava funda”*. Nesses casos, e em tantos outros em que ambiguidade não se pudesse instalar, o problema persistiria.

<sup>12</sup> Sacconi (1990), que pretende compilar os principais problemas gramaticais verificados para a língua portuguesa, registra *Ela agiu rápida*.

<sup>13</sup> Sob inúmeros aspectos, a agramaticalidade de (9) é tão enigmática quanto a gramaticalidade de (21). Por que o sentido de *atrasada* pode estar se referindo a *a menina* mas o de *rápida* não? E por que o sentido de *rápida* deve estar vinculado ao de *chegou* se o mesmo não vale para *atrasado*? Será tão somente porque o advérbio *rápido* existe no dicionário e o advérbio *atrasado* não? Ou terá sido a dicionarização do advérbio *rápido* apenas a consequência do processo que aqui se investiga?

<sup>14</sup> É difícil mensurar exatamente a extensão do problema, mas importa menos aqui o número de ocorrências do que a sua possibilidade. No léxico que serve de base ao Projeto ReGra, que conta hoje com aproximadamente um milhão e meio de formas (incluídas as flexionadas), 78 palavras estão classificadas, simultaneamente, como adjetivos e advérbios. Este dado não é, porém, de todo relevante, na medida em que são igualmente comuns, como o denunciam as sentenças (16), (17), (18), (19) e (20), confusões entre adjetivos e advérbios que, no dicionário, não possuem esta ambiguidade.

<sup>15</sup> Esta subclassificação parece estar de acordo com pelo menos duas outras propostas: a de Bisol (1975) e Carlson (1977) (*apud* FOLTRAN, 1999). No primeiro caso, a autora distingue entre adjetivos que trazem ou não o traço [-inerente]. No segundo, o autor distingue entre predicados de indivíduos (individual level) e predicados de estágios (stage level). Explicam-se, desta forma, a agramaticalidade de sentenças como *O menino chegou inteligente*, por exemplo.

<sup>16</sup> Pode ser evidentemente objeto de controvérsia classificar “desmaiada” como adjetivo no lugar de forma do particípio passado do verbo “desmaiar”. Embora esta

última classificação possa preceder historicamente a primeira, é forçoso observar que a) o particípio passado é uma forma nominal dos verbos, ou seja, funciona morfossintaticamente como adjetivo ou substantivo; b) o falante nem sempre poderá diferenciar o uso particípio e adjetivo das formas da língua portuguesa, sendo esta uma divisão antes metalinguística do que propriamente linguística; e c) em que pesem os problemas relativos ao registro lexicográfico, pelo menos dois dicionários do português (FERREIRA, 1986; HOUAISS, 2001) registram o verbete “desmaiado”, forma de citação de “desmaiada”, exclusivamente como adjetivo, sem referência à forma do particípio.

<sup>17</sup> A adoção desse paradigma permite também isolar, do escopo da descrição, o problema do predicativo do objeto, como em (37) *A viagem fez a menina rápida*, em que “rápida” volta a figurar como predicado primário, uma vez que o alvo da atribuição de caso do verbo “fez” seria todo o constituinte “a menina rápida”.

## Referências bibliográficas

ARISTÓTELES. *Categorias*. Porto: Porto Editora, 1995

BECHARA, E. *Moderna gramática portuguesa*. São Paulo: Companhia Editora Nacional, 1992.

BISOL, L. *Predicados complexos do português: uma análise transformacional*. Porto Alegre: UFRGS, 1975.

BOLINGER, D. Adjectives in English: Attribution and Predication. *Lingua*, v. 18, p. 1-34, 1967.

BORGES, J. L. *Adjetivos: predicados extensionais e predicados intensionais*. 1979. Dissertação (Mestrado em Letras) – Instituto de Estudos da Linguagem, Unicamp, Campinas, 1979.

BRITTO, L. P. L. Em terra de surdos-mudos (um estudo sobre as condições de produção de textos escolares). *Trabalhos de Linguística Aplicada* 2, p.149-167, 1983.

CÂMARA JR, J. M. *Dicionário de Linguística e Gramática*. Petrópolis: Vozes, 1977.

CARLSON, G. A unified analysis of the English bare plural. *Linguistics and Philosophy*, Dordrecht, v. 1, p. 413-457, 1977.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge: The MIT Press, 1965.

CUNHA, C.; CINTRA, L. *Nova gramática do português contemporâneo*. Rio de Janeiro: Nova Fronteira, 1985.

- DUBOIS, J.; GIACOMO, M.; GUESPIN, L.; MARCELLESI, C.; MARCELLESI, J-B.; MEVEL, J-P. *Dicionário de Linguística*. São Paulo: Cultrix, 1973.
- FERREIRA, A. B. H. *Novo dicionário Aurélio da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 1986.
- FOLTRAN, M. J. G. D. *As construções de predicação secundária no português do Brasil: aspectos sintáticos e semânticos*. 1999. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 1999.
- FOLTRAN, M. J. G. D. Relações de predicação. In: MULLER, A. L.; NEGRÃO, E. V.; FOLTRAN, M. J. (Org.). *Semântica Formal*. São Paulo: Contexto, 2003. p.47-60.
- JAKOBSON, R. Linguística e poética. In: *Linguística e comunicação*. São Paulo: Cultrix, s.d.
- KATZ, J. J.; FODOR, J. A. Estrutura de uma teoria semântica. In: LOBATO, L. M. P. *A semântica na linguística moderna: o léxico*. Rio de Janeiro: Francisco Alves, 1963 [1977].
- KURY, A. G. *Novas lições de análise sintática*. São Paulo: Ática, 1993.
- LE MOS, C. T. G. Redações do vestibular: algumas estratégias. *Cadernos de Pesquisa*. São Paulo: Fundação Carlos Chagas, n. 23, 1977.
- McNALLY, L. *A semantics for the English existential construction*. New York: Garland, 1977.
- McNULTY, E. M. *The syntax of adjunct predicates*. 1988. Doctoral (Dissertation) – University of Connecticut, 1988.
- MARTINS, R. T. Análise sintática automática para revisores gramaticais. SEMINÁRIO DO GRUPO DE ESTUDOS LINGÜÍSTICOS DO ESTADO DE SÃO PAULO, Campinas, 26, *Anais...* Campinas, SP, 1997. p.157-162.
- MARTINS, R. T. *et alii*. Linguistic issues in the development of ReGra: a grammar checker for Brazilian Portuguese. *Natural Language Engineering*, v. 4, n. 4, p. 287-307, 1998.
- NUNES, M. G. V. *et alii*. Development of a parser for Brazilian Portuguese. WORKSHOP ON COMPUTATIONAL PROCESSING OF WRITTEN AND SPOKEN PORTUGUESE, 2, CEFET-PR, Curitiba, *Proceedings...* Curitiba, PR, 1996b. p. 71-80.
- NUNES, M. G. V. *et alii*. The design of a lexicon for Brazilian Portuguese: lessons learned and perspectives. WORKSHOP ON COMPUTATIONAL PROCESSING OF WRITTEN AND SPOKEN PORTUGUESE, 2, CEFET-PR, Curitiba, *Proceedings...* Curitiba, PR, 1996b. p. 61-70.

- RAPOSO, E. P. *Teoria da Gramática*. A Faculdade da Linguagem. Lisboa: Editorial Caminho, 1992.
- ROCHA LIMA, C. H. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 1992.
- ROTHSTEIN, S. The syntactic forms of predication. 1983. Doctoral (Dissertation) – MIT, Cambridge, 1983.
- ROTHSTEIN, S. Predicates and their subjects. Bar-Ilan University, 1997. (unpublished manuscript)
- SACCONI, L. A. *Não erre mais!* São Paulo: Atual, 1990
- SPOKEN PORTUGUESE, 2, CEFET-PR, Curitiba, *Proceedings...* Curitiba, PR, 1996a. p. 71-80.
- STOWELL, Timothy. 1987. 'As So, Not so As'. Los Angeles: University of California (unpublished manuscript).
- VANOYE, F. *Usos da linguagem*. São Paulo: Martins Fontes, 1988.
- WILLIAMS, E. Predication. *Linguistic Inquiry*, v. 11. Cambridge: 1980. p. 203-238.
- WILLIAMS, E. Against small clause. *Linguistic Inquiry*, v. 14. Cambridge: 1983. p. 287-308.
- WINGFIELD, A. Sentence processing. In: GLEASON, J. B.; RATNER, N. B. *Psycholinguistics*. Fort Worth: Harcourt Brace Jovanovich College Publishers, 1993.