

SOBRE A CLASSIFICAÇÃO DAS PALAVRAS
(On Word Classes)

Mário A. PERINI, Yara G. LIBERATO, Maria Elizabeth F. SARAIVA &
Lúcia FULGÊNCIO (*Universidade Federal de Minas Gerais*)

ABSTRACT: In this article we discuss some basic questions relative to word classification, namely: the type of unit that should be classified (words, not lexemes); the need to set up clear objectives to classification; and the rigorous distinction between classes (based on paradigmatic relations) and functions (based on syntagmatic relations). We then consider the traditionally accepted distinction between "adjectives" and "nouns", and we show that this is not a class distinction, but rather a functional distinction. Thus, adjectives and nouns are not distinct word classes in Portuguese.

RESUMO: Neste artigo discutimos algumas questões básicas relativas à classificação das palavras, a saber: o tipo de unidade a ser classificada (palavras, e não lexemas); a necessidade de estabelecer objetivos claros para a classificação; e a distinção estrita entre classes (baseadas em relações paradigmáticas) e funções (baseadas em relações sintagmáticas). A partir daí, discutimos a questão da distinção tradicionalmente aceita entre "adjetivos" e "substantivos", e mostramos que essa distinção não é de classe, mas de função. Ou seja, adjetivos e substantivos não se distinguem enquanto classes de palavras em português.

KEY WORDS: Word Classes; Adjectives; Nouns; Noun Phrase.

PALAVRAS-CHAVE: Classes de Palavras; Adjetivos; Substantivos; Sintagma Nominal.

0. Introdução

Este artigo relata parte dos resultados de um projeto em andamento na UFMG, que trata da estrutura interna do sintagma nominal em português, em relação com a classificação das palavras. Desse projeto já resultou a publicação, em 1996, de um número especial da *Revista de Estudos da Linguagem*; o título desse texto, omitido pela *Revista*, é *O Sintagma Nominal em Português: Estrutura, Significado e Função* (autores: Mário A. Perini, Sigrid T. Fraiha, Lúcia Fulgêncio e Regina Bessa Neto).

No presente artigo consideramos especificamente o problema da

classificação das palavras, assim como certo número de questões teóricas que a nosso ver não estão suficientemente esclarecidas na literatura; essas questões serão abordadas a partir de uma discussão da distinção entre “adjetivos” e “substantivos”.

1. Classificar o quê?

A primeira pergunta importante a ser considerada é: ao classificarmos, que espécie de unidade deve ser levada em conta? Classificamos lexemas ou palavras?

Os lexemas correspondem a conjuntos de formas vinculadas paradigmaticamente através da relação tradicionalmente denominada “flexão”. Por exemplo, temos o **lexema** *correr*, que compreende as **palavras** *correr*, *corro*, *corríamos* etc. Podemos então perguntar: ao classificarmos estamos agrupando palavras ou lexemas?

Acreditamos que a classificação, no que nos interessa, é de palavras, e não de lexemas; isso porque as diferentes palavras que compõem os lexemas tradicionalmente definidos podem ter comportamento gramatical muito variado, tanto sob a perspectiva semântica quanto sob a morfossintática. Assim, dentro do lexema chamado “o verbo *correr*”, temos *corríamos*, que tem comportamento muito diferente do de *correr* ou *correndo*. A **palavra** *correr* tem distribuição algo parecida à de um “substantivo” tradicional, e por isso mesmo alguns autores a chamaram “substantivo verbal” (ou “forma nominal do verbo”). Já *correndo* seria antes um tipo de “advérbio”, e assim por diante. Colocando a coisa em termos semânticos, diremos que *correr* tem potencial referencial (é o “nome de uma coisa”), o que *corríamos* e *correndo* não têm.

Concluimos que uma classificação com o objetivo que colocamos para a nossa — a saber, o objetivo de descrever a ordem dos constituintes — deve considerar palavras, e não lexemas. Estritamente falando, *caneta* e *canetas* seriam classificadas separadamente, e com efeito devem sê-lo, se a classificação estiver dirigida para a descrição da concordância nominal, já que sua distribuição difere, ainda que minimamente: um plural não ocorre exatamente nos mesmos ambientes que um singular. Naturalmente, isso é levar o detalhamento a um ponto excessivo; em particular, no caso da descrição da ordem dos termos, não é necessário chegar a tais minúcias.

Isso não quer dizer que seja impossível classificar lexemas. Só que,

pelo que vemos, os objetivos tradicionalmente colocados para a descrição morfossintática não teriam muito o que fazer dessas classes, pois seriam compostas de membros distribucionalmente muito diferentes. De outros pontos de vista, pode fazer sentido classificar lexemas; por exemplo, todas as palavras que compõem o lexema *correr* têm traços semânticos comuns, relacionados com a idéia de “deslocamento rápido”. Essas diferentes classificações não são mutuamente exclusivas pois o sistema de classificação presente na competência do falante-ouvinte é complexo por natureza. Uma forma se associa paradigmaticamente a algumas segundo certos critérios, e a outras segundo outros critérios, ficando essas associações (e conseqüentes classificações) superpostas e simultaneamente disponíveis.

2. As bases da taxonomia

Apesar da muita discussão que tem havido a respeito da noção de “classe” na literatura recente (veja-se, por exemplo, a discussão em torno de protótipos, “squishes” etc.), alguns aspectos ainda estão por explicitar. Vamos expor aqui brevemente alguns pontos que nos parecem básicos.

2.1. Classificando por objetivos

O primeiro deles é que uma classificação só faz sentido se atrelada a um objetivo (no caso, um objetivo descritivo). Isso porque o número de traços distintivos morfossintáticos e semânticos associados a um item léxico é muito grande. Considerando em especial os traços semânticos, se os levarmos todos em conta acabaremos sem dúvida com uma classe para cada item: afinal de contas, *olho* não tem exatamente a mesma matriz semântica de *nariz*. Assim para se fazer uma classificação com base em traços é necessário selecionar os traços que interessam, desprezando os demais.

É o que todo mundo faz, claro. O que falta é uma explicitação dos critérios de seleção dos traços. A maioria das classificações se louva simplesmente na seleção tradicional, que não é preciso criticar aqui. Temos uma proposta que nos parece mais razoável: derivar a seleção dos traços da conveniência descritiva.

No início, nosso horizonte será uma descrição um tanto restrita, ou seja, subordinaremos tudo à conveniência de descrever **a ordem dos termos dentro do sintagma nominal**. Eventualmente, claro, será preciso caminhar para uma descrição muito mais ampla. Não sabemos bem como vai ficar a taxonomia mais adiante, mas no momento o caminho parece claro. Vamos

selecionar como relevantes para a taxonomia todos aqueles traços (semânticos ou morfofossintáticos) que sejam necessários para a descrição da ordem dos termos internos do SN. Isso nos fornecerá um critério não arbitrário para a escolha dos traços.

O resultado será uma classificação na qual **cada classe é composta de elementos cujo comportamento relativo à ordenação no SN é idêntico**. Sustentamos que tais classes devem fazer parte da competência dos falantes, já que estes são capazes de gerar e avaliar corretamente os SNs: de um modo ou de outro, eles classificam os diversos elementos de modo a colocá-los corretamente dentro do SN.

Como se vê, a taxonomia é uma decorrência da análise; ou seja, os elementos que determinam a taxonomia são todos componentes da descrição da ordem dos termos dentro do SN. Por outro lado, há fatores importantes para a descrição da ordem que não podem figurar na taxonomia, por não servirem para distinguir as palavras entre elas. Assim, por exemplo, o caráter restritivo ou não-restritivo de um termo pode, em certos casos, ser correlacionado com a ordem, donde a diferença entre *o bonito barco* (não-restritivo) e *o barco bonito* (restritivo). No entanto, essas duas acepções parecem estar disponíveis a toda e qualquer palavra potencialmente qualificativa, não havendo qualificativas que, por exemplo, não possam ser restritivas. Conseqüentemente, a restritividade não constitui traço distintivo presente no léxico, e não pode servir como uma das bases da taxonomia (o que equivale a dizer que a descrição da ordem dos termos vai além da simples taxonomia das palavras envolvidas).

Além disso, sabe-se que há fatores pragmáticos (textuais) que atuam na determinação da ordem dos termos: foco de contraste, preparação, dadidade (ver Perini *et al.*, 1996, p. 121-125). Tais fatores tampouco podem ser levados em conta no presente estudo, porque não se vinculam a palavras ou a itens léxicos, mas antes a unidades maiores, isto é, ao discurso como um todo, incluindo a situação extra-lingüística. Dentro do universo vocabular que nos interessa, não é possível determinar quais as palavras que podem ser **dadas**, por exemplo. Podemos dizer que o artigo *o* não pode ser dado, mas o efeito da dadidade se observa nos nomes, e é entre os nomes que a oposição *dado/novo* se manifesta. Só que não se trata de **palavras** dadas ou novas, mas de conceitos. Isso vale para os demais fatores de natureza pragmática, de modo que serão deixados de lado neste artigo.

Como se vê, o problema da descrição da ordem dos termos do SN transcende a questão da taxonomia. Não é possível descrever o SN totalmente à maneira dos tagmemicistas, isto é, em termos de uma série de nichos ordenados, cada um deles ocupado por uma classe de elementos.

2.2. Classes e funções

As **classes** são associações paradigmáticas de elementos, e nisso se distinguem das **funções** morfossintáticas ou semânticas. Esse princípio, bastante bem estabelecido, é freqüentemente esquecido na prática da análise, de forma que tanto na gramática tradicional quanto em trabalhos modernos se encontra a cada passo confusões implícitas entre classe e função. A seguir tentaremos explicitar o problema.

Vamos partir de um enunciado qualquer, digamos uma oração como *Assis comprou um gravador japonês*. É possível identificar diversos tipos de relações morfossintáticas ou semânticas que estão realizadas nessa frase; por exemplo, *Assis* é sujeito da frase; e *japonês* tem acepção proventiva (“proveniente do Japão”). Podemos dizer, corretamente, que *Assis* (ou, mais precisamente, o sintagma nominal *Assis*) nessa frase **funciona como** sujeito, e *japonês* **funciona como** proventivo. Note-se que isso só vale para esses itens nessa frase particular; não se pode dizer da palavra *Assis* que é um sujeito (pois pode ser, digamos, objeto direto) nem da palavra *japonês* que é proventiva (pois pode ser agentiva, como em *a invasão japonesa*). Em vez disso, diremos que (como mostra a frase examinada) *Assis* **pode ser** sujeito, e *japonês* **pode ser** proventivo.

Essas potencialidades (“poder ser sujeito”, “poder ser proventivo”), embora sejam depreendidas a partir do exame de enunciados particulares, são o que pode ser atribuído às palavras fora de contexto, isto é, em seu estado de dicionário. Por isso não faz sentido perguntar (sem contexto) se *Assis* é sujeito ou objeto direto; mas faz sentido perguntar (ainda sem contexto) se *Assis* **pode ser** sujeito ou objeto direto. Paralelamente, não se pergunta sem contexto se *japonês* é proventivo (pois pode ter outras acepções), mas pergunta-se se **pode ser** proventivo, agentivo etc.

A idéia é que designações como “sujeito”, “objeto direto” ou “acepção proventiva” se referem a **funções**, ao passo que “poder ser sujeito”, “poder ter acepção proventiva” se referem a traços léxicos que podem ser utilizados como base para uma taxonomia. Em termos saussurianos, relações como

“sujeito de” ou “em acepção proventiva” são sintagmáticas, pois só se definem dentro de um contexto e em dependência deles; ao passo que potencialidades como “poder ser sujeito” e “poder assumir acepção proventiva” são traços paradigmáticos, que relacionam membros de classes cuja associação não aparece nos enunciados, mas na estrutura do léxico, em outras palavras. O que relaciona *Assis* com *comprou* na frase acima é o fato (sintagmático) de que *Assis* é o sujeito de *comprou*; já o que relaciona *Assis* com *Pereira*, fora de qualquer contexto, é o fato (paradigmático) de que ambos os itens têm — no que nos interessa — as mesmas potencialidades morfossintáticas e semânticas.

Uma consequência disso é que não se pode utilizar expressões como “esta palavra está funcionando aqui como adjetivo”, se entendemos “adjetivo” como nome de uma classe. A nosso ver é contraditório afirmar de uma palavra em determinado contexto que “funciona como” membro de uma classe, porque a noção de “classe” é paradigmática, e portanto livre de contexto por definição. O que é um adjetivo, afinal de contas? É simplesmente um item que pode funcionar de determinadas maneiras; se um item está “funcionando como” adjetivo, é um adjetivo, e pronto. Se não é um adjetivo, então não pode funcionar como adjetivo — tudo isso como corolários da própria noção de “classe”.

Quando alguém diz que “esta palavra pode ocorrer como substantivo ou como adjetivo”, provavelmente tem em mente outra coisa: interpretamos essa afirmação como equivalente a “esta palavra pode ocorrer como X ou como Y”, onde X e Y são **funções** sintáticas ou semânticas. Seria bom aplicarmos maior rigor no uso das expressões que usamos, principalmente em um ponto como este, onde a confusão é tão freqüente.

Talvez as coisas fiquem mais claras se estabelecermos de uma vez por todas o que queremos dizer com “adjetivo”: uma função ou uma classe; e se, uma vez estabelecido o valor desse termo, nos mantivermos fiéis a esse valor. Se “adjetivo” é o nome de uma classe, então que fique explícito que se trata de algo que se predica de palavras em seu estado de dicionário, e não de ocorrências de tais palavras em contextos determinados.

Procurando uma terminologia livre de tais confusões, propomos o termo “traço” para designar uma potencialidade léxica — um ou mais traços podem definir uma classe, formada esta pelos elementos que possuem aqueles traços. Para as funções, talvez o melhor fosse utilizar sempre o próprio termo “função”, tradicional em sintaxe; no entanto, em semântica “função” é tão ambíguo que preferimos dizer “acepção”. Assim, diremos da palavra (item léxico) *japonês* que tem os **traços** semânticos <+Proventivo, +Agentivo>, o que significa que **pode**

assumir, segundo o enunciado onde ocorre, as acepções “proventiva” ou “agentiva” (entre outras, evidentemente). Um item pode ter ambos os traços em sua matriz léxica, pois os traços não exprimem mais que potencialidades; mas em um contexto dado o item assume apenas uma acepção cada vez: não é possível construir um enunciado onde *japonês* seja, **ao mesmo tempo**, “proventivo” e “agentivo”. Adotamos também a convenção de colocar os traços entre ângulos, < > (quando são de natureza semântica) ou entre colchetes, [] (quando são morfossintáticos), e as acepções entre aspas, “ ”.

2.3. Classes como feixes de traços

Outro ponto que queremos explicitar aqui é a natureza da classificação baseada em feixes de traços distintivos. Rótulos como “nome”, “preposição”, “verbo” etc., no que pese seu uso quase exclusivo nas análises, só podem ser entendidos como abreviaturas de matrizes de traços. E mais: só podem ser entendidos como termos aproximativos, ou seja, quando classificamos as palavras em dez (ou cem) classes estamos sempre deixando de lado os casos minoritários ou que se consideram, por alguma razão, menos importantes. A única maneira rigorosa de falar das classes de palavras é utilizando o conjunto completo dos traços relevantes.

3. Exemplos

Vamos examinar mais detalhadamente a classificação das seguintes palavras: *mau*, *ruim*, *caneta* e *paternal*.

Em primeiro lugar é preciso definir os traços a serem utilizados na classificação desses itens. A pesquisa relatada em Perini *et al.*, 1996 revelou que alguns traços semânticos são relevantes para a determinação das possibilidades de ordenamento dos termos no SN; por exemplo, <Qualificativo> e <Referencial>, entre outros; aqui vamos limitar-nos a esses dois, para simplificar a discussão. Segundo a análise proposta no trabalho citado, a acepção “qualificativa” está disponível a itens que aparecem tanto antes quanto depois do núcleo do SN; a acepção “referencial” só se realiza (por definição) no próprio núcleo.

Analisando as quatro palavras escolhidas em termos desses dois traços, teremos:

<i>mau</i>	<+Q. +R>
<i>ruim</i>	<+Q. +R>
<i>caneta</i>	<-Q. +R>
<i>paternal</i>	<+Q. -R>

As quatro palavras examinadas se distribuem, em um primeiro momento, em três classes, definidas pelas matrizes de traços <+Q, +R>, <-Q, +R> e <+Q, -R>. Esses traços permitem descrever as possibilidades posicionais de cada palavra, através de certas regras incluídas na análise; essas regras especificam que um elemento marcado <+Q> pode ocorrer antes ou depois do núcleo; um elemento marcado <-Q> não pode ocorrer antes do núcleo; e somente elementos marcados <+R> podem ocorrer como núcleos. Isso prevê as seguintes possibilidades:

mau pode ocorrer como núcleo; e pode ainda ocorrer posposto ou anteposto ao núcleo;

ruim pode ocorrer nas mesmas posições que *mau*;

caneta só pode ocorrer como núcleo;

paternal pode ocorrer posposto ou anteposto ao núcleo, mas não como o próprio núcleo.

Essas previsões são confirmadas para três das palavras em exame. Mas no caso de *ruim* há uma falha, porque essa palavra não pode ocorrer anteposta ao núcleo: *um livro ruim* / **um ruim livro*. E não se conhece nenhum traço semântico de *ruim* que possa ser responsabilizado por esse fato.

Por conseguinte, marcaremos *ruim* como uma exceção à regra que estipula que os elementos marcados <+Q> podem ocorrer antepostos. Ou seja, temos que reconhecer que as possibilidades de ordenação não são inteiramente determinadas pela semântica dos diversos itens envolvidos. Desse modo, admitiremos a necessidade de marcar os itens (ou alguns deles) com traços idiossincráticos, que os identificam como exceções; no caso, marcaremos *ruim* com o traço [-Ant], o que quer dizer que essa palavra é uma exceção à regra ou regras que estabelecem as possibilidades de anteposição. Assim, *ruim* terá a matriz <+Q, +R>, [-Ant], diferenciando-se de *mau*, que é <+Q, +R>, [+Ant]¹.

O que temos em mãos, portanto, é certo número de traços distintivos, cada um dos quais corta o universo dos itens léxicos em duas classes. Os traços, em seu conjunto, definem diversos cortes, e portanto diversas classes;

¹Na verdade há redundâncias nessa notação, mas isso não precisa preocupar-nos no atual estágio da investigação. Por exemplo, não é realmente necessário especificar que *mau* é [+Ant], pois isso já é determinado pela regra mencionada, que estabelece que um item <+Q> pode ser anteposto. Discutir esse tipo de questão só faria sentido se estivéssemos empenhados em desenvolver uma notação formalizada e maximamente econômica, o que nos parece prematuro. No momento trata-se de observar e sistematizar dados, não de construir uma teoria geral; em outras palavras, nosso objetivo é descritivo, não explicativo, e por isso não nos ocuparemos de eliminar eventuais redundâncias da notação utilizada.

como vimos, considerando os quatro itens *mau*, *ruim*, *caneta* e *paternal* e os três traços <Q>, <R> e [Ant], definimos quatro classes. Cada classe compreende certo número de itens cujo comportamento relativo à ordenação dentro do SN é semelhante.

4. A taxonomia e a análise

Essa definição de traços tem uma relação muito íntima com a análise propriamente dita da ordem dos termos no SN. Pode-se perguntar se é a análise que deriva das classes ou as classes que derivam da análise; mas essa pergunta não faz muito sentido, e não pode ser respondida de maneira direta.

Acontece que uma análise é em grande parte feita de generalizações, e as generalizações implicam na postulação de classes; nesse sentido, construir a análise e estabelecer classes são tarefas concomitantes e interdependentes. Digamos que se parta da observação de fatos: o item *mau* pode ocorrer tanto antes quanto depois do núcleo. Outra observação particular é que *mau* possui o traço semântico <+Q>, ou seja, pode ter aceção qualificativa.

Aqui estamos ainda em um estágio de generalização muito incipiente, e no que diz respeito a *mau* não temos mais que duas afirmações particularizadas. Teremos uma análise a partir do momento em que tentarmos uma generalização. Por exemplo, podemos dizer que *mau* pode ocorrer anteposto ou posposto **porque** tem o traço <+Q>. Isso equivale a dizer que “todo item marcado <+Q> pode ocorrer anteposto ou posposto ao núcleo” — e essa afirmação já é mais do que uma simples observação. Engloba uma generalização, e portanto já é uma análise (ou parte de uma análise).

Mas mesmo para formular essa generalização somos obrigados a estabelecer classes: a afirmação “pode ocorrer anteposto ou posposto ao núcleo” se predica não de um item léxico, mas de uma classe de itens, ou seja, de todos os itens marcados <+Q>. O estabelecimento de classes é essencial para a formulação das generalizações, e portanto para a elaboração da análise. E, por outro lado, as classes só fazem sentido dentro de uma análise: senão como justificariamos a escolha do traço <Q> e não, por exemplo, um traço que exprimisse o caráter “desfavorável” ou “desagradável” da semântica de *mau*? Naturalmente, porque o traço <Q> é relevante para a descrição das possibilidades de ordenação da palavra *mau*, ao passo que o fato de que essa palavra significa uma coisa desagradável não tem importância para a descrição das suas possibilidades de ordenação.

Assim, não se pode dizer que a taxonomia precede a análise, nem vice-versa; nem sequer que uma “depende” da outra. Antes, a taxonomia é um dos aspectos da análise. A análise da ordem dos termos do SN lança mão de fatores independentes de qualquer taxonomia léxica (notadamente fatores de ordem textual e pragmática); por isso não se pode identificar a taxonomia com a análise. Mas a taxonomia é um ingrediente básico da análise, já que se observa que certos tipos de itens, definíveis em termos de suas propriedades semânticas ou formais, se comportam de maneira especial do ponto de vista da ordenação.

Voltemos agora à questão do traço formal [-Ant] que, como vimos, identifica o item *ruim* como exceção. Esse traço, ao caracterizar um item segundo seu comportamento dentro do SN, automaticamente estabelece uma nova classe: dentre os itens marcados <+Q> é necessário distinguir dois grupos: aqueles que são proibidos de ocorrer antepostos e aqueles que não sofrem essa restrição. Isso é parte legítima da taxonomia, pois se justifica em termos das conveniências da descrição. O fato de uma das classes definidas por [-Ant] ser muito menos extensa do que a outra não afeta a questão em si; apenas nos autoriza a chamar esses casos minoritários de “exceções”.

Assim, uma pessoa que aprende o item *mau* precisa ficar sabendo uma série de coisas, entre as quais: (a) esse item tem os traços <+Q, +R>; e (b) esse item não é exceção à regra que autoriza os itens <+Q> a ocorrerem antepostos ao núcleo. Naturalmente, certas estratégias de aprendizagem oferecem atalhos, como considerar que um item não é exceção até prova em contrário. Mas o resultado final é o mesmo: as exceções se classificam à parte (dentro da classe mais geral a que pertencem), o que automaticamente coloca as não-exceções também à parte. Consideramos pois uma classe o conjunto de todos os itens cujo comportamento relativo à ordenação no SN seja idêntico, ou seja, que tenham exatamente a mesma matriz de traços — compreendidos aí os traços idiossincráticos do tipo [Ant].

5. Representação das classes em traços

As classes, como vimos, se definem através de matrizes de traços; e esses traços são em parte semânticos, em parte formais (isto é, morfossintáticos). Há uma diferença nítida entre traços semânticos e formais: estes últimos muitas vezes são **ad hoc**, pois não encontram motivação independente em outros setores da gramática; é o caso do traço [Ant], com o qual estamos lidando. Já os traços semânticos não são **ad hoc**: são essenciais não apenas para efeitos de descrição da ordem dos termos no SN, mas ainda

para a caracterização do significado propriamente dito das palavras. Assim, por exemplo, a marca <-Q> no item *caneta* nos informa que esse item não ocorre anteposto ao núcleo, e também que *caneta* não pode exprimir uma qualidade atribuível a um objeto.

A importância dessa distinção vem de que um traço semântico tem, por assim dizer, existência independente, sendo essencial para descrever um **fato**, a saber, o significado de um item. Mesmo se não for relevante para outros propósitos, o traço semântico precisa figurar na análise porque senão certos itens ficarão com sua caracterização semântica incompleta. Já um traço formal tira toda a sua razão de ser de sua utilidade gramatical.

6. Sobre a natureza dos traços

Há um aspecto do comportamento gramatical dos itens que precisa ser sempre levado em conta, que é sua flexibilidade categorial, ou seja, sua grande capacidade de mudar de classe (semântica e/ou morfossintática) segundo as necessidades expressivas do falante. Essa flexibilidade categorial levanta certos problemas para a análise.

Voltemos ao item *caneta*: foi proposto acima que esse item seja marcado, no léxico, com o traço <-Q>, significando que não pode ocorrer modificando (ou qualificando) o núcleo: não tem potencial qualificativo. No entanto, nada impede, em princípio, que a partir de amanhã *caneta* comece a ser usado qualificativamente. É o que acontece às vezes com itens originalmente <-Q>, como *cabeça*, que passou recentemente a ser utilizado (em certa variedade coloquial) qualificativamente: *um filme cabeça* (um filme inteligente). Como é que isso pode acontecer, se o termo era <-Q>, e portanto não tinha potencial qualificativo? Não estaríamos, com essa marca, negando a possibilidade de uma coisa que é comprovadamente possível?

A pergunta tem razão de ser, e pode ser respondida lançando-se mão da distinção entre sincronia e diacronia.

Ao se falar de “potencial” deve-se distinguir duas noções: primeiro, o potencial que um item possui de ser utilizado de determinada maneira **sem que seu uso cause efeito de inovação** (coisas como: efeito humorístico, variedade ou dificuldade de interpretação etc.). Nesse sentido, *caneta* é certamente destituído do potencial expresso pelo traço <Q>, devendo ser marcado <-Q>; e é nesse sentido que a noção de “potencial” nos interessa.

Já a noção diacrônica de “potencial” se refere à possibilidade de um item ser utilizado de determinada maneira **como inovação**. Assim, hoje, o uso de *caneta* como qualificativo é em princípio possível, mas constituirá inovação, acarretando os efeitos mencionados. Foi o que ocorreu com *cabeça* as primeiras vezes que se falou de *papo cabeça*, *filme cabeça* etc. A diferença está em que a primeira noção de “potencial” se refere ao armazenamento de uma possibilidade como parte da convenção lingüística vigente (o que chamamos “língua portuguesa”); a segunda noção se refere às possibilidades de mudança dessa convenção.

Embora não nos interesse neste momento, a noção diacrônica tem relevância e distingue à sua moda os itens léxicos entre eles. Assim, *caneta* pode, concebivelmente, passar a ser utilizado qualificativamente; mas certamente o mesmo não vale para *corríamos*, *sempre* ou *em*. Por isso, é importante distinguir o potencial **de uso** (sincrônico) do potencial **de mudança** (diacrônico); aqui nos interessa o primeiro.

Essa posição tem corolários para a interpretação do que, precisamente, nos dizem os traços utilizados na análise. Assim, quando utilizamos um traço qualquer, digamos <+T>, para marcar um item, estamos dizendo o seguinte:

“esse item (ou, mais exatamente, essa **palavra**; ver a seção 2) pode ser utilizado como “T”, e esse uso é aceito pela comunidade como parte da convenção, e não como uma tentativa de modificar a convenção.”

7. Substantivos e adjetivos

7.1. O núcleo do SN

Vamos exemplificar a aplicação dos princípios discutidos ao caso dos elementos habitualmente colocados sob o rótulo de “adjetivos” e “substantivos”. Para isso começaremos examinando a noção de “núcleo do sintagma nominal” — uma noção que se encontra na base da distinção entre adjetivos e substantivos, mas que não está definida de maneira satisfatória na literatura a que tivemos acesso (por exemplo, nos artigos reunidos em Corbett *et al.* (1993). Aqui apresentaremos a solução proposta em Perini *et al.* (1996).

Mostrou-se nesse texto que a definição de base puramente posicional, como por exemplo a proposta em Perini (1995), não funciona. A razão principal

é que, ao se tentar definir o núcleo posicionalmente, tanto os termos que precedem quanto os que seguem o presumível núcleo podem ocorrer repetidos. Assim, qualquer definição dependente da posição (digamos, “o terceiro termo a partir do final do SN”) esbarra no fato de que não se pode saber de antemão quantos termos repetidos estão presentes em um SN particular. Não obstante, estamos convencidos de que é necessário atacar o problema da definição do núcleo, a fim de permitir a própria formulação das questões de ordenação dos termos: se o objetivo da análise é, por exemplo, discutir a posição do adjetivo no SN, essa posição será definida em relação a quê?

A conclusão a que se chegou é que o núcleo do SN não é uma entidade definível **formalmente** (a partir de sua posição no sintagma, ou de outros fatores, como o controle da concordância); antes, trata-se de uma **função semântica**, a saber, o núcleo é o termo do SN que está tomado em acepção **referencial** — ou seja, como “designação de uma coisa”.

As bases que permitem ao ouvinte determinar qual dos termos de um SN veicula a acepção referencial estão expostas em Perini *et al.* (1996: 75 sqq). Aqui damos apenas um exemplo: seja o sintagma *o carro amarelo*: a palavra *amarelo*, em virtude de sua matriz semântica armazenada no léxico, pode ser referencial (como em *o amarelo está na moda*) ou então qualificativa (*casa amarela*). Mas *carro* só pode ser referencial. Portanto, em *o carro amarelo* a palavra *carro* é referencial (“R”) e *amarelo* qualificativo (“Q”). Há razões para crer que em cada SN há sempre um centro de referência, e apenas um; no caso em pauta, é *carro*, e é esse o núcleo.

Isso explica, entre outras coisas, a facilidade que têm os ouvintes de identificar o núcleo de um sintagma, já que essa identificação é resultado direto da própria compreensão do sintagma. Seria pouco plausível argumentar que os ouvintes encontram o núcleo a partir de um cômputo da posição do elemento dentro do sintagma.

7.2. Adjetivos e substantivos

A análise do núcleo do SN resumida acima acaba inviabilizando a definição formal de “adjetivo” e “substantivo”. Acontece que a definição formal se basearia no potencial funcional de cada palavra: o conjunto de funções sintáticas que cada uma pode ocupar. Por exemplo, diríamos que *paternal* é um “adjetivo” porque não pode ser núcleo do SN, e *João* é “substantivo” porque pode. Se o núcleo do SN pudesse ser definido em termos formais —

digamos, através de sua posição exclusiva dentro do sintagma — teríamos que “adjetivos” e “substantivos” seriam classes morfossintáticas distintas.

Mas a partir do momento em que se nega a diferenciação morfossintática entre o núcleo e os demais termos do SN (em particular os modificadores realizados por “adjetivos” tradicionais), torna-se impossível aplicar essa definição: os “adjetivos” e os “substantivos” tradicionais têm o mesmo potencial funcional sintático, isto é, ocorrem no mesmo conjunto de funções sintáticas. Logo, pertencem, formalmente falando, a uma única classe. O que distingue esses tipos de palavras são suas potencialidades **semânticas**: *paternal* se distingue de *João* por não ocorrer em acepção referencial (*paternal* não é “o nome de uma coisa”). As potencialidades formais observadas desses itens são decorrência automática de suas potencialidades semânticas: cada item tem este ou aquele comportamento formal em virtude de significar isto ou aquilo².

De certo modo isso nos leva de volta à intuição tradicional de que o substantivo seria a palavra que nomeia as coisas. Mas nossa análise se diferencia da tradicional em pontos cruciais: negamos que o substantivo e o adjetivo existam como classes autônomas, morfossintaticamente caracterizáveis. Em vez disso, a diferença tradicionalmente percebida entre “adjetivo” e “substantivo” se interpreta como uma diferença entre palavras que podem ser nomes de coisas (isto é, que podem ter acepção referencial; ou ainda, que são marcadas <+R>) e palavras que não podem.

Essa posição permite capturar um fato extremamente importante, mas que não se encaixa na análise tradicional: a existência de grande número de itens que podem ocorrer em acepção referencial ou qualificativa, como *amigo*:

Meu melhor **amigo** [acepção referencial]

Um gesto **amigo** [acepção qualificativa]

Qualquer sistema que só considere duas alternativas (substantivo X adjetivo) precisa deixar de lado a diferença entre *paternal*, *João* e *amigo*, porque aqui temos **três** tipos de comportamento distinto (semântico e morfossintático).

² Essa não é uma afirmação de valor geral; pode haver, e certamente há, casos em que traços puramente formais distinguem grupos de palavras em classes diferentes. Por exemplo, não vemos maneira de caracterizar semanticamente os substantivos masculinos face aos femininos.

A presente proposta descreve os fatos a partir de fatores semânticos inevitáveis, que precisam de qualquer maneira figurar na análise: *João* é um nome de coisa, e só isso; *paternal* é uma qualidade, e só isso; *amigo* pode ser as duas coisas. A presença desses traços na potencialidade semântica dessas palavras é algo que consideramos um fato, não uma decorrência desta ou daquela teoria.

8. Conseqüências para a taxonomia

8.1. Princípios

As conseqüências desta análise para a questão da classificação das palavras são óbvias: morfossintaticamente, adjetivos e substantivos não se distinguem. Teremos que colocá-los juntos em uma só classe formal, à qual podemos chamar **nomes** (segundo Camara, 1970).

É preciso observar, contudo, que isso não significa que não haja distinções morfossintáticas entre as palavras em geral. Certamente a diferença entre um verbo e um substantivo não pode ser reduzida totalmente a fatores de significado: o verbo se conjuga, ocupa uma função sintática própria etc. Mesmo dentro do SN, muitos elementos poderão ser colocados em classes à parte por razões formais. Por exemplo, não se conhece nenhuma razão semântica para que o artigo ocorra em sua posição característica à cabeça do sintagma. Logo, o **artigo** precisa (pelo que sabemos hoje) ser segregado da classe dos nomes, formando uma classe que se define por um comportamento sintático próprio.

O comportamento dos diversos elementos do SN (exceto o dos nomes) está ainda pouco estudado. Uma tarefa que se coloca de imediato é a de investigar quantos e quais tipos de comportamento formal se podem distinguir dentre os elementos que comparecem no SN. É provável que palavras como *todos, o, meu, outro, cinco* etc. não possam ser incluídas dentro da classe dos nomes. A se confirmar essa hipótese, elas escapariam às regras que governam a posição dos nomes, exigindo regras próprias. Uma conseqüência seria que elas teriam que se colocar em classes diferentes da dos nomes.

8.2. Os nomes

O caso dos nomes, no momento, fica assim:

- (a) parece que existe uma classe, morfossintaticamente distinguível, que

englobaria *grosso modo* as classes tradicionais dos “adjetivos” mais os “substantivos” (com o provável acréscimo de alguns “pronomes”). Os membros dessa classe se caracterizam por sua distribuição: ocorrem na parte final (a “área direita”) do sintagma. A idéia é que a classe dos nomes poderia ser definida posicionalmente. Ainda *grosso modo*, os nomes correspondem aos elementos “lexicais” mencionados na literatura tradicional; os outros seriam os “funcionais”. Evitamos utilizar essa nomenclatura porque nos parece apriorística e baseada em critérios mal delimitados; mas por trás dela há uma intuição que pode ser válida;

- (b) entre os nomes existe uma grande variedade de traços semânticos. Alguns desses traços semânticos são relevantes para o posicionamento dos itens, outros não. Assim, por exemplo, o item *João* denota um objeto concreto, e o item *santidade* um objeto abstrato; mas essa oposição (concreto/abstrato) não funciona na determinação da posição dos itens dentro do SN. Desse ponto de vista, *João* e *santidade* são idênticos, e sua oposição não interessa para efeitos de taxonomia. Por outro lado, o item *japonês* pode ser proventivo (nascido ou proveniente do Japão) ou agentivo (praticado pelos japoneses); já o item *violento* não pode ser nem proventivo nem agentivo, mas pode ser qualificativo. Acontece que essa diferença se reflete nas possibilidades de posicionamento: as acepções agentiva e proventiva só são disponíveis em posição posposta, de maneira que se antepusermos *japonês* o resultado será mal formado (**um japonês carro*, **a japonesa decisão*) porque o item anteposto fica sem acepção possível. Já *violento* pode ocorrer anteposto ou posposto, porque a acepção qualificativa é disponível nas duas posições; donde serem bem formados tanto *temporal violento* quanto *violento temporal*.

Simplificando bastante, a situação é essa. Temos taxonomias nos níveis morfossintático e semântico. E, a partir do momento em que decidimos descrever a ordem, não podemos escapar de nenhum desses níveis: a descrição da ordem dos termos no SN precisa ter uma cara morfossintática e uma cara semântica (além de outras caras, como a funcional, da qual não nos ocupamos neste trabalho).

8.3. A taxonomia que buscamos

Quando falamos de “nomes”, por exemplo, estamos nos referindo a uma classe definida formalmente. E podemos deixar escapar alguma referência a uma subclassificação dos nomes segundo critérios semânticos: alguns podem ser agentivos, outros não etc. Essa maneira de falar pode sugerir que no

fundo a taxonomia que buscamos seria essencialmente formal: os traços morfossintáticos definiriam as classes, depois os semânticos definiriam as subclasses.

Mas não é isso. Rótulos como “nome” (ou, digamos, “verbo”) são apenas abreviaturas de matrizes de traços, e representam aproximações. A única maneira rigorosa de se falar da classificação das palavras é utilizando diretamente as matrizes. Assim, uma palavra como *João* se distingue de *sem* através de certos traços, e de *amigo* através de outros traços. O fato de que no primeiro caso os traços são (também) formais, e no segundo só semânticos, não precisa ser colocado em primeiro plano. Vamos entender a taxonomia como resultado de traços gramaticais, alguns formais e outros semânticos, sem hierarquia de tipos.

REFERÊNCIAS BIBLIOGRÁFICAS

- CAMARA, J. M. Jr (1970) *Estrutura da Língua Portuguesa*. Petrópolis: Vozes.
- CORBETT, G. G.; N. M FRASER. & S. MCGLASHAN. (orgs.) (1993) *Heads in Grammatical Theory*. Cambridge, England: Cambridge University Press.
- PERINI, M. A. (1995) *Gramática Descritiva do Português*. São Paulo: Ática.
- _____ et.alli (1996) O Sintagma Nominal em Português: Estrutura, Significado e Função. da *Revista de Estudos da Linguagem n° Especial*. B. Horizonte: UFMG.