

Lingüística

2.1 Ramificações da lingüística

Como vimos, tanto a linguagem quanto as línguas podem ser estudadas sob diferentes pontos de vista. Portanto, o campo total da lingüística pode ser dividido em diversos subcampos segundo o ponto de vista adotado ou a ênfase especial dada a um conjunto de fenômenos, ou premissas, ao invés de outro.

A primeira distinção a se estabelecer é entre a lingüística geral e a descritiva. É bastante direta em si mesma. Corresponde à que existe entre estudar a linguagem e descrever determinadas línguas. A pergunta "O que é a língua(gem)?" que, no capítulo anterior, dissemos ser a indagação central e definidora de toda a disciplina, é mais adequadamente considerada a indagação central da lingüística geral. A lingüística geral e a descritiva não são absolutamente estanques. Cada uma depende explícita ou implicitamente da outra: a lingüística geral fornece conceitos e categorias em termos dos quais as línguas serão analisadas; a lingüística descritiva, por sua vez, fornece dados que confirmam ou refutam as proposições e teorias colocadas pela lingüística geral. Por exemplo, o lingüista geral poderia formular a hipótese de que todas as línguas possuem nomes e verbos. O lingüista descritivo poderia refutá-la com base em uma comprovação empírica de que houvesse pelo menos uma língua em cuja descrição tal distinção não se verificasse. Porém, para refutar ou confirmar a hipótese, o lingüista descritivo deve operar com determinados conceitos como "nome" e "verbo" que lhe foram fornecidos pelo lingüista geral.

Evidentemente há toda uma gama de motivos pelos quais se desejaria descrever uma certa língua. Muitos dos que trabalham com lingüística descritiva o fazem não porque estão interessados em fornecer dados para a lingüística geral, ou em testar teorias e hipóteses conflitantes, mas porque querem apresentar uma gramática de referência ou um dicionário para fins práticos. No entanto, tal fato não afeta a interdependência dos dois subcampos complementares da lingüística, a geral e a descritiva.

Durante o século XIX os lingüistas ocuparam-se muito da investigação sobre os detalhes do desenvolvimento histórico de línguas específicas

e da formulação de hipóteses gerais sobre as mudanças das línguas. O ramo da disciplina que trata destes assuntos é hoje conhecido, naturalmente, por lingüística histórica. É óbvio que na lingüística histórica, como na não-histórica, o interesse pode estar na linguagem em geral ou nas línguas em particular. É conveniente mencionar agora os termos 'diacrônica' e 'sincrônica', mais técnicos para o caso. Foram inicialmente usados por Saussure (cuja distinção entre 'langue' e 'parole' já foi mencionada no capítulo anterior). Uma descrição diacrônica de uma língua percorre o desenvolvimento histórico da mesma e registra as mudanças que nela ocorreram entre pontos sucessivos no tempo: "diacrônico", portanto, é equivalente a "histórico". A descrição sincrônica é não-histórica: apresenta uma imagem da língua tal qual ela se encontra em determinado ponto no tempo.

Uma terceira dicotomia é a que se verifica entre lingüística teórica e aplicada. Essencialmente, a lingüística teórica estuda a linguagem e as línguas com vistas a construir uma teoria de sua estrutura e funções, independente de quaisquer aplicações práticas que a investigação da linguagem e das línguas possa ter, enquanto que a aplicada ocupa-se da aplicação dos conceitos e descobertas da lingüística a uma série de tarefas práticas, inclusive o ensino de línguas. Em princípio, a distinção entre lingüística teórica e aplicada independe das outras duas até agora estabelecidas. Na prática, pouca diferença há entre os termos "lingüística teórica" e "lingüística geral": a maioria dos estudiosos que utilizam o termo "lingüística teórica" parte do pressuposto de que o objetivo desta é formular uma teoria satisfatória da estrutura da linguagem em geral. No que diz respeito à lingüística aplicada, fica claro que ela se vale tanto do aspecto geral quanto do descritivo.

A quarta e última dicotomia refere-se a uma visão mais estreita ou mais ampla do escopo do assunto. Não há uma distinção terminológica universalmente aceita para tal caso: usaremos os termos 'microlingüística' e 'macrolingüística', dizendo que na microlingüística adota-se a visão mais estreita, e na macrolingüística a mais ampla. De forma mais estreita, a microlingüística trata unicamente de sistemas lingüísticos, sem contar a forma pela qual as línguas são adquiridas, armazenadas no cérebro ou utilizadas em suas várias funções; sem contar a interdependência entre língua e cultura; sem contar os mecanismos fisiológicos e psicológicos envolvidos no comportamento lingüístico; em suma, independente de tudo que não seja um sistema lingüístico, considerado (como Saussure, ou melhor, seus editores, diriam) em si e por si. De forma mais ampla, a macrolingüística trata de tudo o que é pertinente, de alguma forma, seja qual for, à linguagem ou às línguas.

Uma vez que, além da lingüística, muitas outras disciplinas se ocupam da lingua(gem), não causa surpresa que várias áreas interdisciplinares se

tenham identificado com a macrolingüística, recebendo nomes distintivos: sociolingüística, psicolingüística, etnolingüística, estilística etc.

Um ponto que se deve enfatizar é que a distinção entre microlingüística e macrolingüística depende da diferenciação entre lingüística teórica e aplicada. Há, em princípio, um aspecto teórico em cada ramificação da macrolingüística. De tal maneira que em certas áreas da lingüística aplicada, como o ensino de línguas, é essencial adotar-se a visão mais ampla da estrutura e funções das línguas, ao invés da mais estreita. É por isso que alguns autores incorporaram o que aqui denominamos macrolingüística à lingüística aplicada.

Examinaremos algumas áreas da macrolingüística em capítulos posteriores. Poder-se-ia pensar que, em face da reconhecida importância da língua em tantas disciplinas, a lingüística deveria adotar a visão mais abrangente possível de seu objeto de estudo. Em certo sentido é verdade. O problema é que ainda não há, e pode ser que não haja nunca, um quadro teórico satisfatório dentro do qual se possa ver a lingüística, simultaneamente, sob o ponto de vista psicológico, sociológico, cultural, estético e neurofisiológico (para não falar em outros igualmente relevantes). A maior parte dos lingüistas defende hoje que é a microlingüística sincrônica teórica que constitui o cerne de sua disciplina, dando-lhe toda unidade e coerência que tem. Praticamente metade deste livro será dedicada a este núcleo central; o restante tratará da lingüística histórica e de áreas selecionadas da macrolingüística.

2.2 A lingüística é uma ciência?

A lingüística é normalmente definida como ciência da linguagem, ou alternativamente, como estudo científico da linguagem (v. 1.1). O próprio fato de que há uma seção neste livro e em outras introduções à lingüística, explicitamente dedicada à discussão do *status* científico desta disciplina não deve deixar de ser comentado. Afinal, disciplinas cujo *status* científico é inquestionável — a física, a química, a biologia etc. — não têm necessidade de justificar sua reivindicação de se chamar ciência. Por que deveria a lingüística preocupar-se tanto em defender a validade de seu título? E por que, ao defender suas credenciais científicas, o lingüista tantas vezes dá a impressão de protestar em demasia? O leitor tem todo o direito de levantar suspeita.

O primeiro ponto a se salientar é que a palavra inglesa '*science*' tem uma abrangência muito menor do que muitas de suas traduções convencionais em outras línguas: tais como '*Wissenschaft*' em alemão, '*nauka*' em russo e mesmo '*science*' em francês. A lingüística, mais do que a maioria das outras disciplinas, sofre das implicações muito específicas das palavras

inglesas 'science' e 'scientific' que se referem, antes de mais nada, às ciências naturais e aos métodos de investigação que lhes são característicos. Isto ainda é verdade, apesar de que expressões como 'ciências sociais', 'ciências do comportamento' e 'ciências humanas' se façam cada vez mais comuns. Deveríamos, pois, interpretar a palavra 'ciência' no título desta seção como simplesmente 'disciplina acadêmica devidamente constituída'?

Há bem mais nesta questão do que nos sugeria tal interpretação. A maioria dos lingüistas que adotam a definição de sua disciplina como sendo o estudo científico da língua o fazem por ter em mente uma distinção entre uma forma científica e acientífica de realizar a tarefa. Podem discordar sobre algumas das implicações da palavra 'científico', como os filósofos e historiadores da ciência. Mas concordam unanimemente sobre as principais diferenças entre o estudo científico da língua e o estudo não-científico. Começamos, então, com os pontos unânimes.

O primeiro e mais importante deles é que a lingüística é empírica, ao invés de especulativa e intuitiva: opera com dados publicamente verificáveis por meio de observações e experiências. Ser empírica, neste sentido, é para a maioria a própria marca registrada da ciência. Estreitamente relacionada à propriedade de ser empiricamente embasada está a da objetividade. A língua é algo que normalmente não nos preocupa: algo familiar desde a infância, de uma maneira prática e irrefletida. Tal familiaridade prática com a língua tende a representar uma barreira para um exame objetivo. Há diversos tipos de preconceitos sociais, culturais e nacionalistas associados à visão leiga da linguagem e das línguas. Por exemplo, um sotaque ou dialeto de determinada língua pode ser considerado inerentemente mais puro que outro; ou, uma vez mais, certa língua pode ser tida por mais primitiva que outra. A objetividade exige, no mínimo, que se lance um desafio contra tais concepções e que termos como 'puro' e 'primitivo' sejam claramente definidos ou abandonados.

Muitas das concepções sobre a língua que o lingüista coloca em questão, se é que não as abandona totalmente, podem parecer um princípio de mero bom senso. Porém, como observou Bloomfield (1935:3), a respeito do bom senso na abordagem de questões lingüísticas: 'como muito do que se disfarça em bom senso, é na realidade altamente sofisticada e provém, sem grande distanciamento, das especulações dos filósofos da Antiguidade e da Idade Média'. Nem todos os lingüistas têm uma opinião tão desfavorável sobre estas especulações filosóficas da linguagem quanto Bloomfield. Mas seu parecer geral é válido. Os termos usados pelos leigos para falar sobre a língua, bem como suas atitudes em relação a ela, têm a sua história. Muitas vezes pareceriam menos obviamente aplicáveis ou evidentes se soubessem algo sobre suas origens históricas.

Não entraremos neste livro pelos caminhos históricos da lingüística. Alguns comentários gerais, em contrapartida, são pertinentes. As intro-

duções à lingüística costumam traçar uma distinção nítida entre gramática tradicional e lingüística moderna, contrastando o status científico desta com o status não-científico daquela. Há boas razões para haver esta diferenciação e para ressaltar que muitas das concepções erradas sobre a língua, populares em nossa sociedade, podem-se explicar, historicamente, em termos das premissas filosóficas e culturais que determinaram o desenvolvimento da gramática tradicional. Algumas destas concepções serão apontadas e discutidas na seção seguinte. Deve-se frisar, entretanto, que a lingüística, como qualquer outra disciplina, constrói sobre o passado, não só desafiando e refutando doutrinas tradicionais, mas também desenvolvendo-as e reformulando-as. Muitos trabalhos recentes, descrevendo os grandes avanços na investigação científica da linguagem feitos nos últimos cem anos, mais ou menos, deixaram de enfatizar a continuidade da teoria lingüística ocidental desde seus primórdios até os dias de hoje. Muitas vezes foram também anacrônicos, por não tratar a gramática tradicional em termos dos objetivos que ela estabeleceu para si mesma. Não se deve esquecer que os termos 'ciência' e 'científico' (ou seus precursores) foram concebidos de forma diferente em diferentes épocas.

Também deveria ser salientado que o que geralmente se conhece por 'gramática tradicional' — ou seja, a teoria lingüística ocidental voltando pela Renascença e pela Idade Média até a erudição romana e, antes dela, a grega — é muito mais rica e variada do que normalmente se supõe. Além do mais, inúmeras vezes o que se ensinou foi uma versão equívoca e distorcida da gramática tradicional, para muitas gerações de alunos relutantes e desinteressados. Ultimamente os lingüistas vêm adotando uma perspectiva mais equilibrada quanto à contribuição que a gramática tradicional — continuaremos com este termo — vem prestando ao desenvolvimento de sua disciplina. Há ainda muito a se pesquisar sobre as fontes originais que sobreviveram dos períodos mais antigos. Porém há hoje várias histórias da lingüística dando conta mais satisfatoriamente dos fundamentos e desenvolvimento da gramática tradicional, mais do que havia à disposição da geração de Bloomfield e de seus sucessores imediatos.

Voltemos, pois, ao estado atual da lingüística. Sem dúvida alguma ela é hoje mais empírica e objetiva, em suas atitudes e premissas expressas, pelo menos, do que a gramática tradicional. Examinaremos algumas de tais atitudes e premissas na seção seguinte, de uma forma mais detalhada. Entretanto, será que na prática ela é tão empírica e objetiva quanto afirma ser? Eis aqui, por certo, uma possibilidade de dúvida. Existe também a possibilidade de uma contenda, em nível mais sofisticado de discussão, quanto à natureza da objetividade científica e à aplicabilidade ao estudo da língua daquilo que se conhece comumente como método científico.

Na realidade, cientistas e filósofos da ciência já não aceitam mais tão unanimemente a existência de um único método de investigação, apli-

cável a todos os ramos da ciência. O próprio termo 'método científico' tem uma aura antiquada, até mesmo de século passado. Por vezes é dito que a investigação científica deve necessariamente proceder de uma generalização indutiva com base na observação teoricamente não-controlada. Realmente, isto é o que muitos afirmam estar implicado no termo 'método científico'. Mas poucos são os que, mesmo nas ciências naturais, trabalharam desta forma. Seja qual for o significado do conceito de objetividade científica, ele certamente não implica que o cientista deva evitar a elaboração de teorias e a formulação de hipóteses gerais até que tenha levantado uma quantidade suficiente de dados. Os dados científicos, como diversas vezes foi indicado, não são oferecidos gratuitamente pela experiência, mas sim dela extraídos. Observação acarreta atenção seletiva. Não há observação e coleta de dados que não estejam ligados a uma teoria e a uma hipótese. Para usar uma frase da moda, que surgiu com Popper, a observação, por necessidade e desde o início, pressupõe a teoria.

A frase é sugestiva, mas controvertida. Surgiu como reação ao enfoque fortemente empirista dado à ciência, lançado pelos positivistas lógicos no período que precedeu a Segunda Guerra Mundial. Os estudantes de lingüística deveriam conhecer um pouco o empirismo e o positivismo. Sem este conhecimento — embora não precise ser muito detalhado ou profundo — não se pode esperar que compreendam algumas das questões teóricas e metodológicas que atualmente dividem as escolas lingüísticas. Segue-se uma informação mínima necessária, apresentada, na medida do possível, imparcialmente e sem comprometimento com qualquer um dos lados em áreas de sabida controvérsia. Controvérsias, deveríamos acrescentar, são necessárias não só à lingüística mas a toda a ciência. No entanto, elas têm uma relevância especial para o lingüista, na medida em que certos progressos recentes, na lingüística e na filosofia da linguagem, associados aos trabalhos e às idéias de Chomsky, causaram um impacto considerável na discussão mais geral do empirismo e positivismo, tanto da parte de filósofos e psicólogos, quanto de outros cientistas sociais (v. 7.4).

O empirismo é muito mais do que a adoção de métodos empíricos de verificação e confirmação: portanto, há uma distinção crucial a se traçar entre 'empirista' e 'empírico'. O termo 'empirismo' refere-se à concepção de que todo o conhecimento provém da experiência — o vocábulo grego 'empeiria' significa, aproximadamente, "experiência" — e, mais particularmente, de dados da percepção e dos sentidos. Opõe-se, em uma controvérsia filosófica de longa data, ao 'racionalismo' — do latim 'ratio', que significa, neste contexto, "mente", "intelecto" ou "razão". Os racionalistas enfatizam o papel da mente na aquisição do conhecimento. Sustentam, particularmente, a existência de certos conceitos ou proposições apriorísticas ('a priori' significa, numa interpretação tradicional, "conhecido independentemente da experiência") em função dos quais a mente interpreta

os dados da experiência. Voltaremos a certos aspectos mais específicos da controvérsia entre empirismo e racionalismo ao discutirmos o gerativismo (v. 7.4).

Não é preciso, para nossos fins, estabelecer uma distinção entre empirismo e positivismo. O primeiro apresenta um histórico mais longo e é aliado naturais e estreitamente associados em tudo o que nos interessa no momento. O positivismo repousa sobre a distinção entre os chamados dados positivos da experiência e as especulações transcendentais de diversos tipos. Tem uma visão mais secular e antimetafísica, rejeitando qualquer apelo a entidades não-físicas. O objetivo dos positivistas lógicos do Círculo de Viena era chegar a um sistema único de uma ciência unificada, em que todo o corpo do conhecimento positivo estaria representado, em última análise, como um conjunto de proposições precisamente formuladas.

Dois outros princípios mais específicos eram ainda centrais nesta proposta. O primeiro era o agora famoso princípio de verificação: o princípio de que nenhuma formulação tinha significado a menos que pudesse ser verificada pela observação ou por métodos científicos padronizados, aplicados aos dados fornecidos pela observação. O segundo era o princípio do reducionismo: o princípio de que, entre as ciências, algumas eram mais básicas que outras — a física e a química sendo mais básicas que a biologia, e de que na grande síntese da ciência unificada os conceitos e proposições das ciências mais básicas seriam reduzidos aos conceitos e proposições (isto é, reinterpretados em termos destes) das ciências mais básicas. O reducionismo, ao contrário do princípio de verificação, caracterizou um grupo de cientistas muito mais amplo do que o Círculo de Viena há quarenta anos.

O princípio de verificação foi abandonado (embora tenha tido o seu papel na formação da teoria de condições de verdade do significado: v. 5.6) e o princípio do reducionismo é muito menos aceito por cientistas e filósofos da ciência do que foi quando Bloomfield escreveu seu livro clássico da lingüística em 1933. Menciono aqui a figura de Bloomfield porque, sem causar surpresa, ele era um seguidor convicto do empirismo e do positivismo. O que fica muito claro no segundo capítulo de seu livro. Estava, na realidade, estreitamente ligado ao movimento Unidade da Ciência e aceitava irrestritamente o princípio do reducionismo. Foi Bloomfield mais do que ninguém aquele que estabeleceu para a lingüística, especialmente na América, o ideal de ser verdadeiramente científica. Há, portanto, um legado historicamente explicável de empirismo e positivismo na lingüística.

O reducionismo, e de uma forma mais geral, o positivismo, já não é mais tão atraente para a maior parte dos cientistas como foi antigamente. Hoje em dia é amplamente aceito o fato de que não há um método cientí-

fico único aplicável a todos os campos; de que enfoques diversos devem não só ser tolerados, por questão de necessidade a curto prazo, em diferentes disciplinas, como podem ser justificáveis, também a longo prazo, em virtude de diferenças irreduzíveis nos assuntos tratados. Já desde o século XVII — do tempo de Descartes e Hobbes — houve dúvidas, expressas por alguns filósofos da ciência, sobre o programa positivista que propunha dar conta dos processos mentais em termos dos métodos e conceitos característicos das ciências físicas. Muito da psicologia e da sociologia de nosso século, como também grande parte da lingüística do século XX, é positivista em espírito. Mas nas três disciplinas, e mais obviamente na lingüística, o positivismo foi recentemente atacado por ser inaplicável ou estéril.

Em suma, a questão de uma disciplina ser ou não ser científica já não pode ser, se é que foi um dia, solucionada satisfatoriamente fazendo-se referência ao chamado método científico. Toda ciência bem estabelecida emprega seus próprios construtos característicos e seus próprios métodos de obtenção e interpretação de dados. O que apareceu no capítulo anterior como ficção — o sistema lingüístico — pode ser descrito, em termos cientificamente mais respeitáveis, como **construto teórico**. Pode-se questionar a realidade de tais construtos, assim como se podem questionar os construtos teóricos da física e da bioquímica. É mais proveitoso, no entanto, investigar a respeito de cada construto teórico que se postule qual o objetivo explanatório que está preenchendo em relação aos dados.

Tudo o que acaba de se dizer sobre o empirismo, o positivismo e o atual *status* do chamado método científico pretende ser mais ou menos factual e indiscutível. Agora passaremos a pontos controversos.

O primeiro refere-se à implicação da noção popperiana de observação pressupondo teoria. Há controvérsia no uso do termo 'teoria'. O que Popper tinha em mente, e atacava, era a nítida distinção traçada pelos positivistas lógicos entre observação, que em si era tida por neutra teoricamente, e construção de uma teoria, tida por uma questão de generalização indutiva. Ele estava indubitavelmente certo ao desafiar a nitidez da distinção e, mais especialmente, a posição de que a observação e a coleta de dados podem, e devem, acontecer antes da formulação da hipótese. Comumente ocorre que a seleção dos dados é determinada por alguma hipótese que o cientista deseja testar; e não importa como se chegou a tal hipótese. O fato de que a noção positivista de uma observação e coleta de dados não-seletiva seja inválida não significa que não haja absolutamente distinção entre conceitos pré-teóricos e teóricos. Trata-se de um abuso para com o termo 'teoria' agrupar dentro de seu significado todas as concepções e expectativas prévias com que se enfoca o observável e se faz a seleção. Recorreremos à distinção entre conceitos pré-teóricos e teóricos em vários pontos dos capítulos seguintes; e partiremos do princípio de que a observação, embora

necessariamente seletiva, pode sujeitar-se a controles metodológicos satisfatórios, tanto na lingüística como em outras ciências de base empírica.

Um segundo ponto controverso — que é de especial importância na lingüística atual — refere-se ao papel da intuição e aos problemas metodológicos que surgem a este respeito. O termo 'intuição' leva consigo certas associações corriqueiras bastante desfavoráveis. Tudo o que se quer dizer ao mencionar as intuições de um falante nativo sobre sua própria língua são seus julgamentos espontâneos e naturais sobre a aceitabilidade ou inaceitabilidade de enunciados, sobre sua equivalência ou não, e assim por diante. Houve época em que os lingüistas acreditaram que era em princípio possível fugir à necessidade de pedir a falantes nativos que fizessem julgamentos intuitivos sobre sua língua, simplesmente por coletar-se um *corpus* suficientemente grande de dados naturalmente produzidos, submetendo-os em seguida a uma análise exaustiva e sistemática. Pouquíssimos lingüistas adotam hoje tal posição. Ficou claro que muitos enunciados que ocorrem naturalmente são, por motivos lingüisticamente irrelevantes, inaceitáveis e que também nenhum *corpus*, seja qual for sua extensão, poderá conter exemplos de todos os tipos de enunciados aceitáveis. Mas o recurso a evidências intuitivas continua controverso. E há dois aspectos na questão.

O primeiro diz respeito ao problema de as intuições a que o lingüista se refere serem parte da competência lingüística como tal, do falante nativo. Se for o caso, na definição chomskiana de 'competência' e em sua formulação dos objetivos da lingüística, as próprias intuições tornam-se parte do que qualquer descrição lingüística deve diretamente cobrir. A maioria dos lingüistas provavelmente não gostaria de afirmar que a descrição de uma língua deva tratar a intuição dos falantes nativos como dado. Voltaremos a esta discussão ao falar do gerativismo (v. 7.4).

O segundo aspecto da controvérsia diz respeito à confiabilidade dos julgamentos dos falantes nativos, considerados como relatos ou previsões do comportamento lingüístico seu e de outros falantes. O consenso entre os lingüistas parece ser de que tais julgamentos são altamente, pelo menos em determinadas situações, não-confiáveis. Não só os falantes nativos frequentemente discordam uns dos outros, sobre o que é ou não é aceitável, quando não há qualquer outra razão para crer que falam dialetos diferentes, como também seus próprios julgamentos variam com o tempo. Além disso, acontece muitas vezes de um falante nativo rejeitar um enunciado como inaceitável, a ele apresentado pelo lingüista descritivo, e depois ser ouvido produzindo aquele mesmo enunciado em um contexto natural do uso da língua. No que diz respeito às introspecções que o lingüista tem sobre sua própria língua, são pelo menos tão pouco confiáveis, embora muitas vezes por outras razões, quanto às intuições do leigo. O lingüista pode estar menos preocupado que o leigo com os padrões convencionais do uso correto da língua (por exemplo, admitindo abertamente que ele diz 'It's me' ao

invés de 'It's I').[☆] Porém, seus julgamentos têm uma probabilidade maior de ser distorcidos por sua própria consciência das implicações que têm para uma ou outra questão teórica. As introspecções do lingüista sobre o comportamento lingüístico seu e de outros podem muito bem estar impregnadas de teoria, mesmo se a observação direta da conversação espontânea não estiver.

Há na realidade problemas metodológicos bastante sérios quanto à coleta de dados confiáveis, devido a uma série de questões existentes na lingüística teórica. Mas não são mais sérios que os problemas metodológicos que se interpõem no caminho dos que trabalham com psicologia, sociologia ou, em geral, com qualquer ciência social. E sob certos aspectos o lingüista está em melhores condições do que a maior parte dos cientistas sociais, uma vez que é suficientemente clara a parte pertinente ao comportamento lingüístico, e a parte não pertinente, em tudo o que é observável. Acresça-se que há áreas bastante vastas, na descrição de qualquer língua, para as quais a confiabilidade das intuições dos falantes, ou mesmo das introspecções dos lingüistas, não representam maior problema. Não se deve, portanto, exagerar nos problemas metodológicos que surgem ao longo da pesquisa lingüística.

No parágrafo anterior foi feita uma referência à psicologia, à sociologia e a outras ciências sociais. Muitos lingüistas, talvez a maioria deles, classificariam sua disciplina no grupo das ciências sociais. Mas a lingüística não se presta a uma classificação direta dentro de qualquer divisão da pesquisa acadêmica que tome por fundamental, seja a distinção entre ciência e arte, seja a distinção tripartite entre ciências naturais, sociais e humanas. A utilização cada vez maior de expressões como 'ciências da vida', 'ciências do comportamento', 'ciências humanas' ou 'ciências da terra' indicam que muitas disciplinas sentem a necessidade de um reagrupamento estratégico ou tático que pouco tem a ver com as distinções convencionais. Que a lingüística, como cadeira na Universidade, seja colocada em uma área e não na outra é muito uma questão de conveniência administrativa. Como já se frisou anteriormente, ela mantém ligações com uma gama bastante vasta de disciplinas acadêmicas. Dizer que a lingüística é uma ciência não é negar que, em virtude de seu objeto de estudos, ela esteja estreitamente ligada a disciplinas eminentemente humanistas como a filosofia e a crítica literária.

Nas seções seguintes serão discutidos diversos princípios que geralmente são aceitos pacificamente pelos lingüistas de hoje. Na maior parte das vezes poderão ser tomados como produto do ideal científico de objetividade. Como a lingüística moderna, ao afirmar sua objetividade, tantas

vezes proclamou sua individualidade a este respeito, face à gramática tradicional, eles são frequentemente apresentados em contraste com os princípios que determinaram as atitudes e premissas características do gramático tradicional.

2.3 Terminologia e notação

Toda disciplina dispõe de um vocabulo técnico próprio. A lingüística não é exceção. A maioria dos termos técnicos usados pelos lingüistas aparece durante o seu trabalho e é facilmente compreendida pelos que encaram o assunto com simpatia e sem preconceito.

Muitas vezes se objeta que a terminologia, ou o jargão, da lingüística é desnecessariamente complexo. Por que o lingüista tem tanta facilidade de criar novos termos? Por que não se contenta de falar em sons, palavras e partes da fala, ao invés de inventar termos técnicos novos como 'fonema', 'morfema' e 'classe formal'? A resposta é que a maior parte dos termos corriqueiros que se usam com referência à língua — muitos dos quais, aliás, surgiram como termos técnicos da gramática tradicional — é imprecisa ou ambígua. O que não quer dizer que o lingüista, como todos os especialistas, não possa ser acusado, por vezes, de um pedantismo terminológico improcedente. Em princípio, entretanto, o vocabulário lingüístico especializado, se controlado e adequadamente empregado, serve mais para esclarecer do que para mistificar. Elimina muitas ambigüidades e possíveis equívocos.

Com a notação, o mesmo que com a terminologia. Usamos o termo 'language' tanto para falar na linguagem quanto para falar nas línguas. Com isso, precisamos poder identificar exatamente que partes ou traços de uma língua estamos mencionando. O uso de convenções especiais de notação facilita muito esta missão. Por exemplo, poderíamos querer diferenciar o significado de uma palavra de sua forma, e ambos — significado e forma — da palavra em si. Infelizmente não há um sistema de notação universalmente aceito pelo qual se estabeleçam essas e outras distinções. No presente trabalho, faremos um uso distintivo de aspas simples, aspas duplas e itálicos. Por exemplo, diferenciaremos "mesa" de *mesa*, a primeira sendo o significado e a última a forma (ou uma das formas) da palavra 'mesa'. Por meio de tais convenções, poderemos manter distintos, como veremos mais adiante, pelo menos dois dos sentidos da palavra 'palavra': no primeiro, refere-se a algo que esperamos ver listado no dicionário da língua; no segundo, refere-se ao que estaria impresso espaçadamente como uma seqüência de letras em um texto escrito.

Outras convenções serão introduzidas mais tarde, possibilitando a distinção entre formas faladas e escritas; e formas faladas de um tipo (foné-

tas) e formas faladas de outro tipo (fonológicas); e assim por diante. O ponto que se defende aqui é o de que várias convenções de notação são, senão absolutamente essenciais, pelo menos muito úteis para a referência de dados da língua e para o esclarecimento sobre o que se está discutindo. Têm ainda uma outra vantagem, a de forçar o lingüista a pensar cuidadosamente sobre certas diferenciações que de outra forma passariam despercebidas. Muitas vezes é difícil ser absolutamente coerente na aplicação de determinadas convenções; e esta dificuldade leva a uma nova avaliação da distinção teórica para a qual fora estabelecida primeiramente a convenção de notação. Eis uma das vias pelas quais se atingem os progressos em qualquer disciplina.

2.4 A lingüística é ~~descritiva~~, não ~~prescritiva~~

O termo 'descritiva' está sendo empregado aqui em uma acepção diferente do que aquela que se opõe seja a 'geral', por um lado, seja a 'histórica', por outro. O contraste relevante neste caso é o que existe entre *descrever* como as coisas são e *prescrever* o que devem ser. Uma alternativa para 'prescritiva', na acepção que contrasta com 'descritiva', é 'normativa'. Dizer que a lingüística é uma ciência descritiva (ou seja, não-normativa) é dizer que ela tenta descobrir e registrar as regras segundo as quais se comportam os membros de uma comunidade lingüística, sem tentar impor-lhes outras regras ou normas, de correção exógenas.

Talvez seja confuso utilizar o termo 'regra', como acabo de fazer, nestes dois sentidos tão diferentes. Com ou sem razão, os lingüistas falam nestes termos. Portanto, é possível que valha a pena ilustrar a diferença entre os dois tipos de regra — chamemos a uma de imanente e a outra de transcendente, respectivamente — com um exemplo outro que o uso da língua. Vejamos o comportamento sexual em uma determinada sociedade. Se adotarmos o ponto de vista puramente descritivo (isto é, não-normativo) na investigação do comportamento sexual, tentaremos descobrir como as pessoas realmente se comportam: se praticam o sexo antes do casamento, e neste caso, que tipo e desde que idade; se os esposos e esposas são igualmente fiéis ou infiéis uns aos outros; e assim por diante. Na medida em que o comportamento de certos grupos dentro da comunidade é, na prática, governado por princípios determináveis — seja que os membros de tais grupos professem, ou mesmo estejam conscientes de tais princípios, ou não — podemos dizer que seu comportamento é regido por regras: as regras são imanentes a seus comportamentos reais. Mas estas (se têm o direito de serem chamadas de regras) diferem muito quanto ao *status*, senão quanto ao conteúdo das regras de conduta que a lei, a religião ou simplesmente a moral convencional explícita pode haver prescrito. As pessoas podem obe-

decer ou não ao que chamo de regras *transcendentes* (isto é, exteriores e não-imanentes) do comportamento real, e o comportamento que dizem, ou mesmo pensam, ter. Todas estas diferenças têm um correspondente no comportamento lingüístico. Contudo, a distinção mais importante é a que existe entre regras transcendentes (isto é, prescritivas) e regras imanentes (isto é, descritivas). Um *dizer* e um *não-dizer*, se prescritivo, é uma ordem (*Diga/Não diga X!*); se descritivo, é uma afirmação (*As pessoas dizem/não dizem X.*)

A razão pela qual os lingüistas de hoje são tão insistentes com a distinção entre regras descritivas e prescritivas, é simplesmente que a gramática tradicional tinha um caráter acentuadamente normativo. O gramático acreditava que sua missão era formular os padrões de correção e impor, se necessário, aos falantes da língua, tais normas de comportamento. Muitos dos preceitos normativos da gramática tradicional serão conhecidos do leitor falante de inglês: "Nunca se deve usar uma negativa dupla" (*I didn't do nothing* [eu — auxiliar de negação no passado — fazer — nada]); "Não se termina uma sentença por preposição" (*That's the man I was speaking to* [literalmente, "aquele é o homem (que) eu estava falando com"]); "O verbo 'to be' exige o mesmo caso antes e depois dele" (de formas que, pela aplicação desta regra, *It's me* [pronome reto neutro de 3ª pessoa — é — pronome oblíquo de 1ª pessoa] deveria ser corrigido para *It's I* [pronome reto neutro de 3ª pessoa — é — eu]); "*Ain't* [forma coloquial de registro baixo equivalente a '*isn't*': 'não é/está' é errado]; "Não se deve dividir o infinito" (como em *I want you to clearly understand* [eu quero (que) você — partícula de infinitivo — claramente — entender], em que *clearly* está inserido entre *to* e *understand*).[☆]

Considerando tais exemplos, rapidamente notamos que são bastante heterogêneos. Há certos dialetos do inglês em que as chamadas negativas-duplas nunca se empregam (ou seja, em que *I didn't do nothing* nunca é usado como equivalente do inglês-padrão *I didn't do anything* [Eu não fiz nada; literalmente: eu — auxiliar de negação no passado — fazer — algo]); há outros em que, de um ponto de vista meramente descritivo, esta é a construção correta. Quando se apresentam razões para a condenação de uma negativa dupla como incorreta, em termos de algum princípio prescritivo em referência ao qual o uso comum é julgado e considerado devedor, recorremos à lógica em última instância. A lógica diz, supomos, que duas negativas valem uma positiva. O que merece diversos comentários. Primeiro, denuncia uma compreensão errônea do que seja a lógica e de

[☆] As regras da gramática tradicional no português, embora não exatamente as mesmas que em inglês, prescrevem por exemplo: (a) que pronome oblíquo tônico não ocupa posição de sujeito ('Isto é para eu fazer', e não 'Isto é para mim fazer'); (b) que

como ela opere; mas não precisamos entrar na questão da natureza dos axiomas lógicos e na complexidade de como a chamada lógica natural dos comportamentos lingüísticos comuns se relaciona com os sistemas lógicos, construídos e investigados pelos cientistas. O problema é que não há nada de inerentemente ilógico nas chamadas construções com negativas duplas. Nos dialetos em que são regularmente empregadas, operam de forma bastante sistemática, conforme as regras e princípios gramaticais de interpretação, imanescentes no comportamento da comunidade dialetal em questão. Um segundo ponto a ser lembrado é o de que a chamada negativa dupla não pode ser adequadamente descrita, da forma como opera em determinados dialetos do inglês, sem se levar em consideração determinados traços de ênfase e entonação. As regras do inglês-padrão (ou seja, regras imanescentes no comportamento lingüístico dos falantes de um determinado dialeto do inglês) permitem a existência de *I didn't do nothing* (significando, mais ou menos, "*It is not true that I did nothing*" [Não é verdade que eu não fiz nada]) contanto que *didn't* seja enfatizado ou, alternativamente, e com outras implicações ou pressuposições, que *do* ou *nothing* sejam pronunciados com uma ênfase especialmente forte. Nos dialetos em que *I didn't do nothing* (com a ênfase normal da fala) pode significar *I didn't do anything*, há também os significados que o inglês-padrão apresenta, mas a ênfase e entonação evitam que se misturem. Finalmente podemos notar que há muitas línguas em que a chamada negativa dupla ocorre no dialeto literário-padrão: por exemplo o francês, o italiano, o espanhol, o russo — só para mencionar algumas das línguas européias mais modernas e conhecidas. Mesmo o dialeto mais prestigiado do grego antigo — o grego clássico usado por Platão, Sófocles, Tucídides ou pelo próprio pai da lógica, Aristóteles — apresentava uma construção com negativa dupla. E, afinal, a gramática tradicional originou da descrição dos dialetos literários da Grécia Antiga!

Outros preceitos normativos da gramática tradicional — como por exemplo a condenação da cisão do infinitivo (... *to clearly understand*) ou a de *It's me* — provém da aplicação ao inglês de princípios e categorias estabelecidas originalmente para a descrição do grego e do latim. O fato é que as formas a que se aplica o termo 'infinitivo' são formas compostas de um único vocábulo em grego e latim, bem como em francês, alemão, russo, etc. Tradicionalmente as formas compostas de dois vocábulos, como *to understand*, *to go*, etc., também são chamadas de infinitivos, se bem que

a preposição e o fato do pronome ser objeto pedem caso oblíquo e não reto em contextos como 'Este assunto fica entre mim e você' (e não 'entre eu e você'), 'Ontem não o vi na escola' (e não 'não vi ele na escola'); e (c) que não se inicia um período por pronome átono ('Disseram-me que vinhas', e não 'Me disseram que vinhas') entre outras. (N. do T.)

suas funções sejam apenas parcialmente comparáveis às funções, digamos, dos infinitivos latinos. Como veremos adiante, o fato de uma forma poder ser dividida (no sentido que falamos em dividir o infinitivo) é um dos principais critérios que o lingüista aplica para decidir se uma forma é composta de um ou dois vocábulos. Visto que, por outros critérios e por convenções ortográficas da língua escrita, os infinitivos ingleses são compostos de duas palavras, não se pode objetar à sua cisão, em princípio. Quanto a condenar um *It's me*, etc., o fato é que o que se denomina na gramática tradicional de diferenças de casos (*I vs. me*, *she vs. her*, *he vs. him* [respectivamente, caso reto vs. oblíquo] etc.) não é encontrado em todas as línguas; nem tampouco algo que se possa identificar em termos de suas funções e características gramaticais como verbo significando "*to be*". Além disso, nas línguas em que há tanto um caso, quanto um verbo, identificável como sendo equivalente do latim 'esse', ou inglês '*to be*', a diversidade das construções é tamanha que a regra tradicional "O verbo '*to be*' exige o mesmo caso antes e depois dele" imediatamente se denuncia como é — uma regra normativa de base latina que não se sustenta em âmbitos mais gerais.

O que é interessante é que muitos falantes do que os gramáticos tradicionais classificariam de um bom inglês dizem e escrevem formas como *between you and I* [entre você e eu], *He told you and I* [Ele disse a você e eu] etc. Tais construções violam outra regra prescritiva tradicional do inglês: "Verbos e preposições levam o seu objeto para o acusativo". Resulta, presume-se, do que muitas vezes é chamado de **hipercorreção**: a extensão de alguma regra ou princípio, com base em uma má compreensão de seu domínio de aplicação, a uma gama de fenômenos aos quais originalmente não se aplicam. A natureza da regra prescritiva foi mal compreendida — tanto mais que muitos falantes que diriam naturalmente *You and me will go* [literalmente, "você e mim iremos"], jamais diriam *Me will go* ou *He told I* [literalmente, "Mim irei" e "Ele disse a eu", respectivamente]. Interpretamos, pois, como instrução (sob pena de sermos considerados falantes de um inglês ruim) a substituição por *you and I* [você e eu] de *you and me* [você e mim] (ou *me and you* [mim e você]) em todas as posições de ocorrência. Isto resulta na produção de não só o que o gramático tradicional classificaria de correto, *You and I will go together* [Você e eu iremos juntos], etc., mas também o que ele condenaria, *between you and I* [entre você e eu], *He told you and I* [Ele disse a você e eu], etc. Evidentemente não queremos dizer que todos os falantes de inglês que dizem *between you and I*, *He told you and I*, etc., efetuaram a operação de aplicação correta e depois incorreta da regra tradicional. Tais construções são hoje tão comuns na fala de indivíduos de classe média e alta, falantes do inglês-padrão na Inglaterra, que devem ter sido aprendidas naturalmente no processo de aquisição de linguagem, possivelmente pela maioria dos que

as utilizam. Restam poucas dúvidas, quanto ao fato de que se originem num processo de hipercorreção. ☆

Nem a lógica nem a gramática do latim podem servir de árbitros adequados quando vem a questão de decidir se determinado enunciado é correto ou incorreto em inglês. Nem tampouco a autoridade inquestionada da tradição pela tradição ("Foi isso que aprendi, que meus pais e meus avós aprenderam") ou o costume dos considerados melhores autores literários da língua. É uma visão bastante comum de nossa sociedade, ou pelo menos foi até recentemente, dizer que a mudança lingüística necessariamente acarreta um abalo ou corrupção da língua. Tal posição é indefensável. Todas as línguas estão sujeitas a mudanças. É um fato empírico; e é tarefa dos lingüistas históricos investigar os detalhes da mudança lingüística, caso sejam acessíveis, construindo uma teoria explanatória para o fenômeno, a fim de contribuir para a nossa compreensão da natureza da lingua(gem). Os fatores determinantes da mudança lingüística são complexos e, até agora, apenas parcialmente compreendidos. Mas hoje sabemos o suficiente — conhecimento que temos desde meados do século XIX — para que qualquer observador sem preconceitos veja claramente na mudança lingüística que o que é condenado em qualquer época como degeneração e corrupção dos padrões tradicionais do uso, pode sempre ser alinhado com uma modificação anterior da mesma espécie, que fez surgir o uso que os próprios tradicionalistas tratam de inalteravelmente correto.

Quanto ao princípio de conformidade aos padrões estabelecidos pelos melhores literatos da língua, também este é insustentável — insustentável, quer dizer, em relação ao uso que normalmente se faz dos mesmos. Não há razão para crer que um escritor, por mais genial que seja, tenha sido investido, por graça especial, de um conhecimento certo e seguro das regras transcendentais de correção, dom negado a nós outros. O fato é que a gramática tradicional tinha uma inclinação literária muito acentuada. A razão é que em vários períodos importantes da cultura européia — desde o período da erudição alexandrina, no século segundo antes da era cristã, até o humanismo renascentista — a descrição gramatical, primeiro do grego, depois do latim, estava subordinada à tarefa prática de tornar a literatura de uma época mais antiga acessível àqueles que não falavam, e devido à natureza das coisas nem podiam falar, naturalmente o dialeto do grego ou do latim em que se baseava a língua dos textos clássicos. A inclinação literária da gramática tradicional é não só explicável do ponto de vista histórico, como também perfeitamente justificável, no tocante à descrição

☆ Veja-se o caso da sentença, acima citada 'Este assunto fica entre eu e você'. (Nota do Tradutor)

do grego e do latim. Torna-se totalmente injustificável quando trata da descrição gramatical das línguas faladas modernamente.

Não há na língua padrões de correção absolutos. Podemos dizer que um estrangeiro cometeu um erro, se ele disser algo que vá contra as regras imanescentes ao uso dos falantes nativos. Podemos também afirmar, se quisermos, que um falante de um dialeto inglês social ou regional não-padrão falou agramaticalmente, se seu enunciado transgride as regras imanescentes do inglês-padrão. Mas ao dizermos isso, estamos evidentemente partindo do princípio de que ele pretendia, ou pelo menos deveria pretender, usar o inglês-padrão. E esta é uma posição que requer uma justificativa.

Devemos agora frisar — e este ponto é muitas vezes mal entendido — que ao traçarmos uma distinção entre descrição e prescrição, não estamos dizendo que não haja lugar para o estabelecimento e prescrição de normas de uso. Obviamente há vantagens administrativas e educacionais, no mundo moderno, para a padronização do dialeto principal empregado em um determinado país ou região. Este processo de padronização estendeu-se por um longo período de tempo em muitos países do Ocidente, com ou sem a intervenção do governo. Agora vem sendo efetuado aceleradamente, como parte de uma política oficial, em alguns dos países em desenvolvimento da África e da Ásia. O problema de selecionar, padronizar e promover uma determinada língua ou dialeto em detrimento de outros está envolto em dificuldades políticas e sociais. É parte do que tornou-se conhecido por planejamento lingüístico — uma área importante no campo da sociolingüística aplicada.

Nem tampouco se deve pensar que negando que toda mudança na língua seja para pior, o lingüista esteja afirmando que deve ser para melhor. Ele está simplesmente questionando o apelo impensado a critérios empiricamente descreditados. Concorda que, em princípio, pode ser possível avaliar dialetos e línguas em termos de sua flexibilidade, amplitude de expressão, precisão e estética potencial relativas; e com certeza ele aceita que o uso que falantes e escritores individuais fazem de seu dialeto ou língua pode ser mais ou menos eficaz. Entretanto, ele só pode relatar, com base nos trabalhos mais científicos feitos sobre língua e linguagem recentemente, que a maior parte dos julgamentos feitos a este respeito são extremamente subjetivos. Como membro individual de uma comunidade lingüística, o lingüista terá seus próprios preconceitos, sejam pessoais, sejam resultantes de uma formação social, cultural e geográfica; e, por temperamento, ele pode ser conservador ou progressista. Suas atitudes a respeito de sua própria língua não serão menos subjetivas, neste particular, do que as dos leigos. Ele poderá achar um determinado sotaque ou dialeto agradável ou desagradável. Ele poderá mesmo corrigir a fala de seus filhos se os vir usando uma pronúncia, palavra ou construção gramatical desaprovada pelos puristas. Mas ao fazê-lo, se for honesto consigo mesmo, saberá

que o que está corrigindo não é inerentemente incorreto, mas apenas relativamente a um padrão que, por motivos de prestígio social ou de vantagens educacionais, ele quer que os filhos adotem.

No tocante a sua atitude em face da língua literária, o lingüista está simplesmente salientando que a língua é usada para diversos fins e que seu uso em relação a tais fins não deve ser julgado por critérios aplicáveis única ou primeiramente à língua literária. Isto não quer dizer absolutamente que ele seja hostil à literatura, ou contrário ao estudo desta nas escolas e universidades inglesas. De forma alguma; muitos lingüistas, até, têm um interesse especial na investigação dos fins literários aos quais a língua serve e no sucesso que alcança satisfazendo a tais objetivos. Esta é uma parte — e, aliás, muito importante — de um ramo da macrolingüística conhecido por estilística.

2.5 Prioridade da descrição sincrônica

O princípio de prioridade da descrição sincrônica, característico da maior parte da teoria lingüística de nosso século, implica que as considerações históricas são irrelevantes para a investigação de determinados estados temporais de uma língua. Os termos saussureanos 'sincrônico' e 'diacrônico' foram introduzidos anteriormente neste capítulo (v. 2.1). Podemos usar uma das analogias de Saussure para explicar o significado da prioridade do sincrônico sobre o diacrônico.

Comparemos o desenvolvimento histórico de uma determinada língua com um jogo de xadrez que se desenrola perante nós. O estado do tabuleiro está em constante modificação, na medida em que cada jogador movimentava suas peças. Mas o estado do jogo pode ser descrito a qualquer momento em termos das posições que as peças ocupam. (Na realidade não é bem assim. Por exemplo, o estado do jogo é afetado, no tocante às possibilidades de um roque, pelo fato de se retirar o rei de sua posição original e depois devolvê-lo. Mas podemos deixar de lado tais detalhes menores em que a analogia de Saussure não funciona.) Não importa por que caminho os jogadores atingiram um certo estado no jogo. Independente do número, natureza e ordem das jogadas anteriores, o estado atual do jogo é sincronicamente descritível sem qualquer referência a elas. O mesmo acontece, conforme Saussure, com o desenvolvimento histórico das línguas. Todas estão em constante modificação. Mas cada um dos estados sucessivos de uma língua pode, e deve, ser descrito em seus próprios termos, sem referência àquilo a partir do que se desenvolveu, nem àquilo rumo a que, provavelmente, evoluirá.

Tudo isso pode parecer altamente teórico e abstrato. Contudo tem conseqüências bastante práticas. A primeira refere-se ao que chamo de

falácia etimológica. A etimologia é o estudo da origem e desenvolvimento das palavras. Iniciou, no tocante à gramática tradicional ocidental, com a especulação de certos filósofos gregos do quinto século antes de Cristo. O próprio termo 'etimologia' é em si mesmo revelador. Trata-se de uma transcrição latinizada de uma das formas do vocábulo grego 'etumos', que significa "verdadeiro" ou "real". De acordo com uma escola de filósofos do século quinto, na Grécia, todas as palavras associavam-se naturalmente, e não convencionalmente, com aquilo que significavam. É possível que isto não fosse evidente para o grego, diziam; mas poderia ser demonstrado pelo filósofo, capaz de discernir a realidade subjacente às aparências das coisas. Penetrar as muitas vezes enganosas aparências, analisando cuidadosamente as modificações ocorridas no desenvolvimento da forma ou do significado de uma palavra, descobrir a origem do vocábulo e com isto o seu real significado, era revelar uma das verdades da natureza. O que classifico de falácia etimológica é a crença de que a forma ou o significado original de uma palavra é, necessariamente e em virtude deste mesmo fato, sua forma ou significado correto. Esta pressuposição é amplamente divulgada. Quantas vezes não presenciamos uma discussão em que alguém defende que por tal ou tal palavra originar do grego, latim ou árabe, ou seja qual for a língua em questão, o significado correto daquele vocábulo deve ser o que existia na língua de origem! O argumento é falacioso porque a pressuposição tácita de que há uma correspondência originalmente verdadeira ou apropriada entre forma e significado, sobre o qual se sustenta não pode ser substanciada.

A etimologia recebeu um embasamento mais sólido no século XIX. Já não é mais justo dizer, como se afirma que Voltaire teria dito, que a etimologia é uma ciência em que as vogais não servem para nada, e as consoantes para muito pouco! Tal como é hoje praticada, ela constitui um ramo respeitável da lingüística histórica, ou diacrônica. Como veremos no capítulo 6, tem seus próprios princípios metodológicos, cuja confiabilidade depende da qualidade e quantidade dos dados a partir dos quais se formulam. Em condições favoráveis, a confiabilidade das reconstruções etimológicas é sem sombra de dúvida muito alta.

Um ponto que ficou evidente para os etimólogos do século XIX, que é hoje ponto pacífico para todos os lingüistas, é o de que a maioria das palavras no vocabulário de qualquer língua não permite que se trace sua origem. Os vocábulos deliberadamente criados, a partir de formas emprestadas ou pelo uso de algum outro princípio, são atípicos no vocabulário global, e com certeza também do que se pode qualificar de vocabulário básico, não-técnico, da língua. O que o etimólogo atual faz é relacionar palavras de um estado de língua sincronicamente descritível a palavras, atestadas ou reconstituídas, de algum estado anterior da mesma língua ou de alguma outra. Mas as palavras do estado anterior da mesma língua ou

de alguma outra, por sua vez, derivaram de outras, mais antigas. O fato de a forma ou o significado destas palavras ser recuperável pelas técnicas etimológicas, dependerá dos dados que houverem sobrevivido no tempo. Por exemplo, podemos relacionar a palavra inglesa atual 'ten' [dez] à palavra do inglês antigo cujas formas eram *ten* (com vogal longa) ou *tien*. E podemos relacionar o vocábulo do inglês antigo, através de vários estados hipotéticos sucessivos, a uma palavra reconstituída do proto-indo-europeu com a forma **dek̑m*, que também significava "dez". Mas não podemos nos aprofundar mais sem perder a confiança. E mesmo assim, a palavra do proto-indo-europeu **dek̑m* — o asterisco no prefixo indica que trata-se de uma reconstituição, não de um fato atestado (v. 6.3) — obviamente não é a origem, em um sentido absoluto, de todas as palavras que evoluíram a partir dela, em todas as línguas que podemos apontar como pertencentes à família indo-européia. Ela própria deve ter evoluído a partir de outra (que pode ter significado o mesmo que "ten" ou não — não se pode saber), parte do vocabulário de alguma outra língua; e esta palavra, por sua vez, de alguma outra anterior, de outra língua: e assim por diante. De uma forma geral, os etimólogos hoje não se ocupam das origens. Na realidade, eles diriam que em muitos casos (por exemplo o da palavra 'ten') não faz sentido investigar as origens de um vocábulo. Tudo o que o etimólogo nos pode dizer, com maior ou menor segurança, dependendo das provas, é que a forma ou o significado do ancestral mais antigo, ou mesmo hipotético, de uma determinada palavra é tal.

Com isto chegamos ao ponto em que a analogia de Saussure mais obviamente falha. Todo jogo de xadrez, se jogado conforme as regras e se completado, tem início e fim determinado. As línguas não. Não só (pelo que sabemos) não aconteceu de todas as línguas começarem de um mesmo estado do tabuleiro, por assim dizer, e de então desenvolverem de forma diversa, como também é impossível datar o início de uma língua, a não ser por uma convenção arbitrária e muito aproximativa. Não podemos afirmar, por exemplo, quando o latim falado tornou-se francês antigo, ou italiano, ou espanhol. Nem tampouco podemos localizar o ponto em que uma determinada língua deixou de existir — a não ser nos casos de línguas que se extinguíram, mais ou menos repentinamente, quando morreram seus últimos falantes. As línguas, do ponto de vista diacrônico, não têm começo e fim determinados. Em última análise, é por uma questão de convenção e conveniência que dizemos que o inglês antigo e o moderno são dois estados de uma mesma língua, ou duas línguas diferentes.

Há ainda um outro particular em que a analogia saussureana vem falhar. O jogo de xadrez é formulado por regras explicitamente formuladas, e dentro dos limites por elas impostos, os jogadores determinam o curso de qualquer partida que esteja sendo disputada entre os dois, com referência a um objetivo reconhecido. Pelo que se sabe, não há direção

preestabelecida no desenvolvimento diacrônico das línguas. É bem possível que haja certos princípios gerais determinando a transição de um estado lingüístico para outro. Mas se houver tais princípios, não se compararão às regras de um jogo concebido pelo homem, como o xadrez. Examinaremos as chamadas leis da mudança lingüística no capítulo 6.

Normalmente acredita-se que o princípio de prioridade da descrição sincrônica implica que, ao passo que esta independe da descrição diacrônica, a descrição diacrônica pressupõe a análise sincrônica anterior dos estados sucessivos pelos quais as línguas passaram ao longo de sua evolução histórica. É possível que esta não tenha sido a visão de Saussure. Mas é consequência das hoje amplamente aceitas premissas sobre a natureza dos sistemas lingüísticos.

Às vezes de forma enganosa, os lingüistas falam como se o passar do tempo fosse em si suficiente para explicar a mudança lingüística. Há diversos fatores, tanto internos quanto externos à língua, que podem causar a passagem de um estado sincrônico para outro. Alguns destes fatores, provavelmente os mais importantes, são de ordem social. O passar do tempo simplesmente permite que sua complexa interação propicie o que é posteriormente reconhecido como transição entre dois estados de língua.

Além do mais, o conceito de desenvolvimento diacrônico entre estágios sucessivos de uma língua só tem sentido se aplicado a estados lingüísticos relativamente distanciados no tempo. Já me referi anteriormente ao que chamei de ficção da homogeneidade (v. 1.6). Em uma certa medida, ela é útil e necessária. Entretanto, se partimos do princípio de que a mudança lingüística acarreta uma transformação constante ao longo do tempo do que a qualquer momento escolhido é um sistema perfeitamente homogêneo, todo o processo de mudança da língua fica muito mais misterioso do que realmente é. O que caracteriza a fala de uma minoria aparentemente insignificante de uma comunidade lingüística em determinada época pode espalhar-se por quase todos os falantes em questão de uma ou duas gerações. Pode-se considerar bastante legítimo da parte do lingüista que descreva sincronicamente a língua em qualquer um destes dois pontos, que desconte a fala da minoria divergente. Mas se o fizer, e em seguida passar a falar diacronicamente de um sistema lingüístico sincronicamente homogêneo que se transforma em outro igualmente homogêneo, estará distorcendo os fatos. Pior que isso, estará incorrendo no risco de criar para si mesmo pseudoproblemas teóricos insolúveis. Uma vez que compreendamos que nenhuma língua é estável, ou uniforme, teremos dado o primeiro passo para dar conta, teoricamente, da ubiquidade e continuidade da mudança lingüística. Se tomarmos dois estados de uma língua, diacronicamente determinados, que não estejam suficientemente distanciados no tempo, temos grande probabilidade de achar que a maior parte das diferenças entre eles também se apresentam como variações sincrônicas tanto

no estado anterior como no posterior. Do ponto de vista microscópico — que se distingue do macroscópico normalmente adotado na lingüística histórica — é impossível traçar uma distinção nítida entre uma mudança diacrônica e uma variação sincrônica.

Em suma, o princípio de prioridade da variação sincrônica é válido. Porém, na medida em que repousa sobre a ficção da homogeneidade, deve ser aplicado com bom senso e com plena consciência do *status* teórico do sistema lingüístico. É este ponto que agora retornamos.

2.6 Estrutura e sistema

Uma das definições de 'lingua(gem)' que citei no capítulo 1 foi a de Chomsky: "um conjunto (finito ou infinito) de sentenças, cada uma finita em comprimento e construída a partir de um número finito de elementos" (v. 1.2). Adotemos esta formulação como definição parcial do termo 'sistema lingüístico', que foi introduzido, vale lembrar, para eliminar-se a ambigüidade do vocábulo inglês '*language*'.

Na medida em que são por definição estáveis e uniformes, os sistemas lingüísticos não podem identificar-se com as línguas naturais reais: são construtos teóricos, postulados pelos lingüistas para dar conta das regularidades por ele encontradas no comportamento lingüístico dos membros de uma determinada comunidade lingüística — mais precisamente nos sinais lingüísticos que são o produto de seu comportamento lingüístico. Como já pudemos ver, as línguas naturais reais não são nem estáveis nem homogêneas. No entanto, há estabilidade e homogeneidade suficientes na fala daqueles que normalmente são considerados falantes de uma mesma língua para que a postulação do lingüista, de que há um sistema lingüístico comum subjacente, seja útil e cientificamente justificável, a não ser na questão específica da variação sincrônica e diacrônica. Nos próximos três capítulos, pressuporemos a validade do conceito de sistema lingüístico tal qual fica aqui definido e explicado.

Dentre os sinais lingüísticos produzidos por um falante inglês, durante um certo período de tempo, alguns seriam classificados como sentenças da língua, outros não. Precisamos neste estágio investigar quais os critérios em virtude dos quais se faz tal divisão entre sentenças e não-sentenças. Obviamente há princípios determinadores da elaboração de textos e discursos maiores. Além disso, estes princípios são tais que qualquer pessoa que os despreze pode ser, com razão, acusada de transgredir as regras da língua. Embora não tenha passado sem sofrer ataques, recentemente, a premissa tradicional de que a maior parte, senão a totalidade, dos elementos envolvidos no conhecimento de uma língua pode ser descrita em

termos da construção e interpretação de sentenças ainda é aceita pela maioria dos lingüistas.

As sentenças, digamos, são o que seria convencionalmente pontuado como tal na língua escrita. Como vimos, as línguas naturais possuem a propriedade de transferência de meio (v. 1.4). Isto significa que, em geral, qualquer sentença da língua escrita pode ser posta em correspondência com uma da língua falada, e vice-versa. As sentenças faladas, obviamente, não se pontuam como tal com qualquer sinal que seja estritamente equivalente à maiúscula inicial ou ao ponto final, ou ponto simples, das sentenças escritas. Contudo, em face de nossos atuais objetivos, podemos estabelecer uma equivalência aproximativa entre os sinais de pontuação da língua escrita e os padrões entonacionais da língua falada correspondente.

O termo 'estrutura' tem o mesmo destaque na lingüística moderna que em muitas outras disciplinas. Se adotarmos o ponto de vista que foi primeiramente, de maneira muito clara, expresso por Saussure e é atualmente aceito por todos os que se identificam com os princípios do **estruturalismo**, diremos que não só um sistema lingüístico tem uma estrutura, como também que ele é uma estrutura. Por exemplo, na medida em que o inglês escrito e falado são isomórficos (isto é, têm a mesma estrutura), são a mesma língua: não há nada senão a estrutura para compartilharem em comum. O sistema lingüístico em si, é, a princípio, independente do meio em que se manifesta. Neste sentido, trata-se de uma estrutura puramente abstrata.

Os sistemas lingüísticos são estruturas em dois níveis: têm a propriedade da dualidade (v. 1.5). As sentenças faladas não são meras combinações de elementos fonológicos; são também combinações de unidades sintáticas. A definição parcial de Chomsky, de um sistema lingüístico como sendo um conjunto de sentenças, cada uma das quais finita em comprimento e construída a partir de um conjunto finito de elementos, deve ser ampliada para dar conta desta propriedade essencial das línguas naturais. É logicamente possível que dois sistemas lingüísticos sejam isomórficos em um nível sem sê-lo no outro. Na realidade, conforme já foi ressaltado, é pelo fato de os chamados dialetos do chinês estarem suficientemente próximos de serem sintaticamente isomórficos (embora estejam longe de ser fonologicamente isomórficos) que a mesma língua escrita, não-alfabética, pode ser colocada em uma correspondência mais ou menos igual com cada um deles. Também é possível haver um isomorfismo fonológico sem um isomorfismo sintático entre as línguas. Tal possibilidade concretiza-se em maior ou menor escala, quando um falante nativo do inglês, digamos, fala um francês gramaticalmente perfeito com um sotaque inglês especialmente ruim. O mais interessante é que a independência da sintaxe em relação à fonologia é demonstrada muitas vezes de forma bem acentuada no processo de criouliização (v. 9.3).

As línguas naturais, logo, possuem dois níveis de estrutura que são independentes entre si, na medida em que a estrutura fonológica de uma língua não é determinada pela estrutura sintática, e a estrutura sintática não é determinada pela estrutura fonológica. É improvável, na melhor das hipóteses, que haja duas línguas naturais tais que todas as sentenças, escritas ou faladas, de uma possam ser lidas ou ouvidas como sentenças pertencentes à outra (com ou sem o mesmo significado). Entretanto acontece frequentemente, como consequência da independência entre as estruturas sintática e fonológica, de a mesma combinação de elementos (sons na fala, e letras na escrita) realizarem não apenas uma, mas duas ou mais sentenças. Estas podem ser distintas uma da outra pela entoação ou pontuação, dependendo de cada caso. Assim

- (1) *John says Peter has been here all the time*
[John disse (que) Pedro esteve aqui o tempo todo]

é diferente de

- (2) *John, says Peter, has been here all the time*[☆]
[John, disse Pedro, esteve aqui o tempo todo]

na língua escrita por meio da pontuação; e seriam enunciadas normalmente distintos na língua falada por meio do padrão entoacional a que cada um obedece. Mas mesmo sem diferenças de entoação ou pontuação a mesma combinação de elementos pode atualizar mais de uma sentença. Por exemplo,

- (3) *We watched her box*
[Vimos a caixa dela]
[Vimos ela lutar (box)]

poderia ser qualquer uma das seguintes sentenças em inglês; uma em que *her* está na forma adjetiva (v. *his* [dele]) e *box* na forma substantiva (v. *suitcase* [mala]); outra em que *her* está na forma pronominal (v. *him* [o]) e *box* na forma verbal (v. *wrestle* [lutar]).^{☆☆} Não é necessário nos preocuparmos com a análise sintática tradicional de (3) à qual me referi indiretamente. É algo que será retomado mais adiante. Agora basta que estabeleçamos que as sentenças, tais como são tradicionalmente definidas, não podem ser identificadas, e diferenciadas umas das outras, em termos dos elementos fonológicos de que se compõem. Na realidade podemos con-

☆ Vejam-se em português os casos de (a) e (b)

(a) As moças, que são mais educadas, não falam alto.

(b) As moças que são mais educadas não falam alto. (N. do T.)

☆☆ Em português este caso não tem um equivalente direto. (N. do T.)

cluir (3), que elas não podem sequer ser identificadas em termos das unidades sintáticas de que se compõem sem que se considerem outros aspectos da estrutura sintática, inclusive com a designação de unidades sintáticas às tradicionalmente chamadas partes do discurso (nome, verbo, adjetivo etc.).

As unidades sintáticas a partir das quais são constituídas as sentenças, ao contrário dos elementos fonológicos, são muito numerosas. Contudo, assim como os elementos fonológicos, existem em número finito. Digamos que todo sistema lingüístico pressupõe a existência de um inventário finito de elementos e de um vocabulário finito de unidades (simples), juntamente com um conjunto de regras (ou talvez vários tipos) que relacionam os dois níveis de estrutura, dizendo-nos quais as combinações de unidades que constituem sentenças e, por dedução, senão explicitamente, quais não. Devemos notar, como será retomado mais tarde, que o vocabulário de uma língua natural é muito mais do que um conjunto de unidades sintáticas. Nenhuma das modificações ou dos refinamentos terminológicos que serão introduzidos nos capítulos seguintes afetará a essência do que foi dito aqui.

No momento, o que vimos chamando de unidades sintáticas pode ser denominado de **formas**: ou seja, combinações de elementos tais que cada diferente combinação refere-se a uma diferente forma. Mas estas, neste sentido do termo, possuem um significado, e este está longe de ser independente de sua função sintática. Este fato fica claro nas formas *her* e *box* em (3) acima. A abordagem tradicional defenderia que há (pelo menos) duas palavras distintas no vocabulário do inglês; representemos a ambas (com aspas simples) da seguinte forma: '*box*₁' e '*box*₂', respectivamente, diferentes tanto no significado quanto na função sintática, mas compartilhando uma mesma forma, *box*. Mais adiante precisaremos melhor a distinção tradicional entre uma forma e a unidade à qual se refere; e, ao fazê-lo, veremos que o termo 'palavra', usado tanto por lingüistas como por leigos, é extremamente ambíguo (v. 4.1).

Por definição toda sentença é **bem formada**, tanto sintática como fonologicamente, no sistema lingüístico em que se constitui como sentença. A expressão 'bem formada' é mais abrangente do que, mas subordina-se a, o termo mais tradicional que é 'gramatical', bem como este, por sua vez, é mais abrangente do que, mas subordina-se a, o termo 'sintaticamente bem formado'. A natureza e os limites da **gramaticalidade** (ou seja, a boa formação gramatical) serão discutidos no capítulo 4. Aqui basta frisar que o princípio da boa formação (inclusive da gramaticalidade) não se deve confundir com o da aceitabilidade, potencialidade de utilização ou mesmo o da significação. Há um número indefinido de sentenças do inglês ou de outras línguas naturais que, por diversas razões, não ocorreriam normalmente: poderiam conter uma justaposição inaceitável de palavras obscenas

ou blasfêmias; poderia m ser estilisticamente desajeitadas ou excessivamente complexas de um ponto de vista psicológico; poderiam ser contraditórias ou descrever situações que não ocorrem no mundo habitado pela sociedade falante da língua em questão. Qualquer combinação de elementos ou unidades de uma dada língua, L, que não seja bem formada em termos das regras de L é **mal formada**, no que diz respeito a L. Tais combinações de elementos ou unidades poderão ser marcadas como mal formadas através de um asterisco inicial.¹ Assim

(4) * *He weren't doing nothing*
[literalmente, "ele não estavam fazendo nada"]

é mal formada, e de fato agramatical, em relação ao inglês-padrão. É no entanto bem formada em determinados dialetos não-padrão do inglês. Este exemplo ilustra uma questão mais geral; ou seja, a de que diferentes línguas podem ser constituídas dos mesmos elementos e unidades, e o que é uma boa formação para uma pode não ser para a outra. Embora a ilustração se refira a dois dialetos de uma mesma língua, a princípio vale para o que seria considerado duas línguas bastante diferentes.

Poderíamos dizer mais sobre a estrutura dos sistemas lingüísticos. Mas é melhor deixar para os capítulos referentes à fonologia, à gramática e à semântica, onde as questões gerais podem ser introduzidas gradualmente e exemplificadas em maior detalhe.²

Começamos esta seção aceitando a definição chomskiana de *lingua(gem)* (ou seja, de sistema lingüístico) como sendo um conjunto de sentenças. Entretanto é preferível conceber um sistema lingüístico como sendo composto por um inventário de elementos, um vocabulário de unidades e de regras determinantes da boa formação das sentenças em ambos os níveis. E é isto que faremos daqui por diante. Sob a definição apropriada de 'sentença' as duas maneiras de conceber os sistemas lingüísticos coincidem, o que é discutível.

¹ O uso do asterisco para indicar a má formação sintática não deve ser confundido com o uso igualmente comum que dele se faz, já há mais tempo, para indicar, na lingüística histórica, as formas reconstruídas (v. 2.5). O contexto esclarecerá de qual dos dois se trata.

² Os termos 'estrutura' e 'sistema' são freqüentemente utilizados, especialmente por lingüistas britânicos, em um sentido bem especializado: 'sistema' para qualquer conjunto de elementos ou unidades que possam ocorrer em uma mesma posição; 'estrutura' para qualquer combinação de elementos ou unidades que resulte da seleção apropriada em determinadas posições. Assim definidos, os dois termos são complementares: cada um pressupõe o outro. Os sistemas são estabelecidos para determinadas posições de estruturas; e as estruturas são identificadas em termos das seleções feitas a partir dos sistemas (v. Berry, 1975). Neste livro, 'sistema' e 'estrutura' estão empregados em um sentido mais amplo.

LEITURAS COMPLEMENTARES

De uma forma geral, a mesma bibliografia do capítulo 1. Acrescentar Crystal (1971), capítulos 2-3; Lyons (1974).

Dos livros marcados por asterisco na Bibliografia, o de Robins (1979a) é o mais abrangente, e também o mais neutro na apresentação de questões controversas; Lyons (1968) enfatiza a continuidade entre a gramática tradicional e a lingüística moderna, restringindo-se à microlingüística sincrônica, e sendo ligeiramente inclinado para (numa versão hoje ultrapassada) a gramática transformacional; Martinet (1960) é a tradição do estruturalismo europeu; Gleason (1961), Hill (1958), juntamente com Joos (1966), apresentam uma boa abordagem do campo do ponto de vista da chamada lingüística pós-bloomfieldiana; Southworth & Daswani (1974) é especialmente bom no tratamento da lingüística em relação à sociologia e à antropologia, como também na lingüística aplicada; da mesma forma, embora menos abrangente, Falk (1973); Akmajian, Demers & Harnish (1979), Fromkin & Rodman (1974) e Smith & Wilson (1979) são todos coerentemente chomskianos por inspiração e, de uma forma geral, enfatizam o aspecto biológico da língua, mais que o cultural. Para a discussão das várias tendências e escolas da lingüística moderna, e para outras referências, ver o capítulo 7.

A lingüística histórica (isto é, diacrônica) é tratada mais adiante (capítulo 6). Da mesma forma a maior parte das ramificações da macrolingüística (capítulos 8 a 10).

Sobre a lingüística aplicada, ver Corder (1973) e para uma discussão mais detalhada, Allen & Corder (1975a, b, c).

PERGUNTAS E EXERCÍCIOS

1. Em que sentido a lingüística é uma ciência? Isso implica que ela não seja parte das chamadas humanidades?
2. "Como todos os ramos do conhecimento utilizam a língua, pode-se dizer que a lingüística, sob determinados aspectos, está no centro de todos eles, como estudo do instrumento de que devem se valer" (Robins, 1979a: 7). Discuta.
3. "As únicas generalizações úteis sobre a lingua(gem) são as indutivas" (Bloomfield, 1935: 20). Discuta.
4. Por que os lingüistas geralmente criticam tanto a gramática tradicional?
5. "Muitas vezes, tanto filósofos quanto lingüistas acreditam que... as intuições são 'científicas', não passíveis de uma observação direta, variáveis e indignas de confiança. Parece-nos que tal não é uma objeção válida..." (Smith & Wilson, 1979: 40). Discuta.
6. Qual o erro, se é que há, em (a) *entre eu e você* e (b) *Ele pediu para mim fazer isso*. Poderão princípios lógicos ou baseados no latim auxiliar na nossa decisão?
7. Qual a diferença entre uma abordagem descritiva e uma abordagem prescritiva (ou normativa) na investigação da linguagem?
8. Exemplifique, se possível a partir de sua própria experiência, um fenômeno de hipercorreção.
9. "A palavra 'álibi' é hoje comumente mal empregada: trata-se de um termo legal

- vindo do latim, que significa "em algum outro lugar" e não deveria ser usada como sinônimo da palavra 'desculpa'." Discuta.
10. Explique o que se quer dizer por prioridade do ponto de vista sincrônico sobre o diacrônico na lingüística.
 11. Apresente uma apreciação crítica da famosa analogia saussureana entre a língua e o jogo de xadrez.
 12. Uma visão simplificadora de o que seja **tradução literal** seria de que ela consiste na substituição palavra-por-palavra de formas vocabulares da língua-fonte por formas vocabulares da língua-meta. É este o significado normal do termo 'tradução literal'? Você poderia identificar algumas das razões pelas quais a visão simplificadora é utópica no tangente às línguas naturais?
 13. "O sistema lingüístico em si... é uma estrutura puramente abstrata" (p. 65). Considere tal afirmação em relação ao uso de códigos e cifras simples baseados em substituições (a) letra-por-letra e (b) palavra-por-palavra nas mensagens escritas. Tais técnicas criptográficas preservam ou destroem, necessariamente, o isomorfismo?
 14. Você poderia elaborar um código ou um conjunto de cifras simples que explorasse a interdependência dos dois níveis de estrutura em um sistema lingüístico, em que modificações feitas em um não afetassem o outro nível?