6

Mudança Lingüística

6.1 Lingüística histórica

O que hoje denomina-se lingüística histórica desenvolveu-se, pelo menos em suas linhas gerais, no decorrer do século dezenove (v. 2.1).

Os especialistas há muito tinham consciência de que as línguas mudam com o tempo. Sabiam igualmente que muitas das línguas européias descendiam, de certo modo, de línguas mais antigas. Por exemplo, sabia-se que o inglês tinha se desenvolvido a partir do anglo-saxão, e o que hoje chamamos de línguas românicas – o francês, o espanhol, o italiano etc. teve sua origem no latim. Entretanto, antes de se estabelecerem os princípios da lingüística histórica não se tinha consciência, de um modo geral, de que a mudança lingüística é **universal**, **contínua**, e consideravalmente **regular**.

Mais tarde discutiremos em detalhe cada um desses três aspectos da mudança lingüística. Aqui registramos que a universalidade e a continuidade do processo de mudança lingüística – o fato de que todas as línguas vivas são sujeitas a isso e de que o processo em si não pára – foram ofuscados para a maioria das pessoas pelo conservadorismo das línguas literárias-padrão da Europa e pelas atitudes normativas da gramática tradicional (v. 2.4). O *status* do latim é particularmente importante nesse sentido. Tinha sido usado durante séculos na Europa Ocidental como a língua dos sábios, da administração e da diplomacia internacional. A partir do Renascimento o latim foi gradualmente cedendo seu espaço, nessas funções, para as línguas românicas emergentes, bem como para outras não derivadas do latim: o inglês, o alemão, o holandês, o sueco, o dinamarquês etc. No século dezenove o latim já era quase uma língua morta, mas ainda tinha um prestígio que o distinguia da maioria das outras línguas. E ainda goza de tal prestígio junto a gramáticos de tendência tradicionalista. A importância da posição especial do latim no presente contexto é o fato de que até bem depois do Renascimento os especialistas consideravam-no como tendo existido como língua viva, mais ou menos estável, durante uns 2.000 anos, tendo sido preservada de corrupção no decorrer desse período pelo uso por pessoas instruídas, bem como pelas regras e preceitos dos gramáticos. Como vimos, atitudes semelhantes

foram adotadas com respeito às línguas literárias modernas da Europa quando primeiro surgiram – ou melhor, quando se tornaram reconhecidas como línguas que podiam ser utilizadas para fins literários – no período pós-renascentista.

As línguas literárias eram mais consideradas do que as línguas e dialetos não-literários. E quaisquer diferenças que os gramáticos observavam entre o literário e o coloquial, ou entre a língua-padrão e os dialetos não-padrão tendiam a ser condenadas e atribuídas a desleixo ou à falta de instrução. Poucos dentre eles, se é que havia, davam-se conta da importância do fato de que a transmissão das línguas literárias da Europa de geração para geração não é absolutamente típica do modo pelo qual as pessoas adquirem a sua língua nativa enquanto crianças. Tão pouco dava-se a devida atenção ao fato de que, no caso de muitas línguas modernas, especialmente o inglês e o francês, a ortografia, que ainda é baseada na pronúncia de séculos atrás, dissimula a maior parte das mudanças fonéticas e fonológicas que sofreram. Se somos alfabetizados em francês ou em inglês teremos relativamente pouca dificuldade em ler Shakespeare ou Ronsard; no entanto, acharíamos suas obras mais ou menos incompreensíveis se as ouvíssemos como eram pronunciadas por seus autores. Somente depois de muito trabalho detalhado feito durante o século dezenove, no que hoje podemos considerar como o período clássico da lingüística histórica, entre as décadas de 1820 e 1870, os estudiosos chegaram a uma melhor compreensão da relação entre língua falada e escrita, por um lado, e entre línguas-padrão e não-padrão, por outro.

Com base em tal pesquisa detalhada e pela aplicação do chamado **método comparativo** (que será explicado no parágrafo 6.3), foi demonstrado, sem qualquer sombra de dúvida, que não apenas todas as grandes línguas literárias da Europa tinham se originado sob a forma de dialetos falados, mas que também a sua origem e o seu desenvolvimento só poderiam ser explicados em termos de princípios que determinam a aquisição e o uso da língua falada correspondente. A força das atitudes tradicionais e dos hábitos relacionados à escrita é tamanha, que a maioria das pessoas tem dificuldade de pensar nestes termos sem estar treinada para tal.

Por exemplo, freqüentemente achamos difícil apreender a importância do fato de que, embora as línguas possam se extinguir num determinado momento, de modo a podermos nos referir a elas metaforicamente como moribundas, não há nenhum sentido razoável em utilizar a mesma metáfora orgânica ou biológica para falar das línguas como nascendo.[1] Vale a pena chamar atenção para este fato porque, como veremos, a ter-

[1] As línguas *pidgin* e *crioulas* poderiam, talvez, ser consideradas como o resultado da

minologia da lingüística histórica é bastante metafórica. Agrupamos línguas em **famílias** pela sua **descendência comum** de uma **língua-mãe** mais antiga, e chamamos de **geneticamente relacionadas** as línguas que podem ser identificadas como provenientes de uma língua **ancestral** comum: é o caso das línguas românicas que podem ser relacionadas dessa forma ao latim. Quando tal terminologia foi introduzida na lingüística no século dezenove sua interpretação era mais literal do que hoje em geral, devido à influência do romantismo germânico, de um lado, e do evolucionismo darwiniano, de outro. Devemos levar em conta que não existe um ponto a partir do qual o anglo-saxão repentinamente transformou-se em ou deu à luz o inglês, por exemplo, ou o latim deu à luz as línguas românicas, enquanto ele próprio continuava a existir como língua erudita durante vários séculos. No entanto é assim que os não-lingüistas normalmente concebem a origem das línguas.

A verdade é que a transformação de uma língua em outra não é repentina, mas gradual. O que nos leva a dividir a história do inglês em três períodos – inglês antigo (ou anglo-saxão), inglês médio e inglês moderno – e a considerar cada um deles alternadamente como três línguas diferentes ou como três estágios da mesma língua é mais uma questão de convenção e de decisão arbitrária. Existem razões lingüísticas e não-lingüísticas para traçar essas divisões. O que hoje é o inglês-padrão, nas características essenciais de sua fonologia e gramática bem como em grande parte de seu vocabulário, é um descendente do dialeto de Londres que, estando próximo do ponto em que três dos quatro dialetos principais encontraram-se – o merciano, o saxônico ocidental e o dialeto de Kent – contém traços de todos eles. Inclui igualmente alguns traços mais isoladados que derivam do quarto dialeto principal, o dialeto de Northumberland – especialmente as formas *they, their, them* [respectivamente "eles", "deles", "eles" ou "os" (pronome átono)] e a maioria das palavras iniciadas por *sk* – (*'skill'* [habilidade], *'sky'* [céu], *'skin'* [pele], *'skirt'* [saia] etc.) – que sofreu grande influência da língua dos Vikings a partir do século nove.

Durante aproximadamente um século e meio depois da conquista normanda em 1066 a língua das classes dominantes era o francês, tanto para a literatura quanto para a administração; e quando o inglês voltou a ser usado como língua literária, no início do século treze, ele era notadamente diferente do anglo-saxão do período anterior. Além de outros desenvolvimentos ocorridos, o que hoje chamamos de inglês médio tinha sofrido influência do francês normando e sido influenciado por essa língua

união de duas línguas desempenhando o papel de pais, cada uma das quais continua a existir durante o período de vida de seu rebento (v. 9.3). Mas esta interpretação, mais ou menos aceitável como metáfora, de 'pais' e 'nascimento' não é o que estamos focalizando aqui.

tanto em vocabulário quanto em gramática. Chaucer, por exemplo, escreveu no dialeto londrino do inglês médio que começava a emergir como língua nacional padrão, em virtude da importância política e econômica da capital. No final da Guerra dos Cem Anos, no século quinze, a Inglaterra tinha se tornado bastante consciente de sua própria identidade nacional e tinha se transformado, de um Estado Feudal, num Estado com uma burguesia instruída, rica, e cada vez mais poderosa. Este foi um fator importante na formação e na padronização crescente do inglês médio literário.

O Renascimento, que alcançou a Inglaterra no final do século quinze, separa o período do inglês médio do do inglês moderno. Uma das conseqüências mais importantes no campo da educação e da cultura foi o ressurgimento do latim como língua literária. Mas este fenômeno foi de duração relativamente curta. Embora o latim continuasse a gozar de grande prestígio cultural até o decorrer do século dezenove, as maiores obras literárias dos períodos elizabetano e pós-elizabetano, inclusive as peças de Shakespeare e o *Paraíso Perdido* de Milton, foram escritas em inglês. Enquanto isso a Inglaterra começava a desempenhar um papel de importância crescente nos assuntos mundiais. No século dezessete foram instaladas colônias de língua inglesa na América do Norte. E até o século dezenove o inglês era a língua da administração, da instrução superior e dos altos negócios não apenas nos Estados Unidos, Canadá, Austrália e Nova Zelândia, onde já era a primeira língua da maioria dos colonos dominantes política e economicamente, bem como de seus descendentes, mas também na Índia e em outros países da Ásia e da África dentro do Império Britânico. No período pós-Renascimento o inglês tornou-se uma língua internacional do mesmo modo que o latim tinha se tornado uma língua internacional (no chamado Velho Mundo da Europa, África do Norte e partes da Ásia) quase dois mil anos antes, e pelas mesmas razões. Tanto o latim quanto o inglês nada mais foram originalmente do que os dialetos locais de pequenas tribos – o itálico, num caso, e o germânico, no outro – e não diferiam dos demais dialetos itálicos e germânicos relacionados das tribos vizinhas em nenhum detalhe lingüisticamente relevante.

O resumo breve e altamente simplificado que acabamos de fazer sobre a evolução e expansão do inglês servirá para ilustrar a idéia geral de que, embora haja bons motivos para se dividir tanto a história externa quanto a interna de uma língua em períodos mais ou menos distintos, o processo de mudança lingüística em si é contínuo. O que produz a ilusão de descontinuidade entre o anglo-saxão e o inglês médio, por exemplo, ou, em grau menor, entre o inglês médio e o inglês moderno, é a coincidência de vários fatores que incluem, por um lado, lacunas no registro histórico entre períodos identificáveis e, por outro, a estabilidade relativa das línguas literárias por longos períodos. Temos pouquíssimos registros

escritos dos vários dialetos não-literários do anglo-saxão e do inglês médio. Mas podemos ter certeza de duas coisas: em primeiro lugar, de que desde os tempos mais remotos os dialetos do inglês falado eram menos homogêneos e menos suscetíveis de separação nítida um do outro do que são representados por relatos tradicionais da história do inglês, baseados em dados de textos literários; em segundo lugar, de que, se tivéssemos um registro histórico completo de qualquer um dos dialetos falados, fosse o dialeto londrino ou de algum vilarejo numa parte mais remota do país, não seríamos capazes de identificar qualquer momento definido em que o dialeto em questão repentinamente deixou de pertencer a um período para pertencer a um outro. As línguas mudam mais rapidamente em alguns períodos do que em outros. Até as línguas literárias mudam no decorrer do tempo. E as línguas faladas adquiridas na infância e usadas pela vida numa variedade de situações – línguas vivas no sentido completo do termo – mudam muito mais obviamente do que as línguas literárias. Além do mais, nenhuma língua viva é completamente uniforme (v. 1.6). Este fato, como veremos mais tarde, é crucial para a explicação da mudança lingüística.

No que se segue iniciarei com um relato da lingüística histórica do tipo que poderia ter sido feito por um dos chamados neogramáticos ou seus sucessores (exceto no que diz respeito a detalhes que foram esclarecidos mais recentemente ou que estão especificamente relacionados ao presente). Os neogramáticos (*Junggrammatiker* em alemão) eram um grupo de estudiosos, sediados na universidade de Leipzig, no final do século dezenove, que foram os principais responsáveis pela formulação dos princípios e métodos da lingüística histórica que desde então têm regido a maior parte dos trabalhos nesse campo. Tais princípios e métodos eram controvertidos quando apresentados pela primeira vez. Agora sabemos que grande parte da crítica contra eles foi justificada. No entanto, sua supremacia reinou por quase um século e eles subjazem a muito raciocínio utilizado no dia-a-dia do lingüista que trabalha com mudança lingüística e também constituem a base de todos os tratamentos padronizados de famílias de línguas em enciclopédias e outras obras de referência. Em algumas seções seguintes deste capítulo analisaremos um ou outro princípio neogramático e o reformularemos à luz de trabalhos recentes.

6.2 Famílias de línguas

Dizer que uma ou mais línguas pertencem à mesma família – que são relacionadas geneticamente – é dizer que elas são variantes divergentes, descendentes, de uma língua ancestral comum ou **protolíngua**.

Na maioria dos casos não temos conhecimento direto da protolíngua da qual os membros de determinada família ou subfamília descendem. As línguas românicas são bastante atípicas quanto a esse aspecto: embora o dialeto do latim do qual elas descendem certamente diferisse em muitos detalhes de gramática e vocabulário dos textos que nos chegaram às mãos, por mais coloquiais que fossem, temos uma noção muito melhor do que podemos chamar de **proto-romance** do que da maioria das outras protolínguas.

Em termos gerais, as protolínguas são construtos hipotéticos em favor de cuja existência não existem dados diretos, mas postula-se que tenham existido e que são constituídas de tal e tal estrutura a fim de se dar conta da relação genética entre duas ou mais línguas comprovadas. Por exemplo, o protogermânico é postulado como ancestral das línguas germânicas (o inglês, o alemão, o holandês, o dinamarquês, o islandês, o norueguês, o sueco etc.) e o proto-eslavo como ancestral das línguas eslavas (o russo, o polonês, o tcheco, o esloveno, o servo-croata, o búlgaro etc.). Em cada um desses casos temos dados documentados relacionados à história anterior da família. Para o germânico, além de algumas inscrições fragmentadas anteriores, temos uma tradução da Bíblia, datada do século quatro da era comum, em gótico (como era falado pelos vizigodos, estabelecidos naquela época no baixo Danúbio); textos literários bastante extensos nos vários dialetos do anglo-saxão (ou inglês antigo) que representam o período compreendido entre os séculos sexto e onze; os textos das sagas em islandês antigo (ou nórdico antigo) do século doze; textos em alto alemão antigo, datados da segunda metade do século oitavo, e assim por diante. Quanto ao eslavônio, os dados mais antigos de que dispomos são os textos do século nove escritos em eslavônio litúrgico antigo. Em nenhum dos dois casos temos algo tão próximo da protolíngua ancestral postulada quanto a língua dos textos latinos que chegaram até nós; deve ter sido em relação ao dialeto supostamente mais popular do latim (freqüentemente denominado latim vulgar) que estamos chamando de proto-romance.

Com base em todos os dados disponíveis e aplicando os princípios elaborados no decorrer do século dezenove e formulados em sua essência pelos neogramáticos, os estudiosos podem **reconstruir** com um grau razoável de confiança boa parte do sistema sonoro e parte da estrutura gramatical tanto do protogermânico quanto o proto-eslavônio. Eles podem igualmente reconstruir estágios intermediários no desenvolvimento dos membros de determinada família de línguas, confirmados pelos dados, a partir de seu suposto ancestral comum. Por exemplo, a figura 3 fornece uma representação esquemática do desenvolvimento das línguas germânicas reconhecidas oficialmente, faladas hoje e do gótico, cujo declínio data do início da Idade Média, e que finalmente morreu alguns séculos

mais tarde (dando lugar a um ou outro dialeto eslavônio). Observaremos que o inglês, que, como vimos no parágrafo precedente, já era diferenciado como dialeto na época dos nossos documentos sobreviventes mais antigos, aparece mais relacionado ao frisão do que ao holandês ou alemão, e mais relacionado aos dois do que às línguas escandinavas. O frisão já foi mais amplamente falado do que é hoje. Embora não seja uma língua nacional no sentido em que todas as demais línguas germânicas modernas da figura 3 o são, tem *status* oficial na província da Frísia, no norte da Holanda, onde, no entanto, tem sofrido forte influência, pelo menos no vocabulário, do holandês-padrão. Não somente o inglês, mas todas as línguas modernas da figura 3 existem em vários dialetos; e, com muita freqüência, a transição entre um dialeto e outro não é repentina, mas gradual. Como veremos mais tarde, o diagrama convencional da árvore genealógica do relacionamento entre as línguas tende a simplificar demais os fatos, senão a distorcê-los completamente, por não reconhecer o fenômeno da convergência e difusão, e por representar o relacionamento entre as línguas como resultado de divergência necessária e contínua.

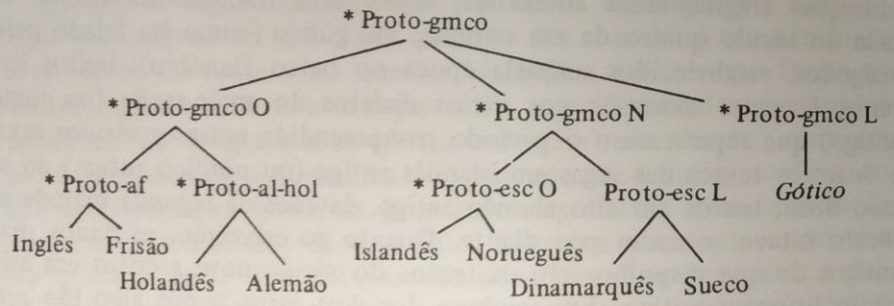

Figura 3. As línguas germânicas. As protolínguas reconstruídas estão precedidas de um asterisco; o nome da língua extinta está em itálico. Gmco = germânico; gmcoO = germânico ocidental; gmcoN = germânico do norte; gmcoL = germânico oriental; af = anglo-frisão; al-hol = alemão da Holanda; escO = escandinavo ocidental; escL = escandinavo oriental. (Foram omitidos muitos detalhes: por exemplo, o alto alemão não está distinguido do baixo alemão, e ancestrais comprovados mais velhos de línguas modernas – o anglo-saxão, o alto alemão antigo etc. – não estão incluídos.)

Retrocedendo ainda mais e levando em consideração uma gama mais ampla de dados, inclusive as inscrições hititas da Ásia Menor (decifradas em 1915), as placas em grego de Míconos (decifradas em 1952) e, para o sânscrito mais antigo, os hinos do Veda – dados que podem ser datados em torno da metade do segundo milênio antes de Cristo – podemos reconstruir parcialmente a fonologia e algumas das características gramaticais e o vocabulário do **proto-indo-europeu**, ancestral hipotético do

protogermânico, proto-eslavônio, protocelta, proto-itálico, proto-indo-iraniano etc., e, em última instância, de todas as línguas indo-européias antigas e modernas.

Podemos até localizar o proto-indo-europeu, com uma certa margem de segurança, no tempo e no espaço – nas planícies do sul da Rússia no quarto milênio antes de Cristo – e, combinando os dados lingüísticos e arqueológicos, dizer algo a respeito da cultura daqueles que o falavam. Por exemplo, a maioria das línguas indo-européias comprovadas primeiro possuem palavras que podem ser remontadas a palavras-fontes hipotéticas que significam "cavalo", "cachorro", "vaca", "carneiro" etc. O fato de o vocabulário reconstruído do proto-indo-europeu conter essas palavras, bem como palavras relacionadas a fiar, tecer, arar e outras atividades agrícolas e de pastoreio, indica com clareza razoável que os falantes do proto-indo-europeu viviam uma vida relativamente estabelecida. Palavras que denotam flora e fauna, condições climáticas etc., permitem-nos identificar, dentro de certos limites, o seu *habitat* geográfico, enquanto o vocabulário comum relativo a instituições sociais e religiosas torna possível inferir acerca das características mais abstratas de sua cultura. Está bastante claro, por exemplo, que sua sociedade era patriarcal e que eles veneravam um deus celeste e outros fenômenos naturais deificados. Quanto aos dados arqueológicos, foi sugerido recentemente que isso indica que os falantes do proto-indo-europeu pertencem à chamada cultura de Kurgan, uma cultura da Idade do Bronze que se espalhou a oeste do sul da Rússia na primeira metade do quarto milênio antes de Cristo e mais tarde pelo leste, penetrando no Irã. Tal hipótese, embora seja a mais plausível dentre as apresentadas até o momento, não é aceita universalmente, e muitos estudiosos são céticos quanto à possibilidade de se afirmar qualquer coisa muito definida sobre o *habitat* e a cultura dos falantes do proto-indo-europeu, com base nos dados disponíveis.

O motivo que me levou a mencionar essa questão é o fato de que a família de línguas indo-européia ocupa um lugar um tanto especial na lingüística histórica. Isto se deve em parte ao fato de que muitas das línguas indo-européias, como vimos, têm registros escritos que remontam a centenas, senão a milhares de anos. Embora muitas das relações existentes no âmbito da família indo-européia pudessem sem dúvida ser estabelecidas com base em línguas faladas modernas, os detalhes de tais relações não poderiam ser determinados – e o proto-indo-europeu não poderia ser reconstruído ao ponto que foi – sem os dados dos textos mais antigos.

Mas afirmar que seria possível agrupar muitas das línguas indo-européias, senão todas, numa única família, mesmo se não tivéssemos registros de estados anteriores dessas línguas, é pressupor que a idéia de agrupar línguas em famílias já deveria ter nos ocorrido e, além disso, que deveríamos ter à nossa disposição um método que seja digno de confiança

para comparar línguas e demonstrar relacionamento genético. Isto nos leva à segunda razão pela qual a família de línguas indo-européia tem lugar de destaque na lingüística histórica: foi a reconstrução do proto-indo-europeu, e das protolínguas intermediárias para as subfamílias do indo-europeu (sobretudo para a subfamília germânica), que forneceu a motivação e, em última análise, a metodologia da qual depende a lingüística histórica tal como a conhecemos. Pode-se discutir que não apenas a lingüística histórica mas a lingüística em si, enquanto disciplina independente e científica, originou-se no que poderia ser descrito, de forma algo romântica, como a busca do proto-indo-europeu no século dezenove.

É hábito datar o começo da erudição indo-européia da afirmação feita em 1786 por Sir William Jones (1746-1794) sobre o sânscrito, a língua antiga sagrada e literária da Índia, e seu relacionamento com o grego, o latim e outras línguas:

> A língua *sânscrita*, qualquer que seja sua antiguidade, tem uma estrutura maravilhosa. É mais perfeita do que o *grego*, mais copiosa do que o *latim*, e mais extraordinariamente refinada do que qualquer uma das duas; no entanto, tem com ambas as línguas uma afinidade tanto nas raízes verbais quanto nas formas gramaticais que não poderia ter sido produzida acidentalmente; tal afinidade é tão forte que nenhum filólogo poderia examinar as três sem acreditar que se originaram de uma fonte comum que, talvez, não mais exista: há uma razão semelhante, embora não exatamente tão contundente, para se supor que tanto o *gótico* quanto o *celta*, apesar de misturados com um idioma bem diferente, tiveram a mesma origem do que o *sânscrito*; e que o *persa* antigo poderia ser acrescentado à mesma família.

Há muito nesta citação famosa que vale a pena observar. Entretanto, deve-se dar ênfase especial ao fato de que o que pareceu uma explicação tão óbvia para Jones, no final do século dezoito, para a semelhança notável que percebeu entre as línguas clássicas da Europa e o sânscrito – a hipótese do relacionamento de família – pode não ter parecido tão óbvio em outra época, ou então para alguém de formação intelectual diferente e com uma perspectiva menos liberal. Idéias evolucionárias pairavam no ar e vinham sendo aplicadas ao estudo das línguas, desde a segunda metade do século dezoito, por estudiosos como Condillac (1715-1780), Rousseau (1712-1778) e Herder (1744-1803), para não mencionar James Burnett (1714-1799) com quem Sir William Jones mantinha correspondência na época. No final do século dezoito, como uma conseqüência da expansão pós-renascentista na Europa, já se sabia muito mais acerca da diversidade das línguas do mundo. Não era mais possível manter que todas as línguas têm que ser estruturalmente semelhantes, com o mesmo grau de plausibilidade das gerações anteriores de estudiosos formados por uma visão clássica. As semelhanças entre o latim e o grego vinham sendo encaradas como

óbvias durante séculos. Mas, no contexto do que era conhecido acerca da diversidade das línguas, o fato de que o sânscrito era impressionantemente semelhante ao grego e ao latim exigia uma explicação; e a explicação que pareceu natural a Sir William Jones, bem como aos seus contemporâneos quando ele a propôs, estava conforme com os movimentos gerais do pensamento europeu da época.

Temos que dizer algo ainda acerca da importância do novo espírito do romantismo, particularmente forte na Alemanha, e sua relação com o nacionalismo. Herder havia defendido a posição de que havia uma ligação íntima entre língua e caráter nacional. Tal idéia enraizou-se na Alemanha e contribuiu para um clima de opinião no qual o estudo de estágios anteriores da língua alemã era encarado como parte integrante da afirmação e autenticação da identidade nacional dos povos falantes de alemão.

Nesse sentido é importante enfatizar a diferença entre língua e raça. Termos como 'germânico' e 'indo-europeu' aplicam-se primordialmente a famílias de línguas. Eles não se aplicam a nada daquilo que um antropólogo físico encararia como raças geneticamente distintas. Não existe, nem jamais existiu algo como uma raça germânica ou indo-européia. Na medida em que o uso desses termos em lingüística histórica implica a existência de uma comunidade lingüística falando protogermânico ou proto-indo-europeu em alguma época e em algum lugar no passado, é razoável supor que os membros dessas comunidades lingüísticas podem ter se considerado como pertencentes aos mesmos grupos culturais e étnicos. A posse de uma língua comum é, e presume-se que sempre tenha sido, um marco importante de identidade cultural e de etnia. Mas não há ligação, a não ser parcial e acidental, entre raça, geneticamente definida, e cultura ou etnia.

É importante enfatizar esse aspecto por dois motivos. O primeiro é que termos como 'germânico' e 'indo-europeu' – alternando com 'nórdico' e 'ariano' – têm recebido freqüentemente uma interpretação racial e até racista. Cabe ao lingüista e ao antropólogo corrigir a base errônea desse tipo de racismo. Não há justificação alguma para a crença na distinção racial entre os falantes das línguas indo-européias; menos ainda para o uso que foi feito do pressuposto de superioridade racial pelos propagandistas nazistas na década de 1930. O mesmo é válido em relação a termos como 'celta', 'eslavônico', ou 'inglês', ou de fato a quaisquer termos que se aplicam primordialmente a famílias de línguas e a línguas.

O segundo motivo para destacar o fato de que não há ligação intrínseca entre raça e língua – e isso reforça o primeiro – é que tal fato nos permite uma melhor compreensão de como são formadas as famílias de línguas e portanto da natureza da linguagem. Não sabemos se alguma

vez existiu uma única protolíngua da qual se originaram todas as línguas humanas, do mesmo modo em que as línguas germânicas descendem do protogermânico, e o protogermânico, por sua vez, do proto-indo-europeu. Não podemos nem mesmo relacionar as línguas indo-européias, com certeza, a nenhuma das outras famílias de línguas principais estabelecidas até agora. É bem possível que todas as línguas remontem, no passado bem distante – talvez há meio milhão de anos – a uma única língua ancestral e sejam, assim, no sentido técnico do termo, membros da mesma família de línguas. Por outro lado, as várias correspondências estruturais bastante impressionantes entre as línguas do mundo, que à primeira vista pareceriam sustentar a hipótese da monogênese, não são menos prontamente analisáveis em termos de difusão e convergência.

A transmissão da linguagem de uma geração para outra é em parte uma questão de biologia e em parte uma questão de cultura. Podemos ser geneticamente programados, enquanto seres humanos, para adquirir a linguagem; mas não somos geneticamente programados para adquirir uma determinada língua. Segue-se que, dadas as condições sociais e culturais certas, não apenas indivíduos, mas comunidades inteiras, podem adquirir uma língua ou dialeto que difere da de seus pais. Os grandes fundadores, no século dezenove, da lingüística histórica, a quem devemos a noção de famílias de línguas com a qual ainda trabalhamos, não atribuíram a esse fato a importância teórica que sabemos que deveria ter. Com demasiada freqüência pressupôs-se que a propalação de línguas por uma grande extensão geográfica implicava grandes movimentos populacionais. Tal pressuposição é, no mínimo, desnecessária. Veremos mais tarde que a difusão e a convergência culturais não são menos importantes para a explicação da mudança lingüística do que a migração e a divergência. O modelo tradicional da árvore genealógica para o relacionamento entre as línguas permite apenas a divergência contínua de línguas de um ancestral comum.

A família de línguas indo-européia é apenas uma entre as numerosas famílias de línguas identificadas até o momento. Alguns estudiosos defenderam uma classificação de todas as línguas do mundo em cerca de trinta famílias principais, das quais algumas das famílias mais amplamente aceitas seriam subfamílias. Mas muito nessas classificação e subclassificação genéticas de línguas ainda é controvertido. Por exemplo, aproximadamente mil línguas diferentes são faladas na África. Todas elas (exceto o inglês, o francês, o espanhol, o africâner etc., que foram levadas para a África no período da colonização européia) foram agrupadas recentemente em quatro grandes famílias, uma das quais, o **camito-semítico** (ou afro-asiático), que compreende todas as línguas nativas faladas ao norte do Saara, contém a família **semítica** tradicionalmente reconhecida, cujos membros mais conhecidos são o árabe, o hebraico e o amárico. De

forma semelhante, as línguas **banto** (que incluem o suaíli, o *xhosa*, o zulu etc.) são atualmente consideradas como uma subfamília **niger-congo** de um modo geral, embora não universalmente. Encontramos o mesmo tipo de situação em relação às línguas faladas em outras partes do mundo. Vamos progredindo gradativamente no agrupamento de um número maior de subfamílias em um número menor do que poderíamos chamar de superfamílias (ou troncos). Como os dados em favor dos agrupamentos maiores são freqüentemente escassos, a classificação genética resultante e que depende deles é respectivamente experimental, e tem que ser tratada como tal. As famílias de línguas identificadas e denominadas pelos lingüistas não são todas igualmente bem estabelecidas.

6.3 O método comparativo

A maneira padrão de demonstrar o relacionamento genético entre as línguas é por meio do chamado método comparativo. Isso foi desenvolvido e refinado durante o que chamamos antes de período clássico da lingüística histórica: entre 1820 e 1870 (v. 6.1). Esse método baseia-se no fato de que a maioria das palavras relacionadas mais obviamente através das línguas podem ser dispostas em correspondência sistemática em termos de sua estrutura fonológica e morfológica. Na década de 1870 os estudiosos haviam alcançado tamanho sucesso na aplicação do método comparativo aos casos mais óbvios de relacionamento genético que eles sentiram-se confiantes na sua eficácia no caso das línguas cujo relacionamento estava longe de ser óbvio.

Exemplificarei o princípio de correspondência sistemática, de início, com as línguas românicas. Com isso temos a vantagem não apenas do fato de que o relacionamento entre elas é inquestionável, mas também de que temos dados diretos da protolíngua da qual elas se derivaram, o latim. Entretanto, como veremos, existem exemplos nos quais as formas do proto-romance que podem ser reconstruídas pelo método comparativo, a partir de dados das próprias línguas românicas, diferem das formas latinas comprovadas.

A tabela 4 compara vários conjuntos de palavras obviamente relacionadas (nas formas em que são normalmente citadas) do latim e de três das línguas românicas: francês, italiano e espanhol. A tabela poderia ser ampliada tanto horizontalmente, incluindo-se as palavras apropriadas de outras línguas e dialetos românicos (romeno, português, catalão, sardenho, ladino etc.), quanto verticalmente, acrescentando-se muito mais conjuntos de palavras relacionadas. Apesar de limitada, a informação fornecida na tabela 4 servirá para ilustrar o que significa o princípio de correspondência sistemática.

Tabela 4
Algumas correspondências sistemáticas formais no latim e três línguas românicas☆

	Latim (L)	*Francês (F)*	*Italiano (I)*	*Espanhol (E)*
(1) "coisa"	causa	chose	cosa	cosa
"cabeça"	caput	chef	capo	cabo
"cavalo"	caballus	cheval	cavallo	caballo
"cantar"	cantare	chanter	cantare	cantar
"cachorro"	canis	chien	cane	
"cabra"	capra	chèvre	capra	cabra
(2) "planta"	planta	plante	pianta	llanta
"chave"	clavis	clef	chiave	llave
"chuva"	pluvia	pluie	pioggia	lluvia
(3) "oito"	octo	huit	otto	ocho
"noite"	nox/noctis	nuit	notte	noche
"fato"	factum	fait	fatto	hecho
"leite"	lacte	lait	latte	leche
(4) "filha"	filia	fille	figlia	hija
"formosa"	formosus			hermoso

O primeiro ponto a observar é que as palavras reunidas em cada linha da tabela 4 são relacionadas não apenas quanto à forma – à qual se aplica o princípio de correspondência sistemática – mas também quanto ao sentido. Mas as palavras podem mudar de significado no decorrer do tempo. Por exemplo, o termo comum do latim clássico para "cavalo" não era *'caballus'*, que significava mais especificamente "cavalo de carga" e também era usado pejorativamente significando "chato" ou "mercenário", mas *'equus'*. Entretanto, *'caballus'* e *'equus'* são obviamente relacionados em termos de significado; e é razoável supor que *'caballus'* perdeu tanto o seu significado específico quanto as suas conotações pejorativas no latim tardio (isto é, no proto-romance) e tornou-se a palavra geral e estilisticamente neutra no lugar de *'equus'*. De maneira inversa, os descendentes da palavra latina *'caput'*, apresentados na tabela, adquiriram todos um conjunto de sentidos mais restritos ou metafóricos: por exemplo, o F. *'chef'* significa "cabeça" no sentido de "chefe", como o italiano *'capo'*; o E. *'cabo'* significa "cabo" (acidente geográfico), "cabo" (na hierarquia

☆ Os termos entre aspas, no original inglês, naturalmente são externos à tabela; em português, no entanto, eles podem ser considerados como membros de uma quarta língua românica e, portanto, integrantes da tabela.

do exército), "fim" etc. Mas novamente existe uma ligação intuitivamente óbvia entre o significado do L. *'caput'* e os de seus descendentes.

Nenhuma das palavras apresentadas na tabela 4 causa problemas quanto ao fato de serem semanticamente relacionadas, apesar de haver margem para discordância sobre a natureza de seu relacionamento, em determinados casos. No entanto, não está claro, com muita freqüência, se duas palavras são ou não semanticamente relacionadas – especialmente no caso de línguas para as quais temos muito menos dados do que no caso das línguas românicas. É por esse motivo que o método comparativo dá prioridade ao relacionamento formal. Devemos observar igualmente que as palavras podem não apenas mudar de significado com o tempo, mas também, por vários motivos, cair em desuso e ser substituídas. Isso explica as lacunas na tabela 4. O espanhol moderno substituiu a palavra derivada do L. *'canis'* por *'perro'*; e nem o italiano nem o francês preservaram em seu vocabulário um descendente do L. *'formosus'*.

Voltemo-nos às correspondências formais exemplificadas na tabela. As palavras estão todas apresentadas na forma em que são normalmente citadas. Devemos ter em mente, portanto, que estamos interessados, em princípio, não em letras, mas em sons. No caso do latim, do espanhol e do italiano, há relativamente pouca discrepância entre ortografia e pronúncia. Temos que lembrar que no espanhol moderno não existe fonema correspondente à letra ⟨*h*⟩; que tanto em espanhol quanto em italiano a letra ⟨*c*⟩ é pronunciada de maneira diferente, em posições diferentes; que ⟨*ch*⟩ é pronunciado como [k] em italiano, mas como [tʃ] em espanhol; e assim por diante. Mas essas discrepâncias são secundárias; podemos, sem violentar demais os fatos, trabalhar a partir do pressuposto de que existe uma correspondência biunívoca entre letras (ou, em alguns casos, grupos de letras: E. ⟨*ll*⟩, ⟨*ch*⟩; I. ⟨*ch*⟩, ⟨*ggi*⟩) e fonemas. A situação do francês é muito diferente. Por exemplo, não há meios de se saber, com base nas convenções ortográficas do francês, que *clef* é pronunciada [kle], mas *chef* é pronunciada [ʃɛf]; ou que *huit* é geralmente pronunciada com um [t] final, enquanto que *nuit* e *lait* não (exceto em algumas expressões cristalizadas); e ainda que existem pronúncias alternativas para *fait*. Entretanto, na medida em que a prática ortográfica francesa baseia-se na pronúncia de alguns séculos atrás (como também é o caso do sistema ortográfico do inglês), podemos considerar as formas escritas do francês, para as nossas finalidades, sem questionamento. Tampouco devemos nos preocupar com o fato de que geralmente não é a forma em que um lexema normalmente é citado que é a fonte das formas diacronicamente relacionadas do francês, espanhol e italiano: é quase sempre a forma do acusativo dos substantivos e adjetivos, e não a forma do nominativo, que fornece a fonte das formas de raiz românicas – *canem, caballum* etc. (perdendo-se o [m] final no latim tardio, ou proto-romance).

Comparando as palavras das diferentes seções da tabela 4, observaremos que existem correspondências regulares entre formas relacionadas (isto é, entre as formas de lexemas relacionados). Essas correspondências estão indicadas em negrito e podem ser representadas em termos de sons em vez de letras, mas, como explicamos acima, considerando a ortografia sem questionamento, da seguinte forma:

(1) L. [k] = F. [ʃ] = I. [k] = E. [k]
(2) L. [pl], [kl] = F. [pl], [kl] = I. [pi], [ki] = E. [ʎ]
(3) L. [kt] = F. [it] = I. [tt] = E. [tʃ]
(4) L. [f] = F. [f] = I. [f] = E. [h]

Tanto o francês ⟨*it*⟩ quanto o espanhol ⟨*h*⟩ estão recebendo aqui um valor fonético que sabemos ser apropriado para períodos anteriores: é isto o que quero dizer quando me refiro ao valor que uma forma escrita aparenta. Poderíamos trabalhar igualmente com transcrições fonéticas (ou com representações fonológicas) das formas faladas modernas. Ainda assim as correspondências formais sistemáticas poderiam ser formuladas. Mas elas seriam menos óbvias à primeira vista. Observaremos que, além de (1)-(4), outras correspondências podem ser extraídas das palavras apresentadas na tabela 4:

(5) L. [b] = F. [v] = I. [v] = E. [b]
(6) L. [a] = F. [ɛ] = I. [a] = E. [a]

e assim por diante. Como explicar essas correspondências sistemáticas?

A resposta fornecida pelos inventores do método comparativo do século dezenove era que as mudanças sonoras que ocorrem em uma língua no curso de sua história são regulares. No entanto, o princípio da regularidade da mudança sonora só veio a ser enfatizado em meados de 1870, quando os neogramáticos proclamaram-no em sua forma mais veemente e inflexível: "As mudanças sonoras que podemos observar na história lingüística documentada procedem de acordo com leis fixas que não sofrem nenhuma perturbação, exceto de acordo com outras leis. "À primeira vista, a tese segundo a qual as **leis sonoras** (como passaram a ser chamadas) funcionavam sem exceção era obviamente falsa. Havia muitos exemplos de palavras evidentemente relacionadas que não apresentavam as correspondências esperadas. Tomemos um exemplo famoso – exemplo esse que se tornou nada mais do que uma aparente exceção quando o problema que levantou foi brilhantemente solucionado pelo erudito dinamarquês Karl Verner, em 1875.

Em 1822, Jacob Grimm (um dos dois irmãos que são mais amplamente conhecidos pelo seu interesse no folclore germânico) revelaram que existe uma correspondência sistemática entre as consoantes das línguas

germânicas, por um lado, e as das outras línguas indo-européias, por outro. Ele não foi o primeiro a notar tal correspondência: o crédito pela observação original tem que ser atribuído ao erudito dinamarquês Rasmus Rask. Mas o trabalho de Grimm, sendo escrito em alemão, era mais imediatamente acessível ao saber internacional; e a lei sonora que foi postulada para dar conta das correspondências observadas é geralmente conhecida como a lei de Grimm. Reformulada em termos da fonética articulatória moderna (e simplificada em alguns aspectos), a lei de Grimm afirma que:

(a) As aspiradas sonoras do proto-indo-europeu (PIE) [$*b^h$, $*d^h$, $*g^h$] tornaram-se oclusivas sonoras [*b, *d, *g] – ou possivelmente fricativas sonoras [*β, *ð, *ɣ] – em protogermânico (PGmco);
(b) as oclusivas sonoras PIE [*b, *d, *g] tornaram-se oclusivas surdas [*p, *t, *k] em PGmco;
(c) as oclusivas surdas PIE [*p, *t, *k] tornaram-se fricativas surdas [*f, *θ, *h] em PGmco.

Os asteriscos indicam, segundo a convenção há muito estabelecida em lingüística histórica, que os sons em questão são reconstruídos, e não diretamente confirmados. Vamos agora nos concentrar na noção de reconstrução. Tanto o proto-indo-europeu quanto o protogermânico são, naturalmente, construtos hipotéticos (v. 6.2).

A lei de Grimm, formulada dessa maneira, engloba um grande número de correspondências observadas. Por exemplo, ela dá conta do fato de que em inglês (In) temos [b] onde em latim (L.), grego (Gr.), sânscrito (S.) etc., temos [p]: cf. I. *father*: L. *pater*, Gr. *pater*, S. *pitar-*; I. *foot*: L. *pes/pedis*, Gr. *pous/podus*, S. *pāt/padas*. Ela também dá conta das consoantes iniciais e médias do gótico *taihun*: L. *decem*, Gr. *deka*, S. *daśa* – In. *ten* não conserva nenhum traço da consoante média, mas compare-se o *zehn* do alemão moderno, bem como o *zehan* do alto alemão antigo e o *tehan* do saxão antigo (o [t^s] inicial do alemão representado pela letra <z> na ortografia resulta da chamada mudança sonora do alto alemão, que ocorreu provavelmente por volta do século sexto da era comum). O som [ʃ] do S. *daśa*, aqui representado por <s>, resulta de uma palatalização do PIE [*k], que, no início da pré-história, afetou muitas das subfamílias orientais das línguas indo-européias, incluindo todas as línguas indo-iranianas, bálticas e eslavas, bem como as línguas armênias e albanesas. Existem algumas complicações ligadas à reconstrução do que aqui considerei como oclusivas velares, por uma questão de simplicidade, [$*g^h$, *g, *k], no PIE; mas elas não são relevantes para a formulação geral nem para a validade da lei de Grimm. Levando em conta desenvolvimentos subseqüentes em determinadas línguas, ou protolínguas intermediárias, a lei de Grimm, tal como foi resumida no parágrafo anterior, é sustentada por

um número impressionantemente grande de exemplos de correspondência sistemática.

Mas havia muitas exceções aparentes. O próprio Grimm havia comentado sobre algumas delas: "A mudança sonora ocorre na maioria dos casos, mas nunca se esgota completamente em cada caso individual; algumas palavras permanecem na forma que tinham no período anterior; a corrente da inovação passou por cima delas." Por exemplo, o [*p*] do In. *spit, spew* corresponde a um [p] em outras línguas, numa violação aparente da lei de Grimm: L. *spuo* etc. Coisa semelhante ocorre com Gmco [t], Gr. [t], S. [t]: cf. In. *stand*: L. *sto/stare* etc. Aqui, de fato, a corrente da inovação, como disse Grimm, deixou intactas as consoantes germânicas. Observaremos, entretanto, que em cada caso, as oclusivas surdas, [*p, *t, *k], ocorrem na posição de segundo segmento de encontros consoantes de dois segmentos. Tudo o que temos a fazer, portanto, é modificar a formulação da lei de Grimm dada acima, de modo a que se suponha não ter se aplicado aos **reflexos** germânicos (isto é, descendentes) de PIE [*p, *t, *k] nesse tipo de ambiente fonético (ou fonológico). Com efeito, estamos afirmando – para introduzir alguma terminologia moderna – que a lei de Grimm dá conta de uma mudança sonora **condicionada foneticamente**. Assim formulada, a preservação de uma oclusiva surda em palavras como In. *spit/spew, stand, eight* etc., pode ser encarada como um desenvolvimento regular.

Mais interessante ainda é uma outra classe de exceções aparentes. Se consideramos as palavras para "pai" e "irmão" em várias línguas germânicas diferentes do alemão, vemos que elas diferem entre si quanto à consoante média: Go. *fadar: brôθar*, A. *Vater*: *Bruder* etc. E o inglês antigo apresenta a mesma diferença: In. antigo *faeder: brōþor*. O fato de o alemão, como o seu ancestral alto alemão antigo (*fater: bruodar*), ter uma oclusiva surda na palavra para "pai" e uma oclusiva sonora na palavra para "irmão" pode ser explicado, mais uma vez, como sendo a conseqüência da mudança sonora do alto alemão. Vamos admitir que, com base em todos os dados, possamos reconstruir como as origens protogermânicas das palavras em questão, **fader-* e **brôθar-*, o que é representado por <*d*> como sendo uma oclusiva [d] ou fricativa [ð], mas sonora em ambos os casos, diferindo, assim, da fricativa surda [θ] da palavra para "irmão". Como as palavras correspondentes das línguas indo-européias não-germânicas não revelam tal diferença (L. *pater: frater*, S. *pitar-: bhrātar-* etc.) e, segundo a lei de Grimm, PIE [*t] deveria fornecer PGmco [*θ], a palavra para "pai" parece irregular no tocante à sua consoante média, mas não à inicial.

Foi este o problema que Verner solucionou. Ele demonstrou que, supondo-se que as palavras PIE para "pai" e "irmão" tenham diferido entre si quanto ao lugar do acento de palavra, como ocorre em sânscrito (*pitár-: bhrâtar-*), o *status* aparentemente excepcional do Pgmco. **fader* poderia ser satisfatoriamente explicado em termos do que agora chamamos de **lei de Verner**: as fricativas surdas intervocálicas, por exemplo, [θ], tornam-se sonoras a não ser que estejam precedidas de um acento de palavra. O que se supõe é uma seqüência de estágios que podem ser representados da seguinte maneira:

(i) PIE *pətér-: *bʰráter-
(ii) *faθér-: *bróθar-
(iii) *faðér-: *bróθar-
(iv) Pgmco *fáder-: *bróθar-

A lei de Grimm é tida tradicionalmente como responsável pela transição de (i) para (ii), e a lei de Verner pela transição de (ii) para (iii). Considera-se que ambas as leis, como se pode observar, agiram antes do período que identificamos como o do protogermânico, que se caracteriza por ter o acento na primeira sílaba de todas as palavras. As mudanças sonoras de responsabilidade das leis de Grimm e de Verner juntas poderiam ser explicadas de maneira diferente hoje. Isto não traz conseqüências para a nossa discussão. A questão é que toda uma classe de exceções aparentes à lei de Grimm situa-se no âmbito de uma generalização adicional: uma outra lei sonora, conforme demonstrado por Verner.

Várias outras chamadas lei sonoras foram formuladas por volta da mesma época que a lei de Verner. Consideradas juntas, elas permitiram aos estudiosos ter uma idéia bem melhor da cronologia relativa dos desenvolvimentos dentro dos diferentes ramos da família de línguas indo-européia. Mais importante ainda, elas tornaram o famoso princípio neogramático da regularidade absoluta da mudança sonora muito mais plausível do que teria parecido a uma geração anterior de lingüistas especializados em lingüística histórica. Esse princípio era altamente controvertido quando foi apresentado na década de 1870. No entanto, logo tornou-se aceito, pela maioria daqueles que participavam do que podemos considerar como a principal corrente da erudição, como a base mesma, não apenas do método comparativo, mas de toda a disciplina da lingüística histórica. Mais tarde teremos a oportunidade de observar de modo mais crítico o princípio da regularidade da mudança sonora, bem como o uso que os neogramáticos faziam do termo 'lei' a ele relacionado. Mas nada do que se afirma de suas qualidades deve ser subtraído de sua importância metodológica. Ele obrigou aqueles que a ele aderiram a distinguir entre mudança sonora condicionada e não condicionada e a formular tão precisamente quanto possível as condições sob as quais determinada mudança sonora concidionada ocorreu. E ainda colocou-lhes sobre os ombros o ônus de fornecer uma explicação para as formas que não tinham se desenvolvido de acordo com as leis sonoras a cujas condições tais formas pareciam obedecer. Dois dos fatores explicativos aos quais os neogramáticos e seus seguidores apelavam nesse sentido eram a analogia e o empréstimo, como veremos (v. 6.4).

Nesta seção estamos concentrados na técnica da reconstrução histórica pelo método comparativo. Devemos chamar a atenção do leitor para a importância de não interpretar erradamente as formas marcadas por asterisco☆ (isto é, as formas hipotéticas: por exemplo, PIE **pətér-* ou Pgmco **fader-*) que resultam da aplicação da técnica da reconstrução. Elas não podem ser identificadas com as formas comprovadas do proto-indo-europeu ou de qualquer outra protolíngua. Há várias razões para isso.

Em primeiro lugar, o método comparativo tende a exagerar o grau de regularidade num sistema lingüístico reconstruído. Esta observação pode ser ilustrada com referência às diferenças entre certas formas comprovadas do latim e o que consideramos como formas pertencentes ao proto-romance, das quais se derivam as formas correspondentes em francês, italiano, espanhol etc. A palavra latina comprovada para "cabeça" tinha a forma *caput* para citação e *capit-* como radical. Nenhuma das línguas românicas conserva qualquer prova do [t] final do radical. Em vez disso, elas indicam que a forma do proto-romance era **capu(m)*: ver tabela 4 anterior. É bastante provável, no entanto, que o substantivo irregular 'caput' tenha sido regularizado no latim tardio. Mas não temos nenhuma comprovação direta de que tenha sido. Pode ter sido regularizada independentemente em ramos diferentes da família românica, mas em data relativamente prematura. A questão é que as irregularidades tendem a desaparecer no decorrer do tempo e, de um modo geral, o método comparativo é incapaz de reconstruí-las.

Em segundo lugar, o método comparativo funciona a partir do pressuposto de que cada membro de uma família de línguas relacionadas encontra-se numa linha direta de descendência da protolíngua e permaneceu intocada, durante esse tempo, pelo contato com outras línguas e dialetos relacionados. Tal pressuposto é, no mínimo, irreal. Todas as línguas são, em grau maior ou menor, diferenciadas quanto a dialetos. Não há razão para crer que o proto-indo-europeu, o protogermânico, o proto-eslavo e as outras protolínguas que tomamos hipoteticamente como a fonte de famílias e de subfamílias de línguas comprovadas eram homogêneas do ponto de vista dialetal. Sempre que possível, o método comparativo reconstrói uma protoforma única para todas as formas comprovadas. Segue-se que o sistema lingüístico reconstruído tende a ser, não apenas morfologicamente mais regular, mas também dialetalmente mais uniforme do que qualquer sistema lingüístico real. Além disso, não temos meios de saber se todos os sons que ocorrem sob forma com asterisco ocorreram de fato ao mesmo tempo e no mesmo dialeto da protolíngua.

☆ No original, ***starred forms***, literalmente "formas estreladas", em alusão ao asterisco.

Por essas e outras razões, as protolínguas reconstruídas têm que ser consideradas como construtos hipotéticos, cuja relação com as línguas de fato faladas no passado é um tanto indireta. Não podemos aprofundar os detalhes técnicos dessa questão – nem os vários critérios que devem ser utilizados no processo de reconstrução. Basta, para as nossas finalidades, ter chamado atenção para o fato de que toda reconstrução histórica tende a idealizar e a simplificar os fatos. No que diz respeito às formas com asterisco, algumas partes da reconstrução podem ter bases mais seguras do que outras; e nenhuma parcela pode ser mais segura do que os dados que a comprovam. Os dados são altamente variáveis.

Concluindo, devemos mencionar que, embora tenhamos nos concentrado nessa seção na reconstrução lexical, é possível, em casos favoráveis, reconstruir características da estrutura gramatical de protolínguas. Foram certamente as correspondências morfológicas entre línguas relacionadas que impressionaram a primeira geração de lingüistas especializados em lingüística histórica dessa maneira, já que traços gramaticais como sufixos, segundo eles, dificilmente passariam de uma língua para outra por empréstimo (v. 6.4).

6.4 Analogia e empréstimo

O conceito de **analogia** remonta à antiguidade. O termo em si vem da palavra grega 'analogia', que significava "regularidade" e mais especialmente, no uso dos matemáticos e gramáticos, "regularidade proporcional". Por exemplo, a regularidade proporcional entre 6 e 3, por um lado, e 4 e 2, por outro, é uma analogia, no sentido desejado do termo: é uma relação de quatro quantidades (6, 3, 4, 2) de modo a que a primeira dividida pela segunda é igual à terceira dividida pela quarta (6 : 3 = 4 : 2). O raciocínio analógico foi largamente utilizado por Platão e Aristóteles e por seus seguidores, não apenas em matemática, mas também no desenvolvimento de outros ramos da ciência e da filosofia, incluindo a gramática. Sem levar isto em conta, é impossível compreender um dos princípios básicos da gramática tradicional: o de paradigma. Dado, digamos, o paradigma *jump, jumps, jumping* e *jumped* (i.e., as formas do verbo modelo do inglês 'jump': o termo 'paradigma' vem da palavra grega que significa "modelo" ou "exemplo"), podemos construir equações proporcionais como as seguintes: *jump : jumps = help : x; jump : jumped = help : y*; etc. E podemos resolver essas equações atribuindo às quantidades desconhecidas (*x, y*, etc.) os seus valores apropriados (*helps, helped*, etc.).

É isto, então, o que significa 'analogia' na gramática tradicional, e mais particularmente na controvérsia entre analogistas e anomalistas,

surgida no segundo século antes de Cristo e que durou, de uma forma ou de outra, até a atualidade, e exerceu uma influência profunda no desenvolvimento da teoria lingüística. *Grosso modo* podemos dizer que os analogistas eram de opinião que a relação entre forma e significado era governada pelo princípio da regularidade proporcional, e que os anomalistas adotavam o ponto de vista oposto. Não precisamos entrar nos detalhes dessa controvérsia confusa e perturbadora. No entanto, é importante ter em mente que tudo isto faz parte das bases sobre as quais os neogramáticos estabeleceram a sua própria noção de analogia e o papel desta no desenvolvimento das línguas.

Tomemos um exemplo. O inglês, como o alemão, distingue entre o que se chama convenientemente de verbos fracos e fortes. Os primeiros, que constituem a grande maioria de todos os verbos da língua, formam o passado pelo acréscimo de um sufixo ao radical do presente (cf. In. *jump-s, jump-ed;* A. *lieb-t, lieb-te*); os fortes apresentam uma diferença, de um tipo ou de outro, nas vogais dos radicais correspondentes do presente e do passado, e geralmente não possuem o sufixo de passado característico dos verbos fracos (cf. In. *ride-s, rode; sing-s, sang;* Al. *reit-et, ritt; sing-t, sang*). Os verbos fortes classificam-se em várias subclasses segundo a natureza da alternância vocálica que distingue as suas formas de presente e passado; e são comumente considerados irregulares. Eles são certamente menos regulares do que os verbos fracos, que vêm aumentando durante muitos séculos e que obedecem à regra sincrônica produtiva, existente de longa data. Os dados em favor da produtividade sincrônica da regra em questão provêm em parte da aquisição da linguagem pelas crianças e em parte da habilidade do falante adulto de formar o passado de verbos novos que encontra pela primeira vez na forma do presente (ou do particípio presente, em inglês). Quanto à aquisição da linguagem, o fato de que a criança já domina a regra de formação do passado das formas regulares pela sufixação é comprovado pela sua produção, não apenas de um grande número de formas corretas (por exemplo, *jumped* [pulou], *walked* [caminhou], loved [*amou*]), mas também ocasionalmente da forma incorreta como em *rided* [cavalgou] ou *goed* [foi].☆ Na realidade, por mais paradoxal que possa parecer à primeira vista, a produção de formas incorretas como essas por analogia com algum membro típico da classe regular dos verbos fracos (*jump* : *jumped* = *ride* : *x*; então *x* = *rided*) constitui prova mais convincente de que a criança está aplicando uma regra do que a produção de um número qualquer de formas corretas do passado, que poderiam ter sido, em princípio, recordadas e imitadas como todos inanalizáveis (8.4).

☆ O equivalente em português seriam formas como "fazi", por "fiz" etc.

Existem alguns exemplos de verbos fracos que se transformaram em verbos fortes por força da analogia na história do inglês. Por exemplo, em alguns dialetos americanos a forma do passado de *'dive'* [mergulhar] é *dove*, e não *dived*; e, ao contrário do que se poderia supor, a forma *dove* é que é o resultado da inovação. Na grande maioria dos casos, entretanto, a analogia operou na direção contrária, aumentando o número dos verbos fracos em detrimento dos fortes: por exemplo, *holp* (v. *got*) do inglês médio foi substituído por *helped* do inglês moderno. Devemos observar que tanto *dived* → *dove* quanto *holp* → *helped* são resultados de pressão analógica. A língua apresenta dois padrões de formação, qualquer um dos quais poderia servir de paradigma para a extensão analógica.

Vale observar aqui que o fato de que muitos verbos do inglês e do alemão obviamente relacionados, apresentam o mesmo fenômeno de alternância vocálica constitui um dado particularmente impressionante em favor da hipótese segundo a qual essas duas línguas são, de fato, geneticamente relacionadas: cf. In. *begin-s, began, begun* : A. *beginn-t, begann, begonn-en;* In. *bring-s, brought;* A. *bring-t, brach-te, ge-brach-t;* In. *find-s, found* : A *find-et, fand, ge-fund-en;* In. *give, gave, giv-en* : A. *gib-t, gab, ge-geb-en.* (Acrescentei a forma do particípio passado tanto para o alemão quanto para o inglês quando ela difere da forma do passado, como quase sempre acontece em alemão.) A analogia funcionou independentemente tanto no alemão quanto no inglês, durante várias centenas de anos, no sentido de reduzir a incidência da alternância vocálica: assim, por exemplo, enquanto *'help'* é fraco no inglês moderno, o verbo relacionado *'hilfen'* em alemão é forte (*hilf-t, half, ge-holf-en*). As mudanças sonoras que ocorreram independentemente nas várias línguas germânicas também surtiram efeito, aumentando o número de conjuntos diferentes de alternâncias vocálicas e tornando a correspondência entre as formas de determinados verbos isolados menos sistemática, quanto a detalhes, do que era em períodos anteriores. Mas ainda existem dezenas de verbos que apresentam uma alternância vocálica semelhante. O mesmo ocorre com o holandês, que, como vimos antes (v. 6.2), é relacionado mais de perto ao alemão do que ao inglês: cf. *begin-t, begon, begonn-en; breng-t, brach-t, ge-brach-t* ("trazer"); *vind-t, vond, ge-vond-en* ("encontrar"); etc. Até as línguas germânicas do norte têm verbos fortes cujas formas de passado e de particípio passado podem ser relacionadas às formas do presente em termos de alternâncias vocálicas mais ou menos regulares: v. do sueco, *skriver, skrev, skrivit* ("escrever"); *kryper, krop, krupit* ("rastejar"). Na realidade, esse padrão de alternância vocálica remonta em última instância ao período do proto-indo-europeu: cf. Gr. *peith-ō, pe-poith-a, e-pith-on* ("persuadir"); *leip-ō, le-loip-a, e-lip-on* ("sair"); etc. Como foi mencionado no final da seção precedente, foi este o tipo de correspondência – à qual Sir William Jones se referiu como "uma afinidade tanto nas raízes verbais quanto nas formas

gramaticais que não poderia ter sido produzida acidentalmente" (v. 6.2) – que tanto interessou os fundadores da filologia comparativa. Mas aqui estamos focalizando a analogia; e a esse respeito há duas colocações a fazer.

A primeira foi especialmente enfatizada pelos neogramáticos: a de que a analogia freqüentemente inibe (ou inverte subseqüentemente o efeito das) mudanças sonoras que, não fosse isso, seriam regulares. Por exemplo, depois do funcionamento da lei de Verner (v. 6.3), mas anteriormente aos nossos textos mais antigos, o [s] intervocálico tornou-se [r] em germânico. É esta mudança sonora que dá conta da letra ⟨*r*⟩ – ainda pronunciada como [r] em alguns dialetos – no plural do passado do verbo 'to be' [ser, estar] em inglês, contrastando com o que a ortografia mostra ter sido um [s] no singular: *were : was.* O holandês apresenta o mesmo contraste (mas sem diferença nas vogais): *ik was* "eu era/fui/estava/estive": *wij waren* "nós éramos/estávamos/fomos". O alemão, por outro lado, alterou o radical do singular por analogia com o do plural: *ich war : wir waren.* Neste exemplo o [s] final historicamente regular foi substituído pelo [r] historicamente irregular. É interessante notar que no latim bem remoto o [s] intervocálico também tornou-se [r]. Daí o contraste, no latim clássico, entre a forma do nominativo singular *honos* "honra" e as demais formas do mesmo nome: *honorem, honoris,* etc. (de **honosen, *honosis,* etc.). E mais tarde no latim *honos* foi substituído por *honor,* de modo que *honor-* generalizou-se como o radical por todo o conjunto de formas flexionadas. Vale também acrescentar que a analogia é responsável pelo fato de que o verbo *'to be'* é o único verbo no inglês moderno padrão em que há uma diferença entre um radical do singular e um radical do plural, no passado. No inglês médio, muitos dos verbos fortes apresentavam uma diferença semelhante. A analogia mais uma vez generalizou um dos dois radicais (ou, em alguns casos, a forma do particípio passado); e isto responde pela flutuação considerável através dos dialetos do inglês e até mesmo no uso espontâneo de falantes isolados.

A segunda colocação a ser feita com respeito à analogia é a de que tal fator é muito mais poderoso na mudança lingüística do que supunham os neogramáticos. Na realidade, eles tendiam a invocar a influência da analogia somente quando isto lhes permitia justificar exceções aparentes a uma das leis sonoras que postulavam. Além disso, alguns deles fizeram uma distinção entre mudança sonora como um processo explicável fisiologicamente, e analogia como algo resultante da intervenção esporádica e imprevisível da mente humana. Para aqueles que adotavam esse ponto de vista as leis sonoras eram tidas como comparáveis às chamadas leis da natureza. Hoje em dia está mais claro, em primeiro lugar, que não se pode fazer uma distinção tão nítida, tratando-se de línguas, entre o físico e o psicológico; e, em segundo lugar, que a analogia funciona tanto no nível fonológico quanto no nível gramatical da estrutura lingüística – contanto que o termo tradicional "analogia" seja interpretado segundo o espírito, e não a letra, da tradição. O que era tradicionalmente descrito em termos de regularidade proporcional pode ser incluído no princípio mais geral da regularização com bases nos padrões de correspondência entre forma e significado. De fato, seria razoável identificar tanto a noção saussureana de estrutura quanto a noção gerativista de criatividade regida por regra com uma versão modernizada do conceito tradicional de analogia. Mas esta é uma questão mais ampla e controvertida (v. 7.4).

Um outro fenômeno para o qual os neogramáticos apelaram para explicar algumas das exceções aparentes às leis sonoras foi o **empréstimo.** Por exemplo, além da palavra *'chef'*, que identificamos acima como o descendente francês do latim *'caput'*, cuja forma em que normalmente é citado em proto-romance pode bem ter sido **capu(m)* (ver tabela 4), encontramos também, no francês moderno, a palavra *'cap'* (cf. *'de pied en cap'*, literalmente "do pé à cabeça"). A forma *cap* viola claramente todas as três leis sonoras (além da perda da vogal final) necessárias para derivar *chef* de **capu.* A explicação é que a forma foi tomada emprestada pelo francês (em época relativamente remota) ao provençal, ao qual não se aplicaram as leis sonoras em questão. De modo semelhante, existem muitas palavras em inglês que começam por *sk-* em sua forma escrita (cf. *sky* [céu], *skill* [destreza], *skirt* [saia], etc.) que são exceções aparentes à lei sonora que transforma [sk] em [ʃ] antes de vogais anteriores em inglês (cf. *shirt* [camisa], *ship* [navio], *shed* [barraca], etc.). Essas palavras foram tomadas emprestadas de um outro dialeto escandinavo trazido à Inglaterra na época das invasões dos Vikings, dialetos esses que tiveram uma influência considerável na fala da região dominada pelos dinamarqueses. (Até hoje grande parte do vocabulário dos dialetos locais do norte da Inglaterra e do sul da Escócia é de origem escandinava, facilmente identificável; mas aqui estamos concentrados naquilo que podemos considerar como empréstimos no inglês-padrão.) Pares de palavras análogas nativas ou emprestadas, co-existentes, como *'skirt'* e *'shirt'*, em inglês, ou *'cap'* e *'chef'*, em francês, são freqüentemente denominadas **duplas.** As duplas lexicais, como podemos notar, são muito raramente sinônimas, mesmo do ponto de vista descritivo (cf. *'skirt' : 'shirt', 'skipper'* [capitão de navio] : *'shipper'* [embarcador].

Podemos afirmar o mesmo, acerca de empréstimo, que afirmamos sobre analogia: trata-se de um fator muito mais importante para a mudança lingüística do que supunham os neogramáticos (e muitos dos seus seguidores). Particularmente, o empréstimo, como a analogia, não deve ser visto como fonte de mera explicação para exceções aparentes às leis sonoras. Se o inglês é considerado uma língua puramente germânica ocidental – e ela o é, convencionalmente (v. 6.2) – temos que dizer que, no decorrer de

sua história, tomou emprestados de outras línguas e dialetos não apenas aspectos de seu vocabulário, mas também de sua gramática e fonologia.

Mas cabe perguntar se faz sentido falar em termos tais que impliquem uma distinção nítida entre formas nativas e não nativas. É evidente, há muito tempo, que os diagramas de árvores genealógicas para ilustrar o desenvolvimento das línguas e o relacionamento entre elas podem levar a erro, se interpretados como modelos realistas dos processos históricos. Trabalhos mais recentes em dialetologia e em sociolingüística tornaram clara a importância da variação dialetal e estilística dentro de uma comunidade lingüística como fator causador de mudança lingüística. Sob condições de variação sincrônica – e mais especialmente de bilingüismo e diglossia (v. 9.4) – o conceito tradicional de empréstimo talvez não seja aplicável.

Seja como for, é certo que os neogramáticos estabeleceram uma distinção nítida demais entre o que poderia ser tratado em termos de leis sonoras e o que deveria ser explicado por analogia ou empréstimo. Entretanto, a maior parte das explicações sobre o desenvolvimento histórico das línguas ainda seguem os neogramáticos quanto a esse aspecto.

6.5 As causas da mudança lingüística

Por que as línguas mudam no decorrer do tempo? Não há consenso na resposta a esta pergunta. Muitas teorias sobre a mudança lingüística foram propostas. Mas nenhuma delas dá conta de todos os fatos. O máximo que podemos fazer aqui é mencionar, e comentar brevemente, alguns dos principais fatores aos quais os estudiosos se referiram na explicação da mudança lingüística.

É costume, na discussão dessas questões, trabalhar com duas distinções em separado: (a) entre mudança lingüística, por um lado, e mudanças gramatical e lexical, por outro; (b) entre fatores internos e externos. Mas nenhuma dessas duas distinções deve ser exagerada. Como vimos, o enfoque dos neogramáticos de que a mudança sonora é radicalmente diferente de outros tipos de mudança lingüística é, na melhor das hipóteses, uma meia verdade. Mesmo aqueles processos explicáveis mais ou menos fisiologicamente como a **assimilação** (que resulta na transformação de sons para idênticos ou semelhantes a outros, em função de situação ou modo de articulação: cf. I. *otto, notte,* etc., na tabela 4 da seção 6.3), ou a **haplologia** (perda de uma ou duas sílabas foneticamente semelhantes em uma seqüência: por exemplo, **Engla-land* "país dos anglos" > *England* [Inglaterra]) requerem o apoio de outros fatores mais gerais para produzirem mudanças permanentes no sistema sonoro de uma língua. Quanto à distinção entre fatores internos e externos, que depende da abstração do sistema lingüístico, enquanto tal, da matriz cultural e social em que

funciona, essa também falha, em última instância: a função comunicativa da linguagem, que interrelaciona forma e significado dentro de um sistema lingüístico, também relaciona tal sistema à cultura e à sociedade a cujas necessidades ele serve.

Dois dos fatores mais gerais da mudança lingüística foram mencionados na seção precedente: analogia e empréstimo. Podemos agora enfatizar que muito do que os neogramáticos resolveram em termos de leis sonoras pode ser incluído no âmbito da ação conjunta desses dois outros fatores. As leis sonoras em si não têm valor explanatório: elas constituem apenas um resumo do que sucedeu em determinada área (mais precisamente, em determinada comunidade lingüística) entre dois pontos na linha do tempo. Fazendo um retrospecto, e adotando um ponto de vista macroscópico, a mudança em questão pode parecer suficientemente regular (no sentido em que o princípio da regularidade era compreendido pelos neogramáticos e seus seguidores). No entanto, a investigação das mudanças sonoras que estão ocorrendo atualmente mostrou que elas podem se originar em uma ou mais palavras emprestadas e a partir daí espalhar-se por analogia por um período de tempo.

Um dos sintomas desse processo de mudança lingüística é o que se chama comumente de **hipercorreção**. Um exemplo disso é a extensão analógica da vogal do inglês sulista de *butter* [manteiga] em palavras como *butcher* [açougueiro] por falantes do norte da Inglaterra que adquiriram (i.e., tomaram emprestada) a pronúncia RP daquela classe de palavras. A hipercorreção fonética desse tipo não difere, quanto ao que a causa, da hipercorreção que resultou, na fala da classe média e freqüentemente na de pessoas instruídas, falantes do inglês sulista padrão, em dizer *between you and I* [entre você e eu].☆ Logo veremos que o primeiro tipo de hipercorreção, mas não o segundo, poderia levar eventualmente ao que poderia ser descrito, macroscópica e restrospectivamente, como uma mudança sonora regular.

Não estamos sugerindo, naturalmente, que toda mudança sonora pode ser explicada dessa forma. Temos ainda que admitir a possibilidade de **deriva fonética** gradual e imperceptível através do tempo por todas as palavras em que determinado som ocorre. O que queremos dizer é que há uma variedade de fatores que podem interagir de modo a produzir o mesmo tipo de resultado final: algo que normalmente é tido como mudança sonora regular e que, pelo menos na tradição neogramática, é contrastado com fenômenos alegadamente esporádicos como analogia e empréstimo.

Aqueles que enfatizam a distinção entre fatores internos e externos – e mais especialmente os que se identificam com as doutrinas do estrutu-

☆ Ver a segunda nota do tradutor no capítulo 2.

ralismo e do funcionalismo (v. 7.2, 7.3) – tendem a atribuir a mudança lingüística tanto quanto possível ao que se classifica como fatores internos: principalmente às readaptações contínuas que são efetuadas por um sistema lingüístico na passagem de um estado de equilíbrio (ou quase equilíbrio) para outro. Um dos proponentes mais influentes deste ponto de vista foi o erudito francês André Martinet, que tentou dar conta da mudança lingüística, e mais particularmente da mudança sonora, com referência à sua concepção das línguas como sistemas semióticos auto-reguladores, regidos pelos princípios complementares do menor esforço e da clareza comunicativa. Aquele princípio (sob o qual podemos incluir fenômenos explicáveis fisiologicamente como a assimilação e a haplologia, e também a tendência a encurtar formas altamente predizíveis) terá o efeito de reduzir o número de distinções fonológicas e de maximizar a ação de cada uma delas. Ele será controlado, no entanto, pela necessidade de se manter um número suficiente de distinções com a finalidade de separar emissões que de outra forma, seriam confundidas nas condições acústicas em que as línguas faladas são normalmente usadas. Tal noção tem uma boa dose de apelo intuitivo e um certo número de mudanças sonoras foram explicadas nesses termos. Até o momento, no entanto, não foi demonstrado que ela tem todo o poder explanatório que seus proponentes alegam em seu favor.

A principal contribuição que os estruturalistas e os funcionalistas fizeram para a lingüística histórica está na sua insistência de que cada mudança postulada em um sistema lingüístico tem que ser avaliada em termos de suas implicações para o sistema como um todo. Por exemplo, eles tornaram claro que várias partes da lei de Grimm (ou da grande mudança vocálica, ocorrida na transição do inglês médio para o inglês moderno, no início) têm que ser consideradas juntas. E levantaram questões interessantes sobre os tipos de **reações em cadeia** que parecem ocorrer em determinados períodos no desenvolvimento histórico das línguas. Vamos tomar a lei de Grimm como exemplo: as aspiradas sonoras do PIE [*b^h, *d^h, *g^h], ao perderem a aspiração, fizeram com que as oclusivas sonoras não aspiradas do PIE [*b, *d, *g], perdessem a sua sonoridade e isto, por sua vez, fez com que as oclusivas surdas do PIE [*p, *t, *k], se tornassem fricativas? Ou foram as oclusivas surdas do PIE que iniciaram o processo, puxando as demais, por assim dizer, para o lugar que elas vagavam? Tais perguntas podem não ter resposta. Mas pelo menos reconhecem o fato de que várias mudanças resumidas na lei de Grimm podem estar ligadas por uma relação causal.

O que hoje é denominado **reconstrução interna** (contrastando com reconstrução através do método comparativo) também pode ser creditado ao estruturalismo. Isto baseia-se na convicção segundo a qual regularidades parciais sincronicamente observáveis e assimetrias podem ser explicadas

com referência ao que teria constituído processos produtivos num período anterior. Por exemplo, mesmo se não tivéssemos dados comparativos nem registros de estágios anteriores do desenvolvimento do inglês, poderíamos inferir que as regularidades parciais evidentes nos verbos fortes (cf. *drive* : *drove* : *driven* [dirigir], *ride* : *rode* : *ridden* [cavalgar]; *sing* : *sang* : *sung* [cantar], *ring* : *rang* : *rung* [tocar]; etc.) eram as relíquias, por assim dizer, de um sistema mais globalmente regular de flexão verbal. A reconstrução interna é hoje uma parte reconhecida da metodologia da lingüística histórica e tem provado o seu valor em várias ocasiões.

Como veremos mais tarde, o gerativismo desenvolve-se a partir de, e em parte continua, uma versão determinada do estruturalismo. É característico do gerativismo encarar a mudança lingüística em termos do acréscimo, da perda ou da reordenação das regras que determinam a competência lingüística de um falante. Na medida em que a distinção competência/desempenho pode ser identificada com a distinção *langue/parole* do estruturalismo saussureano (v. 7.2), a contribuição feita pelos gerativistas para a teoria e metodologia da lingüística histórica pode ser encarada como um refinamento e um desenvolvimento da concepção estruturalista de mudança lingüística. Em ambos os casos dá-se preferência ao que se classifica como fatores internos. A noção estruturalista de auto-regulação foi substituída pela de reestruturação das regras do sistema lingüístico e de uma tendência à simplificação. É difícil perceber quaisquer diferenças fundamentais entre essas duas noções.

Entretanto, uma diferença entre a distinção chomskiana competência/desempenho e a distinção saussureana *langue/parole* é a de que a primeira presta-se mais facilmente do que a segunda a uma interpretação psicológica. Os gerativistas, como veremos, têm se preocupado bastante, por várias razões, com o problema da aquisição da linguagem pelas crianças. Eles têm enfatizado o fato de que a criança, quando começa a adquirir a sua língua materna, não aprende as regras do sistema subjacente, mas tem que inferi-las dos padrões de correspondência entre forma e significado que ele detecta nos enunciados proferidos à sua volta. O que tradicionalmente é considerado como falsa analogia (por exemplo, a tendência da criança de dizer *goed* ["iu"] ao invés de *went* ["foi"]) é encarada pelo gerativista como parte do processo mais geral da aquisição de regras.

Os gerativistas não são os primeiros a terem procurado uma explicação para a mudança lingüística na transmissão da linguagem de uma geração para outra. Mas eles observaram com mais cuidado do que outros o processo de aquisição da linguagem em termos da natureza das regras exigidas em estágios identificáveis nesse processo. Além disso, eles começaram a investigar em detalhe a mudança sintática, bem como as mudanças fonológica e morfológica: até recentemente a mudança sintática dificilmente era estudada, a não ser ocasional e desordenadamente. Mais importante

que qualquer outra coisa, entretanto, é o fato de que o gerativismo forneceu à lingüística histórica uma concepção mais precisa de **universais** formais e substantivos, em relação aos quais podem ser avaliadas como mais ou menos prováveis as mudanças de uma língua, postuladas em estágios pré-históricos e não comprovados.

Do ponto de vista de seus débitos, tanto o estruturalismo quanto o gerativismo tendem a prestar atenção insuficiente à importância da variação sincrônica como fator na mudança lingüística. Independentemente de qualquer outra coisa, isto deu origem a pseudoperguntas do tipo: a mudança lingüística é repentina ou gradual? A mudança lingüística origina-se na competência ou no desempenho? Quanto à primeira, já faz mais de cem anos que Johannes Schmidt desafiou o conceito da árvore genealógica do relacionamento lingüístico, favorecido pelos neogramáticos, e chamou atenção para o fato de que as inovações de todos os tipos, e mais particularmente as mudanças sonoras, podem se espalhar a partir de um centro de influência, como ondas em um lago, perdendo força à medida que elas atingem pontos cada vez mais distantes do centro. Nas décadas que se seguiram foi demonstrado, especialmente por eruditos que trabalhavam com a história das línguas românicas, onde os dados sincrônicos e diacrônicos eram abundantes, que aquilo que veio a ser denominado de **teoria das ondas** de mudança lingüística, fornecia um relato dos fatos mais satisfatório, em muitos casos, pelo menos, do que a teoria mais ortodoxa da árvore genealógica, com o seu pressuposto inerente de divergência repentina e, a partir de então contínua, entre dialetos relacionados. Os dialetólogos também demonstraram que, longe de se aplicarem simultaneamente a todas as palavras às quais eram aplicáveis, as mudanças sonoras poderiam se originar em apenas uma ou duas palavras e então espalhar-se por outras palavras e, pelos caminhos da comunicação, por outras áreas. Se isto é o que ocorre em geral, está claro que a pergunta se a mudança lingüística é gradativa ou repentina perde muito de sua razão de ser. E, assim como os indivíduos podem vacilar entre o uso de uma forma mais velha ou mais nova, também podemos vacilar diante da pergunta se as mudanças lingüísticas originam-se na competência ou no desempenho.

Mais recentemente, os sociolingüistas demonstraram que o que era válido para a difusão geográfica das variantes fonológicas, gramaticais ou lexicais, também é válido para a sua difusão pelas classes socialmente distingüíveis de uma dada comunidade. Em geral, tornou-se claro que fatores sociais (do gênero que consideraremos no capítulo 9) são muito mais importantes na mudança lingüística do que se percebia antes. Afinal de contas, não são apenas as fronteiras geográficas, ou até mesmo políticas, que estabelecem limites ao grau de intercomunicação entre as pessoas que moram na mesma região. Os dialetos sociais podem se distinguir uns dos outros tanto quanto os dialetos geograficamente determinados. Por outro

lado, dadas as condições sociais certas (rutura numa sociedade tradicionalmente estratificada, a imitação de formas ou expressões utilizadas pela classe alta, etc.), um dialeto social pode ser modificado pelo contato com outro. Com efeito, agora está se tornando aceito o fato de que o **bilingüismo** e a **diglossia** – e até mesmo a **pidginização** e a **crioulização** – podem ter desempenhado um papel muito mais importante na formação de famílias de línguas do que se imaginou alguma vez (cf. 9.3, 9.4).

Iniciamos esta seção com a pergunta: Por que as línguas mudam no decorrer do tempo? Podemos concluí-la repetindo o que foi dito num capítulo anterior (v. 2.5): a ubiqüidade e a continuidade da mudança lingüística são bem menos intrigantes quando nos damos conta de que nenhuma língua natural é jamais estável ou uniforme, e que muito do que é descritível, macroscopicamente, como mudança lingüística, é o produto de variação sincrônica socialmente condicionada. Não estamos com isso afirmando que toda mudança lingüística deve ser tratada dessa forma, mas apenas que fatores sociais são sem dúvida mais importantes do que já se imaginou.

LEITURAS COMPLEMENTARES

A maioria dos livros didáticos e das introduções à lingüística contêm capítulos sobre mudança lingüística. Particularmente os capítulos 18-27 de Bloomfield (1935) valem a pena ser lidos pela sua colocação essencialmente neogramática, com muitos exemplos, agora clássicos, do inglês e de outras línguas.

Introduções recentes à lingüística histórica como tal incluem Aitchison (1981); Bynon (1977); Lehmann (1973). Bynon (1977: 281-2) fornece mais referências, por assunto, para tudo o que foi discutido neste capítulo; Aitchison (1981) enfatiza o papel dos fatores sociais na mudança lingüística.

Sobre a história do inglês (em vários níveis de detalhe e tecnicidade): Barber (1972); Baugh (1965); Francis (1967); Lass (1969); Potter (1950); Strang (1970); Traugott (1972). Para outras línguas e famílias de línguas, a *Encyclopaedia Britannica*, 15ª edição (1974), é a obra de referência mais apropriada.

PERGUNTAS E EXERCÍCIOS

1. Quais são as bases para o reconhecimento de três períodos diferentes na história do inglês: inglês antigo (anglo-saxão), inglês médio e inglês moderno?
2. Com base em informação fornecida por enciclopédias ou outras obras de referência, relacione os principais membros oriundos das famílias de línguas germânica, românica e eslava.
3. O que é uma **protolíngua**?
4. Explique a finalidade da **reconstrução** em lingüística histórica.
5. Diz-se do inglês, como do francês, do russo, do hindi etc., que são línguas indo-

européias. O que significa isto? É verdade? O que dizer do finlandês, do húngaro, do basco, do turco, do tâmil?

6. Por que você imagina que Sir William Jones atribuiu tanta importância ao que chamou de "raízes verbais" e "formas gramaticais" (v. p. 178)?
7. "Uma das razões mais fortes para adotar o pressuposto da mudança fonética regular é o fato de que a constituição dos resíduos... projeta bastante luz sobre a origem de formas novas" (Bloomfield, 1935:405). Discuta.
8. Formule a **lei de Grimm** (utilizando exemplos diferentes dos do texto) e mostre como a **lei de Verner** a ela se relaciona.
9. Explique o que se quer dizer quando se invoca a regularização **analógica** de formas sincronicamente irregulares.
10. Que inferências podem ser feitas sobre a história de uma língua a partir da existência de **duplas lexicais**? Faça uma lista de dez pares de lexemas desse tipo em português moderno. Que distinção poderia ser feita, se possível, entre duplas lexicais e formas gramaticalmente equivalentes, co-existentes, do mesmo lexema (*got* : *gotten* [do verbo *get* : "obter"], *learnt* : *learned* [de *learn* "aprender"], *dived* : *dove* [de *dive* "mergulhar"]? Como você classificaria *brothers* : *brethren* [irmãos : irmandade], *mediums* : *media* [ambientes], *struck* : *stricken* [de *strike* "atingir"] em relação a tal distinção?
11. Demonstre de que maneira o **empréstimo** pode dar conta de exceções aparentes ao funcionamento regular de uma lei sonora.
12. "A mudança lingüística, assim, oferece dados importantes acerca da natureza da linguagem humana – a saber, o fato de que é regida por regras" (Akmajian, Demers e Harnish, 1979: 226). Discuta esta afirmação com relação à noção gerativista de **reestruturação**.
13. "O que talvez constitua a contribuição mais importante para uma compreensão do mecanismo real da mudança lingüística originou-se na investigação sociolingüística detalhada de comunidades lingüísticas vivas" (Bynon, 1977:198). Discuta.
14. Explique e exemplifique a noção de **reconstrução interna**.
15. Compare e confronte a **teoria da árvore genealógica** e a **teoria das ondas** (*Wellentheorie*) do desenvolvimento lingüístico.
16. Avalie a contribuição do estruturalismo e do gerativismo à teoria e à metodologia da lingüística histórica.
17. Qual a contribuição que o estudo de: (a) **aquisição da linguagem** e (b) línguas **pidgin** e **crioulas** pode trazer à lingüística histórica? (Há mais dados referentes a esta pergunta nos capítulos 8 e 9.)