

Capítulo 1

A língua como sistema de representação mental

1. Introdução

Neste capítulo, procuramos caracterizar as questões de fundo (em particular, questões de natureza epistemológica) que têm definido o programa da gramática generativa, independentemente dos mecanismos técnico-teóricos propostos no âmbito específico da investigação sobre a estrutura das línguas humanas. Insistimos em particular na natureza «mentalista» da teoria, isto é, na concepção de que o seu objecto de estudo consiste num sistema de regras e princípios radicados em última instância na mente humana, e não em propriedades absolutas das expressões linguísticas consideradas em si mesmas, ou consideradas como um aspecto particular do comportamento humano independente das propriedades mentais subjacentes à sua produção e compreensão.

Esta preocupação da gramática generativa com o aspecto psicológico da linguagem, e sobretudo com o problema da aquisição da linguagem (em última instância com o seu aspecto biológico), não aparece apenas com a TRL; pelo contrário, tem estado no centro das preocupações de Chomsky (e de outros generativistas) desde os seus primeiros trabalhos, como se pode ver, por exemplo, pela data de publicação da célebre revisão crítica de Chomsky (1959) a Skinner (1957), um trabalho sem dúvida tão revolucionário no impacto que teve no desenvolvimento das ciências da cognição e dos estudos sobre a aquisição da linguagem como Chomsky (1957) o foi relativamente à área dos estudos gramaticais. É no entanto na TRL que se torna possível (talvez pela primeira vez) ancorar mais solidamente na teoria gramatical as investigações relativas à aquisição e desenvolvimento da linguagem na infância, através do modelo de «princípios e parâmetros».

2. A tensão entre a natureza e a convenção nos estudos da linguagem

As línguas naturais (com as propriedades que os linguistas e filósofos lhes têm reconhecido) são adquiridas e faladas espontaneamente apenas pelos membros da espécie humana, isto é, por organismos com um determinado tipo específico de estrutura e organização mental. Parece pois difícil escapar à conclusão de que as propriedades essenciais da linguagem são directamente determinadas por propriedades mentais da linguagem são determinadas por propriedades mentais dos seres que as falam, e que estudar a linguagem humana consiste em estudar determinadas propriedades da mente humana, radicadas em última instância na organização biológica da espécie.

Ao longo da história dos estudos sobre a linguagem, esta conclusão (que nos parece indiscutível) tem sido no entanto objecto de intensos debates, e muitas das teorias desenvolvidas por filósofos e linguistas a têm negado, explícita ou implicitamente. Como Chomsky o assinala várias vezes, o pensamento científico e humanista ocidental tem uma extrema dificuldade em assumir que os produtos do pensamento (entre os quais a linguagem) possam radicar na natureza biológica dos seres humanos tal como as estruturas anatómicas. A posição antimentalista tem usualmente uma fundamentação social: os seres humanos vivem em comunidades sociais; a linguagem é um instrumento essencial na vida social das mulheres e dos homens; logo, a explicação última das propriedades da linguagem tem a ver com o seu funcionamento social; em última instância, é um produto *conventional* da cultura dos seres humanos vivendo em sociedade, e não um produto *natural* da sua organização mental.

Não é este o lugar para apresentar detalhadamente as diferentes escolas e correntes que se sucederam na história da linguística e a posição que tomaram relativamente a esta dicotomia ⁽¹⁾. A teoria da Gramática Generativa inscreve-se decididamente na corrente naturalista dos estudos sobre a linguagem e a natureza humana (o que não significa, como Chomsky accentua várias vezes, que seja o produto ou o seguimento directo de nenhuma das tradições históricas que a antecederam) ⁽²⁾.

⁽¹⁾ Ver, a este respeito, Chomsky (1966a; 1972; 1975), Robins (1967) e Lyons (1968). Uma versão contemporânea da posição convencionalista reduz a linguagem humana (e as suas propriedades centrais) a um simples instrumento comunicativo estruturalmente optimizado com vista à comunicação mais eficaz possível. Como Chomsky tem insistido várias vezes, nem a comunicação constitui o único uso que os seres humanos fazem da linguagem, nem é viável efectuar uma relação das complexas propriedades estruturais da linguagem a requisitos comunicativos. Sobre este tópico, ver, por exemplo, Chomsky (1975, capítulo 2).

⁽²⁾ Entre os exemplos da perspectiva naturalista sobre a linguagem, assinalamos a

3. O programa de investigação da gramática generativa

Na base do pressuposto de que a linguagem é um sistema de conhecimentos interiorizado na mente humana, Chomsky define o programa de investigação da Gramática Generativa como o desenvolvimento das quatro seguintes questões (para uma exposição extremamente clara deste programa, ver Chomsky (1988a) e as entrevistas dedicadas à linguística em Chomsky (1988b)) ⁽¹⁾:

- (1) Qual é o conteúdo do sistema de conhecimentos do falante de uma determinada língua particular, por exemplo do Português? O que é que existe na mente deste falante que lhe permite falar/compreender expressões do Português e ter intuições de natureza fonológica, sintáctica e semântica sobre a sua língua?
- (2) Como é que este sistema de conhecimentos se desenvolve na mente do falante? Que tipo de conhecimentos é necessário pressupor que a criança traz *a priori* para o processo de aquisição de uma língua particular para explicar o desenvolvimento dessa língua na sua mente?
- (3) Como é que o sistema de conhecimentos adquirido é utilizado pelo falante em situações discursivas concretas?
- (4) Quais são os sistemas físicos no cérebro do falante que servem de base ao sistema de conhecimentos linguísticos?

gramática especulativa dos escolásticos do período medieval, a qual distinguiu entre os aspectos universais e os aspectos accidentais das línguas, realçando que a gramática é essencialmente universal e radicada no conhecimento humano, e que os aspectos accidentais residem no vocabulário diferente das línguas particulares; e a tradição da Gramática de Port-Royal dos séculos XVII e XVIII, com a sua insistência que a linguagem é um produto da natureza mental humana, reflectindo directamente a organização e funcionamento do raciocínio. Sobre esta última corrente, ver em particular Chomsky (1966a).

⁽¹⁾ O termo «língua», tal como utilizado na formulação das questões (1)-(4), no decorrer deste livro, e de um modo mais geral na literatura generativista, refere-se a um sistema de conhecimentos mental, e não ao conjunto de objectos abstractos (frases ou expressões) determinado por esse sistema. Nesta acepção, o termo «língua» e sinónimo de «gramática» (interiorizada) (ver a secção 4) ou de «competência» (ver a secção 5). Chomsky (1966b) utiliza nesta acepção o termo técnico «língua-I» (de «língua interiorizada»), opondo-o ao termo «língua-E» (de «língua exteriorizada»), que refere o conjunto de frases e expressões determinadas pela língua-I. O objecto de estudo da gramática generativa é a língua-I, não a língua-E. Para mais detalhes sobre esta dicotomia convencional, ver Chomsky (1986b, 21-46; 1988b, 580-590).

O empreendimento generativista atribui um lugar central à questão (2), tanto do ponto de vista filosófico/epistemológico como do ponto de vista da teoria gramatical propriamente dita. Em particular, o cuidado atribuído à interacção entre (1) e (2) é a pedra-de-toque da gramática generativa: qualquer proposta relativa ao tipo de conhecimentos iniciais que a criança traz para o processo de aquisição tem de poder explicar adequadamente o carácter dos conhecimentos adquiridos relativamente a uma língua particular; e inversamente qualquer proposta quanto ao carácter dos conhecimentos sobre uma língua particular tem de ser compatível com os conhecimentos iniciais da criança e com o facto de a aquisição e o desenvolvimento dessa língua serem feitos a partir dos conhecimentos iniciais. Por outras palavras: nem todas as gramáticas que descrevem adequadamente os dados de uma língua particular são psicologicamente possíveis. É necessário, para além disso, que possam ter sido desenvolvidas pela criança com base no sistema de aquisição inicial.

Nas próximas secções, abordamos os problemas levantados por cada uma destas questões com mais ou menos pormenor consoante se inscrevem ou não nas preocupações centrais deste livro. Como o nosso objectivo central é a apresentação de uma teoria linguística que serve de guia à elucidação dos problemas (1) e (2), é a estes que dedicamos mais espaço, especialmente à questão central (2). Começamos portanto com algumas observações sobre (1), (3) e (4), para nos ocuparmos a seguir mais longamente de (2).

4. A gramática como sistema computacional

O primeiro feixe de problemas leva ao desenvolvimento da gramática de uma língua particular, isto é, de um modelo que represente o sistema de conhecimentos particular do falante, capaz de explicar as suas intuições sobre a forma e a significação das expressões linguísticas, nomeadamente se são ou não admitidas pela sua língua (isto é, se são aceitáveis ou não) (1). As investigações dos generativistas (sobretudo nos últimos 20 anos) mostraram que a gramática interiorizada consiste por um lado num «dicionário mental» das formas da língua e por outro num sistema de princípios e regras actuando de forma computacional sobre essas formas, isto é,

(1) Seguindo uma tradição de longa data na gramática generativa, utilizamos o termo «gramática» ambigualmente, para nos referirmos ao objecto real na mente do falante de uma determinada língua particular e ao modelo científico que o linguista constrói desse objecto. Utilizamos por vezes a expressão «gramática interiorizada» para nos referirmos exclusivamente ao objecto mental.

construindo representações mentais constituídas por combinações categorizadas das formas linguísticas (2). Estas representações determinam de uma forma muito explícita as propriedades fonológicas e sintácticas das expressões da língua, assim como aquelas propriedades semânticas que são derivadas directamente a partir das propriedades sintácticas. A gramática determina igualmente o modo como estas representações se articulam com outros sistemas conceptuais da mente humana ou com o sistema neuromuscular que determina a pronúncia das expressões (3).

Para dar um exemplo simples do funcionamento da gramática, tomemos uma expressão do Português como, por exemplo, a oração interrogativa **que livro o Luís leu?** A informação contida no dicionário mental sobre **ler** indica que esta forma pertence à categoria verbo, e que é necessariamente acompanhada de um complemento de objecto directo (uma expressão de categoria NP). Com base nesta informação, uma regra geral do sistema constrói a representação (5), determinando em particular as seguintes propriedades da expressão: (i) o produto da combinação V-NP pertence à categoria VP; (ii) dentro do VP, V precede NP (7):

(5) $[_{VP} [_{V} \text{leu}] [_{NP} \text{que livro}]]$

Uma outra regra do sistema determina que a representação (5) se combina com uma expressão de categoria NP com a função de sujeito, e que a representação resultante pertence à categoria Frase (S):

(6) $[_S [_{NP} \text{o Luís}] [_{VP} [_{V} \text{leu}] [_{NP} \text{que livro}]]]$

(2) Podemos assimilar as formas primitivas do dicionário às palavras, o que não é tecnicamente correcto mas é suficiente para a exposição neste ponto. Ver o capítulo 3 para uma discussão técnica da forma e conteúdo deste «dicionário».

(3) A linguagem é, para Chomsky, um sistema formal interpretado, no sentido da lógica (isto é, as expressões são construídas por um sistema de regras exclusivamente formais e são posteriormente investidas de significação). A interpretação semântica atribuída às expressões derivadas pelo sistema formal podem ser determinadas por regras «tardias» do próprio sistema, ou por regras de outros sistemas de representação mental, como o sistema de crenças e de pressuposições acerca do mundo que nos rodeia, o sistema de significações parcialmente arbitrárias e convencionais que a sociedade atribui às palavras, e um sistema de representação semântica parcialmente autónomo em relação à gramática. Ver o capítulo 5 para uma discussão de algumas destas questões.

(4) Ver o capítulo 2.3.2 para uma explicação da notação dos parênteses aqui utilizada. Ver o capítulo 9 para uma caracterização do princípio que determina a construção sintáctica da representação (5) a partir das propriedades do verbo **ler**.

Finalmente, outra regra do sistema determina, com base na presença da forma que no NP objecto directo, que esta categoria é deslocada para a posição inicial da oração, e que a expressão recebe uma interpretação interrogativa (o travessão indica a posição a partir da qual o NP é deslocado):

(7) [_{NP} que livro] [_S [_{NP} o Luís] [_{VP} [_V leu] —]]

A gramática interiorizada do falante é um sistema autónomo, cujos princípios e representações lhe são específicos (isto é, não são partilhados por outros sistemas ou capacidades humanas). Para além disso, a gramática é independente dos outros sistemas conceptuais da mente humana (como, por exemplo, princípios de conversação, princípios de formação de conceitos, etc.), embora mantendo com eles uma interacção complexa em pontos específicos. Esta visão da organização da mente humana tem sido chamada de modular (*). Qualquer fenómeno linguístico, nesta perspectiva, é um produto complexo da acção da gramática e de outros mecanismos conceptuais. Uma das tarefas do linguista consiste precisamente em determinar quais são aqueles aspectos de uma expressão linguística que devem ser representados no nível da gramática formal (ver igualmente a discussão na secção seguinte).

(*) A concepção modular da mente humana defende que esta é formada por módulos autónomos, cada um deles caracterizado por princípios e representações específicas. Estes módulos «comunicam» entre si em pontos determinados, mantendo uma interacção complexa que determina as propriedades dos fenómenos mentais humanos. Assim, para além do módulo linguístico, a mente humana possui por hipótese um módulo matemático, um módulo musical, um módulo espacial (que lhe permite compreender mapas, por exemplo), um módulo de formação de conceitos, entre outros cujo descobrimento e caracterização pertence às ciências cognitivas. Ver, a este respeito, Fodor (1983) e Gardner (1983; 1985). Esta posição opõe-se a uma visão da mente como sendo uma estrutura uniforme, não diferenciada no que respeita aos seus princípios de funcionamento, nomeadamente postulando uma «inteligência generalizada» aplicável a todos os domínios mentais e determinando o desenvolvimento cognitivo geral da criança, incluindo o seu desenvolvimento linguístico. Esta concepção da estrutura mental era defendida pelo psicólogo Jean Piaget (e continua a ser defendida pelos seus seguidores). Para um debate entre Chomsky e Piaget relativamente a estas questões (sem como numerosos comentários e discussões por outros linguistas, psicólogos e neurologistas), ver Piattelli-Palmarini (ed.) (1979). (Recomenda-se a versão inglesa desta obra, mais fiel ao texto das publicações submetidas originalmente em inglês.)

5. Competência e performance

A questão (3) está na base da distinção efectuada por Chomsky (1965, 3-15) entre *competência* e *performance*. A competência é o conhecimento mental «puro» de uma língua particular por parte do sujeito falante, isto é, a sua gramática interiorizada. A «performance», por sua vez, designa o uso concreto da linguagem em situações de fala concretas (Chomsky (1965, 3)). Como os actos de fala põem em jogo variáveis de natureza social e psicológica independentes do conhecimento gramatical da língua, a estrutura, organização e conteúdo de qualquer expressão linguística (ou sequência discursiva de expressões) pronunciada em condições naturais é determinada por uma combinação muitas vezes complexa de factores que tem a ver apenas parcialmente com a competência. Os problemas levantados por (3) dizem pois respeito ao modo como a competência é utilizada (ou distorcida) em situações de performance através da sua interacção com esses sistemas (*).

Para dar apenas alguns exemplos, a atenção e a memória podem provocar «erros» de competência (por exemplo, quando entre o sujeito e o verbo se interpõe material suficiente para fazer esquecer ao falante o número do sujeito e induzir desse modo uma forma verbal com a concordância errada). As falsas partidas, as mudanças de plano discursivo, as hesitações, o gaguejar, etc., são outros factores que afectam o resultado gramatical das expressões enquanto reflexo da competência. Num plano diferente, existem convenções linguísticas de natureza discursiva que determinam significações para as expressões que não são aquelas atribuídas pela gramática. Um caso interessante é o dos actos de fala indirectos como (8) (ver Seuren (1975)):

- (8) a. Podes passar-me a manteiga?
b. Está um frio de rachar.

A expressão (8a) é gramaticalmente uma oração interrogativa, caracterizada pelo sistema formal como um pedido de informação. No entanto, num acto de fala em que seja normalmente pronunciada (por exemplo, à mesa, durante o almoço), esta expressão é interpretada como um pedido para

(*) A trajectória real de um projectil constitui uma analogia possível, na medida em que é o produto da trajectória ideal determinada pela balística com as condições atmosféricas do mundo real (vento, etc.) que a afectam. Do mesmo modo, a estrutura, conteúdo e organização das expressões da linguagem são o produto da interacção da competência ideal com factores psicológicos ou sociais que afectam essa competência.

efectuar uma determinada acção (passar a manteiga), e não como um pedido de informação. O sentido convencional, neste caso, anula o sentido formal. Do mesmo modo, (8b), uma oração declarativa exprimindo *a priori* uma proposição assertiva, pode ser utilizada como um pedido ou mesmo como uma ordem (para fechar a janela, por exemplo), num contexto discursivo apropriado e se o falante tiver um ascendente de qualquer tipo sobre o ouvinte.

Outras áreas de natureza geral que caem no âmbito da performance são a estrutura das conversações e a relação entre o significado gramatical das expressões e o contexto do discurso (isto é, o modo como esse contexto desambigua expressões semanticamente ambíguas ou atribui novas significações a expressões que *a priori* não as possuem). Qualquer destes tópicos põe em jogo a competência em interacção com fenómenos complexos de natureza social, psicológica e discursiva (por exemplo, estratégias de comunicação) ⁽¹⁰⁾.

Num sentido mais restrito, o termo «performance» é utilizado (sobre tudo na literatura generativista) para referir exclusivamente os mecanismos psicológicos de percepção e processamento da linguagem que «facilitam» o funcionamento da gramática interiorizada (ver, por exemplo, Berwick e Weinberg (1984)). Neste sentido, a teoria da performance tem como um dos seus objectivos centrais a construção de um «analisador» (em Inglês, um «parser»), o qual consiste num mecanismo capaz de analisar a cadeia falada linear (que é tudo a que o ouvinte tem acesso) de acordo com os princípios de uma gramática generativa, os quais envolvem noções mais abstractas como estrutura de constituintes e movimentos transformacionais, entre outras (ver os capítulos 2 e 4, respectivamente). Note-se a propósito que a gramática enquanto modelo da competência é neutra relativamente à produção/compreensão da linguagem, sendo subjacente a ambas. Em particular, não é um modelo da produção linguística pelo falante, erro de interpretação por vezes cometido devido ao termo «generativo», o qual significa simplesmente neste contexto «caracterizar explicitamente os

⁽¹⁰⁾ Dada a natureza diversa dos factores extragramaticais que determinam os actos de fala, não podemos falar num único modelo generalizado de performance, mas sim em vários modelos, de acordo com a especificidade e autonomia dos factores em jogo. Por exemplo, o estudo do modo como o contexto discursivo desambigua o sentido das expressões semanticamente ambíguas está na base de uma disciplina específica da performance, a *pragmática*. (Ao atribuir a pragmática à performance, não estamos de modo nenhum a negar a existência de uma possível «competência pragmática» interiorizada na mente do falante.)

objectos de um determinado conjunto», neste caso particular as frases de uma língua ⁽¹¹⁾.

Em síntese, o estudo da competência enquanto puro sistema de conhecimentos mental implica que o linguista proceda a uma abstracção das diversas variáveis em jogo nos actos de fala concretos, ou seja, implica que o objecto do seu estudo seja «um falante-ouvinte ideal, situado numa comunidade linguística completamente homogénea, que conhece a sua língua perfeitamente, e que, ao aplicar o seu conhecimento da língua numa performance efectiva, não é afectado por condições gramaticalmente irrelevantes tais como limitações de memória, distrações, desvios de atenção e interesse, e erros (casuais ou característicos)». (Chomsky (1965, 3), citado da tradução portuguesa, p. 83.) ⁽¹²⁾

6. A realização física do sistema da competência

Enquanto sistema mental, a língua tem necessariamente um suporte material no cérebro humano. A questão (4) tem a ver com a relação língua-cérebro, em particular com os mecanismos neuronais que suportam o reconhecimento gramatical. Pouco teremos aqui a dizer sobre esta área da investigação, a qual, se bem que esteja ainda numa fase prematura, tem sido objecto de esforços renovados no âmbito do interesse crescente pelas ciências da cognição em geral. Estes estudos, escusado será dizê-lo, terão de pôr em jogo uma colaboração estreita entre linguistas, psicólogos, patologistas da fala e neurologistas, os quais, com base em resultados pelo menos parciais obtidos pelos linguistas relativamente às questões (1)-(3), deverão investigar os mecanismos físicos específicos que estão na base dos diversos sistemas da competência e da performance. Uma área privilegiada nesta investigação é o estudo das afasias, em particular o estudo do modo se-

⁽¹¹⁾ Chomsky (1988a) coloca igualmente entre as questões de (3) o problema da produção da linguagem, que considera no entanto o seu aspecto mais misterioso e sobre o qual o pensamento científico pouco ou nada tem a dizer. Entre os problemas levantados pela produção encontram-se o uso criativo e inovador da linguagem, isto é, a possibilidade de pronunciarmos pensamentos que nunca exprimimos ou ouvimos anteriormente, a liberdade da linguagem relativamente a estímulos externos ou internos (por exemplo, se tenho fome, não digo necessariamente *tenho fome!*), e a sua adequação (salvo casos patológicos) às situações discursivas. Critérios deste tipo eram frequentemente adiantados pelos Cartesianos para a postulação da noção de «mente», oposta à noção de «corpo» como autónomo.

⁽¹²⁾ Para uma discussão das contensões metodológicas e conceptuais que a distinção competência/performance tem suscitado na linguística contemporânea, ver Newmeyer (1983).

lectivo como o conhecimento gramatical é afectado por tipos diferentes de lesões cerebrais⁽¹⁾.

A possibilidade de sucesso teórico na caracterização neuronal detalhada do sistema gramatical levanta a questão do estatuto epistemológico dos conceitos utilizados pelo linguista na caracterização desse sistema, em particular, se estes conceitos devem ou não ser substituídos pelos do neurologista no momento em que este seja capaz de oferecer uma caracterização igualmente detalhada e profunda em termos de células, neurónios, etc. A questão mais geral que se coloca é a de saber se o estudo da mente — entendida como um conjunto de sistemas cognitivos — deve ou não ser reduzido ao estudo do cérebro no momento em que os neurologistas tiverem um conhecimento completo do modo como os sistemas físicos suportam os vários sistemas mentais.

Como o assinala Chomsky (1988a) (e outros cientistas da cognição relativamente aos seus domínios particulares), é pouco provável que a compreensão dos fenómenos linguísticos possa ser atingida sem referência a noções especificamente gramaticais como Frase, Grupo Nominal, Grupo Verbal, pronome, anáfora, ligação, etc. O mesmo se passa, de um modo mais geral, relativamente aos restantes sistemas cognitivos da mente humana. Os dois tipos de estudos — da mente e do cérebro — sem dúvida que se complementam, mas situam-se em planos diferentes da realidade, sem que nenhum se possa reduzir ao outro. A mente, nesta perspectiva, é o conjunto das propriedades abstractas dos sistemas físicos do cérebro, e a compreensão destas propriedades passa necessariamente pelo uso de conceitos (em particular o conceito de *representação*) independente dos mecanismos materiais postulados ao nível do cérebro.

7. A questão da aquisição da linguagem

Como mencionámos acima, a questão central do programa de investigação da gramática generativa é sem dúvida (2), ou seja, o problema de saber como é que a gramática se desenvolve na mente do sujeito falante, e em que base. Por outras palavras, é a questão da aquisição da linguagem, que foi, e continua a ser, uma das questões mais debatidas na história

⁽¹⁾ A investigação das relações língua-cérebro tem antecedentes importantes (que remontam ao século XIX) no estudo das localizações cerebrais de aspectos diferenciados da linguagem (por Broca e Wernicke, entre outros). Broca foi o primeiro a apresentar evidência experimental de que a língua se situa predominantemente no hemisfério esquerdo do cérebro.

do pensamento filosófico e linguístico ocidental integrada ou não ao âmbito mais vasto do debate epistemológico sobre a aquisição do conhecimento em geral⁽¹⁾.

O problema central nesta questão é o do papel específico da mente humana neste processo. De um lado situam-se aqueles que acreditam ser o desenvolvimento linguístico (e a aquisição do conhecimento em geral) basicamente determinado por causas externas à mente humana, concretamente pelas experiências e interacções da criança com o meio ambiente. No caso da linguagem, o meio ambiente é representado pela fala das pessoas que convivem com a criança, e pelas interacções linguísticas em que a criança intervém. Segundo os defensores desta posição, a linguagem é essencialmente uma questão de *aprendizagem*, no sentido específico de aquisição pela mente, através de práticas adequadas, de um sistema exterior a ela. Segundo o psicólogo B. F. Skinner, por exemplo, a aquisição de uma língua consiste fundamentalmente numa aprendizagem de hábitos de «comportamento verbal» através de processos de observação, memorização, generalização indutiva, associação, etc. Esta posição, por outro lado, realça o papel do ensino explícito e da prática como sendo essencial tanto na implantação como na solidificação dos conhecimentos adquiridos. Finalmente, o papel da mente humana em todo este processo é diminuído. Quando muito, a mente possui princípios de inteligência extremamente gerais, que suportam a capacidade de efectuar generalizações e associações (na área da linguagem e do conhecimento em geral), mas não existem princípios ou estruturas especificamente dirigidos para a aprendizagem das línguas.

Diferentes versões desta posição têm sido defendidas ao longo da história do pensamento epistemológico (filosófico e científico) ocidental, constituindo aquilo a que é usual chamar a tradição empirista (da filosofia, psicologia ou linguística). Na psicologia do século XX, estas ideias foram defendidas explicitamente pelo behaviorismo Norte-Americano e, relativamente à aquisição da linguagem, especificamente por Skinner (1957).

Para a tradição racionalista, na qual Chomsky se inscreve (e da qual é sem dúvida o expoente máximo na psicologia e na linguística do século XX), a mente humana desempenha um papel fundamental na aquisição da linguagem. Segundo esta perspectiva, as propriedades centrais da linguagem são determinadas por princípios e estruturas mentais de conteúdo especificamente linguístico, as quais funcionam como uma espécie de «planta» arquitectónica no processo de aquisição, dirigindo o desenvolvimento

⁽¹⁾ Sobre os tópicos desenvolvidos nesta e nas seguintes secções, ver, entre outros, Baker (1979), Gleitsman e Wanner (1982), Chomsky (1965; 1972; 1986b; 1988a; 1988b), Cook (1988), Aitchison (1989) e Lightfoot (1989).

linguístico num sentido predeterminado. Estas estruturas mentais pertencem exclusivamente à espécie humana e são geneticamente determinadas, ou seja, radicam na organização biológica da espécie. Segundo esta concepção, adquirir uma língua é mais uma questão de *maturação* e de *desenvolvimento* de um «órgão» mental biológico do que uma questão de aprendizagem (no sentido que explicitámos atrás). A este conjunto de princípios e estruturas mentais especificamente linguísticos, Chomsky (1966b) chama *Meccanismo de Aquisição da Linguagem* (em Inglês, «Language Acquisition Device», usualmente abreviado em «LAD»). Nos modelos linguísticos propostos pelos generativistas, o LAD recebe o nome de «Gramática Universal» (abreviamos em UG, do Inglês «Universal Grammar»).

A concepção racionalista não nega o papel do meio ambiente na aquisição da linguagem. Em primeiro lugar, a fala das pessoas que rodeiam a criança e as suas experiências verbais são determinantes para iniciar o funcionamento do mecanismo de aquisição, sem no entanto determinar as propriedades finais atingidas pelo sistema gramatical. Ou seja, sem estar imersa num ambiente linguístico, uma criança não aprende a falar⁽¹⁵⁾. Em segundo lugar, os meios linguístico, emocional e educativo são factores que determinam o grau de desenvolvimento da linguagem pela criança sem que isso signifique, de novo, que determinam a direcção do desenvolvimento ou o conteúdo final do sistema⁽¹⁶⁾.

(15) Não é só o meio ambiente linguístico que é determinante no desenvolvimento da linguagem (ou no desenvolvimento das outras capacidades intelectuais de uma criança): o equilíbrio, a estabilidade emocional e a saúde física são condições ambientais tão importantes como a interacção verbal. Estes truismos tem sido inicialmente confirmados pelas crianças mantidas em isolamento durante os primeiros anos da sua vida. Para um dos casos mais bem estudados no âmbito da psicologia e da linguística moderna (o caso de Genie), ver Curtiss (1977).

(16) Chomsky (1988b: 502) dá como exemplo o crescimento de uma flor, um processo sem dúvida geneticamente determinado. O desenvolvimento final atingido pela flor depende crucialmente dos cuidados com que é tratada, embora estes não possam alterar o curso específico do crescimento ou o seu resultado final no que respeita às características particulares da espécie (por exemplo, se plantarmos a semente de uma rosa não obtemos um cravo).

Estas observações têm consequências importantes relativamente à elaboração de currículos escolares apropriados para o desenvolvimento da linguagem nas crianças (preliminamos o termo «desenvolvimento» ao termo «aprendizagem» pelos motivos apontados no texto). Os ataques feitos à gramática generativa de uma perspectiva pedagógica atribuído-lhe a ideia de que as capacidades linguísticas se desenvolvem «por si mesmas» e não necessitam de cuidados escolares (ou outros) revela ignorância ou desonestidade. No extremo oposto, a ideia de que o desenvolvimento das capacidades linguísticas da criança passa pela aprendizagem da teoria gramatical, em particular de uma pseudogramática generativa reduzida a uma análise das expressões em árvores, representa uma posição de

Existe um aspecto, no entanto, em que o «meio ambiente» linguístico determina de um modo óbvio o conteúdo parcial daquilo que é aprendido: se uma criança cresce numa comunidade linguística em que se fale o Japonês, por exemplo, a língua que aprende é o Japonês e não o Português. Voltamos mais à frente (quando da discussão da noção de parâmetro) a esta observação trivial mas tão importante na determinação das características exactas do modelo de aquisição subjacente ao desenvolvimento da linguagem.

8. O problema da projecção

O problema da aquisição da linguagem pode ser teoricamente colocado do seguinte modo: qual é a relação que existe entre os «dados primários» a que a criança tem acesso durante a fase de aquisição da linguagem (o seu meio ambiente linguístico) e o sistema de conhecimentos final que caracteriza a competência linguística do adulto (incluindo os seus juízos de gramaticalidade e as suas intuições fonológicas, sintácticas e semânticas)? Baker (1979) chama a esta questão o «problema da projecção», na medida em que implica uma projecção quantitativa e qualitativa dos dados primários (necessariamente finitos e consistindo em expressões relativamente simples) sobre o conjunto infinito de expressões da língua. Isto significa que a criança adquire a partir de uma experiência finita um sistema de competência que se aplica sobre um conjunto infinito de expressões e que incorpora propriedades complexas (nomeadamente recursivas, ver o capítulo 2.5.) que os dados primários certamente não exigem. Em resumo, o sistema da competência final (a gramática do adulto) é quantitativa e qualitativamente muito mais complexo do que o sistema simples necessário para caracterizar os dados primários a partir dos quais o sistema final é adquirido.

Colocado nestes termos, o problema da aquisição leva logicamente à conclusão de que existe um mecanismo mental inato de aquisição que medeia

efeitos absolutamente nefastos e revela uma incompreensão total tanto dos fundamentos básicos da gramática generativa como do processo de aquisição e desenvolvimento da linguagem pela criança. Tal como uma flor, uma criança precisa de um meio ambiente rico que lhe permita desenvolver ao máximo as suas potencialidades linguísticas (isto é, precisa de leitura, de conversação, de prática da escrita, etc.). A última coisa que a criança necessita é que lhe ensinem a estrutura da sua língua. Essa estrutura, ela própria encarrega-se de a desenvolver sem necessidade de instrução, e de um modo muito mais rico, sofisticado e articulado do que aquilo que é proposto por qualquer teoria linguística actual, incluindo a TRL.

entre os dados primários e a gramática final, e que procede à projecção quantitativa e qualitativa que caracteriza o sistema final. Em síntese, a gramática final (o sistema da competência) é o resultado da interacção entre os dados primários e o mecanismo mental de aquisição. A resolução do problema da projecção consiste pois em determinar tão precisamente quanto possível as responsabilidades respectivas do mecanismo de aquisição e dos dados primários no processo de aquisição e desenvolvimento do sistema final da competência do sujeito falante.

Nos termos deste modelo, a tradição empirista pressupõe (pelo menos implicitamente) que a relação entre os dados primários e a competência final do falante é relativamente directa, e que não existem diferenças notáveis de natureza qualitativa entre os dois. Para desenvolver o sistema de conhecimentos final a partir dos dados primários, apenas são necessários princípios mentais intuitivos simples como a capacidade de estabelecer analogias entre diferentes expressões ou paradigmas e de proceder a generalizações de vária ordem com base nessas analogias⁽¹⁷⁾. Para além destes processos mentais, a aquisição da linguagem é uma questão de treino, imitação e memorização por parte da criança, e de correcção, aprovação ou reprovação por parte dos pais ou das pessoas responsáveis pela sua educação.

Chomsky tem apontado várias vezes que este modelo é irrealista porque ignora (ou nega) o abismo não só quantitativo, mas essencialmente qualitativo, que existe entre os dados linguísticos primários e o sistema de conhecimentos final do adulto, a sua gramática interiorizada. Chomsky nota que o indivíduo falante possui toda uma série de conhecimentos sobre a sua língua cujo desenvolvimento não pode provir da aplicação de mecanismos intuitivos de generalização sobre os dados primários, e muito menos por imitação ou memorização. Nos termos de Chomsky, os estímulos primários são «pobres», isto é, não contêm a informação necessária para explicar o sistema rico e complexo de conhecimentos finais (através de generalizações, imitações, etc.). Em textos recentes (ver, por exemplo (1986b; 1988a)), Chomsky refere-se a esta situação como sendo um dos casos particulares daquilo a que chama o «problema de Platão»: concretamente, o problema de saber como é que um adulto possui um sistema de conhe-

(17) Um exemplo simples de generalização analógica é o seguinte: com base na observação de que um número finito de verbos cujo infinitivo termina em -ar forma o pretérito imperfeito com a terminação -va, o falante conclui que *qualquer* verbo cujo infinitivo termine em -ar forma o pretérito imperfeito com aquele sufixo. Em síntese, no entanto, exemplos simples de generalizações são muito mais difíceis, se não mesmo impossíveis, de construir.

cimentos específico tão complexo e tão rico sobre a sua língua (e outros sistemas mentais, sem dúvida) dada a pobreza dos estímulos iniciais aos quais é exposto durante a fase de aquisição. Os estímulos iniciais são pobres porque não contêm as informações pertinentes para o desenvolvimento das propriedades específicas do sistema final.

A solução para o problema de Platão, segundo Chomsky, só pode ser uma: se os dados primários são insuficientes para explicar o sistema de conhecimentos final, então é necessário concluir que a mente da criança põe à sua disposição um conjunto de princípios linguísticos complexos (o LAD, ou, do ponto de vista do linguista, UG) que guia de um modo determinado e extremamente restringido a aquisição e desenvolvimento da linguagem. A este argumento em favor de um sistema UG inato, Chomsky chama o «argumento da pobreza dos estímulos». Em síntese, o falante adulto tem determinados conhecimentos sobre a sua língua para os quais não existe evidência (linguística, psicológica, ou de qualquer outro tipo) que possam ser determinados directamente pelo meio ambiente linguístico inicial durante o curso normal do processo da aquisição, com ou sem a aplicação de mecanismos de associação, generalização, etc. Estes conhecimentos têm pois de ser atribuídos à «planta arquitectónica» inata que guia e determina a aquisição e desenvolvimento da linguagem.

9. Os dados primários e a não existência de informação negativa no processo de aquisição

Alguns estudos recentes sobre a natureza dos dados primários e o seu papel na aquisição e desenvolvimento da linguagem apoiam a tese de Chomsky sobre a pobreza dos estímulos iniciais e a necessidade de postular um mecanismo mental inato de aquisição.

Estes estudos mostram que o meio ambiente linguístico de uma criança típica durante a fase de aquisição é essencialmente formado por expressões gramaticais da língua, e que é com base nesta *informação positiva* que a criança desenvolve uma gramática interiorizada⁽¹⁸⁾. Por outro lado,

(18) A questão de saber se os dados primários contêm igualmente expressões «negadas» e semiagramaticais é objecto de debates na literatura psico-linguística. Note-se que nesse caso o problema da projecção é ainda mais complicado para a criança. Outra questão debatida é a de saber se os dados primários consistem essencialmente numa fala simplificada e estruturada de modo apropriado às capacidades psico-linguísticas das crianças em cada etapa do seu desenvolvimento (o termo inglês para esta fala é «motherese»). Para uma síntese de várias posições sobre este problema (nomeadamente a ideia de que este tipo de fala não é pertinente para a aquisição), ver Gleitman e Wanner (1982) e Newmeyer (1983).

«quilo a que os psicólogos e os linguistas chamam *informação negativa*, ou seja, informação sobre as expressões inaceitáveis da língua, desempenha um papel nulo ou diminuído no processo de aquisição: as crianças tipicamente não recebem instrução gramatical, quer sob a forma de correção de erros quer sob a forma de explicações explícitas (ver Baker (1979), Cook (1988), Lightfoot (1989) e os trabalhos sobre psicologia do desenvolvimento da linguagem af citados) (19). Como Lightfoot (1989) o assinala, a criança não é um «pequeno linguista» construindo a sua gramática com base simultaneamente em informações sobre expressões gramaticais e não gramaticais (20).

Uma das questões mais prementes postas pelo problema da aquisição é então a de explicar como é que o sistema final inclui o conhecimento de que determinadas expressões ou frases são inaceitáveis, ou seja, como é que a criança desenvolve o conhecimento essencialmente negativo de que tal ou tal expressão não existe na sua língua ou não pode ter uma determinada significação *a priori* logicamente possível. O problema pode ser sucintamente caracterizado do seguinte modo: a partir de informações unicamente positiva a criança desenvolve toda uma série de conhecimentos negativos sobre a língua. Quais são então os mecanismos de aquisição que permitem este processo? O modelo empirista de aquisição não parece ter instrumentos suficientes para responder a esta questão: não se entende bem como é que a generalização analógica permite a construção de um conhecimento negativo a partir de informação unicamente positiva (ver a secção seguinte para alguns exemplos concretos). Para o modelo racionalista, pelo contrário, este tipo de processo não é problemático, reforçando a conclusão de Chomsky de que os estímulos iniciais da criança são «pobres», e de que apenas um mecanismo inato suficientemente complexo pode explicar a aquisição e desenvolvimento da linguagem (21).

(19) A correção pelos pais, por exemplo, ainda que ocorra durante o processo de aprendizagem, tende mais a ser sobre a adequação do conteúdo da fala das crianças relativamente à situação discursiva, e menos sobre a forma gramatical das expressões. Em qualquer caso, a correção parece não ter um efeito decisivo sobre o rumo do processo de aprendizagem (muitas vezes a criança corrigida pura e simplesmente ignora as correções e continua sistematicamente a fazer o mesmo erro, até à altura própria, biológica, de incorporar a forma ou as formas correctas).

(20) Nem os dados primários da criança se encontram organizados em paradigmas claros e arrumados como os do linguista, isto é, dirigidos para a formação de hipóteses científicas.

(21) Cook (1988) mostra igualmente que os dados primários têm de obedecer a uma exigência de uniformidade. Por exemplo, se a correção pelos pais parece predominante no processo de aprendizagem de algumas crianças, mas é totalmente inexistente no caso de outras crianças, e se ambos os grupos adquirem aproximadamente a mesma compe-

10. Alguns exemplos concretos: aspectos sintácticos e semânticos da forma se

Para tornar mais concreta a discussão das duas últimas secções, ilustremo-la com algumas propriedades da sintaxe do Português relacionadas com a forma *se*, mostrando quão implausível é tentar encontrar a sua origem na experiência linguística primária da criança durante a fase de aquisição.

Em Português, bem como nas outras línguas Românicas, o morfema *se* tem duas funções (entre outras não pertinentes para esta discussão): de pronome «anafórico» reflexivo ou recíproco (tomando a sua referência de outro Grupo Nominal (NP) presente na oração) e de «pronome impessoal» com referência indeterminada («arbitrária», na terminologia da TRL) e significação próxima da do pronome indefinido *alguém*. No exemplo (9), o pronome *se* tem uma interpretação primariamente anafórica (recíproca); no exemplo (10), pelo contrário, a sua interpretação é primariamente «arbitrária»:

(9) Nesta penitenciária, os presos agriem-se frequentemente.

(10) Nesta penitenciária, agriem-se os presos frequentemente.

Abstraindo das expressões adverbiais **nesta penitenciária e frequentemente**, o exemplo (9) recebe primariamente a interpretação anafórica recíproca «para cada par {x, y} do conjunto dos presos, x agride y (podendo x=y)» (a interpretação reflexiva «cada membro x do conjunto dos presos agride x» embora gramaticalmente possível, é pragmaticamente menos normal); e o exemplo (10) recebe a interpretação «um conjunto de pessoas indeterminadas (alguém) agride os presos», em que o pronome *se* corresponde ao sujeito da oração, ou seja, ao Agente da acção de agredir, como o NP **os guardas** na expressão **nesta penitenciária, os guardas agriem os presos frequentemente**.

Se chamarmos a sua atenção para esse facto, os falantes do Português

tência linguística final, então a aquisição não pode depender da correção de um modo crucial. A questão mais geral levantada pela exigência de uniformidade é o facto de as crianças de uma dada comunidade linguística desenvolverem sistemas finais de competência linguística muito semelhantes quanto ao seu escopo e qualidade, mau grado terem experiências iniciais diferentes no domínio não só linguístico, mas também social, educativo, emocional e cultural (ver Chomsky (1965) para uma primeira discussão deste problema). Citamos uma passagem de Cook (1988, 61) que nos parece ser uma das melhores sínteses sobre a questão: «So long as the environment contains a certain amount of language, it appears not to be crucial to the acquisition of grammatical competence what this sample consists of; any human child learns any human language, whatever the situation.»

reconhecem (embora talvez de um modo não tão imediato) que a forma se em (9) pode possuir igualmente uma referência indeterminada «arbitrária», sendo a expressão nesse caso parafraseável por «quanto aos presos, alguém os agride frequentemente nesta penitenciária», em que de novo se corresponde ao sujeito (ao Agente) da oração (12).

Contudo, nenhum falante do Português aceita para (9) a interpretação apresentada em (11), com a forma se indeterminada correspondendo ao objecto directo:

- (11) Nesta penitenciária, os presos agriem pessoas indeterminadas (alguém) frequentemente (os guardas, por exemplo).

Esta interpretação, no entanto, seria perfeitamente normal do ponto de vista da consistência semântica, lógica e pragmática (isto é, não seria difícil imaginar uma situação que tornasse a interpretação (11) adequada para (9)). A sua completa impossibilidade significa que a gramática interiorizada dos falantes do Português contém princípios formais independentes da semântica e da pragmática que permitem a referência «arbitrária» para se no caso de esta forma corresponder ao *sujeito* da oração (ao Agente), *mas não no caso de corresponder ao objecto directo*, como em (9) (apenas a interpretação analógica é então possível).

Esta propriedade da gramática interiorizada dos falantes do Português não se desenvolve certamente com base nos dados primários aos quais estes são expostos enquanto crianças, durante o período de aquisição da linguagem. Muitas crianças provavelmente não são nunca expostas a frases com se impessoal durante este período, e mesmo que o sejam não recebem normalmente instrução sobre a interpretação, em particular que o seu uso impessoal é excluído quando se corresponde a um objecto directo (o que

(12) Esta interpretação é facilitada fazendo uma pausa entre o NP os presos e o verbo. Devidamente instados, os falantes do Português aceitam igualmente para (10) uma interpretação recíproca «para cada par {x, y} do conjunto dos presos, x agride y». Esta interpretação, no entanto, é de certo modo marginal quanto à sua aceitabilidade.

No universo discursivo introduzido pelo exemplo (10), a «referência» de se pode ser estabelecida pragmaticamente (é «normal», por exemplo, que sejam os guardas que agriem os presos), algo que é irrelevante para o argumento que desenvolvemos no texto. Repare-se que são sempre possíveis outras referências para se, de acordo com a sua interpretação gramaticalmente «arbitrária».

Em (10), o verbo concorda em pessoa e número com o NP os presos, o que faz com que este exemplo seja minimamente diferente de (9). Certos falantes poderão preferir a interpretação arbitrária com o verbo no singular (nesta penitenciária, agride-se os presos frequentemente), o que é igualmente irrelevante para o argumento que desenvolvemos no texto.

constituiria um caso de informação negativa). Esta informação não é mencionada (explícita ou implicitamente) em manuais de gramática escolar e muito menos por pais, assistentes de infância ou professores de educação primária.

Por outro lado, não se vê bem como é que o mecanismo empirista de generalização por analogia pode estar na base da impossibilidade da interpretação (11) para a frase (9). Na realidade, uma aplicação plausível deste mecanismo leva neste caso a predições incorrectas. Existe sem dúvida uma analogia de significado entre o pronome se «arbitrário» e o pronome alguém, como referimos acima. Plausivelmente, as crianças adquirem um conhecimento (pelo menos passivo) do pronome **alguém** antes de adquirirem o uso impessoal, «arbitrário», de se. Segundo os esquemas de aprendizagem propostos pelos empiristas, esta analogia deveria então levar a criação e a derivar a gramaticalidade de (13b) com uma interpretação simultânea de (12b) através de uma generalização baseada na analogia entre (12a) e (13a) e na gramaticalidade de (12b):

- (12) a. Nesta penitenciária, alguém agride os presos frequentemente.
b. Nesta penitenciária, os presos agriem alguém frequentemente.
- (13) a. Nesta penitenciária, agriem-se os presos frequentemente.
b. * Nesta penitenciária, os presos agriem-se frequentemente.

A impossibilidade de (13b) com a significação de (12b) mostra que a noção de analogia é de facto irrelevante para a aquisição do sistema de conhecimentos que o falante possui sobre a função impessoal da forma se. Este conhecimento não parece ser pois derivado, nem directamente nem através de generalizações indutivas, a partir dos dados primários a que as crianças são expostas durante a aquisição. Como esse conhecimento faz no entanto indubitavelmente parte do sistema linguístico dos adultos, temos de concluir que é derivado a partir de princípios linguísticos inatos pertencentes ao mecanismo mental de aquisição da linguagem, princípios esses independentes dos dados primários da aquisição (2).

Consideremos agora o uso anafórico do pronome se (ver os capítulos 8 e 15 para uma exposição detalhada dos princípios em jogo). Nesta função, se exige a presença na mesma oração de um NP independente do qual possa retirar o seu valor referencial (na terminologia da linguística, um *antecedente*). Em (14), se tem como antecedente o NP as atrizes (nas expres-

(2) Ver Burzio (a publicar) para uma proposta concreta relativa aos princípios que estão na base da impossibilidade de (13b) com a significação de (12b).

sões que se seguem, o NP em avaliação como antecedente potencial de se encontra-se em itálico>:

(14) *As actrizes insultaram-se.*

Nas expressões de (15), no entanto, se não pode tomar como antecedente o NP **as actrizes**:

- (15) a. * *As amigas das actrizes insultaram-se.*
b. * *As actrizes disseram que se tinham insultado as amigas antes da cerimónia de atribuição dos Óscares.*
c. * *As actrizes disseram que as amigas se tinham insultado.*

A expressão (15a) não pode significar que as amigas das actrizes insultaram as actrizes, tendo apenas a interpretação em que se toma como antecedente o NP completo **as amigas das actrizes**. A expressão (15b) não pode significar que as actrizes tinham dito que elas, actrizes, tinham insultado as amigas; nessa expressão, se pode unicamente (um tanto ou quanto marginalmente) tomar como antecedente o NP **as amigas**. Finalmente, (15c) não pode significar que as actrizes tinham dito que as amigas as tinham insultado a elas, actrizes. Tal como em (15b), são as amigas que se insultam a si próprias.

De novo, as noções empiristas sobre a aquisição parecem ser de pouca utilidade na explicação deste conjunto de intuições semânticas. Nenhuma criança recebe treino sobre os valores referenciais possíveis ou impossíveis para se em expressões como (15). Mais especificamente ainda, nenhuma criança (ou adulto) comete erros semânticos relativamente a frases como (15), atribuindo-lhes uma interpretação impossível a se para ser depois corrigida pelos pais ou professores. Na realidade, a complexidade dos factos neste caso realça ainda mais a implausibilidade de qualquer instrução. Por exemplo, nenhuma noção simples de proximidade linear entre se e o seu antecedente é suficiente para caracterizar os factos, como uma inspecção rápida de (15a) mostra (as actrizes é o NP linearmente mais próximo de se e no entanto não pode ser o seu antecedente). Também aqui, temos de concluir que são princípios inatos da Gramática Universal que estão em jogo na determinação das interpretações possíveis do uso anafórico da forma se (para uma explicação destes factos, ver os capítulos 8 e 15).

Para terminar este conjunto de aplicações (simplificadas) do argumento da pobreza dos estímulos, consideremos de novo a forma se na sua interpretação impessoal. Nas línguas Românicas em que o sujeito pronominal pode ser nulo (como o Português, o Espanhol e o Italiano), se impessoal pode ocorrer em orações simples ou em orações subordinadas, desde

que os verbos respectivos pertençam a uma das conjugações do indicativo ou do conjuntivo:

- (16) a. Telefonou-se ao João.
b. Ele disse [que se tinha telefonado ao João].
c. Ele quer [que se telefone ao João].

Em Espanhol e em Italiano, se impessoal não pode ocorrer numa oração em que o verbo seja infinitivo:

- (17) a. * E necessario [telefonarsi a Giovanni] (Italiano).
b. * Es necesario [telefonarse a Juan] (Espanhol).

Uma vez mais, não é plausível manter que estas distinções estão presentes nos dados primários das crianças que adquirem o Espanhol ou o Italiano, sob a forma de instrução negativa explícita ou de correção por parte de outros falantes. Mas mais perturbante para as teorias empiristas da aquisição é o facto de em Português a expressão correspondente a (17) ser gramatical:

- (18) É necessário [telefonar-se ao João].

Neste caso, não só é implausível, mas pura e simplesmente absurdo, manter que os pais italianos e espanhóis treinam sistematicamente os seus filhos de modo a não aceitar (17), ao passo que os pais portugueses treinam os seus filhos de modo a aceitar (18).

Existe uma diferença entre o Italiano e o Espanhol, por um lado, e o Português, por outro, que pode ser relacionada com a diferença de gramaticalidade entre (17) e (18). O Português, contrariamente às outras duas línguas, possui uma conjugação de «infinitivo flexionado», ou «pessoal», a qual consiste na forma do infinitivo simples modificada por sufixos de concordância em pessoa e número (ver, entre outros, Cunha e Cintra (1984), Mateus *et al.* (1989) e, na perspectiva da TRL, Raposo (1987a); ver também o capítulo 7.1). Podemos assim propor a hipótese de que em (18) estamos na realidade face ao infinitivo flexionado na terceira pessoa do singular (uma forma idêntica à do infinitivo simples), e que é este infinitivo que «licencia» o pronome se.

A ser correcta, esta hipótese coloca a seguinte questão relativamente à aprendizagem: como é que a criança infere a gramaticalidade de (18) da existência na sua língua de uma conjugação de infinitivo flexionado? Dados primários a que é exposta não trazem consigo nenhuma indicação milagrosa nesse sentido, em particular porque em (18) a concordância não se

manifesta morfológicamente. Por outro lado, essa conexão não é objecto de nenhuma instrução por parte de pais, professores ou manuais de gramática. De novo a conexão é (indirectamente) estabelecida pelo mecanismo inato de aquisição da linguagem, a partir de uma cadeia causal de princípios que determinam uma ligação necessária entre a forma e o significado verbal, através da noção de «Caso Abstracto»⁽²¹⁾.

Em síntese, uma teoria empirista da aquisição da linguagem, defendendo que a competência linguística do falante adulto é adquirida directamente a partir do «meio ambiente», através de instrução, memorização, formação de hábitos e aplicação de mecanismos de generalização indutiva, é inadequada para dar conta do carácter extremamente rico e complexo do conhecimento final adquirido. Em particular, é incapaz, na esmagadora maioria dos casos, de apresentar uma conexão linguisticamente plausível entre os dados primários e a natureza específica do conhecimento final, como vimos pela análise (breve) de vários aspectos do comportamento sintáctico da forma se. A única conclusão plausível é de que a mente possui um «plano arquitectónico» para a construção do sistema de conhecimentos linguísticos final, o qual opera de um modo altamente específico na base de dados primários limitados, guiando de um modo quase rígido o desenvolvimento de certas estruturas e proibindo outras (por vezes através de cadeias causais complexas e não triviais, e que compete à investigação gramatical descobrir). As estruturas permitidas, por sua vez, determinam os juízos de gramaticalidade e as intuições do falante adulto sobre a sua língua.

11. A Gramática Universal

A teoria do linguista sobre o estado inicial do mecanismo de aquisição da linguagem é a *Gramática Universal* (UG), entendida como a soma dos princípios linguísticos geneticamente determinados, específicos à espécie humana e uniformes através da espécie⁽²²⁾. Podemos conceber a Gramática Universal como um órgão biológico, que evolui no indivíduo

(²¹) Para a noção de Caso Abstracto, ver o capítulo 12. Para uma análise da forma pessoal se em italiano onde se mostra a sua dependência relativamente a essa noção, ver Burzio (1986) e Cinque (1989) (que contém igualmente uma análise do contraste entre paradigmas semelhantes a (17) e frases semelhantes a (18) em Português).

(²²) A especificidade de UG à espécie humana é hoje em dia incontroversa mau grado as tentativas (falhadas) de atribuir sistemas comparáveis a macacos, abelhas ou golfinhos (para uma síntese recente destas tentativas, ver Atkinson (1989, capítulo 3)). A uniformidade de UG através da espécie humana constitui certamente a hipótese nula e é consistente com os resultados da investigação gramatical bem como das investigações no domínio da psico-linguística.

como qualquer outro órgão. O resultado dessa evolução e a gramática final que caracteriza os conhecimentos linguísticos do falante adulto. Nos termos de Chomsky, a Gramática Universal é o estado inicial da faculdade da linguagem (S₀), e a gramática do indivíduo adulto constitui o seu estado final, firme ou estável (S, do Inglês «Steady State»).

Em qualquer fenómeno gramatical minimamente complexo, as propriedades da Gramática Universal desempenham um papel fundamental na determinação (logo, na explicação) das propriedades da gramática final, como concluímos na secção anterior através de alguns exemplos ilustrativos do argumento da pobreza dos estímulos. O estudo da natureza e das propriedades exactas da Gramática Universal é pois o objectivo central do empreendimento generativista.

Nesta secção, apresentamos em primeiro lugar os problemas gerais que se colocam a uma teoria adequada da Gramática Universal. Em seguida, apresentamos uma análise esquemática da organização geral dos modelos de Gramática Universal que se sucederam na história da gramática generativa. O modelo recente desenvolvido no âmbito da TRL (o modelo de «princípios e parâmetros») é apresentado na secção 12 (²³).

11.1. Tensão entre descrição e explicação na Gramática Universal

O problema central que se coloca na construção de uma Gramática Universal adequada é o de conciliar a diversidade das línguas com o caminho rígido e altamente específico tomado pelo desenvolvimento das gramáticas individuais na base de dados primários limitados. A Gramática Universal tem de ser suficientemente flexível para acomodar a variação entre as diferentes línguas, mas tem ao mesmo tempo de possuir a rigidez necessária para explicar as propriedades altamente específicas que caracterizam o conhecimento final dos falantes (do tipo observado na secção anterior).

A diversidade linguística leva naturalmente os linguistas a multiplicar os mecanismos descritivos da gramática, sem preocupação com o estabelecimento de princípios que os restrinjam. Contudo, para explicar a natureza específica do conhecimento final, torna-se necessário restringir esses mecanismos através de princípios universais, diminuindo assim drasticamente o número de gramáticas que podem ser derivadas do sistema inicial.

Qualquer dos fenómenos apresentados na secção anterior ilustra este

(²³) Sobre a história da gramática generativa e das causas que levaram à formação, desenvolvimento e abandono dos modelos de UG que antecederam a TRL, ver Newmeyer (1980).

ponto. Por exemplo, podemos deixar ao critério das gramáticas particulares a tarefa de especificar os antecedentes possíveis para a forma *se* anafórica, sem qualquer menção mais geral do fenómeno na Gramática Universal. Se a nossa preocupação é descrever, e não explicar, é possível construir uma gramática na qual (14) e todas as expressões de (15) são gramaticais, outra gramática na qual apenas (15a) é gramatical, outra em que (14) e (15b) é gramatical, e assim sucessivamente (27). Nesta perspectiva, a gramática do Português é aquela em que (acidentalmente) (14) é gramatical, mas as expressões de (15) não o são. Todas as outras gramáticas são no entanto possíveis, e uma tal teoria faz a predição de que podem existir tantas línguas diferentes quantas as combinações possíveis de gramaticalidade/não gramaticalidade para (14) e (15a,b,c).

Nesta perspectiva, no entanto, é um mistério absoluto o facto de estas combinações arbitrárias de expressões não ocorrerem nas línguas humanas, e de a interpretação das anáforas ser na realidade altamente restringida (como em Português) e não ter a sua fonte em nenhum tipo de instrução explícita. Para explicar a natureza altamente específica do sistema de conhecimentos final, temos pois de atribuir a UG restrições tais que a única gramática permitida seja aquela que admite (14) mas admite (15) (na interpretação referida) (28). Ao construir uma UG assim restringida, o linguista diminui as suas propriedades descritivas: assim, nenhuma das gramáticas acima indicadas (com excepção daquela que admite (14) mas exclui (15)) poderá ser derivada, ou descrita, a partir dessa UG.

Na medida em que UG representa as propriedades biologicamente determinadas do sistema inicial de aquisição da linguagem, ao propor uma tal UG, o linguista faz a previsão extremamente forte de que nenhuma gramática humana permitirá frases do tipo de (15). Em Islandês, no entanto, expressões semelhantes a (15c), com a anáfora reflexiva *sig*, são gramaticais. Assim, em (19), *sig* pode tomar como antecedente o NP *Jón* (exemplo de Manzini e Wexler (1987)):

- (19) *Jón* segir að María elski *sig*.
Jón diz que María ame (conjuntivo) Refl

(27) Temos em mente a interpretação impossível das expressões de (15) em que o NP *as actrizes* é o antecedente da forma *se*.

(28) Ver no entanto a discussão sobre o exemplo (19). Essas restrições incidem sobre as possibilidades interpretativas da forma *se*, em particular sobre a relação estrutural que existe entre *se* e o seu antecedente. Ver o capítulo 8 para uma discussão da natureza dessas restrições. Recordemo-nos igualmente que dada a pobreza inicial dos estímulos, o sistema de conhecimentos final do falante sobre (14) e (15) não pode provir do meio linguístico ambiente.

Face a estes dados, o linguista terá de reconsiderar as propriedades que atribuiu à Gramática Universal de modo a que esta permita uma anáfora com as propriedades de *sig*, mas sem perder de vista que em muitas línguas (como o Português) as propriedades das anáforas são consideravelmente mais restritivas. Isto é, o linguista terá de apresentar uma visão mais flexível de UG (aumentando minimamente a sua capacidade descritiva, de modo a permitir o desenvolvimento de gramáticas em que expressões como (19) são possíveis), mas sem sacrificar a sua capacidade explicativa (29). Esta tensão entre descrição e explicação tem caracterizado a grande maioria dos trabalhos realizados nos últimos anos em gramática generativa, e constitui certamente um dos seus aspectos mais interessantes.

Em síntese, restringindo UG, diminuímos as suas capacidades descritivas, possibilitando a explicação da natureza altamente restritiva das gramáticas finais e o facto de muitas das suas propriedades serem adquiridas sem instrução explícita, correcção de erros ou explicação gramatical. Relaxando UG, por outro lado, acomodamos a variação linguística mas corremos o risco de perder a sua capacidade explicativa. A arte reside pois em chegar a um equilíbrio exacto entre estes dois pólos, ou seja, a um sistema ao mesmo tempo suficientemente flexível para permitir a variação linguística mas suficientemente restringido para explicar a rigidez final do sistema de conhecimentos adquiridos.

11.2. A Teoria Standard

A primeira fase da gramática generativa foi essencialmente descritivista. Tratava-se então de aplicar a noção nova de «regra transformacional» aos dados linguísticos, e a riqueza e diversidade dos dados que iam sendo progressivamente analisados levava à multiplicação e relaxamento dos mecanismos descritivos, sem tentativas significativas de restringir o enorme número de gramáticas que a teoria permitia. A preocupação com os problemas de aprendizagem existia a um nível teórico, mas não se traduzia na prática concreta dos linguistas através de análises que procurassem explicar como é que os fenómenos em consideração poderiam ser adquiridos, ou seja, que restrições específicas seria necessário impor sobre a Gramática Universal para explicar a sua aquisição.

A teoria da Gramática Universal desta fase da gramática generativa é

(29) Um modelo de UG com as propriedades necessárias para resolver o problema específico da variação de comportamento das anáforas reflexivas é proposto por Manzini e Wexler (1987).

conhecida sob o nome de *Teoria Standard* e recebe a sua formulação mais explícita e desenvolvida em Chomsky (1965) ⁽²⁰⁾. A Teoria Standard é uma teoria de *regras gramaticais*, na medida em que concebe as gramáticas essencialmente como sistemas de regras, funcionando de uma certa maneira e mantendo entre si um número restrito de interacções possíveis.

A hipótese central desta teoria é de que as línguas humanas utilizam dois tipos de regras: as regras de reescrita categorial e as regras transformacionais. Estas regras organizam-se em componentes e aplicam-se do seguinte modo: as regras de reescrita derivam a «estrutura profunda» das frases; sobre esta aplicam-se em sucessão as regras transformacionais até se chegar à «estrutura de superfícies», considerada um reflexo praticamente directo da manifestação auditiva concreta das frases.

O formato a que essas regras obedeciam (especialmente no caso das regras transformacionais) era extremamente flexível, sendo possível construir regras capazes de descrever qualquer fenómeno imaginável, mesmo aqueles cuja plausibilidade de ocorrência em qualquer língua humana seria nulo. Por exemplo, não existem línguas com frases formadas através da aplicação de uma regra transformacional que movimente uma preposição em qualquer posição frásica para o início da oração principal. Assim, em nenhuma língua é possível derivar uma sequência como *sem o João saiu a carteira* através de uma regra que mova a preposição *sem* para o início da oração a partir de uma expressão como *o João saiu sem a carteira*. O modelo de UG da Teoria Standard, no entanto, não continha nenhum mecanismo universal que impedisse a formulação de uma tal regra. Na perspectiva do modelo Standard, portanto, a inexistência de uma tal regra na gramática de qualquer língua humana era um puro acidente. Podemos sintetizar esta situação dizendo que tal teoria tinha um *poder expressivo* demasiado elevado.

As regras propostas nesta fase são também extremamente especializadas: os generativistas descrevem separadamente, e de um modo relativamente atomista, as construções individuais (como a passiva, as interrogativas, as relativas, as pseudoclivadas, etc.), sem procurar estabelecer as propriedades gerais ou as condições comuns a que estas pudessem obedecer ⁽²¹⁾. Para além disso, as diferenças encontradas de língua para língua

⁽²⁰⁾ Ainda que a Teoria Standard não fosse suficientemente restritiva como base de uma teoria plausível da aquisição, é em Chomsky (1965) que o programa racionalista da gramática generativa relativamente ao problema da aquisição recebe uma das suas formulações mais explícitas. No que respeita aos elementos centrais deste programa, não é fácil encontrar nos trabalhos mais recentes de Chomsky algo que não tenha já sido dito em *Aspects*.

⁽²¹⁾ A própria noção de «construção gramatical», por sua vez, era directamente her-

em cada construção individual motivavam quase sempre a elaboração de descrições gramaticais independentes e autónomas, uma vez mais sem procurar estabelecer de modo sistemático se os fenómenos tinham algo em comum.

Devido à sua enorme riqueza descritiva, um modelo deste tipo permitia a construção de gramáticas diferentes compatíveis com um mesmo conjunto de dados linguísticos. Para decidir em última instância qual destas gramáticas é correcta, a Teoria Standard continha uma *medida de avaliação* destinada a determinar uma classificação dessas gramáticas, sendo a gramática com classificação mais elevada mantida, e as restantes rejeitadas. Um dos critérios propostos para a avaliação das gramáticas consistia numa medida numérica de simplicidade segundo a qual a gramática correcta era a mais «curta», isto é, aquela fazendo uso de um número menor de símbolos (ver Chomsky (1965, 37-47) para uma das poucas discussões detalhadas sobre a medida de avaliação).

Na perspectiva deste modelo, a criança é, durante a fase de aquisição, um «pequeno cientista» construindo hipóteses sobre as regras da sua língua que sejam compatíveis com os dados gramaticais a que vai sendo progressivamente exposta, e eliminando aquelas que se revelam incompatíveis com dados novos. Se num dado momento dois sistemas de regras se revelam igualmente compatíveis com os dados, a criança utiliza a medida de avaliação, descartando-se do sistema mais complexo.

O pendor acentuadamente descritivista deste modelo, o facto de permitir uma incrível variedade de gramáticas, o seu enorme poder expressivo e o insucesso em construir uma medida de avaliação psicológica e linguisticamente plausível contam-se entre os factores que determinaram, por volta de finais dos anos 60, uma insatisfação mais ou menos generalizada da parte de linguistas e psicólogos com a Teoria Standard, em última instância o seu abandono, e o desenvolvimento subsequente de um modelo alternativo, a *Teoria Standard Alargada* ⁽²²⁾.

dada da gramática tradicional, servindo de base não questionada para a aplicação dos novos mecanismos formais. É apenas em Chomsky (1970) que assistimos ao começo da utilidade entre as construções tradicionais e as regras transformacionais, com a ideia de que a construção passiva é objecto de duas regras transformacionais independentes na sua formulação e aplicação (a «anteposição do objecto» e a «posposição do Agente»).

⁽²²⁾ Ver Baker (1979) para algumas críticas específicas ao modelo de *Aspects* do ponto de vista de uma teoria da aquisição. Nesta breve síntese histórica ignoramos o desenvolvimento, nos últimos anos da década de 60, da chamada *Semântica Generativa*, uma

11.3. A Teoria Standard Alargada

A partir da segunda metade da década de 60, e especialmente durante os anos 70, a redução da capacidade descritiva e do poder expressivo da teoria e o aumento da sua capacidade explicativa constituíram a preocupação central de Chomsky e dos generativistas que partilhavam as suas concepções sobre a linguagem humana. Uma das estratégias para atingir este objectivo consistiu em retirar da formulação das regras particulares aqueles aspectos que podiam ser convertidos em princípios gerais da linguagem, e atribuí-los à Gramática Universal. Um dos primeiros exemplos desta estratégia é a dissertação de doutoramento de John R. Ross, *Constraints on Variables in Syntax* (Ross (1967)), a qual descreve uma série de contextos linguísticos («ilhas», na sua terminologia) para fora dos quais nenhuma regra transformacional pode mover um elemento (as chamadas «condições de Ross», ver o capítulo 14.2.1) ⁽³⁾.

Progressivamente, a natureza do trabalho em gramática generativa vai ganhando dimensões diferentes. Uma preocupação central dos generativistas nesta fase é a de restringir drasticamente as possibilidades descritivas dos sistemas de regras e diminuir o seu poder expressivo, ao mesmo tempo simplificando o seu formato, uma tarefa facilitada pela estratégia descrita acima. De facto, ao retirar das regras aqueles aspectos que podem ser convertidos em condições gerais, os linguistas encontravam-se com uma versão simplificada dessas regras, com uma espécie de esqueleto que constitui de facto a sua natureza fundamental. O trabalho efectuado durante esta fase reduz assim a enorme variedade de sistemas de regras a um número muito pequeno de opções, as quais actuam em conjunto com princípios reprodutivos universais na caracterização das gramáticas possíveis.

A redução das regras categoriais de rescrita a um número drasticamente reduzido de opções é efectuado em Chomsky (1970) através dos princípios da *teoria X-barra* (ver os capítulos 6 e 7). Neste artigo, Chomsky diminui também a capacidade descritiva das regras transformacionais ao propor que determinados fenómenos linguísticos (concretamente, as construções de nominalização como a *leitura do livro*, a *construção da ponte*, etc.) deixem de ser tratados transformacionalmente e passem a ser caracterizados atra-

corrente (de curta duração) oposta aos modelos desenvolvidos por Chomsky, e certamente sem as preocupações psicológicas relativas à aquisição que guiavam, no fundamental, os trabalhos de Chomsky e dos seus colaboradores mais próximos (ver Newmeyer (1980)).

⁽³⁾ O próprio Chomsky (1962; 1964) propusera já durante a fase da Teoria Standard condições gerais destinadas a restringir a aplicação das regras transformacionais (entre as quais a Condição de A-sobre-A). Estas condições, no entanto, revelaram-se inadequadas, sendo substituídas pelas Condições de Ross (ver Newmeyer (1980, 177-180)).

vés de uma interacção do léxico com o sistema das regras de rescrita. De um modo mais geral, Chomsky (1970) propõe que as regras transformacionais não podem ter acesso aos morfemas derivacionais que compõem as palavras ⁽⁴⁾.

A obra mais marcante do novo estilo de linguística praticado pelos generativistas durante os anos 70 é o artigo de Chomsky «Conditions on Transformations» (Chomsky (1973)). Neste trabalho, Chomsky propõe uma série de condições sobre as regras transformacionais (a Condição da Subjacência, a Condição do Sujeito Especificado e a Condição da Oração Temporalizada, entre outras), que vão determinar consideravelmente o trabalho efectuado pelos generativistas no resto dos anos 70 (para a subjacência, ver o capítulo 14.2; para as outras duas condições, ver o capítulo 8). De facto, em certa medida, grande parte dos artigos escritos depois de Chomsky (1973) consistem na aplicação destas condições a novos dados (não necessariamente de natureza transformacional, como as anáforas e os pronomes), o que determina, inversamente, a necessidade constante de rever e reformular as condições no sentido de as adequar a esses dados (por vezes de modo drástico, como no caso da Condição do Sujeito Especificado e da Condição da Oração Temporalizada).

O modelo que emerge progressivamente de todo este trabalho (chamado *Teoria Standard Alargada*, abreviado em EST, do Inglês «Extended Standard Theory») pode ser caracterizado como um modelo de *regras e princípios* (por oposição ao modelo Standard, um modelo unicamente de regras). Segundo a concepção da Teoria Standard Alargada, a gramática contém um número distinto e autónomo de componentes, cada uma delas com uma organização própria caracterizada por um pequeno número de regras e por princípios que restringem a aplicação dessas regras. Existe uma componente lexical, uma componente das regras de base, uma componente transformacional, uma componente para a interpretação de pronomes e de anáforas, e assim sucessivamente. A organização interna de cada componente é simples, mas as suas interacções podem ser relativamente complexas.

Na perspectiva deste modelo, a criança é ainda de certo modo um «pequeno cientista» construindo hipóteses sobre as regras da sua língua. No entanto, o conjunto de princípios e de condições da Gramática Uni-

⁽⁴⁾ Por exemplo, nenhuma regra transformacional pode neste modelo derivar (a) construção da ponte a partir da expressão (alguém) constrói a ponte através da substituição da morfologia verbal na forma constrói pelo sufixo derivacional -ção (entre outras operações necessárias para converter a segunda expressão na primeira). Ver o capítulo 3.2.

versal determina crucialmente que o número de hipóteses acessível à criança é agora drasticamente reduzido, facilitando o processo de aquisição (15).

12. O modelo de Princípios e Parâmetros

O modelo de Princípios e Parâmetros, proposto pela primeira vez em Chomsky (1981), incorpora em grande parte os resultados teóricos da Teoria Standard Alargada, bem como a sua concepção da organização da gramática em subteorias (ou módulos, ou componentes) autónomas, cada uma delas com uma organização e princípios independentes, e tendo como objecto domínios diferenciados da linguagem. Mais ainda do que no modelo da EST, existe uma tendência para eliminar as regras, ficando o modelo composto no essencial por princípios extremamente gerais, distribuídos pelas várias componentes, e cuja interacção determina representações com um alto grau de complexidade. De um modelo de regras e princípios passa-se pois a um modelo unicamente de princípios. Cada uma das componentes da gramática é extremamente simples na sua organização interna, e mantém uma rede de interacções com as outras componentes (16).

12.1. A noção de parâmetro

Esta organização modular, no entanto, tem agora como pano de fundo uma concepção diferente da natureza dos princípios universais de UG. Nesta concepção, a Gramática Universal é constituída por dois tipos de princípios. Tal como na Teoria Standard Alargada, UG contém princípios rígidos, invariáveis, e que qualquer gramática final terá de incorporar. Entre estes encontram-se, por exemplo, o princípio de Projecção (ver o capítulo 9.7), o princípio de que as orações das línguas humanas possuem necessariamente um NP sujeito e um VP predicado (ver o capítulo 9.8.2), e o princípio que determina que as regras de movimento apenas podem mover constituintes sintácticos (o chamado Princípio de Dependência Estrutural).

Para além destes princípios rígidos, contudo, existe agora igualmente um sistema de princípios abertos, os *parâmetros*. Estes são uma espécie

(15) A Teoria Standard Alargada abandona virtualmente qualquer preocupação com o sistema de avaliação de gramáticas da Teoria Standard. Isto deve-se certamente ao número extremamente reduzido de hipóteses possibilitadas pela Gramática Universal.

(16) Esta concepção da organização de UG em módulos é o espelho do tipo de organização que Chomsky propõe para a mente humana em geral (ver a nota 8 e Chomsky (1975; 1980b; 1988b)).

de «comutadores linguísticos» cujo valor final e definitivo apenas é atingido durante o processo de aquisição, através da sua fixação (ou ligação) numa de duas posições possíveis com base na informação obtida a partir do meio linguístico ambiente. Neste modelo, a aquisição pela criança da gramática final da sua língua consiste essencialmente em dois aspectos: a aprendizagem das formas lexicais da língua, com as propriedades fonológicas, sintácticas e semânticas determinadas pelo «dicionário mental» e a atribuição aos vários parâmetros da Gramática Universal do valor que possuem nessa língua. Estamos pois longe do modelo da criança como um «pequeno linguista» construindo hipóteses, e que caracterizava a Teoria Standard e parcialmente a EST. Neste modelo, as opções possíveis em UG são dadas pelos parâmetros, sendo portanto muito limitadas, visto que estes são em número finito e possuem por hipótese apenas dois valores. A natureza do sistema final é ainda mais limitada, visto que são os dados primários acessíveis à criança que determinam de um modo rígido o valor atribuído a cada um dos parâmetros.

Quando todos os parâmetros estão ligados, a criança adquire uma «gramática nuclear» (em Inglês, «Core Grammar»), isto é, um sistema complexo de conexões entre os princípios universais rígidos e os parâmetros, o qual determina de um modo altamente específico as propriedades de cada língua particular. A aquisição é assim completamente identificada com o crescimento e a maturação de UG, que passa de um estado apenas parcialmente especificado (com parâmetros por fixar) a um estado completamente especificado (com os parâmetros fixados), funcionando então como um sistema computacional (ver a secção 4) (17).

12.2. A fixação do valor dos parâmetros no processo da aquisição

A ligação dos parâmetros numa dada posição é determinada pelos dados linguísticos primários simples ao alcance da criança. Ou seja, a informação linguística contida nesses dados permite à criança decidir qual a posição a atribuir a cada parâmetro. Podemos ilustrar este processo pela

(17) A gramática nuclear é a componente fundamental do sistema final adquirido. Chomsky propõe que para além deste núcleo, o sistema final do falante de uma língua particular possui resíduos históricos, empressimos de outras línguas, expressões parcialmente fixadas pela convenção social (como as expressões idiomáticas), em síntese, construções e elementos marcados, formando uma espécie de «periferia» linguística, possivelmente submeida de modo parcial a alguns dos princípios de UG.

consideração de um dos parâmetros mais discutidos na literatura generativista, o chamado «parâmetro do sujeito nulo» (ver também o capítulo 16.3 e as referências aí indicadas).

A Gramática Universal contém um princípio rígido que determina a existência da posição de sujeito nas orações das línguas humanas (ver o capítulo 9.8.2). A Gramática Universal, no entanto, não determina que essa posição seja necessariamente preenchida por um NP com conteúdo fonético. Assim, em línguas como o Português, o Italiano e o Espanhol, é possível deixar essa posição «vazia» (o travessão indica a posição onde deveria normalmente ocorrer o sujeito fonético):

(20) — já chegaram da escola.

Em línguas como o Inglês ou o Francês, a posição de sujeito tem de ser preenchida foneticamente:

- (21) a. * — sont déjà arrivés de l'école.
b. * — already arrived from school.

- (22) a. Ils sont déjà arrivés de l'école.
b. They already arrived from school.

A Gramática Universal põe assim à disposição da criança uma escolha, entre a realização fonética obrigatória vs. a realização fonética opcional do sujeito da oração, ou seja, um parâmetro com duas posições possíveis que a criança terá de fixar durante o processo de aquisição⁽³⁹⁾. Esta fixação é feita com base naquilo que a criança ouve, ou seja, é feita com base nos dados linguísticos primários. Por exemplo, uma criança que cresça numa comunidade em que se fala Português não tem dificuldade em fixar o parâmetro na posição pro-drop, ao ouvir expressões como (20) nos seus dados linguísticos primários.

Se os parâmetros são fixados pela criança com base nos seus dados linguísticos primários, a natureza destes (simplicidade, ausência de informação negativa, ver a secção 9) impõe por sua vez condições extremamente restritivas sobre a natureza dos parâmetros. Concretamente, o seu conteúdo tem de ser tal que a fixação possa ser feita com base em dados simples e facilmente acessíveis à criança. Um hipotético parâmetro cuja fixação ne-

(39) Adaptando uma terminologia utilizada nos primeiros trabalhos sobre este fenómeno linguístico, estas escolhas são designadas através das expressões «valor pro-drop» (para a realização fonética opcional do sujeito) e «valor não-pro-drop» (para a realização fonética obrigatória).

cessite o acesso pela criança a frases «exóticas» e pouco frequentes na fala dos adultos é deste modo altamente suspeito como candidato a parâmetro efectivo de UG.

Suponhamos por exemplo um hipotético parâmetro cuja fixação só possa ser decidida com base em frases em que co-ocorrem um pronome interrogativo e um pronome relativo, ou duas expressões interrogativas, como em (23):

- (23) a. Eu encontrei os livros que tu perguntaste a quem o Luís tinha emprestado.
b. Que livros é que tu perguntaste a quem o Luís tinha emprestado?

Embora (relativamente) gramaticais, estas frases são suficientemente «exóticas» para as descartarmos do conjunto dos dados linguísticos iniciais, sobretudo dada a exigência de uniformidade (ver a nota 21). Logo, nenhum parâmetro de UG pode ser tal que necessite o acesso a estas frases para a sua fixação⁽³⁹⁾.

A hipótese relativa ao papel nulo da informação negativa no processo de aquisição tem também consequências importantes para as propriedades dos parâmetros, em particular para a questão de determinar a posição em que se encontram no estado inicial (S_0) da aquisição. Existem a este respeito duas possibilidades lógicas. Um parâmetro pode estar numa posição «neutra», independente das duas posições possíveis em que pode ser fixado durante a aquisição (para simplificar a discussão, identifiquemos estas posições pelos símbolos «+» e «-»). Neste caso a criança tem de decidir, a partir da sua experiência linguística, se a fixação é feita em «+» ou em «-», ou seja, num determinado momento, o parâmetro é «ligado» num destes dois valores a partir de uma posição previamente neutra.

A outra possibilidade é de que um parâmetro tenha um valor primitivo não marcado («+» ou «-»), ou seja, que o processo de aquisição comece imediatamente (em S_0) com o parâmetro fixado numa das duas posições possíveis, antes mesmo que a criança tenha sido exposta a quais-

(39) Esta discussão tem um interesse histórico porque o primeiro parâmetro sugerido na literatura generativista (em Rizzi (1982, 73-74)) necessitaria de dados deste tipo para a sua fixação, sendo portanto extremamente implausível do ponto de vista da aquisição (ver o capítulo 14.2.2, para mais pormenores).

Note-se, no entanto, que o estatuto (relativamente) aceitável destas frases em línguas como o Português implica que os princípios e os parâmetros (fixados) da Gramática Universal têm de possuir as propriedades necessárias para permitir o desenvolvimento de uma gramática que as caracterize como gramaticais nestas línguas.

quer dados linguísticos iniciais. Neste caso, se a língua à qual a criança é exposta tiver para esse parâmetro valor idêntico ao da ligação inicial, a criança não procede a nenhuma alteração; se, pelo contrário, a língua aprendida tem o valor oposto para o parâmetro, a criança terá de mudar o valor inicial com base nas expressões que ouve. Concretamente, na ausência de evidência linguística contraditória nos dados primários, a criança não toca no parâmetro; apenas se os dados linguísticos contiverem evidência contraditória mudará então a criança o valor do parâmetro.

Consideremos agora, à luz da hipótese sobre a ausência de informação negativa no processo de aprendizagem (cf. a secção 9), as implicações que ca- da uma destas possibilidades tem relativamente ao parâmetro do sujeito nulo.

Suponhamos que no estado inicial S_0 este parâmetro é neutro relativamente a qualquer dos valores «+» ou «-» que pode tomar. Para uma criança que aprende Português, a tarefa de fixar o parâmetro no valor «pro-drop» não oferece problema: orações como (20), sem sujeitos fonéticos, são suficientemente abundantes nos dados iniciais para permitir à criança decidir que a sua língua é efectivamente «pro-drop», e ligar o parâmetro com esse valor. Note-se que neste caso a criança procede à fixação do parâmetro unicamente com base em informação positiva (isto é, a existência de expressões como (20) no seu meio linguístico ambiente).

Para a criança que aprende Inglês, contudo, a tarefa é mais complicada. É certo que orações como (21b) não são gramaticais nesta língua, mas crucialmente, a criança não tem por hipótese acesso a esta conclusão, visto que se trata de informação negativa. Neste caso, para permitir a fixação correcta, temos de abandonar parcialmente a ideia de que a informação negativa não desempenha nenhum papel na aquisição, embora continuando a não permitir que as correcções, as instruções explícitas, as explicações gramaticais, etc., possam ter um papel relevante. Chomsky (1981, 9) admite que em certas circunstâncias, o facto de uma determinada estrutura não ocorrer nos dados linguísticos primários (aquilo a que chama informação negativa *indirecta*) é suficiente para fixar o valor de um parâmetro, sem que esse processo inclua correcção ou instruções explícitas (ou seja, informação negativa *directa*). Aceitando esta conclusão, a criança que aprende Inglês acabará eventualmente por fixar o valor do parâmetro na posição «não-pro-drop», não porque teve acesso à informação de que (21b) não é gramatical, mas simplesmente porque, num determinado momento da maturação da sua UG, o facto de nunca ter ouvido tais expressões é suficiente para fixar o parâmetro nessa posição ⁽⁴⁰⁾.

⁽⁴⁰⁾ Note-se que a criança terá de «ignorar» a existência de frases imperativas («eat your soup!») nos seus dados primários, o que significa que possui informação suficiente

Suponhamos agora que o parâmetro tem um determinado valor não marcado, ou seja, que no estado inicial S_0 se encontra desde logo ligado nesse valor. Como veremos, esta hipótese permite manter a versão forte da ideia de que a informação negativa (incluindo a informação negativa indirecta) não desempenha nenhum papel na aquisição da linguagem, se o valor não marcado do parâmetro for a posição não-pro-drop.

Suponhamos então, de um modo mais concreto, que as crianças entram (universalmente) no processo de aquisição com o valor do parâmetro do sujeito nulo fixado na posição não-pro-drop. Se uma criança cresce numa comunidade em que se fala Inglês, não toca no parâmetro, visto que no seu meio linguístico ambiente não existem dados positivos que contradigam a fixação predeterminada do parâmetro (mas ver a nota 10). Se a criança cresce numa comunidade onde se fala Português, tem acesso a frases como (20) ou seja, informação positiva que contradiz directamente a fixação original do parâmetro. Com base nesta informação, a criança coloca o comutador na posição oposta. Neste modelo, a informação negativa indirecta não desempenha nenhum papel na aquisição. Em resumo, se adoptarmos o modelo da teoria dos parâmetros em que estes são inicialmente neutros quanto à sua ligação final, temos de permitir que a informação negativa indirecta seja relevante na aquisição do sistema linguístico final. A manutenção da versão forte da hipótese sobre a não pertinência da informação negativa exige que adoptemos um modelo específico da teoria dos parâmetros, em que estes possuem um valor inicial não marcado e que apenas é alterado no decurso da aquisição se houver evidência positiva em contrário nos dados primários. Os psico-linguistas que trabalham no quadro teórico da Teoria de Princípios e Parâmetros têm usualmente favorecido esta última hipótese, embora seja ainda prematuro, nesta fase das investigações sobre a aquisição, adiantar conclusões rígidas (ver sobre estas questões os vários trabalhos em Roeper e Williams (eds.) (1987)).

Berwick (1982) propôs uma formalização parcial, em termos da teoria dos conjuntos, da hipótese sobre a não pertinência dos dados negativos na selecção do valor dos parâmetros. Esta formalização é expressa no Princípio do Subconjunto (adaptado de Chomsky (1986b, 146)):

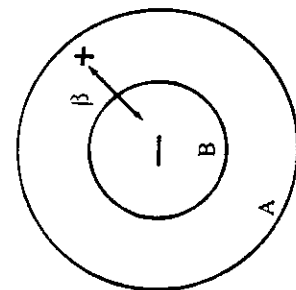
(24) *Princípio do Subconjunto*

Se um parâmetro tem dois valores, «+» e «-», e o valor «-» determina uma linguagem que é um subconjunto da linguagem

sobre a estrutura destas frases que lhe indicam que não deve generalizar a ausência do sujeito às outras orações da língua.

determinada pelo valor «+», então o valor «-» é o valor não-marcado, seleccionado por UG na ausência de qualquer evidência linguística.

Ilustremos a situação descrita pelo princípio (24) através de um diagrama de Venn (a linguagem A inclui propriamente todas as frases da linguagem B):



(25)

O princípio (24) determina que o parâmetro β tem o valor não-marcado que determina a linguagem B («-»), visto que esta é um sub-conjunto próprio da linguagem A. Suponhamos que a criança cresce numa comunidade onde se fala B. Nesse caso, a ligação primitiva do parâmetro é consistente com os dados primários (positivos) a que tem acesso. Suponhamos agora que a criança cresce numa comunidade onde se fala A. Nesse caso, ela tem igualmente acesso a dados positivos (nomeadamente as frases de A não pertencentes a B) inconsistentes com o valor inicial do parâmetro, e que lhe permitem facilmente mudar esse valor de «-» para «+», determinando a linguagem A.

Suponhamos agora, contrariamente ao princípio (24), que o valor não-marcado, inicial, do parâmetro é aquele que determina a linguagem A (isto é, a linguagem maior). Se a criança cresce numa comunidade onde se fala A, a ligação inicial é consistente com os dados primitivos, e é portanto mantida. No entanto, surgem agora problemas no caso de a criança crescer numa comunidade onde se fala B. Nesse caso, os dados primários a que tem acesso são certamente consistentes com a fixação inicial do parâmetro, *dado que a linguagem B faz parte da linguagem A*, e, logo, tem de ser também determinada por esta fixação. Para efectuar a mudança do valor «+» para o valor «-», a criança necessitaria neste caso de informação *negativa* sobre a sua língua, isto é, da informação de que as frases em A mas não em B não são gramaticais na língua (B) que está a adquirir. Con-

cluímos pois que o Princípio do Subconjunto, proibindo que o valor inicial do parâmetro seja aquele que determina a linguagem A, é de facto a formalização da hipótese de que a informação negativa não desempenha nenhum papel na aquisição da gramática.

O valor não-marcado que assumimos para o parâmetro do sujeito nulo (a posição não-pro-drop) é uma consequência directa do Princípio do Subconjunto. Numa língua pro-drop, a omissão do sujeito é apenas *opcional*. Assim, a cada frase de uma língua não-pro-drop correspondem *duas* frases numa língua pro-drop:

(26) They arrived late.

(27) a. Eles chegaram tarde.
b. Chegaram tarde.

Qualquer língua pro-drop é assim «maior» que uma língua não-pro-drop, incluindo-a propriamente. Logo, o Princípio do Subconjunto determina que o valor inicial do parâmetro do sujeito nulo é a posição não-pro-drop, aquela que corresponde à linguagem «menor».

Se atribuirmos o Princípio do Subconjunto à estrutura cognitiva da criança, fazemos então uma hipótese extremamente forte sobre o mecanismo de aquisição inicial, e em particular sobre o decurso da aquisição, que é possível pôr à prova através de estudos empíricos sobre o desenvolvimento linguístico da criança (41).

(41) Relativamente ao parâmetro do sujeito nulo, o Princípio do Subconjunto prevê que as crianças que aprendem Português, por exemplo, iniciam a aquisição com uma gramática não-pro-drop que não permite sujeitos nulos, e que apenas mais tarde começam a utilizar estas expressões, depois do valor do parâmetro ter sido mudado. Construir experiências psico-linguísticas que provem estas predições não é no entanto uma tarefa fácil, pela simples razão de que existem outros factores em jogo (como, por exemplo, factores de performance) que podem mascarar total ou parcialmente os efeitos que esperamos de uma dada fixação do parâmetro.

Hyams (1986; 1987), baseada em dados de aquisição de crianças italianas e inglesas, desenvolve uma análise em que o valor inicial do parâmetro do sujeito nulo é a posição pro-drop (Hyams mostra que as crianças inglesas produzem frequentemente e sistematicamente frases sem sujeitos antes de passarem a uma fase em que incluem o sujeito). Esta análise não é à primeira vista consistente com o Princípio do Subconjunto. Hyams no entanto sugere que as línguas não-pro-drop não são na realidade um subconjunto próprio das línguas pro-drop, e que o Princípio do Subconjunto não se pode pois aplicar no seu caso. Não cabe no âmbito deste livro discutir esta questão, e reenviamos o leitor para os trabalhos citados acima. Para uma visão discordante de Hyams, ver Bloom (1990).

12.3. Os parâmetros como princípios abstractos

Um modelo tentador da noção de parâmetro consiste na ideia de que os parâmetros são o reflexo directo de propriedades linguísticas «concretas», directamente observáveis nos sistemas gramaticais. Nesta concepção, o valor de cada parâmetro corresponderia biunivocamente a uma propriedade simples de um sistema linguístico, e o valor oposto corresponderia à propriedade contrária, existente noutra sistema linguístico. Nesta visão ingénua, qualquer diferença superficial entre duas línguas seria motivo suficiente para postular um parâmetro na Gramática Universal, a qual seria um número tão elevado de parâmetros que a aquisição do sistema final seria uma tarefa extremamente difícil, se não mesmo impossível, de atingir.

Na TRL, no entanto, os parâmetros são propriedades abstractas da Gramática Universal que não se encontram em correspondência biunívoca com propriedades simples (directamente observáveis) dos sistemas linguísticos. A fixação de um parâmetro num valor α determina (em interacção com outros princípios de UG) feixes complexos de propriedades concretas em pontos diversos da gramática, podendo mesmo algumas dessas propriedades possuir uma relação bastante remota e indirecta com a natureza mais imediata do parâmetro. Plausivelmente, UG contém apenas um pequeno número de parâmetros deste tipo distribuídos pelas várias componentes do sistema.

Exemplificando de novo com o parâmetro do sujeito nulo, esta perspectiva implica em primeiro lugar que a formulação exacta do parâmetro não refere necessariamente a propriedade concreta da omissão/não omissão do sujeito, mas sim uma propriedade mais abstracta de UG que tem como efeito a possibilidade (ou impossibilidade) de omitir o sujeito (para uma discussão concreta, ver o capítulo 16.3).

Em segundo lugar, este modelo da noção de parâmetro prediz que a omissão do sujeito é apenas uma das propriedades concretas que se encontram associadas à fixação do parâmetro na posição pro-drop, o que é confirmado pela análise linguística (ver também a discussão no capítulo 16.3). Este modelo abstracto da noção de parâmetro tem implicações importantes para os estudos de gramática comparativa e de gramática histórica. Em gramática comparativa, torna-se possível reduzir feixes de diferenças superficiais entre dois sistemas linguísticos à fixação de valores opostos num único parâmetro (ou num número muito reduzido de parâmetros); em gramática histórica, torna-se possível conceber as mudanças históricas em termos da mudança do valor de um ou mais parâmetros num dado sistema linguístico. Muitas das considerações feitas na subsecção anterior relativamente à fixação dos parâmetros durante a aquisição são igual-

mente aplicáveis neste último caso, algo que não podemos desenvolver aqui.

13. Conclusão

Como assinalámos no início deste capítulo, o modelo de Princípios e Parâmetros, e em particular a noção de parâmetro, têm possibilitado a junção de estudos puramente gramaticais sobre a estrutura de línguas individuais com estudos psico-linguísticos de natureza «experimental» sobre a aquisição e o desenvolvimento da linguagem, e com estudos de gramática comparativa. Esta interdisciplinaridade está ainda no seu começo, e não temos dúvidas que trará desenvolvimentos e modificações consideráveis à visão específica do modelo de UG proposta na TRL, e em particular à noção de parâmetro que apresentámos nesta secção.