

CONSTRANGIMENTOS INTERPRETATIVOS NA COMPREENSÃO DA INTENCIONALIDADE DO PONTO DE VISTA DA 3ª PESSOA.

António Zilhão (PhD KCL)
Universidade de Lisboa

"When we talk of language as a symbolism used in an exact calculus, that which is in our mind can be found in the sciences and in Mathematics. Our ordinary use of language conforms to this standard of exactness only in rare cases. Why then do we in philosophizing constantly compare our use of words with one following exact rules? The answer is that the puzzles which we try to remove always spring from just this attitude towards language."

Wittgenstein, L., *The Blue Book*, pp. 25/26.

0. INTRODUÇÃO

Esta comunicação tem como ponto de partida as seguintes palavras recentes de Donald Davidson:

"Se eu alguma vez desenvolvi, ou recordei, pensamentos wittgensteinianos, foi em grande medida em virtude de ter abordado os problemas da intencionalidade pela perspectiva da 3ª pessoa; e esta é uma abordagem que sempre me pareceu requerer (juntamente com muito mais coisas) o enquadramento fornecido pelas estruturas da semântica formal e da teoria da decisão" (D.Davidson, 1998, 97-98).

Que latitude devemos dar a estas palavras? Tal como Davidson, não tenho dúvidas de que parte substancial do esforço filosófico de Wittgenstein consistiu em pôr em evidência os becos sem saída a que nos conduz a abordagem filosófica tradicional dos

problemas da intencionalidade baseada na perspectiva da 1ª pessoa. Tampouco me parece haver dúvidas de que esse esforço destrutivo de Wittgenstein foi desenvolvido em simultâneo com a tentativa de esboçar uma perspectiva alternativa acerca do mesmo problema claramente baseada na perspectiva da 3ª pessoa. Parece-me que Davidson tem igualmente razão quando associa com Wittgenstein a ideia de que a compreensão da intencionalidade de acordo com a perspectiva da 3ª pessoa exige a construção de um enquadramento conceptual próprio radicalmente distinto do enquadramento tradicional. Obviamente, esse enquadramento tem que incluir, para além de outras, as duas vertentes que Davidson menciona: uma teoria linguística e uma teoria comportamental.

Mas será que, para além deste ponto de partida comum, se pode detectar um movimento de continuidade, convergência e desenvolvimento no trabalho filosófico de Wittgenstein e Davidson? Há uma linha da exegética wittgensteiniana contemporânea, representada pelo filósofo norte-americano Jim Hopkins, que defende que sim. Para Hopkins, a melhor estratégia a seguir para se alcançar uma compreensão adequada dos pontos de vista de Wittgenstein acerca da intencionalidade é procurar entender o seu trabalho à luz do trabalho mais preciso e rigoroso de Davidson, i.e., como se o primeiro constituísse uma espécie de antecipação *avant la lettre* do segundo(cf. Hopkins, 1999).

Na comunicação que a seguir apresento irei dispor perante vós argumentos por meio dos quais tentarei convencer-vos de que o enquadramento conceptual construído por Davidson para lidar com o problema da compreensão da intencionalidade do ponto de vista da 3ª pessoa não é adequado; em particular, a teoria semântica que ele elege como sendo aquela no interior da qual a compreensão da intencionalidade linguística deve ser desenvolvida não é a teoria apropriada para esse efeito. Se os meus argumentos forem convincentes, então duas opções ficam em aberto no que diz respeito à avaliação do trabalho de Wittgenstein. Ou Hopkins tem razão e os argumentos apresentados contra Davidson deverão ser extensíveis a Wittgenstein; ou esses argumentos não são extensíveis a Wittgenstein e a abordagem de Hopkins não é apropriada. A opção correcta é, do meu ponto de vista, a segunda. Deste modo, num segundo momento desta comunicação, tentarei mostrar que o enquadramento conceptual construído por Davidson se afasta em aspectos cruciais tanto da letra como do espírito da abordagem de Wittgenstein. E também que esta última sai incólume dos problemas que comprometem aquele.

1. CONSTRANGIMENTOS A PRIORI DE RACIONALIDADE

De acordo com a Filosofia da Mente de Davidson, as crenças são estados informacionais que apenas seres dotados de linguagem podem ter. O argumento no qual esta contenção se baseia é o seguinte. Ao contrário dos estados informacionais simples, as crenças são, por definição, estados informacionais que são sujeitos a avaliação semântica pelo seu possuidor. Ora, a capacidade de proceder a uma tal avaliação não pode estar presente num possuidor de estados informacionais a menos que este possua os conceitos de verdadeiro e falso. O carácter metalinguístico destes conceitos implica, por sua vez, que eles possam emergir apenas num contexto de interpretação linguística. Portanto, ter crenças implica ser-se dotado de linguagem.

Esta definição de crença implica que as crenças sejam as atitudes proposicionais básicas. Com efeito, todas as outras atitudes proposicionais partilham com as crenças a característica de terem um conteúdo semanticamente avaliável; todavia, cada uma delas acrescenta a esta característica comum uma característica atitudinal própria. Logo, ter crenças é uma condição necessária para ter quaisquer outras atitudes proposicionais. Por dedução em cadeia, segue-se que ser-se uma criatura à qual podem ser atribuídas quaisquer atitudes proposicionais depende de se ser uma criatura dotada de linguagem.

Ainda de acordo com Davidson, uma criatura é um agente racional se o seu padrão de comportamento for enquadrável no contexto da teoria bayesiana da decisão; mas um tal padrão de comportamento só é enquadrável no contexto de uma tal teoria se for possível atribuir com sentido crenças e desejos a essa criatura. Como acabámos de ver, só podem atribuir-se com sentido crenças e desejos a uma criatura se ela for dotada de linguagem. Logo, a detecção de capacidades linguísticas num grupo de criaturas fundamenta a sua consideração como dinamicamente racionais e, por conseguinte, como autoras de acções susceptíveis de serem interpretadas de acordo com linhas de força teórico-decisionais. Deste modo, para Davidson, a racionalidade não é apenas uma questão de propósito; ela é também uma questão de uso de linguagem.

Todavia, a detecção de que uma actividade aparentemente comunicacional tem um carácter linguístico não é algo que possa resultar de uma simples observação. Um comportamento deste género é uma linguagem, de acordo com Davidson, se for possível apreendê-lo no contexto do enquadramento estrutural fornecido por uma teoria de verdade que implemente o aparato da lógica de 1ª ordem com identidade. Deste modo, assim como a conformidade com os

axiomas da teoria bayesiana da decisão é considerada como constitutiva da racionalidade de um agente, também a conformidade de um padrão de comportamento *prima facie* comunicacional com a estrutura de uma teoria de verdade que implemente o aparato da lógica de 1ª ordem com identidade é considerada como apresentando a condição constitutiva da atribuição de uso de linguagem a uma qualquer criatura.

O ponto de vista de acordo com o qual a atribuição de capacidades linguísticas a quaisquer criaturas vem a par com a pressuposição de que o seu comportamento comunicacional obedece às regras de um cálculo lógico mergulha as suas raízes em Quine. Com efeito, é esta tese que se deixa descobrir como subjazendo ao famoso retrato feito por Quine de um Livingston linguista em acção junto de uma comunidade de prolores de formas fonéticas do género de 'Gavagai' (cf. Quine, 1960, cap. II). A primeira etapa do trabalho do Livingston linguista de Quine, cujo objectivo seria produzir um manual de tradução Selvagem-Inglês, consistiria em levar a cabo dois empreendimentos cruciais. Em primeiro lugar, usando uma capacidade intuitiva para distinguir atitudes de assentimento de atitudes de dissensão, o linguista procederia à tradução sem resíduo para Inglês dos sentidos-estímulo das frases observacionais de Selvagem; em segundo lugar, apoiando-se nessa primeira tradução, ele procederia à tradução, igualmente sem resíduo, das conectivas proposicionais de Selvagem para Inglês, considerando as tabelas de verdade destas últimas na Teoria das Funções de Verdade como tendo um valor definitório e fazendo corresponder os valores de verdade tradicionais aos comportamentos de assentimento e dissensão.

Uma vez que a capacidade de produzir um número reduzido de frases observacionais independentes umas das outras não teria ainda a complexidade estrutural suficiente para merecer a designação de "linguagem", seria o assentimento pelos nativos às tautologias do Cálculo Proposicional que constituiria a verdadeira condição constitutiva da sua apreensão como falantes de uma linguagem. Dada a identificação behaviourista feita por Quine do pensamento com disposições verbais, as condições transcendentais por ele requeridas para a identificação de um dado padrão de comportamento comunicacional como uma linguagem são as mesmas que aquelas que são requeridas para a identificação dos actores desse comportamento como seres pensantes. Ora, esta associação estreita entre linguagem e pensamento é claramente partilhada por Davidson, como o atestam inúmeras passagens dos seus escritos.

Deste modo, tanto para Quine como para Davidson, as condições transcendentais necessárias para a identificação de um

dado padrão de comportamento comunicacional como uma linguagem e, portanto, para a detecção da presença de pensamento, encontram-se na possibilidade de enquadrar a estrutura desse putativo sistema linguístico no quadro determinado pelas regras de um cálculo lógico. Pensar seria assim, antes do mais, implementar um tal cálculo. Aquilo que os separa a este respeito é apenas a opinião acerca de qual o âmbito do cálculo lógico em questão. Para introduzir uma distinção feita por Frege (cf. Frege, 1883), enquanto que Quine exige que uma criatura seja um utente de um *calculus ratiocinator* para poder ser vista como pensante, Davidson exige que, para assim poder ser considerada, essa criatura tenha que ser um utente de uma *lingua characterica*. Apesar desta diferença, porém, podemos considerar que ambos estes pontos de vista se deixam caracterizar pela ideia de que a linguagem é pensamento em acto e que o pensamento é, em primeiro lugar, uma manifestação empírica de um sistema dedutivo.

Neste contexto, a lógica de 1ª ordem com identidade e a teoria bayesiana da decisão aparecem ligadas por um alto grau de parentesco psicológico. Este parentesco transparece no facto de ambas estas teorias serem consideradas como aquelas que apresentam as condições constitutivas da racionalidade nos domínios de aplicação que elas regem, nomeadamente, a linguagem/pensamento e a decisão/acção. Elas constituem, portanto, aquilo que poderíamos designar como o domínio da 'racionalidade pneumática', por um lado, e como o domínio da 'racionalidade prática', por outro lado. Juntas, a racionalidade pneumática e a racionalidade prática exibiriam as condições necessárias mínimas para a atribuição de racionalidade a uma criatura. As teorias acima mencionadas serviriam assim de fundamento a toda a actividade interpretativa.

Duas questões essenciais surgem naturalmente quando nos vimos confrontados com a atribuição de uma tal proeminência a estas teorias, a saber: Porque é que a racionalidade pneumática é "constituída" pela lógica de 1ª ordem com identidade (ou pelo seu segmento proposicional)? Porque é que a racionalidade prática é "constituída" pela teoria bayesiana da decisão? Curiosamente, estas questões essenciais nunca são, que eu saiba, tratadas nem nos textos de Davidson nem nos textos de Quine.

Mas estas questões têm que ter uma resposta. Penso que esta é a seguinte. Como é sabido, a teoria bayesiana da decisão pressupõe que os agentes racionais agem de acordo com o princípio da maximização da utilidade esperada. Este princípio, por sua vez, formaliza o modo por meio do qual o comportamento dirigido de um agente pode ser apercebido como optimizando a sua vantagem subjectiva. Assim, a atribuição de um carácter alegadamente

constitutivo à teoria bayesiana da decisão como meio de identificação da acção racional é, na realidade, o resultado de se ter implicitamente estabelecido uma equação entre o conceito de racionalidade agencial e um princípio de optimização da vantagem subjectiva. Penso que o mesmo se passa no caso da atribuição à lógica de 1ª ordem com identidade de um carácter constitutivo da racionalidade no domínio da linguagem e do pensamento. Esta segunda asserção carece de alguma fundamentação. Passo a expô-la.

Um sistema inferencial regulamenta os modos por meio dos quais é legítimo passar-se de certas concatenações de símbolos para outras concatenações de símbolos. Cada uma destas concatenações pode ser tomada como uma representação. Ora, a geração de sequências lógicas de representações pode ser vista como um procedimento de optimização em dois sentidos que, no entanto, se complementam. Por um lado, se uma sequência de representações se desenrola por meio da implementação sucessiva de regras lógicas de inferência, então tem-se a garantia de que, se a representação inicial era verdadeira, então a representação de chegada também será verdadeira; poderemos chamar 'preservação da verdade' a uma tal característica das sequências lógicas de representações. Deste modo, um aparato inferencial que implemente um sistema de regras lógicas é, precisamente, um aparato que optimiza a preservação da verdade. Por outro lado, os procedimentos implementados por um aparato de maximização da preservação da verdade podem ser considerados como procedimentos de optimização num outro sentido. Como vimos acima, as atitudes intencionais que são portadoras de valores de verdade são as crenças. Uma crença é uma cristalização cognitiva de informação acerca do mundo. Ora, essa informação pode ser de melhor ou pior qualidade. A ideia subjacente à consideração de uma crença como verdadeira é, precisamente, a de que ela contém a melhor informação possível acerca do aspecto do mundo acerca do qual ela é. Por conseguinte, as crenças verdadeiras deverão ser as melhores crenças que um sistema cognitivo pode ter; o que quer dizer que um procedimento que tem como finalidade maximizar a preservação da verdade é, também, um procedimento de optimização no sentido em que é um procedimento que permite à criatura que o implementa preservar maximamente aquela característica de um conjunto de crenças que o torna no **melhor** conjunto de crenças que ela pode ter acerca do aspecto do mundo acerca do qual elas são. Deste modo, a teoria inferencial codificada na Lógica de 1ª Ordem com identidade pode ser vista como formalizando os procedimentos por meio dos quais o comportamento de um sistema cognitivo se estrutura como um comportamento de optimização da vantagem

subjectiva.

E assim, creio eu, se justifica porque é que, de acordo com Davidson, tal como o agente racional é aquele cujo comportamento prático pode ser apreendido como implementando a teoria bayesiana da decisão, também o detentor de racionalidade pneumática seja aquele cujo comportamento linguístico possa ser apreendido como implementando a Lógica de 1ª Ordem com identidade (ou o seu fragmento proposicional). Três pressupostos essenciais subjazem, portanto, a esta justificação. O primeiro é o de que a condição necessária para que um agente possa ser considerado como racional é a de que ele possa ser visto como um otimizador da vantagem subjectiva; o segundo é o de que linguagem e pensamento são as duas faces de uma mesma moeda; o terceiro é o de que, dada a sua associação essencial com o pensamento, a linguagem é parte integrante de uma faculdade geral e unificada de cognição, i.e., constitui a sua parte pneumática.

Repare-se que, na medida em que constituem o fundamento do trabalho interpretativo, as definições davidsonianas de agente racional e de detentor de racionalidade pneumática não podem ter um estatuto puramente normativo; elas têm também que ter um papel explicativo na ciência que estuda as acções humanas. O facto de tanto Quine como Davidson se colocarem numa postura interpretativa invalida a possibilidade de se fazer uma distinção essencial entre sistema linguístico, entendido como um sistema abstracto e impessoal, e comportamento linguístico, o qual, uma vez que consistiria no uso daquele sistema por criaturas independentes dele, poderia obedecer a uma lógica própria, diferente da do sistema objecto de uso.

Na realidade, uma vez que, seja num empreendimento de tradução radical, seja num empreendimento de interpretação radical, a linguagem-objecto falada pelo detentor de racionalidade pneumática é extensionalmente concebida como uma totalidade de frases cujo sentido é inferido a partir do comportamento linguístico do falante, para que a estrutura do Cálculo Proposicional e/ou da Lógica de 1ª ordem com identidade possa ser discernível nessa putativa linguagem-objecto, é necessário que o comportamento linguístico do falante da mesma dê efectivamente corpo à exigência formal que estamos habituados a associar a cálculos lógicos, habitualmente concebidos como entidades abstractas e independentes. Isso significa, entre outras coisas, que esse comportamento tem que decorrer de um modo tal que, descontando os eventuais episódios de distração, lapso, etc., se possa considerar que satisfaz uma das propriedades essenciais de um tal cálculo lógico, nomeadamente, a propriedade da

consistência. Esta consequência é claramente assumida por estes autores, os quais a colocam à cabeça como condição *a priori* da interpretabilidade de qualquer comportamento linguístico. Ela recebe mesmo uma designação específica, a saber 'princípio de caridade'.

Pode assim considerar-se que todo o empreendimento filosófico davidsoniano da interpretação radical se apoia na validade do princípio de caridade. Em consequência, se se mostrar que este princípio é inválido todo o edifício da interpretação radical cairá por terra. Senão vejamos: dado que este princípio constitui apenas uma outra designação para a exigência de consistência lógica; dado que a satisfação desta exigência é uma condição necessária para a atribuição de racionalidade pneumática; dado que a atribuição de racionalidade pneumática é, por sua vez, considerada como equivalente à atribuição de competência linguística e capacidade pensante aos indivíduos objecto do trabalho interpretativo; e ainda dado que esta última atribuição é condição necessária para que os alvos do trabalho interpretativo possam ser considerados como agentes; segue-se que, se o princípio de caridade for inválido, não existirá um modo efectivo por meio do qual se possa ancorar na realidade comportamental aquela teoria interpretativa por meio da qual essa realidade deveria supostamente ser compreendida e explicada.

Na secção que se segue pretendo precisamente mostrar que não há qualquer forma de interpretar o princípio de caridade que, simultaneamente, preserve a conexão semântica com a propriedade da consistência lógica (por referência à qual ele é introduzido) e seja empiricamente razoável. Concluo esta secção defendendo que devemos ser consequentes e abandonar o princípio de caridade.

Nas secções 3. e 4. apresento um conjunto de considerações por meio das quais pretendo mostrar que o fracasso do apelo ao princípio de caridade radica num fracasso mais geral da perspectiva que lhe deu origem. Esta perspectiva pode ser caracterizada como decorrendo da adesão, explícita ou implícita, não importa, aos três pressupostos acima mencionados. Na secção 3. defendo que o primeiro destes pressupostos é inverosímil, devendo, por isso, ser abandonado. Na secção 4. mostro que a evidência empírica actualmente disponível aponta para a provável falsidade tanto do segundo como do terceiro destes pressupostos; em consequência, defendo também o seu abandono. Finalmente, tentarei mostrar que a defesa desse abandono não é inconsistente com os propósitos filosóficos manifestados por Wittgenstein.

2. CONSTRANGIMENTOS A *POSTERIORI* DE HUMANIDADE

A forma mais imediata de tentar aplicar o princípio de caridade é aquela que decalca irreflectidamente a condição de racionalidade que ele institui da condição de racionalidade que se encontra associada aos cálculos lógicos entendidos como objectos matemáticos. Neste caso, essa condição encontrar-se-ia associada a uma condição de maximização perfeita da preservação da verdade. Uma tal condição poderia ser entendida de duas maneiras. A primeira seria como uma condição de fechamento dedutivo, isto é, esta seria a condição de que uma criatura seria cognitivamente racional se e somente se fosse capaz de, simultaneamente, deduzir de um determinado conjunto de crenças de que dispusesse num dado momento todas as suas consequências lógicas e nada delas deduzisse que não fosse uma sua consequência lógica.

Ao mesmo tempo que é a forma mais imediata de entender a condição davidsoniana de racionalidade cognitiva e o princípio de caridade que a acompanha, esta é também a mais manifestamente absurda. Existe um número infinito de consequências lógicas que um dado conjunto de crenças tem, pelo que não faz sentido colocar como condição de racionalidade cognitiva para criaturas espaço-temporalmente limitadas que essas criaturas sejam capazes de deduzir um número infinito de consequências num tempo necessariamente finito. Por outro lado, a esmagadora maioria das infinitas consequências lógicas dedutíveis seriam, debaixo de qualquer critério de avaliação minimamente realista, completamente inúteis. Perder tempo e gastar recursos a deduzir um tal conjunto de consequências deveria, por conseguinte, ser mais facilmente tomado como uma manifestação de irracionalidade, num sentido intuitivo do termo 'racionalidade', do que uma condição de racionalidade, em qualquer sentido do termo.

A segunda forma de entender a condição de racionalidade cognitiva como uma condição de maximização perfeita da preservação da verdade é entendê-la como a capacidade de decidir para cada conteúdo de crença dado à consideração do indivíduo putativamente racional se esse conteúdo seria ou não dedutível do conjunto de crenças já efectivamente existente.

Se estivermos a mover-nos no contexto da lógica de 1ª ordem com identidade, como Davidson pretende, é fácil de ver que a posse de uma tal competência por uma criatura seria equivalente à posse pela mesma da capacidade algorítmica de decidir num número finito de passos se uma expressão arbitrária desse cálculo seria ou não um teorema. Isso é algo que, como é sabido, o Teorema de Church prova não ser possível de fazer.

Se estivermos a mover-nos no contexto do Cálculo

Proposicional, como Quine pretende, é verdade que uma tal restrição não se encontra presente. De facto, no contexto do Cálculo Proposicional, existem procedimentos algorítmicos para detectar tautologias, nomeadamente, as tabelas de verdade. Todavia, embora neste caso não se possa falar em impossibilidade lógica é possível falar-se em impossibilidade prática. Um algoritmo para detectar tautologias é um algoritmo computável num tempo exponencial e este género de algoritmos é normalmente considerado em ciência da computação como o exemplo de algoritmos de complexidade impraticável. Não parece, portanto, fazer muito sentido colocar à cabeça, como condição de inteligibilidade interpretativa, que os interpretados tenham um comportamento linguístico que seja tal que só possa ser originado a partir de mecanismos de realização material impossível.

A ideia de que o princípio de caridade deva ser entendido em associação com um princípio de racionalidade cognitiva que projecta sobre os alvos da interpretação uma condição de maximização perfeita da preservação da verdade não faz, portanto, qualquer sentido. Se se pretende que a ideia consubstanciada no primeiro princípio desempenhe algum papel relevante no trabalho interpretativo o segundo princípio tem que ser consideravelmente agudo.

A forma de o aguar que tende a apresentar-se-nos primeiro à consideração é aquela que consiste na selecção de um núcleo duro de procedimentos inferenciais que têm que ser encontrados no comportamento linguístico de quaisquer criaturas putativamente racionais. Só que esta forma de aguar o princípio em causa revela-se, na realidade, extremamente problemática. Com efeito, que critérios poderemos nós usar para efectuar essa selecção? De um ponto de vista estritamente lógico, não se pode operar qualquer distinção quanto à importância intrínseca de quaisquer duas tautologias. Como Wittgenstein afirmou no *Tractatus*, elas são todas de igual patente (cf. Wittgenstein, 1922, 6.127).

Nenhumas considerações de simplicidade intrínseca são, por isso, de aceitar. Em que poderia consistir a simplicidade intrínseca de uma inferência? No menor tamanho da sua expressão simbólica? Na sua proximidade dedutiva em relação aos axiomas? Nenhum destes critérios parece ser aceitável. Se nos tomarmos a nós próprios como referência, sabemos que nos relacionamos mais facilmente com certas conectivas do que com outras, independentemente do tamanho das inferências de que elas fazem parte; sabemos que uma fórmula curta com quantificação múltipla pode aparecer-nos como mais complexa do que uma fórmula mais longa com apenas um quantificador, etc. Por outro lado, existem

diferentes axiomatizações possíveis dos cálculos lógicos e, por conseguinte, a proximidade ou distância de um dado teorema em relação aos axiomas depende do sistema de que ele faz parte.

Apelos à nossa intuição também me parecem ser um critério a rejeitar. Pessoas diferentes têm, comprovadamente, intuições diferentes a respeito de qual é a ordenação apropriada da complexidade das fórmulas inferenciais. Por outro lado, se o enquadramento do comportamento comunicacional objecto de interpretação na grelha fornecida pela lógica de 1ª ordem com identidade é uma condição *a priori* de racionalidade cognitiva, porque é que as nossas ordenações *a posteriori* de complexidade, supondo que seria possível chegar a acordo a seu respeito, deveriam constituir uma condição necessária para a aplicação dessa condição *a priori*? Esta questão é tanto mais premente quanto, tal como Chris Cherniak demonstrou, para qualquer proposta de identificação da ordenação da complexidade inferencial que esteja de acordo com a nossa intuição, é sempre possível imaginar uma ordenação da complexidade inferencial que não esteja de acordo com a nossa intuição e que, todavia, satisfaça integralmente os requisitos do princípio de caridade (cf. Cherniak, 1986, 34-39). Isto significa, por conseguinte, que a condição transcendental da atribuição de racionalidade pneumática, tal como se encontra formulada por Quine e por Davidson, não nos permite, na realidade, identificar qualquer núcleo duro de processos inferenciais por meio do apelo aos quais a condição de maximização perfeita da preservação da verdade possa ser aguada.

Falhada a possibilidade de entender do modo descrito acima o critério davidsoniano de racionalidade pneumática, resta apenas uma hipótese. Trata-se da hipótese de que o único requisito em associação com o qual o princípio da caridade é imposto seja o de que tem que ser possível identificar no comportamento comunicacional objecto de interpretação **algum** subconjunto próprio do conjunto de padrões inferenciais codificados nos cálculos lógicos atrás mencionados. À partida não se fariam quaisquer exigências quanto à extensão desse conjunto; apenas se exigiria a satisfação da condição de consistência lógica consubstanciada no princípio de caridade.

Chegado aqui, gostaria, uma vez mais, de relembrar que, de acordo com Quine e Davidson, a única evidência empírica susceptível de consideração para o desenvolvimento do trabalho de interpretação é constituída pelo comportamento, comunicacional e não comunicacional, dos interpretados. Esse comportamento desenrola-se, porém, no tempo. O que significa que quaisquer detecções de padrões inferenciais e avaliações da sua consistência

pressupõem a comparação entre prolações presentes e prolações passadas dos interpretados. Para que uma expectativa de consistência faça sentido é, portanto, necessário pressupor que haja algum controle pelos mesmos do conteúdo das suas prolações passadas. Todavia, considerações mínimas de temporalidade e finitude levam-nos a considerar que esse controlo não pode ser perfeito. Deste modo, é necessário que introduzamos alguma expectativa, mesmo que implícita, quanto à capacidade e à estrutura mnésicas das criaturas objecto de interpretação.

Se nos tomarmos a nós próprios como padrão de medida, então aquilo que os livros de Psicologia nos ensinam a respeito da nossa memória é que ela se encontra dividida em dois sub-sistemas: a memória de longo prazo e a memória de curto prazo. A memória de longo prazo é normalmente concebida como o sub-sistema mnésico no qual grandes quantidades de informação são armazenadas durante um longo período de tempo; a memória de curto prazo é normalmente concebida como o sub-sistema mnésico no qual se encontram armazenadas pequenas quantidades de informação por períodos de tempo substancialmente mais curtos, nomeadamente, enquanto um processo activo de pensamento está a decorrer em associação com a execução de uma dada tarefa (cf. Baddeley, 1990).

Uma primeira hipótese mais timorata de interpretação do princípio de caridade é, então, a de que a condição de consistência deveria incidir apenas sobre o confronto da prolação actual com aquele conjunto de prolações cujas representações poderíamos considerar estarem ainda presentes na memória de curto prazo dos interpretados, pressupondo que a estrutura e as capacidades mnésicas destes fossem essencialmente similares às nossas. Se pensarmos apenas em termos de Psicologia da Dedução esta hipótese parece ter alguma plausibilidade; ela é, porém, demasiado vulnerável ao género de considerações que são interpretativamente relevantes. De facto, de acordo com os psicólogos cognitivos, a capacidade da nossa memória de curto prazo é assaz limitada. O que quer dizer que isolar, para efeitos do trabalho de interpretação, episódios comportamentais associados à vida da memória de curto prazo de um indivíduo equivaleria a isolar conjuntos de putativas expressões de crenças sem a massa crítica necessária para fundamentar a extracção de quaisquer ilações de carácter semântico.

Deste modo, para fazer sentido interpretativo, o princípio de caridade tem que poder ser aplicável, senão à totalidade, pelo menos a uma porção considerável dos registos do comportamento comunicacional dos interpretados. Ora, a contraparte mental desses registos só pode encontrar-se na memória de longo prazo dos

interpretados. Parece-me, porém, que a experiência do intercâmbio humano efectivo mostra que a contenção de que a organização interna da memória de longo prazo e, portanto, o comportamento linguístico resultante da sua consulta, obedeceriam a esta última versão mais aguada do princípio de caridade é uma contenção extremamente implausível. Para o mostrar, peço a vossa atenção para uma pequena história verídica, a qual tem como único mérito poder servir como exemplo paradigmático, susceptível de ser substituído por um sem número de outras histórias do mesmo género que qualquer um de vós indubitavelmente conhece.

Suponham que há um certo indivíduo S que viaja regularmente na District Line do Metro de Londres para se dirigir para o Campus do Strand do King's College. Suponham que, para o fazer, ele sai do comboio na estação de Embankment e faz a pé pelo caminho que bordeja a margem norte do Tamisa a meia milha que separa esta estação do Campus. Coerente com este comportamento, ele diz crer ser essa estação a que fica mais perto do seu destino e ser essa a razão pela qual ele a utiliza. Suponhamos ainda que S afirma também saber que no seu percurso no centro de Londres a District Line acompanha o curso do Tamisa pela margem norte. E que, noutra ocasião ainda, quando perguntado se sabe qual a distância a que as estações de Metro se encontram umas das outras no centro de Londres, ele responde, correctamente, "meia milha". Por outro lado, ele também responde acertadamente à pergunta "A que distância se encontra a estação de Embankment do Campus do Strand?" com a resposta "Cerca de meia milha".

Ora bem, de acordo com o ponto de vista de acordo com o qual haveria uma conexão necessária entre linguagem, pensamento e maximização da verdade, S teria que ser considerado como alguém que seria dotado de uma outra crença que se segue obviamente das três últimas crenças, de acordo com regras de dedução que tendemos a considerar elementares e que, portanto, teriam que fazer parte do subconjunto próprio de processos inferenciais que **nos** definem como criaturas pensantes. Esta seria a crença de que haveria, como de facto há, uma estação da District Line muito mais perto dos portões do campus do Strand do que a estação de Embankment (trata-se da estação de Temple). O problema aqui é que a consideração desta última crença como verdadeira é inconsistente com a consideração como verdadeira da primeira das crenças de S acima mencionadas. Isto é, se se considerar que S tem esta última crença, então ele terá em simultâneo as crenças 'A estação de Embankment é a que fica mais perto do campus do Strand' e 'A estação de Embankment não é a que fica mais perto do campus do Strand'. Ora, de acordo com Quine, se alguém parecesse considerar

como verdadeiras duas crenças tão obviamente contraditórias como estas isso significaria que a palavra 'não' não teria no seu vocabulário o mesmo sentido que no nosso e, portanto, o seu discurso teria que ser reinterpretado por forma a evitar a inconsistência (cf. Quine, 1986, 81). Isto a menos que uma das duas seguintes possibilidades fosse o caso: ou S teria mudado de ideias entre a manifestação da primeira crença e a manifestação das outras crenças de cuja conjunção se inferiria obviamente a crença contraditória da primeira ou as crenças contraditórias apresentar-se-iam a S sob modos de apresentação que disfarçariam a sua contraditoriedade.

Esta é, como disse, uma história verdadeira e S é uma pessoa real e, por sinal, bastante inteligente. Quando comuniquei a S a existência da estação de Temple ele revelou-se, aliás, extremamente surpreendido. Todavia, não me parece que S tenha que ser considerado irracional e, portanto, ininterpretável, ou que se tenha que considerar que ele mudou de ideias, ou que se tenha que considerar que as crenças em questão se lhe tivessem apresentado sob modos de apresentação diferentes ou que as suas frases anteriores devessem ser reinterpretadas para o seu discurso poder ser compreensível. Tal como no caso das experiências para testar o axioma da transitividade da Teoria da Decisão, também aqui me parece que há uma solução que explica os dados comportamentais e que, ao mesmo tempo que é cognitivamente plausível e humanamente familiar, contradiz o princípio de caridade, isto é, admite a inconsistência, sem necessidade de ter de apelar a falhas cognitivas como lapsos ou distrações para a justificar.

Para apresentar esta solução regressemos ao que os manuais de Psicologia dizem sobre a memória. Segundo estes, a memória de curto prazo desempenha as duas seguintes funções, ambas de importância crucial: por um lado, mantém à disposição e trata a informação necessária ao desempenho de uma tarefa específica que o agente tenha que efectuar; por outro lado, filtra a informação nova a qual será, posteriormente, arquivada na memória de longo prazo. Quer dizer, a memória de curto prazo serviria assim de sistema de controlo através do qual a informação teria que ser canalizada para entrar ou sair da memória de longo prazo. Esta concepção da memória de curto prazo é normalmente associada com a ideia de que a memória de longo prazo seria um armazém de informação compartimentado de forma extremamente complexa mas que, de um ponto de vista inferencial, ela seria, ao contrário da memória de curto prazo, basicamente inerte. Para ser eficaz, este tandem memória de curto prazo/memória de longo prazo necessitaria ainda de ser servido por dois outros dispositivos. Por um lado, um

dispositivo de chamada à memória de curto prazo da parte da informação armazenada na memória de longo prazo necessária no momento; por outro lado, um dispositivo tipo catálogo de organização da compartimentação da informação armazenada na memória de longo prazo para permitir a eficácia das chamadas (cf. Baddeley, 1990).

Armados deste conhecimento básico acerca do modo como se pensa que a nossa memória funcione, regressemos à história de S. Uma interpretação plausível desta história seria então a seguinte: S teria de facto as crenças contraditórias 'A estação de Embankment é a mais próxima do campus do Strand' e 'A estação de Embankment não é a mais próxima do campus do Strand', só que elas se encontrariam armazenadas em compartimentos diferentes da sua memória de longo prazo. Por exemplo, podemos imaginar que a primeira delas se encontraria no arquivo constituído a partir do registo da informação originada na experiência directa do uso do Metro, enquanto que a segunda se encontraria no arquivo constituído a partir do registo da informação obtida numa experiência de visita ao Museu dos Transportes. Dado o carácter relativamente ocioso da maior parte da informação contida no segundo destes arquivos, podemos perfeitamente imaginar que ele não seria instintivamente consultado pela memória de curto prazo nas circunstâncias concretas de utilização diária do Metropolitano de Londres. Estas duas crenças contraditórias encontrar-se-iam assim ambas albergadas na memória de S sem que daí viesse grande mal ou ao mundo ou a S a não ser o de que S não poderia ser considerado como um maximizador da preservação da verdade das suas crenças, uma vez que o seu comportamento nestas circunstâncias deveria ter que ser considerado como normal e não como resultante de um lapso cognitivo momentâneo.

O aspecto da história *supra* que é de realçar é o de que eu não faço efectivamente a mínima ideia se S visitou ou não o Museu dos Transportes de Londres. O apelo a esta possibilidade explicativa tem apenas como finalidade mostrar como não é difícil tornar cognitivamente inteligíveis violações flagrantes do princípio de caridade sem ter que apelar para lapsos cognitivos. De facto, o dado comportamental bruto consiste, pura e simplesmente, na violação clara do princípio de caridade por S. E exemplos como o de S podem ser repetidos até à exaustão.

Para além de inteligível, a explicação que apresentei para o comportamento de S transmite igualmente uma clara sensação de plausibilidade. Esta sensação de plausibilidade tem, por conseguinte, que ter a sua origem na concordância da explicação com um princípio outro que não o de caridade. Daí que tenhamos a

intuição clara de que o aparelho cognitivo de S tem mais o que fazer do que procurar que o comportamento global, linguístico ou outro, do agente no qual está instalado satisfaça o desiderato da consistência. Qualquer que seja o seu conteúdo substantivo exacto, este princípio parece-me ser um princípio de eficácia local. Parece-me igualmente que este é o género de princípio que implicitamente associamos ao funcionamento do nosso aparelho cognitivo quando o encaramos como uma máquina biológica específica realizada numa estrutura material particular. Quando, porém, falamos em racionalidade e pensamento independentemente de um aparelho cognitivo particular essa associação implícita de facto não ocorre.

Note-se que, nesta história, S é um cidadão das ilhas britânicas que fala uma língua considerada à partida inteligível pelo seu interlocutor; se a situação em que ambos se encontrassem fosse uma situação de interpretação radical, não é difícil imaginar os epíclis interpretativos a que a aplicação do princípio de caridade ao conjunto das prolações efectuadas por S no decorrer de um período de tempo necessariamente longo conduziria o interlocutor-intérprete de S.

Disse acima que a hipótese de que o princípio de caridade deveria ser interpretado como incidindo apenas sobre os comportamentos linguísticos dependentes do controle da memória de curto prazo seria, apesar de inútil para o ponto de vista em que a interpretação radical se coloca, uma hipótese *prima facie* plausível, se considerada em termos de Psicologia da Dedução. Gostaria de retornar agora a esta afirmação. Em particular, gostaria de mencionar o facto de que têm sido obtidos resultados empíricos na investigação em Psicologia da Dedução que parecem contradizer mesmo esta hipótese, excessivamente timorata, de interpretar o princípio de caridade.

Esta investigação desenvolve-se em torno da apresentação de problemas dedutivos com apenas algumas premissas a grupos de sujeitos sem educação lógica formal explícita para que estes os resolvam. Estas experiências encontram-se detalhadamente descritas em obras de Evans (cf. Evans, 1982), Goldman (cf. Goldman, 1986), e Evans, Newstead & Byrne (cf. Evans, Newstead & Byrne, 1993), entre outros, pelo que eu me dispensarei de as reproduzir. O que é importante reter das mesmas é que, dado o reduzido número de premissas envolvido em cada um destes problemas, parece seguro poder afirmar-se que a sua resolução não implica a necessidade de os sujeitos recorrerem a informação guardada na memória de longo prazo. Embora os resultados sofram alguma variação de experiência para experiência, o seguinte padrão parece ser relativamente

estável: a esmagadora maioria dos sujeitos faz correctamente deduções em *Modus Ponens*; uma grande percentagem, normalmente acima de 50%, falha a resposta correcta em problemas que envolvam o uso da regra de *Modus Tollens* (na realidade, mais de 80% falham a regra de *Modus Tollens* se o antecedente estiver negado, enquanto que um pouco mais de 70% acertam nesta regra se o antecedente estiver afirmado); uma percentagem não negligenciável de sujeitos (rondando os 50% nalguns casos, mas sempre acima de 25% em todos os casos) comete falácias clássicas como a da afirmação do conseqüente e a da negação do antecedente; uma maioria esmagadora, bem próximo dos 100%, falha a resposta correcta na famosa tarefa de selecção de Wason, a qual constitui basicamente uma forma engenhosa de verificar se as pessoas raciocinam espontaneamente de acordo com a tabela de verdade da implicação.

Por muito incompletos que possam ser, estes resultados constituem um problema extremamente sério para os defensores da ideia de que a adequação de um padrão de comportamento comunicacional à estrutura formal própria de um cálculo lógico constitui uma condição necessária para a atribuição de pensamento e competência linguística aos autores desse comportamento. Como é fácil de ver, os padrões de erro inferencial acima mencionados rapidamente geram contradições e, logo, violações do princípio de caridade (por exemplo, se um sujeito da posse das premissas ' $A \rightarrow B$ ' e ' $\sim B$ ' não acerta na regra de MT, isso significa que não nega a proposição A; mas se 'A' continua disponível no seu stock de premissas, ele pode derivar 'B' por MP de ' $A \rightarrow B$ ' e 'A'; ao fazê-lo, porém, ele terá inferido as proposições contraditórias 'B' e ' $\sim B$ '). Estes padrões de erro inferencial são, todavia, demasiado estáveis para que possam ser interpretados como resultando de lapsos, distrações, ou curto-circuitos cognitivos momentâneos. Por outro lado, estes resultados foram obtidos testando pessoas que falam a nossa língua e que estão inseridas no mesmo contexto cultural em que nós estamos. Imaginem agora o que aconteceria se os seus padrões de respostas aos testes tivessem sido tomados como padrões de frases não interpretadas, o sentido das quais deveria ser revelado pela projecção sobre os mesmos seja de todo o aparato lógico do Cálculo Proposicional, como Quine defende, seja de todo o aparato inferencial da Lógica de 1ª Ordem com identidade, como Davidson defende.

A seguinte conclusão parece-me portanto incontornável: o princípio de caridade, tal como apresentado por Quine e Davidson, o qual desempenha um papel imprescindível no aparato teórico-interpretativo erigido por estes autores, não é susceptível de uma

clarificação empiricamente satisfatória. Deste modo, o edifício teórico erigido sobre ele não pode ter um verdadeiro alcance explicativo. Princípio e edifício devem assim ser abandonados.

3. RACIONALIDADE PSICOLÓGICA E CONDIÇÃO DE APRENDIBILIDADE

Se a argumentação apresentada na secção anterior está correcta, pelo menos nas suas linhas gerais, ela pode ser tomada como evidência a favor da tese de que a associação conceptual entre racionalidade pneumática e maximização da preservação da verdade é uma associação ilegítima, pelo menos na medida em que se queira considerar os seres humanos como pneumaticamente racionais. Este resultado é, por sua vez, consistente com inúmeros resultados obtidos em estudos empíricos de Psicologia da Decisão, os quais favorecem igualmente a conclusão de que a equação implicitamente estabelecida por Davidson entre racionalidade agencial e observância do princípio da maximização da utilidade esperada tão-pouco seria uma associação legítima, pelo menos na medida em que se quisesse considerar os seres humanos como agentes racionais.

Como disse acima, o que une estes dois requisitos conceptuais e explica o facto de ambos serem apresentados por Davidson como condições transcendentais da atribuição de racionalidade é a acima referida equação implícita entre os conceitos de racionalidade e de maximização da vantagem subjectiva. A este respeito gostaria de defender aqui a seguinte tese de fundo: **se quisermos conservar o princípio da racionalidade do senso comum humano, teremos que abandonar a caracterização substantiva do conceito de racionalidade psicológica em termos de optimização da vantagem subjectiva.** Evidentemente, esta tese não foi demonstrada. Mas penso que o que ficou dito é já suficiente para mostrar que é necessário dissociar os princípios de eficácia cognitiva local inerentes ao funcionamento de qualquer sub-sistema cognitivo, incluindo os que constituem o senso comum humano, de princípios particulares de optimização da vantagem subjectiva, como é o caso do princípio da maximização da preservação da verdade. Os primeiros têm que levar em linha de conta, de forma essencial, as limitações de tempo e de recursos que estão efectivamente ao dispor de um agente numa qualquer situação real; os segundos caracterizam-se, precisamente, por fazerem abstracção de tais limitações. A necessidade de se proceder a esta dissociação milita fortemente a favor da tese que defendo.

O abandono desta caracterização implícita do conceito de racionalidade psicológica permite-nos desde logo desfazer as

associações conceptuais que subjazem às condições de racionalidade que Davidson estipula serem condições sintéticas *a priori* da interpretação. Uma vez que tenhamos dado este passo, encontramos livres para prestar uma atenção menos preconceituosa à evidência comportamental efectivamente disponível. E o que ela nos mostra é não apenas que comportamentos "irracionais", não explicáveis à custa de quaisquer lapsos cognitivos, ocorrem com bastante mais frequência do que seria de esperar, mas também que, com frequência, esses comportamentos "irracionais", uma vez colocados no contexto cognitivo apropriado, têm um maior êxito local que os comportamentos "racionais" correspondentes inferidos a partir da aplicação indiscriminada do princípio de caridade.

Dito isto, convém todavia acrescentar duas considerações. A primeira revolve em torno da caracterização positiva dos princípios interpretativos alternativos ao princípio de caridade. A segunda revolve em torno da chamada condição de aprendibilidade. Começemos pela primeira consideração.

Se os utentes de uma linguagem não são, necessariamente, maximizadores da preservação da verdade, então, como Davidson claramente o viu, torna-se muito mais difícil adquirir um controlo conceptual sobre os princípios interpretativos por meio do uso dos quais é realmente possível compreender a intencionalidade do interior da perspectiva da 3ª pessoa. O princípio de caridade teria que ser substituído por princípios cujo conteúdo substantivo está ainda por explicitar. A caracterização mais precisa do conteúdo destes princípios de eficácia cognitiva parece-me, todavia, ser, antes de mais, um empreendimento fundamentalmente empírico. Como o mostram as dificuldades que o modelo de Davidson encontra quando confrontado com a sua aplicabilidade empírica, o alcance de considerações apriorísticas a este respeito parece-me, nestas circunstâncias, ser bastante limitado.

Em particular, parece-me que a dificuldade em formular adequadamente tais princípios não pode ser iludida com apelos à empatia. A empatia pode, sem dúvida, ser um auxiliar precioso na realização de trabalho de campo; mas ela nem constitui nem substitui uma caracterização substantiva da especificidade da estrutura dos comportamentos humanos, linguísticos ou outros. A mesma observação se aplica ao princípio de Humanidade, tal como formulado por Grandy (cf. Grandy, 1973). Este princípio é basicamente o de que uma tradução será tanto mais inteligível quanto mais ela tornar o comportamento objecto de interpretação num comportamento semelhante àquele que nós teríamos nas mesmas condições. Mas o estabelecimento de uma analogia com o nosso próprio caso, mesmo que nos permita prever com sucesso o

comportamento dos agentes alvo da interpretação, não nos fornece qualquer caracterização substantiva dos princípios por meio dos quais se poderia obter uma explicação do nosso próprio comportamento e, *a fortiori*, do comportamento daqueles que nós interpretamos analogicamente. Além disso, a aplicação deste princípio pressupõe um grau de auto-conhecimento que me parece estar longe de ser o caso. Por outro lado ainda, esta posição de Grandy parece-me ser desnecessariamente restritiva.

Repare-se que a discussão mais recente entre a maioria dos adeptos da 'theory-theory' e os adeptos da 'simulation theory' limita-se a reproduzir o debate Quine/Davidson vs. Grandy no contexto de um paradigma internalista. De facto, para a maioria dos 'theory-theorists', a capacidade que os membros da nossa espécie revelam para interpretar e prever com sucesso o comportamento intencional dos outros membros da espécie só admite ser explicada como resultando do exercício da competência específica de um dos módulos da nossa mente, módulo esse no qual se encontraria inatamente inscrita uma teoria interpretativa do género daquela que é proposta por Davidson, a qual seria verdadeira a respeito dos seres humanos (cf. Gopnik & Wellman, 1994). Mas se o género de objecção acima levantado contra o uso do princípio de caridade como princípio interpretativo está correcto, então, por maioria de razão, a tese de que a nossa mente estaria dotada de um módulo inatamente programado para interpretar os membros da nossa espécie de acordo com uma teoria realizadora de um princípio explicativo de aplicabilidade mais que duvidosa resulta extremamente implausível.

Para os 'simulation-theorists', por outro lado, a nossa capacidade interpretativa admite ser explicada como resultando de um exercício de simulação por meio do qual o intérprete compreenderia o interpretado apenas na medida em que conheceria as atitudes e comportamentos intencionais que ele próprio geraria nas circunstâncias em que o interpretado se encontraria e pressuporia que o interpretado geraria nessas circunstâncias essas mesmas atitudes e comportamentos (cf. Gordon, 1986 e Goldman, 1992). Esta tese, porém, é, no mínimo, tão vulnerável quanto a de Grandy à objecção que põe em causa a legitimidade do pressuposto de auto-conhecimento. Mas, mesmo que suponhamos que há uma resposta apropriada a esta objecção, levanta-se de imediato uma segunda objecção. Trata-se da seguinte. A pressuposição de auto-conhecimento parece implicar que existe, de facto, uma teoria interpretativa implícita nalgum módulo da nossa mente, a activação do qual, no caso em que o intérprete se toma a si mesmo como objecto de interpretação, é suficiente para a obtenção do

resultado apropriado. No caso, porém, em que o interpretado é outra pessoa essa activação precisa de ser suplementada com um exercício extra de simulação. Mas esta restrição do conhecimento teórico ao próprio caso não elimina o apelo a um género de conhecimento teórico, o que nos remete de novo para o ponto de vista dos 'theory-theorists', ao qual o ponto de vista dos 'simulation theorists' pretendia ser uma alternativa e não apenas uma adenda.

Não há, portanto, modo de fugir à questão fulcral. E esta é a de que a explicitação teórica da estrutura do comportamento cognitivo humano comum está ainda por fazer. Em particular, o conceito de racionalidade psicológica necessita ainda de ser clarificado. Clarificar este conceito é conseguir alcançar uma formalização aceitável do nosso senso comum. Como inúmeros investigadores em Inteligência Artificial o têm repetidamente constatado nos últimos anos, esta é uma tarefa tremendamente complexa. Isso não significa, porém, que ela seja impossível, ao contrário do que filósofos como Dreyfus pensam (cf. Dreyfus, 1992). Davidson e Dreyfus parecem-me precisamente constituir o Cila e o Caríbdis filosóficos entre os quais é necessário navegar.

Passemos agora à segunda consideração mencionada acima. A ideia de que a linguagem e o pensamento seriam, necessariamente, mecanismos de maximização da preservação da verdade surge igualmente associada na filosofia de Davidson a uma outra condição completamente independente das condições *a priori* de racionalidade, nomeadamente, àquela a que ele chama a **condição de aprendibilidade**. A ideia subjacente à formulação desta condição pode resumir-se do seguinte modo. Uma característica que parece ser comum a todas as linguagens naturais é a de que elas oferecem a possibilidade de construir um conjunto infinito de frases, a partir de um conjunto de expressões semanticamente primitivas. Ora, uma estrutura com esta característica só pode ser aprendida ou dominada pelos seus utentes se aquilo que eles aprendem ou dominam for representável por meio de uma descrição construtiva, i.e., uma descrição que represente a aptidão para o manuseamento de uma totalidade infinita como gerável por processos finitos. Acontece que um candidato ideal para constituir uma tal descrição construtiva de uma linguagem é, precisamente, uma teoria recursiva de verdade. Por sua vez, um sistema de maximização da preservação da verdade constitui um candidato *prima facie* ideal para realizar uma teoria recursiva de verdade.

Vimos acima o género de problemas empíricos a que a apreensão de manifestações de comportamento linguístico como lances de maximização da preservação da verdade dá origem.

Todavia, o género de considerações de composicionalidade subjacentes à formulação da condição de aprendibilidade tem que ser levada em linha de conta como um constrangimento incontornável na explicitação da estrutura do género de conectividade que subjaz à nossa linguagem. Não há, todavia, razão para pensar que chegados aqui deparámos com um beco sem saída. De facto, existe pelo menos um modelo alternativo ao de Davidson que igualmente satisfaz a condição da recursividade sem para isso ter que pressupor que o comportamento linguístico dos humanos só é interpretável no âmbito de uma teoria de verdade modificada que implemente todo o aparato inferencial da Lógica de 1ª Ordem com identidade. Trata-se do modelo chomskyano da Gramática Universal. A Teoria da Gramática Universal descreve um sistema sintáctico recursivo, o qual, de acordo com Chomsky, se encontraria inscrito na mente dos seres humanos desde a nascença. Este sistema dividir-se-ia num conjunto de princípios universais rígidos e de parâmetros sujeitos a variação dentro de certos limites. A aprendizagem da língua materna por uma criança consistiria assim fundamentalmente num trabalho de colocação na posição apropriada do conjunto de parâmetros não rigidamente regulados pela Gramática Universal. Deste modo, qualquer língua humana satisfaria, por definição, a condição da aprendibilidade (para os portadores de cérebros humanos nas condições de maturação adequadas). Ela seria simplesmente uma das possíveis projecções da gramática universal, a qual formataria à partida a faculdade da linguagem de que todos os seres humanos disporiam.

Uma descrição construtiva de uma qualquer linguagem humana seria assim uma descrição da forma como ela se deixaria derivar da Teoria da Gramática Universal por meio da fixação apropriada dos parâmetros apropriados. O elemento determinante para a elaboração de uma tal descrição construtiva para uma linguagem particular seria assim a atribuição automática aos falantes dessa linguagem dos universais linguísticos que caracterizariam a Gramática Universal e não a postulação de que o seu comportamento linguístico e não linguístico obedeceria ao princípio de caridade. Uma vez elaborada, a teoria construtiva da linguagem particular em questão descreveria, não o emparelhamento de um conjunto infinito de frases, extensionalmente definido por meio de uma descrição recursiva, com as suas condições de verdade, mas antes uma função considerada em intensão, i.e., esta teoria descreveria a gramática dessa linguagem. A compreensão de uma linguagem poderia assim reduzir-se à compreensão das diferentes modalidades de uso pelos seus falantes das estruturas formais determinadas pela investigação sintáctica. Neste sentido, a semântica seria

dispensável. A compreensão teórica de uma linguagem exigiria a compreensão teórica da sua sintaxe recursiva e da sua pragmática (cf. Chomsky, 2000, 132).

Este ponto de vista pode estar certo ou errado. Mas a sua consideração como possivelmente verdadeiro permite-nos abandonar sem problemas de maior a ideia central subjacente ao pensamento de Davidson de acordo com a qual linguagem e pensamento seriam necessariamente sistemas de optimização da verdade sem nos deixar reféns da condição de aprendibilidade. Nenhuma necessidade conceptual nos obriga, portanto, a aceitar o ponto de vista "transcendental" de Davidson. Pelo contrário, a possibilidade de que ele possa ser empiricamente verdadeiro parece até ser, como acabámos de ver, bastante remota. Assim, nenhuma razão nos pode impedir de, com, pelo menos, igual plausibilidade, sustentar que não é porque somos sistemas optimizadores da verdade que falamos, mas que, pelo contrário, é porque falamos que podemos, entre muitas outras coisas, exprimir-nos de acordo com princípios de optimização da verdade e que podemos, inclusive, construir toda uma forma de vida assente nesses princípios.

A necessidade de dissociar a interpretação da linguagem natural da observação de princípios de maximização da preservação da verdade é uma das linhas de pensamento que se encontra mais abundantemente expressa nos escritos de Wittgenstein. A sua obra maior, as *P.U.*, encontra-se recheada de exemplos. É o caso das §§23 e 65, onde Wittgenstein alerta o seu leitor para a multiplicidade de jogos de linguagem que podemos detectar na observação atenta dos usos linguísticos, assim como para a dependência dos mesmos das formas de vida às quais estão associados. Ou o caso da §81, onde Wittgenstein contesta explicitamente o ponto de vista de Ramsey (não por acaso o fundador da moderna Teoria da Decisão e uma das principais referências teóricas de Davidson), de acordo com o qual a Lógica se encontraria com a linguagem natural no mesmo género de relação de abstracção idealizada com que as Ciências da Natureza se encontrariam com os fenómenos naturais que elas descreveriam. É, finalmente, o caso também das §§ 107 e 108, onde Wittgenstein intui claramente que o género de resultados empíricos relatados nesta comunicação deveria ter que vir a obter-se quando afirma que quanto maior for a precisão com que observarmos a linguagem efectivamente falada, maior será o contraste entre o resultado dessas observações e as exigências teóricas *a priori* com que tendemos a abordá-la.

A estes exemplos, outros poderiam ser acrescentados. Todos eles parecem-me deixar claro que as perspectivas de Wittgenstein e

Davidson acerca de quais são os constrangimentos interpretativos requeridos para se poder abordar com êxito a intencionalidade do ponto de vista da 3ª pessoa divergem, pelo menos, numa questão crucial. A de qual deve ser o enquadramento semântico com base no qual se pode interpretar a linguagem natural com sucesso. Em particular, a de saber se as estruturas de uma semântica formal tarskiana desempenham algum papel essencial nesse enquadramento. Ao contrário do que parece depreender-se do ponto de vista exegético de Hopkins esta **não** é uma questão de detalhe.

4. LINGUAGEM NATURAL, SENSO COMUM, LÓGICA E CIÊNCIA

Acima, correlacionei a defesa por Davidson de um ponto de vista logicista acerca da linguagem com três pressupostos. A sua adesão implícita a uma concepção que equaciona a racionalidade humana com um princípio de maximização da vantagem subjectiva; a sua adesão à ideia de que linguagem e pensamento seriam duas faces da mesma moeda; e, finalmente, a sua adesão implícita a uma visão unificada da mente humana, i.e., a uma visão que determina que os mesmos princípios de base determinam, necessariamente, o funcionamento dos diferentes aspectos do nosso sistema cognitivo, linguagem incluída. Mostrei já que a equação entre racionalidade e maximização da vantagem subjectiva não só não é necessária como nem sequer é empiricamente adequada. Mostrei também como o seu abandono não nos deixa reféns da condição de aprendibilidade. Gostaria agora de tentar mostrar-vos que parecem existir boas razões para tão-pouco aceitar como correctas as ideias de que linguagem e pensamento não seriam senão duas faces da mesma moeda e de que a mente humana seria um mecanismo basicamente unificado. Em seguida, procurarei mostrar-vos que, ao contrário do que sucede com a filosofia de Davidson, a incorrecção de tais ideias não entra em choque com a perspectiva de Wittgenstein acerca da intencionalidade.

Para tratar da questão da relação que se supõe existir entre linguagem e pensamento vou novamente usar como referência contemporânea o trabalho de Chomsky. De acordo com este linguista, para além de inata, a faculdade da linguagem seria uma faculdade independente das outras faculdades cognitivas, nomeadamente, da inteligência geral ou da faculdade do senso comum e do pensamento dela decorrente. Ora, esta contenção, que contradiz directamente o segundo dos pressupostos que identifiquei como subjacentes ao pensamento de Davidson, **não é apriorística**. Ela é formulada como a melhor hipótese disponível para dar conta de um conjunto de resultados empíricos que se têm acumulado nos últimos tempos.

O raciocínio subjacente à formulação desta hipótese é o seguinte. Se os princípios por meio dos quais se rege o uso da linguagem fossem os mesmos que aqueles pelos quais se rege a actividade cognitiva de um modo geral, então deveriam observar-se as seguintes consequências: primeira, lesões e doenças cerebrais que afectam as competências cognitivas dos seres humanos deveriam afectar de igual modo a sua capacidade linguística; segunda, lesões e doenças cerebrais que comprovadamente afectam a competência linguística deveriam de igual modo interferir com o uso das competências cognitivas. A pesquisa neurológica tem vindo a detectar casos que, sistematicamente, contradizem tanto a primeira como a segunda das consequências acima mencionadas. Com efeito, existem doenças cerebrais (e.g., hidrocefalia, Síndrome de Williams) que causam problemas graves de atraso mental sem afectarem minimamente as competências linguísticas dos pacientes; existem lesões cerebrais que privam os pacientes da sua faculdade linguística (lesões nas áreas de Broca e Wernicke) sem parecerem afectar as suas capacidades cognitivas de uma forma geral; e existem doenças cerebrais que afectam severamente o desenvolvimento linguístico de certas crianças (a chamada Deficiência Linguística Específica) sem lhes causarem quaisquer problemas sócio-cognitivos graves como o atraso mental ou o autismo.

Qualquer um destes 3 casos parece constituir uma instância que falsifica a tese de que as competências linguísticas e cognitivas, i.e., a linguagem e o pensamento, se organizariam em torno de princípios comuns. Em particular, o primeiro destes 3 casos parece contradizer frontalmente a tese de acordo com a qual a obediência ao princípio de caridade constituiria uma condição necessária para a interpretação linguística ser possível. O discurso dos pacientes que sofrem de hidrocefalia ou do Síndrome de Williams é não só sintacticamente irrepreensível como claramente compreensível, apesar das constantes inconsistências que o afligem, ou do comportamento errático ao qual se encontra associado (cf. Pinker, 1995).

A estes 3 casos pode, aliás, juntar-se ainda um quarto. Trata-se do caso sobejamente conhecido de que o período do desenvolvimento humano no qual é mais fácil aprender uma língua é o período entre os 2 e os 5 anos. Como é sabido, este é um período da vida humana no qual as capacidades cognitivas gerais estão menos desenvolvidas do que no período adulto, o período no qual o desenvolvimento das faculdades cognitivas gerais atinge o seu auge. Mas neste último período, a dificuldade que os seres humanos revelam em aprender adequadamente uma outra língua é igualmente um

fenómeno conhecido e sofrido por todos. Até ver, a interpretação mais plausível para estes factos parece ser a de que as faculdades da linguagem e da cognição são distintas, obedecem a princípios diversos de funcionamento e têm períodos de maturação distintos.

Deste modo, mesmo que, por hipótese absurda, fosse o caso que a posse de racionalidade cognitiva estivesse indissoluvelmente ligada à obediência a um princípio de maximização da vantagem subjectiva, daí só se seguiria que a atribuição de capacidade linguística teria que estar condicionada por um tal princípio se a premissa adicional de que linguagem e pensamento seriam apenas duas faces da mesma moeda fosse verdadeira. Não sendo esta premissa verdadeira, de acordo com a evidência empírica disponível, a extrapolação feita por Davidson de que falar seria essencialmente um procedimento de maximização da preservação da verdade carece do carácter de necessidade transcendental que ele lhe atribui.

Em diferentes secções dos seus múltiplos escritos, existem referências de Wittgenstein ao modo como devemos entender o uso do verbo 'pensar' e de palavras dele derivadas. Essas múltiplas referências têm duas características comuns, uma de carácter negativo, outra de carácter positivo. A primeira é a negação de que o uso comum desse verbo refira os elementos do fluxo da experiência fenomenológica, por referência aos quais as palavras efectivamente pronunciadas ganhariam sentido, como pretendia a semântica empirista tradicional. A segunda é a ideia de que aquilo que a realização de uma análise fisionómica da experiência do uso comum desse verbo nos revela é que esse uso consiste num modo peculiar de descrever certas actividades (linguagem, trabalho, comportamento prático, etc.), nomeadamente, um modo que associamos à sua classificação como humanas (cf. Wittgenstein, 1979). Em todos estes casos, o que está sempre em causa é o modo como estas palavras são efectivamente empregues no seu uso comum e o modo como, na actividade filosófica, esse uso comum pode ser distorcido. Que podem essas referências de Wittgenstein dizer-nos acerca daquilo que ele pensava sobre a questão que agora nos ocupa? Do meu ponto de vista, nada de substancial.

De facto, em nenhuma circunstância Wittgenstein toma posição acerca do modo como um empreendimento explicativo do género, e.g., da Psicologia Cognitiva poderia ou deveria reconstruir o uso do conceito, por forma a que ele pudesse desempenhar um papel relevante nesse empreendimento científico. Ora, este é precisamente o contexto no qual o trabalho de Chomsky deve ser inserido. Os termos "pensar" e seus derivados são usados por ele num sentido técnico, relevantemente definido em função das

necessidades teórico-explicativas do programa de investigação no qual ocorrem. Neste contexto, não vejo como a evidência atrás apresentada poderia não dar razão a Chomsky. Obviamente, Wittgenstein não teria apresentado quaisquer objecções ao procedimento reconstrutivo levado a efeito por Chomsky. Esse é o procedimento normal da Ciência; a este respeito, ele limitar-se-ia, provavelmente, a dizer que o propósito do seu trabalho filosófico seria outro.

Todavia, algumas das passagens efectivamente escritas por Wittgenstein dizem-nos coisas relevantes sobre o modo como este pressuposto do pensamento de Davidson afecta a sua filosofia. Ao contrário do que acontece no caso de Chomsky, no contexto da filosofia de Quine e Davidson os termos "pensar" e seus derivados não são introduzidos num sentido explicitado tecnicamente. De facto, a compreensão do seu uso por estes autores remete **simultaneamente** para uma reelaboração teórica da ideia positiva associada por Wittgenstein aos usos do termo na linguagem natural e para um fim teórico-explicativo. Ora, nem a reelaboração teórica do termo faz jus ao modo como ele é efectivamente usado na linguagem natural nem a sua aplicação num contexto explicativo parece ser, como vimos, empiricamente adequada. Mas cada uma destas vertentes do sentido da palavra ampara a outra impedindo um efectivo controle pelos seus utentes do sentido com que é usada. O ancoramento na linguagem natural impede que o termo seja redefinido tendo em vista considerações de utilidade explicativa; o ancoramento numa finalidade teórico-explicativa impede que o uso do termo seja contestado com base em considerações suscitadas por uma análise fisionómica da linguagem natural. Acontece que este género de procedimento se choca directamente com inúmeras observações de Wittgenstein acerca do carácter nefasto da confusão entre diferentes planos de linguagem. Segundo ele, uma tal confusão estaria na origem de muitos dos nós conceptuais nos quais os filósofos se têm deixado enredar.

Chegado aqui, gostaria ainda de acrescentar que, ao contrário do que é tido como adquirido pela generalidade dos comentadores que se reivindicam da ortodoxia wittgensteiniana, se é necessário encontrar algum intelectual contemporâneo filosoficamente relevante do qual o espírito do empreendimento filosófico de Wittgenstein se encontre próximo, esse intelectual é Noam Chomsky e não Donald Davidson. De facto, parece-me que o espírito de investigação naturalista que encontramos no trabalho linguístico de Chomsky se compatibiliza mais facilmente com a postura filosófica de Wittgenstein do que o espírito apriorista que domina o pensamento de Davidson. É verdade que ao trabalho

publicado de Wittgenstein falta o aspecto construtivo, científico, que encontramos na obra de Chomsky. Como vimos já, o próprio Wittgenstein foi bastante claro acerca da natureza do seu trabalho, considerando-o como um empreendimento de demolição das construções gasosas erigidas pela filosofia tradicional e não como um empreendimento de construção científica. Mas foi igualmente claro na afirmação de que o seu objectivo último era o de deixar livre de um entulho conceptual secular o terreno sobre o qual futuras investigações empíricas acerca da linguagem e da mente pudessem ter lugar (cf. Wittgenstein, 1952, §118). Ora, o que caracteriza a actividade de Chomsky é precisamente o facto de ele ter construído sobre o solo conceptual terraplanado por Wittgenstein um edifício científico no interior do qual podemos começar a vislumbrar como é possível, de facto, compreender a intencionalidade do ponto de vista da 3ª pessoa. E fê-lo de tal forma que os fundamentos sobre os quais este edifício se apoia são apenas aqueles sobre os quais qualquer programa de investigação científica se apoia.

Esta última observação permite-me agora clarificar um pouco melhor o sentido no qual me parece que é Chomsky e não Davidson quem transporta o testemunho passado por Wittgenstein. É que, tal como qualquer outro programa científico, o programa de Chomsky pode estar errado. Davidson, ao invés, blindou o seu programa de investigação contra a refutação empírica estipulando que a inteligibilidade dos próprios dados empíricos pressupõe a aceitação prévia de princípios *a priori*, como o princípio de caridade ou os axiomas da teoria da decisão, dos quais a teoria pretensamente explicativa se deixa derivar. Ao fazê-lo, Davidson colocou-se na linha daquelas teorias que Wittgenstein caracterizou na *P.G.* como assemelhando-se a construções complexas que têm o aspecto de máquinas aparentemente destinadas a executar um determinado trabalho mas que, na realidade, dado o modo como as suas peças se encontram dispostas, funcionam em circuito fechado (cf. Wittgenstein, 1969, 194). Este género de teorias constitui, precisamente, o cerne da tradição filosófica que Wittgenstein se esforçou por desmistificar. Este é também um aspecto divergente da relação filosófica entre Wittgenstein e Davidson que Hopkins deixa passar em claro. Tal como a divergência em torno de qual deve ser o enquadramento semântico do trabalho de interpretação, tão-pouco esta divergência entre Wittgenstein e Davidson me parece ser uma questão de detalhe.

Para terminar, gostaria ainda de acrescentar algumas notas finais a respeito do terceiro pressuposto – o da unidade da mente humana – e do modo como a sua negação por Chomsky pode ser vista

como uma vindicação de certos pontos de vista de Wittgenstein.

Parafraseando S. Pinker (Pinker, 1994, 419), o retrato-robot das ideias de Chomsky que prevalece nos meios académicos é o de que ele proporia que a mente humana se encontraria dividida numa enorme multiplicidade de módulos independentes, os quais gerariam faculdades tão díspares e improváveis quanto a faculdade para andar de bicicleta sem cair, para usar gravatas que não destoem das camisas ou para construir carburadores. Evidentemente, a divisão da mente humana em faculdades distintas efectivamente proposta por Chomsky nada tem a ver com este retrato jocoso. Para os meus propósitos neste ensaio, gostaria de salientar apenas três características desta proposta. A primeira é a de que cada uma das faculdades cuja existência é postulada por Chomsky é associada a mecanismos próprios de funcionamento, os quais não são automaticamente transponíveis para as outras faculdades; deste modo, não haveria qualquer estrutura unitária essencial da mente humana na base da qual pudessem ser compreendidos, por exemplo, todos os processos de aprendizagem humanos independentemente do seu objecto; a segunda é a de que, para além das faculdades da linguagem e do senso comum, uma das outras faculdades cuja existência independente Chomsky postula é a faculdade da formação de ciência; finalmente, a terceira é a de que esta postulação é feita com base no mesmo género de considerações empiricamente sustentadas que legitimam a dissociação entre linguagem e pensamento ou, se se preferir, entre faculdade da linguagem e faculdade do senso comum.

Não vou agora entrar em detalhes a respeito destas considerações. Limitar-me-ei a fazer a seguinte sugestão. Se a hipótese de Chomsky acerca da estrutura da mente humana estiver correcta, i.e., se a hipótese de que, para além de distinto da faculdade da linguagem, o pensamento se divide, ainda, ele próprio, entre, pelo menos, pensamento decorrente da faculdade do senso comum e pensamento decorrente da faculdade da formação de ciência, então esta hipótese pode ser consistentemente desenvolvida na seguinte direcção. O princípio da maximização da preservação da verdade constitui o princípio central a partir do qual a faculdade de formação de ciência se constitui, sendo, por isso, um dos elementos essenciais da *diferença específica* que, simultaneamente, a individualiza e a *distingue* da faculdade do senso comum.

Se esta extensão do ponto de vista de Chomsky é aceitável, duas consequências se seguem. A primeira é a de que a interpretação do uso da faculdade da linguagem pelos falantes de linguagens naturais com base nos princípios da maximização da

preservação da verdade subjacentes à semântica formal defendida por Davidson constitui uma dupla violação do princípio da independência funcional das faculdades mentais. Uma violação da independência da faculdade da linguagem em relação a qualquer uma das faculdades do pensamento e uma violação da independência da faculdade do pensamento do senso comum em relação à faculdade de formação de ciência. Usando o modelo chomskyano da mente humana, poderíamos então dizer que Davidson interpreta erradamente o uso da linguagem natural pelos seres humanos como se este fosse um produto do funcionamento da faculdade de formação de ciência. Os princípios de funcionamento desta seriam assim tomados como os princípios essenciais de funcionamento de uma inexistente mente humana unitária. Esta dupla violação do princípio da independência funcional das faculdades constituiria assim a justificação explicativa para o choque observado nas secções anteriores entre aquilo que a evidência empírica sugere e os preceitos determinados pelos princípios da teoria supostamente interpretativa.

A segunda consequência é a seguinte. Existem inúmeras observações de Wittgenstein, de acordo com as quais a linguagem natural não deve ser encarada como uma realização empírica de um sistema dedutivo do género da Lógica de 1ª Ordem com identidade, devendo antes ser compreendida nos seus próprios termos. Estas afirmações, que, como vimos, contradizem directamente o cerne do programa de Davidson e de que um bom exemplo, para além da já mencionada §81 das *P.U.*, se pode encontrar na citação do *Blue Book* que serve de epígrafe a este ensaio, sustentam-se apenas sobre uma análise fisionómica da experiência efectiva de uso da linguagem natural. Mas, se a hipótese de Chomsky e a extensão da mesma que aqui proponho vierem a ser confirmadas pela investigação científica, estas afirmações de Wittgenstein podem igualmente encontrar nelas a sustentação empírica que actualmente lhes falta.

BIBLIOGRAFIA:

- Baddeley, A., 1990, *Human Memory – Theory and Practice*. Hillsdale (NJ), Lawrence Erlbaum Assoc.
- Cherniak, C., 1986, *Minimal Rationality*. Cambridge (MA), The MIT Press.
- Chomsky, N., 1986, *Knowledge of Language – its Nature, Origin and Use*. Westport (CT), Praeger.
- Chomsky, N., 2000, *New Horizons in the Study of Language and Mind*. Cambridge, Cambridge University Press.

- Davidson, D., 1980, *Essays on Actions and Events*. Oxford, Clarendon Press.
- Davidson, D., 1982, "Rational Animals" in *Dialectica*, 36, 4, 317-27.
- Davidson, D., 1984, *Inquiries into Truth & Interpretation*. Oxford, Clarendon Press.
- Davidson, D., 1995, "Could There Be a Science of Rationality" in *International Journal of Philosophical Studies*, 3, 1-16.
- Davidson, D., 1998, "Replies" in *Crítica – Revista Hispanoamericana de Filosofía*, XXX, 90, 97-112.
- Dreyfus, H., 1992, *What Computers Still Can't Do – A Critique of Artificial Reason*. Cambridge (MA), The MIT Press.
- Evans, J. St., 1982, *The Psychology of Deductive Reasoning*. Hillsdale (NJ), Lawrence Erlbaum Assoc.
- Evans, J.St., Newstead, S.E., Byrne, R.M.J., 1993, *Human Reasoning – The Psychology of Deduction*. Hillsdale (NJ), Lawrence Erlbaum Assoc.
- Frege, G., 1883, "Über den Zweck der Begriffsschrift" in *Begriffsschrift und andere Aufsätze*. Hildesheim, G. Olms, 1964.
- Goldman, A., 1986, *Epistemology and Cognition*. Cambridge (MA), Harvard University Press.
- Goldman, A., 1992, "In Defence of the Simulation Theory" in *Mind & Language*, 4, 161-185.
- Gopnik, A. & Wellman, H.M., 1994, "The Theory-theory" in Hirshfield & Gelman (eds.), *Domain-Specificity in Cultural Cognition*. Cambridge, Cambridge University Press.
- Gordon, R.M., 1986, "Folk-Psychology as Simulation" in *Mind & Language*, 1, 158-171.
- Grandy, R., 1973, "Reference, Meaning and Belief" in *The Journal of Philosophy*. LXX, 439-452.
- Hopkins, J., 1999, "Wittgenstein, Davidson and Radical Interpretation" in Hahn, F., (ed.), *The Library of Living Philosophers – Donald Davidson*. Chicago, Chicago University Press.
- Pinker, S., 1994, *The Language Instinct – The New Science of Language and Mind*. Harmondsworth, Penguin.
- Pinker, S., 1995, "Language Acquisition" in Osherson, D.N. (ed.), *An Invitation to Cognitive Science vol. 1 – Language* (ed. by Gleitman, L. and Liberman,

- M.), 135-182. Cambridge (MA), The MIT Press.
- Quine, W.v.O., 1960, *Word and Object*. Cambridge (MA), The MIT Press.
- Quine, W.v.O., 1986, *Philosophy of Logic*. Cambridge (MA), Harvard University Press.
- Wittgenstein, L., 1922, *Tractatus Logico-Philosophicus*. London, Routledge & Kegan Paul.
- Wittgenstein, L., 1952, *Philosophische Untersuchungen. Werkausgabe – Band 1*. Frankfurt a.M., Suhrkamp, 1984.
- Wittgenstein, L., 1958, *The Blue and Brown Books*. Oxford, Blackwell's.
- Wittgenstein, L., 1969, *Philosophische Grammatik. Werkausgabe – Band 4*. Frankfurt a.M., Suhrkamp, 1984.
- Wittgenstein, L., 1979, *Zettel. Werkausgabe – Band 8*. Frankfurt a.M., Suhrkamp, 1984.