

2ª PARTE
MENTE E LINGUAGEM:
AS BASES DO INTERNALISMO LINGUÍSTICO

Capítulo IV

A concepção computacional de *mente*

Nos três capítulos da parte anterior, estive ocupado em esclarecer os termos usados na discussão clássica sobre os nomes próprios e em relatar e comentar as disputas filosóficas que a pontuaram. Examinei algumas noções importantes pressupostas nessas disputas, apresentei algumas teorias de nomeação, tanto descritivistas como causalistas, e apontei alguns problemas e dificuldades que as mesmas enfrentam. A partir desse momento, entretanto, começo a expor as ideias que fundamentarão a minha própria explicação do funcionamento semântico dos nomes próprios. Ela se distinguirá da parte majoritária das explicações clássicas por ser uma explicação mentalista ou internalista. Isso significa que ela parte da premissa de que as propriedades semânticas dos nomes próprios só podem ser esclarecidas se levarmos em conta os fenômenos mentais que ocorrem nos falantes que usam esses nomes. Com base nisso, fica claro que uma das noções mais importantes para a minha explicação é a noção de *mente*, consequentemente, antes de expor minha visão sobre nomes próprios, preciso esclarecer qual a minha concepção de *mente*.

Para ser bem direto, declaro logo de início que a concepção de *mente* que adoto é a concepção computacional. Apesar de minha descrição da estrutura do mental ter suas peculiaridades, meu modo de conceber a mente é basicamente o mesmo dos cientistas cognitivos. Para mim, como para eles, a relação entre mente e cérebro pode ser comparada à relação entre *software* e *hardware*. Em outras palavras, para mim, a mente é um tipo de sistema que é implementado pelo cérebro.

Neste capítulo, vou tentar justificar minha concepção de *mente* mostrando os problemas da concepção anterior – a concepção substancial de *mente* – e evidenciando como a concepção baseada na metáfora do computador se impôs para a maior parte dos pesquisadores da mente a partir da metade do século passado. Ao final, ainda examinarei algumas críticas que Searle faz à abordagem cognitivista da mente e apresentarei argumentos para mostrar que as mesmas não se justificam.

4.1. A mente como uma substância

Quando me lembro de episódios da minha infância, uma cena que ocorreu há muito tempo é de alguma forma reeditada. A cena pertence ao meu passado, mas em certo sentido ela ainda está ali comigo. Como isso acontece? Pode-se dizer que o passado volta a existir? De que forma ele volta? Onde ele está quando me lembro dele?

Às vezes, depois que ouço uma música, fico com ela tocando insistentemente na minha cabeça, mas só eu a ouço. É estranho. A música não se propaga pelo ar em forma de ondas sonoras, ela simplesmente não está no ambiente. Todavia eu a ouço. Como isso é possível? Onde está a música quando só eu a ouço?

Quando leio uma boa história, sempre acontece algo espantoso, os personagens ganham vida para mim. Por exemplo, quando li *Guerra e Paz*, fui apresentado a vários personagens incríveis e apaixonantes, cada um com sua personalidade, com seus desejos, medos, dúvidas. Uns faziam ações notáveis, outros ficavam só nas boas intenções; uns se apaixonavam e sofriam por isso, outros caíam mortos no campo de batalha. Todos eles estavam na história. Mas onde estava a história? No pesado bloco de papel e tinta que eu tinha nas mãos, ou em algum outro lugar?

Qualquer um que se encontre desarmado das sutilezas filosóficas contemporâneas dirá sem nenhum pudor que meu passado, minha música e minha história estavam na minha mente. O que há de peculiar nesse modo de falar é que a mente é retratada como um local. Coisas como lembranças, ideias, pensamentos, raciocínios, fantasias, crenças, dúvidas e outras que tais não podem ser encontradas em lugar nenhum do espaço, mas podem ser localizadas na mente. Essa imagem da mente faz sentido? Muitos filósofos do passado acharam que fazia, tanto que trataram de criar suas próprias metáforas espaciais para explicar o modo como a mente se relaciona com as coisas mentais. Platão, por exemplo, no *Teeteto*, disse que as ideias estão na mente assim como passarinhos estão numa gaiola. E, semelhantemente, muitos outros filósofos trataram a mente como o lugar onde se abrigam ideias e outras coisas do tipo.

Não pretendo relatar toda a história dessas metáforas, antes quero me deter em uma que julgo especialmente influente: a metáfora do suporte. A imagem

que essa metáfora nos oferece da mente é a imagem de uma tela na qual se pintam as ideias ou de um palco no qual se representam os eventos mentais. Em outras palavras, nessa metáfora, a mente é retratada como algo que tem a função de sustentar e ligar os fenômenos mentais de um sujeito.

É interessante observar que, no início da filosofia moderna, os principais divulgadores dessa metáfora costumavam afirmar que a mente é uma substância, e frequentemente usavam o termo “substância” com o sentido de “substrato”. Isso pode ser verificado tanto em Descartes quanto em Locke.

Um substrato é literalmente um suporte. Quando se diz que a mente ou alma é uma substância, e liga-se ao termo “substância” o sentido de “substrato”, quer-se dizer antes de tudo que a mente ou alma é algo que sustenta fenômenos como pensamentos, desejos, ideias etc. Esses fenômenos inerem, vale dizer, estão afixados na mente ou alma.

Descartes e Locke eram dualistas, ou seja, eles acreditavam que há pelo menos dois tipos de substância: substâncias materiais, que funcionam como suporte de fenômenos físicos, e substâncias espirituais, que funcionam como suporte de fenômenos mentais. Nossos corpos são substâncias do primeiro tipo e nossas mentes são substâncias do segundo, e, apesar dessas substâncias possuírem naturezas essencialmente diferentes, elas interagem causalmente. Vejamos mais de perto como Descartes e Locke argumentam em favor do dualismo.

Seria um erro grosseiro pensar que Descartes inventou o dualismo. Bem antes dele, Platão já era dualista, e talvez Sócrates, e Pitágoras antes de todos. O que Descartes tem de novo é o modo de justificar a tese dualista. Sua argumentação em favor do dualismo está apoiada essencialmente no procedimento da dúvida metódica. Efetivamente, o que ele faz é colocar à prova todas as suas crenças, pondo em dúvida tudo o que for dubitável. Fazendo assim, ele descobre que há pelo menos uma verdade da qual ele não pode duvidar, a saber, a de que ele pensa e existe enquanto coisa pensante. Essa verdade não pode ser posta em dúvida porque, se uma pessoa duvidasse de que pensa, no exato momento em que a dúvida se apresentasse, ela já estaria pensando. Temos da existência de uma coisa pensante uma ideia clara e distinta, e isso é o mesmo que dizer que tal existência é evidente.

Em contrapartida, a ideia que temos dos corpos físicos não é clara e distinta, o que significa que é possível que ela seja falsa. É possível, por exemplo,

que todas as coisas que eu vejo e chamo de físicas sejam apenas ilusões e não tenham nenhuma existência fora da minha mente. Segundo Descartes só podemos saber que as coisas físicas têm uma existência real porque Deus, por ser veraz, garante que o que vemos no mundo não são ilusões. Dessa forma, podemos afirmar que existem coisas físicas fora de nossas mentes. Para Descartes, o que caracteriza essencialmente essas coisas é o fato de elas serem extensas, ou seja, o fato de ocuparem lugar no espaço.

Em todo caso, apesar das coisas físicas serem reais, permanece uma diferença essencial entre a existência delas e a existência de nossa mente. A existência de nossa mente é indubitável, enquanto que a existência das coisas físicas pode ser posta em dúvida pelo menos em um primeiro momento. Desse modo, fica claro que a mente não pode ser uma coisa física. Deve-se admitir, portanto, que os fenômenos físicos e os fenômenos mentais são suportados por substâncias de tipos diferentes. Vê-se assim que o dualismo é uma consequência inevitável da aplicação da dúvida metódica.

Como foi dito acima, para Descartes, a característica essencial das coisas físicas é a extensão, por isso mesmo ele chama a substância material de *res extensa*, isto é, coisa extensa. A característica essencial da mente, por outro lado, é pensar (ter ideias), o que leva Descartes a chamá-la de *res cogitans*, isto é, coisa pensante. Enquanto coisa não física, a mente não tem extensão, em todo caso, isso não impede Descartes de tratar a mente como o lugar onde as ideias residem. De fato, em várias passagens ele afirma que as ideias estão na mente. No parágrafo 15 da terceira meditação, por exemplo, ele afirma que as ideias estão em nós como figuras ou imagens. Esse tipo de metáfora espacial se encontra com frequência nas obras de Descartes. Nessas passagens, fica claro que a substância espiritual, embora inextensa, funciona como o substrato que suporta nossas ideias. Para Descartes a mente ou alma é como um livro e as ideias são como os caracteres.

Apesar de toda a argumentação de Descartes, sua concepção de *mente* suscitava alguns problemas bem sérios. O maior deles, sem dúvida, é o problema mente-corpo, isto é, o problema de explicar como mente e corpo interagem. Se a mente é algo não-físico, fica difícil explicar como ela causa efeitos físicos (por exemplo, quando levanto o braço após decidir levantá-lo) e sofre a ação de causas físicas (por exemplo, quando sinto dor após levar uma topada). Como é possível que substâncias essencialmente diferentes mantenham relações causais entre si?

Descartes tentou responder essa questão fazendo certas especulações sobre o funcionamento da glândula pineal, mas, pelo que se vê em sua correspondência com a princesa Elizabeth da Boêmia, nem mesmo ele estava muito convencido dessa resposta. As gerações posteriores de filósofos se convenceram menos ainda, e o problema mente-corpo permaneceu problemático até hoje.

Locke, como Descartes, também admitia a existência de dois tipos de substância e claramente usava “substância” no sentido de “substrato”¹. Eis o que ele escreve no *Ensaio Acerca do Entendimento Humano*:

Nós temos uma noção tão clara de substância espiritual como nós temos de corpo; uma nós supomos ser (sem saber o que ela é) o *substratum* daquelas ideias simples que temos de fora; e a outra supomos (igualmente ignorando o que ela é) ser o *substratum* para aquelas operações que nós experienciamos dentro de nós mesmos (LOCKE, 1973: 213).

Nessa passagem fica clara a metáfora do suporte. E, inegavelmente, todo *Ensaio* é pleno de metáforas espaciais da mesma cepa. A conhecida caracterização da mente do recém-nascido como uma folha em branco (*tabula rasa*) é provavelmente a metáfora mais representativa de todas. A partir dela, Locke apresenta a tese básica de seu empirismo: tudo o que temos na mente adquirimos através da experiência, seja ela experiência sensível ou reflexiva. Quando vemos um objeto, ouvimos um som, provamos um sabor ou temos qualquer experiência do tipo, uma ideia consentânea se forma na nossa mente e fica lá vivamente impressa por algum tempo. Depois, a ideia tende a se desvanecer, mas, se não se apagar completamente, sempre poderá ser avivada pela faculdade da memória. De qualquer forma, é na mente e só na mente que as ideias existem. A mente assume assim para Locke a função precípua de repositório de ideias. Não obstante, ela não é um repositório que simplesmente recebe as ideias e as acolhe passivamente. A mente tem faculdades ou poderes (dentre as quais estão o entendimento e a vontade) que lhe permitem fazer várias operações com as ideias. Pensar, perceber, recordar, imaginar e desejar são algumas dessas operações. É a essas operações que Locke se refere na passagem supracitada, é delas que a mente é considerada um *substratum*.

¹ Deve-se notar que, apesar de defender a existência de uma substância espiritual, para Locke, a identidade pessoal de um indivíduo não é garantida por essa substância, e sim por sua atividade reflexiva, em particular, por sua memória (cf. LOCKE, 1973: livro II, cap. XXVII).

O uso da metáfora do suporte é o que aproxima Locke de Descartes. É isso que está na origem do dualismo de ambos, e é isso que os conduz ao problema mente-corpo. Ao refletir sobre esse problema, porém, Locke toma um rumo diferente do de Descartes. Para ele, o problema não tem solução. Pelo menos não tem nenhuma solução científica. Segundo Locke, tudo o que podemos dizer a respeito do modo como a substância espiritual se relaciona com a substância material é que, assim como Deus escolheu anexar ao movimento efeitos que nós não podemos conceber como são produzidos, Ele pode ter escolhido adicionar à matéria uma faculdade de pensar (citado em CHOMSKY, 2000: 1). Em outras palavras, a mente trabalha de comum acordo com o corpo porque Deus determinou que fosse assim. Na opinião de Locke, nenhuma outra explicação é possível.

Como essas rápidas considerações demonstram, no início da filosofia moderna, a metáfora do suporte foi empregada reiteradamente tanto nas teorias da mente de orientação racionalista como nas de cunho empirista. Sua grande vantagem é que, em certa medida, ela se ajustava à intuição popular de que a mente é um tipo de lugar. Uma vez que a mente era vista como substrato de eventos mentais, tornava-se possível dizer que esses eventos tinham uma localização, eles ocorriam na mente. É possível descrever as coisas dessa maneira mesmo quando se fala de uma substância espiritual sem extensão (se bem que assim a descrição perde grande parte da sua literalidade). Por causa disso, é possível continuar falando da mente da forma natural como falamos cotidianamente. Eu continuo podendo dizer que a cena que eu lembro, que a música que eu ouço e que a história que eu afiguro estão na minha mente. Essas coisas estão na minha mente porque minha mente é a substância que as congrega e suporta. Tudo isso realmente parece tornar o emprego da metáfora do suporte muito vantajoso para uma explicação dos fenômenos mentais. Todavia, ao longo do tempo, essas vantagens se mostraram insuficientes e a ideia de uma substância espiritual que serve de suporte para o mental perdeu a sua credibilidade.

Vários fatores ajudaram a desacreditar essa ideia. A intratabilidade do problema mente-corpo quando se adota o dualismo de substância foi um deles. A crítica de Hume à noção de *substância* também teve o seu peso. Mas o fator decisivo, na minha opinião, foi uma mudança na visão científica do mundo. Nos tempos de Descartes e Locke, ainda havia lugar para substâncias espirituais dentro

dessa visão, mas logo depois isso mudou drasticamente. Para mim, a noção de *substância espiritual* começou a ser mal vista porque ela parecia sugerir que a mente é imortal. De fato, os argumentos cartesianos mostravam que a substância espiritual tem uma existência independente da substância material. Assim, não há nenhuma razão para que a alma pereça junto com o corpo. Se isso não dá certeza da imortalidade da alma, pelo menos dá uma esperança. De todo modo, quando começou a faltar espaço na visão científica do mundo para a crença na imortalidade da alma, a noção de substância espiritual também perdeu seu lugar.

Foi assim que a metáfora do suporte entrou em processo de falência. Sem ela, novas metáforas se fizeram necessárias para a investigação do mental. Uma metáfora adequada, no entanto, só se tornou disponível no século XX, com o advento do computador.

4.2. A mente como um programa de computador

Em meados do século passado, uma nova metáfora explicativa da mente surgiu e, junto com ela, uma nova ciência. A nova metáfora era a metáfora do computador e a nova ciência era a ciência cognitiva. A metáfora consistia na afirmação de que a mente é um tipo de programa que, em vez de rodar em um computador eletrônico, roda no cérebro. Foi nessa metáfora que os cientistas cognitivos basearam todo o seu programa de investigação dos fenômenos mentais.

Entre os vários fatores que contribuíram para o aparecimento da metáfora do computador, dois foram cruciais. O primeiro e mais óbvio foi a invenção do computador digital. O segundo foi a necessidade de explicar comportamentos humanos complexos, comportamentos que são realizados em passos serialmente ordenados, como tocar piano, por exemplo. A combinação desses dois fatores possibilitou o que ficou conhecido como “a revolução cognitiva”. Vou tentar detalhar o papel que cada fator desempenhou nessa revolução.

Logo que o computador surgiu, vários pensadores tentaram mostrar paralelos entre o seu funcionamento e o funcionamento do cérebro e da mente. John Von Neuman, Warren McCulloch, Norbert Wiener, Alan Turing, Allen Newell, Herbert Simon e outros participaram dessa empreitada. Alguns deles tentaram humanizar o computador, e outros, por assim dizer, tentaram

computadorizar o cérebro. Pelo menos uma das observações feitas pelo grupo dos humanizadores precisa ser mencionada aqui.

A observação se refere às capacidades cognitivas dos computadores e foi originalmente feita por Turing. Em seu célebre artigo *Computing Machinery And Intelligence*, de 1950, ele propôs a questão sobre se computadores podem pensar. Logo de início, ele identificou algumas dificuldades para se responder essa questão e resolveu substituí-la por outra que considerou estreitamente relacionada com ela. A nova questão seria a seguinte: um computador poderia jogar o jogo de imitação? Para entendermos essa questão, precisamos esclarecer o que é o jogo de imitação ².

Conforme a descrição de Turing, o jogo é jogado por três pessoas: um homem (A), uma mulher (B) e um interrogador (C). Primeiramente, C é isolado em uma sala e, utilizando um *teleprinter* ³, tem de fazer perguntas a A e B com o intuito de descobrir qual é o homem e qual a mulher. Acontece que A deve responder como se fosse mulher e B deve responder normalmente. Depois de algumas perguntas, o homem sai do jogo e entra uma máquina que também finge ser a mulher. E o entrevistador continua a fazer suas perguntas. O que Turing afirma então é que, se a máquina se sair pelo menos tão bem quanto o homem na tarefa de enganar seu interrogador, ela terá passado no teste. Ele estima ainda que, no final do século XX, o uso das palavras e a opinião das pessoas esclarecidas estariam tão mudados que seria possível afirmar que uma máquina assim pensa.

O artigo de Turing se tornou um dos textos seminais da *Inteligência Artificial*. A principal lição que os representantes da IA tiraram do artigo foi esta: “se vocês querem criar uma máquina inteligente, basta criar uma máquina que faça alguma coisa inteligente, como provar teoremas lógicos, jogar xadrez ou jogar o jogo de imitação”. Com efeito, foi isso que Allen Newell e Herbert Simon, dois pioneiros da IA, fizeram. Eles foram os primeiros que efetivamente tentaram desenvolver programas inteligentes. E o primeiro desses programas foi o *Logic Theorist LT*, o primeiro provador automático de teoremas. Ao que parece, Newell e Simon ficaram tão impressionados com a “inteligência” do LT que mandaram

² O jogo de imitação é o que depois ficou conhecido como “Teste de Turing”.

³ Instrumento da época usado para mandar e receber mensagens via teclado e folhas impressas.

um artigo relatando sua descoberta para o *Journal of Symbolic Logic* e incluíram LT como co-autor. Infelizmente, essa inclusão impediu que o artigo fosse aceito.

O que é importante destacar aqui dos trabalhos de Turing, Newell e Simon é a contribuição que eles deram para a consolidação da metáfora do computador. Foi a partir das ideias e dos resultados que eles demonstraram que se disseminou a crença de que poderíamos entender melhor a mente humana se prestássemos atenção em certas características do computador. Afinal, é com base nessas características que podemos explicar o comportamento inteligente de um computador. Examinando-as, poderíamos explicar também o nosso próprio comportamento inteligente.

Uma das características mais importantes é sem dúvida a de que o computador tem dois níveis, um nível lógico-representacional, o nível do *software*, e um nível físico, o nível do *hardware*. O nível lógico-representacional é composto das operações abstratas que são feitas quando um programa é executado e dos elementos simbólicos que são usados nessas operações. Os programadores são os responsáveis por criar o nível lógico-representacional. Já o nível físico é composto do equipamento concreto, com seus dispositivos eletrônicos, elétricos, mecânicos etc. Os engenheiros da computação são responsáveis por criar o nível físico.

Uma outra característica interessante do computador é a de que o nível lógico-representacional é realizável em qualquer *hardware* que tenha a capacidade computacional da máquina de Turing ⁴. Dessa forma, fica evidente que os dois níveis são logicamente independentes e, por conseguinte, o nível dos programas pode ser descrito sem nenhuma referência ao nível físico. Foi Putnam que chamou a atenção para esse fato (cf. GARDNER, 2003: 46).

Uma vez que tais características fundamentais eram encontradas nos computadores, a metáfora adotada pelos cognitivistas lhes permitia transportá-las para o ser humano. Com efeito, Howard Gardner mostra isso em *A Nova Ciência da Mente*. De acordo com ele, quando tais características do computador foram observadas:

⁴ A máquina de Turing é uma máquina teórica que, mesmo executando apenas operações bem simples, é capaz de executar qualquer algoritmo.

A analogia com o sistema humano e os processos humanos de pensamento era patente. O cérebro humano (ou “estados corpóreos”) correspondia ao *hardware* computacional; padrões de pensamento ou solução de problemas (estados mentais) podiam ser descritos de forma totalmente independente da constituição específica do sistema nervoso humano. Além disto, os seres humanos, não menos que os computadores, armazenam programas; e a mesma linguagem simbólica podia ser invocada para descrever programas de ambas as entidades (Idem, ibidem).

Em outras palavras, no humano, a mente faz o papel do *software* e o cérebro faz o papel do *hardware*, e assim como o *software* pode ser descrito sem referência ao *hardware*, é possível em princípio enfocar e descrever as operações da mente sem recorrer a uma descrição dos fenômenos neurofisiológicos do cérebro.

De fato, essa possibilidade de descrição independente é vista como uma necessidade teórica. É preciso que o cientista cognitivo possa levantar hipóteses sobre as operações mentais envolvidas na realização de determinada tarefa ou comportamento mesmo se não tiver nenhuma evidência de como o cérebro trabalha em tais ocasiões. É claro que o pesquisador da mente não está proibido de buscar e usar evidência fisiológica, não há nenhuma restrição metodológica a isso. O ponto é que ele não depende disso essencialmente. De qualquer forma, o mais comum é que as descobertas no nível do processamento orientem a pesquisa no nível da neurofisiologia. Uma vez que certas operações abstratas tenham sido identificadas na mente, é tarefa do neurocientista investigar quais fenômenos neurofisiológicos são responsáveis por aquelas operações.

A descrição das operações também não é restringida por protocolos excessivamente rígidos. Ela pode ser feita tanto em linguagem natural como em linguagem de programação, na forma de um programa de computador. No primeiro caso, o pesquisador conta com um vocabulário muito mais rico do que o antigo vocabulário behaviorista. Agora é possível até usar termos típicos do jargão mentalista clássico, termos como “crença”, “desejo”, “expectativa”, “ideia” etc. No segundo caso, há a vantagem de o programa elaborado poder ser efetivamente executado em um computador. Uma vez executado, os resultados da simulação poderão ser estudados e comparados com o caso humano. Tempos de resposta, padrões de reação a variações nos parâmetros do programa e outras medidas são evidências que podem ser usadas para corroborar a hipótese de que as operações executadas pela mente são mesmo aquelas descritas. E nunca é demais lembrar

que todo esse trabalho de elaboração e teste de hipóteses só se tornou possível graças à metáfora do computador.

De qualquer forma, a metáfora foi uma condição necessária mas não suficiente para a revolução cognitiva. A outra condição, como antes já havia mencionado, foi a necessidade de explicar de uma forma verossímil o comportamento humano serialmente ordenado, em especial o comportamento linguístico. Se os psicólogos cognitivistas não tivessem sentido essa necessidade, talvez a metáfora do computador não tivesse despertado tanto interesse.

Mas porque eles sentiram essa necessidade? Obviamente, porque as explicações que antes eram dadas para o comportamento serialmente ordenado começaram a ser consideradas insatisfatórias. Como essas explicações, em geral, eram fornecidas por psicólogos behavioristas, somos obrigados a examinar as linhas gerais do behaviorismo para entendermos porque elas foram consideradas insatisfatórias.

Pode-se dizer que o programa behaviorista foi proposto por John Watson com o propósito de tornar a psicologia realmente científica. Até então o principal método de investigação usado pelos psicólogos era a introspecção. Um investigador treinado, ao realizar alguma tarefa simples num ambiente controlado, prestava a máxima atenção aos seus próprios processos mentais e em seguida os relatava da forma mais objetiva e minuciosa possível. Em termos simples, isso era a introspecção. Esse método foi idealizado por Wilhelm Wundt, um dos fundadores da psicologia moderna. Na opinião de Watson, porém, a introspecção não era um método científico. Para ele, os resultados de uma verdadeira ciência não poderiam se basear em relatos de primeira pessoa. Por mais rigoroso que seja o investigador, seu relato sempre será intrinsecamente subjetivo. Além disso, a introspecção só pode ser aplicada a seres humanos, o que necessariamente deixa todos os outros animais fora do alcance da investigação psicológica. Para Watson, isso não era aceitável. Mas, acima de tudo, era inaceitável para ele que eventos mentais e conscientes fossem tomados como o objeto da psicologia. Na sua perspectiva, esses eventos não podem ser objeto de investigação científica porque não podem ser apreendidos por métodos públicos de observação. Para que a psicologia se tornasse uma ciência respeitável era preciso que ela adotasse os métodos de investigação das ciências naturais e que se limitasse a estudar eventos alcançáveis por esses métodos. Nomeadamente, ela deveria se limitar a estudar o

comportamento e suas causas ambientais. É nesse sentido que Watson define a psicologia como a ciência do comportamento, em oposição à definição anterior segundo a qual a psicologia era a ciência da mente. A tarefa do psicólogo behaviorista seria descobrir as leis que correlacionam estímulos ambientais a comportamentos específicos. Uma vez descobertas essas leis, o psicólogo, como qualquer cientista das ciências duras, poderá prever e até determinar comportamentos. Certamente, se ele puder controlar certas variáveis ambientais, também poderá controlar o comportamento. Dessa forma, o *modus operandi* da psicologia se tornava semelhante ao de todas as ciências naturais. A visão que Watson tinha dessa ciência do comportamento pode ser resumida nesta sua conhecida declaração: “A psicologia, tal como o behaviorista a vê, é uma ciência natural puramente objetiva. Sua meta teórica é a previsão e o controle do comportamento” (Idem, *ibidem*: 123).

A argumentação de Watson teve um impacto surpreendente sobre a comunidade científica e, em pouco tempo, o behaviorismo se tornou uma corrente de pesquisa psicológica extremamente influente, chegando a dominar o cenário da psicologia norte-americana por décadas. No entanto, seu domínio começou a declinar quando alguns pesquisadores começaram a se sentir paralisados pelas suas restrições metodológicas.

A principal dessas restrições era a de que o psicólogo não poderia recorrer a eventos internos do indivíduo para explicar o seu comportamento. De fato, para alguns behavioristas, como Skinner, por exemplo, o psicólogo não deveria nem mesmo mencionar termos mentalistas como crença, desejo, pensamento, intenção etc. Era sempre desejável que esses termos fossem substituídos por termos ‘científicos’, o que Skinner também ensinou a fazer.

Essa restrição se tornou intolerável porque em certo momento ficou evidente que a única forma de explicar o comportamento serialmente ordenado era exatamente recorrendo aos eventos mentais dos indivíduos que apresentam o comportamento. Isso ficou muito claro, por exemplo, para um ex-aluno e ex-cooperador de Watson chamado Karl Lashley. Como Gardner relata:

Lashley percebeu que antes que novos *insights* sobre o cérebro, ou sobre os computadores, pudessem ser aplicados nas ciências psicológicas, seria necessário confrontar o behaviorismo diretamente. Por conseguinte, nas suas observações iniciais, Lashley expressou sua convicção de que qualquer

teoria da atividade humana teria de dar conta de comportamentos complexamente organizados como jogar tênis, tocar um instrumento musical, e – acima de tudo – falar (Idem, *ibidem*: 27).

Lashley apresentou sua crítica ao modelo behaviorista de explicação do comportamento serialmente ordenado em um artigo intitulado *O Problema da Ordem Serial no Comportamento* (1951) ⁵. Ele começa explicando que comportamento serialmente ordenado é qualquer comportamento realizado em uma sequência de passos ordenados. Tocar flauta, jogar tênis, amarrar o cadarço dos sapatos e escrever um texto são comportamentos desse tipo. Virar a cabeça, apertar a campainha e dar petelecos na orelha de alguém normalmente não são. Os behavioristas explicavam o comportamento serialmente ordenado com base na ideia de cadeia associacionista. Uma cadeia associacionista é uma cadeia de estímulos e respostas; nela, cada elo é a resposta para o elo anterior e o estímulo para o elo seguinte. O reforço e a repetição condicionariam as pessoas a fazerem as associações e a manifestarem o comportamento.

Lashley se baseou em uma série de estudos de caso para demonstrar a implausibilidade da explicação behaviorista e indicar um caminho diferente. Ele mostrou, por exemplo, que quando um pianista toca um arpejo o movimento que ele faz é literalmente mais rápido do que seria necessário para os estímulos elétricos irem dos dedos ao cérebro e voltarem para servirem de estímulo para os movimentos seguintes. A evidência fisiológica sugeria que a sequência completa do arpejo era determinada por um tipo de ideia ou intenção na cabeça do pianista. O arpejo era apenas a realização concreta dessa ideia abstrata. Casos de lapsos de escrita e de interpretação fonética também indicariam a incorreção da explicação behaviorista. Esses casos mostrariam duas coisas. Primeiro: quando uma pessoa escreve uma palavra, a ordem das letras é dada por um tipo de esquema de ordem existente na sua mente. A escrita da palavra é simplesmente o resultado do encaixe das letras no esquema preestabelecido. Vários lapsos de escrita comuns são o resultado do encaixe errado das letras; o esquema de ordem, no entanto, é preservado. Segundo: todas as partes da sequência de um comportamento ordenado estão em certo sentido disponíveis antes de serem usadas. É como se a sequência completa fosse antecipada (mesmo que inconscientemente) antes da execução.

⁵ Lashley já havia apresentado o artigo em 1948, no famoso Simpósio Hixon.

Lashley ainda deu vários outros exemplos e todos eles apontavam para a mesma conclusão, a saber, a organização do comportamento serialmente ordenado não é determinada por uma cadeia de estímulos e respostas, ela é determinada antes por uma hierarquia de processos psicológicos internos. O indivíduo pode utilizar *input* ambiental nesses casos, mas toda a forma organizacional é orquestrada internamente.

Como se vê, essa conclusão era absolutamente desastrosa para o behaviorismo, e, de fato, o artigo de Lashley foi um golpe contundente na ciência concebida por Watson. Entretanto, o golpe que realmente derrubou o gigante foi dado pelo linguista Noam Chomsky.

Como Lashley, Chomsky estava interessado no comportamento serialmente ordenado, mais especificamente, no comportamento linguístico. E, da mesma forma que Lashley, ele percebeu que a explicação behaviorista para a linguagem estava longe de ser convincente. Ele percebeu que a única explicação razoável para a linguagem tinha que ser uma explicação internalista ou mentalista, e foi uma explicação assim que ele formulou. Não foi a primeira do tipo, pois Locke já tinha uma teoria mentalista da linguagem, e mesmo mais recentemente, alguns linguistas, como Ferdinand Saussure e Zellig Harris (que foi professor de Chomsky), tinham teorias internalistas. Não obstante, Chomsky foi o primeiro que efetivamente mostrou que a nossa capacidade de desenvolver e usar linguagem é baseada em procedimentos algorítmicos que são implementados nos nossos cérebros.

Neste capítulo, não vou dar detalhes nem da crítica de Chomsky ao behaviorismo nem da sua explicação para a linguagem. Dada a importância de Chomsky para a história do internalismo linguístico, vou dedicar o próximo capítulo inteiramente ao exame de suas ideias. No momento, quero apenas mencionar a inestimável contribuição que suas ideias deram para a consolidação do ponto de vista cognitivista. Chomsky contribuiu, por um lado, ao combater o behaviorismo, ao denunciar as deficiências de seus métodos e fundamentos, e, por outro lado, ao propor uma nova via para o estudo da linguagem, uma via que recolocava a mente e suas estruturas constitutivas no centro das preocupações da pesquisa linguística. Com isso, ele ajudou não só o campo da linguística a se desenvolver mais, mas a ciência cognitiva como um todo.

Com a derrocada do behaviorismo, caía a última grande barreira para o avanço da ciência cognitiva. Doravante, o estudo da mente estava liberado. Os pesquisadores ganhavam uma licença para olhar dentro da caixa-preta da mente e descobrir como ela funciona. Até mesmo a fala mentalista, que antes tinha o estigma de ser anticientífica, agora podia ser usada sem embaraço. Essa suspensão da censura behaviorista, junto com o aumento da confiança na metáfora do computador, garantiu o triunfo final da revolução cognitiva.

Uma questão que se pode fazer ainda é a seguinte: por que a metáfora do computador é melhor do que a metáfora do suporte? Bem, uma vantagem inegável da metáfora do computador é o fato dela não sugerir que a mente é uma substância ou um lugar que serve de palco para os eventos mentais. Com isso, evita-se em grande parte os problemas a respeito da existência independente da mente. Efetivamente, como vimos no final da seção anterior, um dos principais motivos do descrédito da concepção substancial da mente é que, em última instância, ela implicava que a mente podia existir independentemente do corpo. Essa conclusão era inevitável desde que a mente era vista como uma substância autônoma. A partir do momento em que a crença na imortalidade da alma passou a ser encarada como não-científica, a concepção substancial da mente ficou desacreditada. A metáfora do computador, por outro lado, não tem nenhuma implicação semelhante. Não faz sentido pensar que um programa de computador pode rodar sem um computador, por isso, quando comparamos a mente com um programa, evitamos o problema da existência independente da mente. É verdade que um programa pode ser implementado em diferentes *hardwares*, mas não pode ser implementado absolutamente fora de qualquer *hardware*.

Outra aparente vantagem da metáfora do computador, que também está ligada com a primeira, é a de tornar o problema mente-corpo menos assustador. De fato, muitos filósofos continuam falando do problema mente-corpo e realmente não se pode dizer de forma alguma que ele esteja resolvido. Entretanto, ele não tem sido um empecilho para a pesquisa empírica efetiva sobre a mente. Aparentemente, a metáfora do computador faz os cientistas cognitivos se sentirem menos incomodados com o problema mente-corpo. Computadores também têm dois níveis de descrição, o nível do *software* e o nível do *hardware*, mas ninguém acha misterioso o modo como os dois níveis se relacionam. Simplesmente não há um problema *software-hardware*. As pessoas sabem como

esses dois níveis interagem. Dado o paralelo entre cérebro e computador, parece natural pensar que a interação entre mente e corpo também não tem nada de extraordinário. O que acontece é que ainda não entendemos bem o mecanismo que interpreta os sinais do cérebro e os transforma em experiências conscientes, mas muitos cientistas acreditam que entender esse mecanismo é uma questão de tempo. Enquanto isso, a pesquisa empírica corre célere sem precisar mostrar nenhuma consideração especial pelo problema mente-corpo. Mesmo que ainda não possamos explicar exatamente como o nível mental interage com o nível cerebral, essa ignorância não interfere em nada na nossa capacidade de descrever as operações mentais que um indivíduo faz quando realiza uma determinada tarefa. Pelo que se vê na prática, essa capacidade descritiva é o que mais importa para os cientistas cognitivistas. É somente disso que eles precisam para continuar elaborando teorias, testando hipóteses e fazendo o conhecimento avançar. O problema mente-corpo é remetido aos filósofos.

Por fim, é preciso dizer que, a metáfora do computador não explica tudo, e ainda há muita gente que acha que a mente não deveria ser tratada como um sistema de programas rodando no cérebro (o próprio Lashley era reticente quanto às similaridades entre cérebro e computador). De qualquer forma, até aqui a metáfora do computador tem se mostrado heurísticamente muito mais fértil do que a metáfora do suporte, possuindo a vantagem de evitar ou pelo menos minimizar certos problemas que a metáfora anterior suscitava. Isso parece suficiente para justificar sua adoção.

Com isso, concluo minha exposição sobre a revolução cognitiva e sobre a história da metáfora do computador. Resta agora esclarecer de que modo me posiciono em relação às ideias cognitivistas.

Devo dizer que concordo com as linhas gerais da abordagem cognitivista da mente. Concebo a mente como um sistema de computação e representação, ou seja, para mim, a mente é composta de programas em execução (processos) e dos códigos desses programas (sistema de representação). Aqui vai uma exposição rápida e panorâmica da minha concepção de mente:

Os processos mentais se resolvem em operações tais como leitura, escrita e busca de informação ⁶, inferência, associação, recordação etc. Qualquer

⁶ No capítulo 6, vou definir o termo “informação”. No momento, a palavra pode ser entendida em seu sentido coloquial.

ação humana que não é um simples automatismo fisiológico está baseada em operações mentais. Isso significa que essas operações devem ser vistas como os passos de um programa em execução que dá como resultado a tal ação. O conjunto dessas operações constitui o que eu chamo de “nível processual da mente”. O sistema de representação mental, por outro lado, é a linguagem da mente ⁷. Como qualquer linguagem, ele é gerado a partir de um alfabeto e de uma gramática. O alfabeto dessa linguagem é constituído pelos *qualia*. Nesse sentido, um *quale* deve ser visto como um tipo de símbolo mental. Sensações de cores, de sons, de gostos, de dor, de prazer e até de coceiras são *qualia*. As regras da gramática da linguagem mental não são regras no sentido de expressões ou definições normativas, o que significa que elas não se encontram escritas em lugar nenhum. Elas devem ser vistas antes como leis naturais, mais especificamente como leis biológicas. Contudo, se as regras da gramática mental tivessem de ser expressas em palavras, elas diriam coisas como “toda forma deve ser preenchida com uma cor” ou “nenhum som abaixo da frequência x deve ser representado”. Quando os *qualia* são organizados de acordo como a gramática da linguagem mental o resultado são as representações. Toda essa linguagem mental é inata no sentido de que não aprendemos nem o alfabeto nem a gramática da mente. A maneira como produzimos representações é resultado direto do que somos essencialmente, de nossas capacidades intrínsecas, e não do que aprendemos. Com efeito, eu nunca aprendi a ver azul, embora eu veja, e nunca ninguém precisou me ensinar que dores são desagradáveis, embora eu tenha essa sensação. Por fim, gostaria de chamar o conjunto das representações efetivamente produzidas de “nível representacional da mente” ⁸.

Note-se que a distinção entre o nível processual e o nível representacional da mente é uma distinção secundária. A distinção primária é a distinção entre mente e cérebro. A distinção secundária é uma subdivisão dentro da esfera do mental. Deve-se notar, além disso, que o nível representacional da mente requer uma análise muito mais minuciosa e possivelmente mais controversa

⁷ O que eu chamo de “linguagem da mente” tem uma grande proximidade com o que Fodor chama de “linguagem do pensamento”.

⁸ Não estou bem certo, mas creio que minha distinção entre nível processual e nível representacional corresponde à distinção que Ray Jackendoff faz entre “mente computacional” e “mente fenomenológica” (cf. SEARLE, 2006: 218).

do que o nível processual. Muitas coisas a respeito do nível representacional não são claras para mim, e é possível que em algum momento possa vir a revisar algumas afirmações que fiz no parágrafo anterior.

Para quem ainda não observou, é preciso mencionar que o nível representacional não é nada mais nada menos do que o nível da consciência. Não é de se admirar, portanto, que a análise desse nível seja tão árdua. A consciência é reconhecidamente um tema árduo. No meu modo de entender, o maior problema em analisar a consciência é que a metáfora do computador não nos ajuda muito com ela. Podemos dizer que a consciência é a linguagem de programação da mente e, nesse caso, estaremos usando a metáfora, mas isso não nos ajuda a resolver os problemas mais importantes sobre a consciência. Alguns desses problemas podem ser expressos em indagações como estas: Como devemos analisar a ontologia de primeira pessoa da consciência? Como podemos comprovar a eficácia causal dos *qualia*? No final das contas, a experiência consciente é acessível a quem (ou a o quê)? Esta última indagação, para mim, expressa o problema mais sério de todos, o problema do sujeito da experiência consciente. Daniel Dennett apresenta esse problema na introdução de *The Mind's I* e responde o seguinte:

Dizer que é acessível a algum subsistema do cérebro não é distingui-la das atividades e eventos inconscientes, que são também acessíveis a vários subsistemas do cérebro. Se algum subsistema particular e especial é constituído de tal forma que seu tráfico com o resto dos sistemas faz com que haja mais um *eu* no mundo, uma “coisa que se sente como sendo algo”⁹, isso está longe de ser óbvio (DENNETT & HOFSTADTER, 1985: 13).

A metáfora do computador não pode nos ajudar a solucionar esses problemas precisamente porque os problemas são sobre a natureza da linguagem do programa e não sobre os processos do programa. Quando se trata de analisar os processos, não há problema. Nós podemos descrever ou simular os processos mentais usando uma linguagem natural ou mesmo uma linguagem de programação. Na simulação nós temos os mesmos processos vestidos com outra linguagem. E com isso não perdemos nada, pois o que nos interessa nesse caso é o algoritmo que a mente executa e não a linguagem em que ela o executa. Em contrapartida, quando se trata de analisar a linguagem da mente, não há

⁹ No original, “thing it is like something to be”.

possibilidade lógica de simulação. Não existe isso de simular uma linguagem. Nós podemos traduzir uma linguagem, explicar sua gramática, dar exemplos de construções linguísticas e coisas do tipo, e de fato tudo isso nós podemos fazer com a linguagem da mente. A questão é que nada disso nos ajuda a resolver os problemas que enunciei acima. Para resolvê-los precisaríamos entender o que efetivamente causa as representações mentais. Precisaríamos de algo como uma física de nosso mundo interior, ou melhor ainda, precisaríamos de uma física que unificasse a física do mundo exterior e a física do mundo interior. Particularmente, não vejo nenhum indício de como essa ciência seria possível.

De qualquer forma, estou fazendo esses comentários apenas para esclarecer um pouco melhor minha visão a respeito da mente. A verdade é que, nos capítulos seguintes, principalmente nos dois últimos, não vou precisar resolver nenhum problema do âmbito da representação. Vou precisar efetivamente descrever algumas operações mentais, mas isso é tudo. Ou seja, doravante vou me limitar a análises do nível processual da mente. Nisso sigo a ortodoxia da ciência cognitiva. Normalmente, os cientistas cognitivos concentram seus estudos no nível processual da mente e é isso que vou fazer também. Não faço isso por causa de um “pavor da consciência” (para usar uma expressão de Searle). Não tenho esse pavor; pelo contrário, tenho um grande interesse pelas questões relativas à consciência. Faço isso simplesmente porque, para os propósitos que tenho em mente, realmente não preciso falar nada sobre o nível representacional. O que eu quero é explicitar o que acontece na mente dos falantes quando eles usam nomes próprios e, para isso, só preciso descrever as operações mentais que são feitas nessas ocasiões.

Antes de seguir para os capítulos em que essa explicitação será feita, porém, preciso examinar algumas objeções que Searle faz à concepção cognitivista da mente. Ele as apresenta nos capítulos 9 e 10 de seu livro *A Redescoberta da Mente*. Vamos a elas.

4.3. As objeções de Searle

De acordo com a minha interpretação, em *A Redescoberta da Mente*, Searle apresenta duas críticas principais contra a ciência cognitiva. Uma contra a sua fundamentação, que ele expõe no capítulo intitulado “A crítica da razão

cognitiva”, e uma contra um ponto de sua metodologia, que ele expõe no capítulo intitulado “O estudo propriamente dito”. Suas críticas são pormenorizadas e uma discussão completa delas exigiria uma análise igualmente minuciosa de cada ponto. Meu objetivo aqui, porém, é apenas examinar os argumentos mais gerais e indicar os seus pontos falhos. Como tentarei mostrar, toda a discordância de Searle parece ser motivada pelo fato de ele achar que a mente não pode ser estudada à parte da consciência. Começemos então examinando a crítica à fundamentação da ciência cognitiva, quer dizer, a crítica à metáfora do computador.

Para Searle, a afirmação de que o cérebro é um computador não diz nada relevante. Para mostrar isso, ele começa corretamente examinando a definição de “máquina de Turing”. Ele explica que uma máquina de Turing é uma máquina ideal que faz operações de leitura e escrita de símbolos em uma fita móvel de acordo com as instruções que ela recebe. Searle não chega a declarar explicitamente que um computador é uma máquina real que faz o mesmo que uma máquina de Turing, que é de fato a definição de “computador”. Em vez disso, ele se concentra em uma consequência secundária da definição. A consequência é que um computador pode ser feito de qualquer coisa, com tanto que simule uma máquina de Turing. Searle demonstra especial prazer em ilustrar esse ponto. Por exemplo, ele diz que um sistema de ratos, gatos e queijos poderia ser um computador, um grupo de pombos treinados para bicar também poderia e um sistema de canos e alavancas idem. Em suma, tudo poderia ser um computador. Destarte, quando alguém diz que o cérebro é um computador, não estaria dizendo nada relevante.

Já aqui, ao meu ver, a crítica searliana revela uma falha. Não é porque um computador pode ser feito de qualquer coisa que eu digo que qualquer coisa é um computador. Algumas afirmações e exemplos de Searle dão a entender que eu poderia dizer de qualquer coisa que ela é um computador. Ele diz, por exemplo, que de acordo com a teoria da computação, o cérebro é um computador digital “no mesmo sentido em que estômago, fígado, coração, sistema solar e o estado do Kansas são todos computadores digitais” (SEARLE, 2006: 297). Acontece que na prática não dizemos que fígados e estados americanos são computadores digitais, de alguma forma fazemos uma diferença entre essas coisas e os computadores eletrônicos. Na minha avaliação a diferença é clara: computadores rodam qualquer

programa (respeitadas as configurações mínimas de *hardware*), fígados e estados americanos não. A resposta é óbvia, mas mesmo assim, Searle tem uma resposta para ela. O ponto que ele quer enfatizar é que nada é intrinsecamente um computador e por isso tudo pode ser visto subjetivamente como um computador. Subjetivamente, posso interpretar os pares de entradas e saídas de um fígado de modo que esses pares correspondam a uma função qualquer. Assim, torna-se possível dizer que um fígado roda qualquer programa. Com efeito, Searle argumenta que a definição didática padrão de “computação” implica nos seguintes resultados:

1. Para qualquer objeto há uma descrição tal desse objeto que, sob esta descrição, o objeto seja um computador digital.
2. Para qualquer programa e para qualquer objeto suficientemente complexo, há alguma descrição do objeto segundo a qual ele está implementando o programa. Assim, por exemplo, a parede às minhas costas está agora mesmo implementando o programa Wordstar, porque há algum padrão de movimentos moleculares que é isomórfico em relação à estrutura formal do Wordstar. Mas, se a parede está implementando o Wordstar, então, se for uma parede grande o bastante, estará implementando qualquer programa, inclusive qualquer programa implementado no cérebro (Idem, *ibidem*: 298-9).

Na minha opinião, esse argumento de Searle é simplesmente absurdo, mas isso era de se esperar, já que ele tem a pretensão de refutar uma resposta óbvia. Mesmo que possamos dar um número ilimitado de interpretações aos padrões de movimentos das moléculas que compõem uma parede, não faz sentido dizer que a parede roda qualquer programa. Se fizesse, não precisaríamos de engenheiros da computação, os pedreiros fariam o trabalho deles com igual competência e a custo bem menor. Mas a verdade é que há uma diferença clara entre computadores e paredes. A diferença está na atribuição semântica que é feita em cada caso.

Chamo de “atribuição semântica” a uma relação entre estados físicos e símbolos de uma linguagem. Por exemplo, a atribuição semântica que permite a programação de computadores é a que estabelece que um circuito aberto significa “1” e um circuito fechado significa “0”. Além dessa atribuição semântica, temos apenas sintaxe, e isso é suficiente para a programação universal de um computador eletrônico. No caso da parede, se bem compreendo o exemplo de Searle, uma atribuição semântica seria algo como isto: o padrão de movimento

molecular da parede no instante t_1 significa o parâmetro p_1 do Wordstar, o padrão de movimento molecular da parede no instante t_2 significa o parâmetro p_2 do Wordstar e assim por diante. Dada essa atribuição semântica, a parede estaria implementando o Wordstar.

O problema é que, para que a parede implementasse outro programa qualquer, teríamos que mudar a atribuição semântica. Em contrapartida, no caso do computador eletrônico, quando o programa muda, a atribuição continua a mesma. Em outras palavras, computadores rodam qualquer programa dada uma única atribuição semântica, paredes rodam qualquer programa mas, para cada um deles, elas precisam de uma atribuição semântica diferente. Essa é a diferença básica entre computadores eletrônicos e paredes. Se entendemos as coisas assim, podemos dizer que cérebros se encaixam na mesma categoria de computadores eletrônicos, pelo menos no tocante ao fato de cérebros também poderem implementar qualquer programa dada uma única atribuição semântica. No caso do cérebro, a atribuição semântica relacionaria níveis de voltagem nas sinapses e operações ou estados mentais. Podemos supor que essa atribuição é constante como no caso do computador eletrônico, embora ainda não saibamos descrevê-la completamente.

Em resumo, embora tudo possa ser visto como um computador, não dizemos que tudo é um computador. Só chamamos de “computador” coisas que podem rodar qualquer programa e fazem isso mantendo uma atribuição semântica constante, o que não é o caso nem da parede nem do fígado e nem dos estados americanos, mas é o caso dos computadores eletrônicos e dos nossos cérebros.

Searle ainda apresenta outros argumentos para mostrar que um cérebro não deve ser visto como um computador. Mas, como ficou demonstrado, há um sentido claro com base no qual podemos fazer uma analogia entre o cérebro e os computadores eletrônicos. Dessa forma, pode-se concluir que os outros argumentos estão igualmente equivocados.

É importante esclarecer que Searle não argumenta contra a metáfora do computador devido a motivações reacionárias. Ele não quer que a investigação da mente retorne ao estado que tinha antes da revolução cognitiva. Pelo contrário, na minha opinião, a argumentação de Searle é motivada por uma crença ultrarrevolucionária. Ele acredita, e acertadamente, diga-se de passagem, que a metáfora do computador permite o estudo da mente sem o estudo da consciência,

e isso para ele é errar o alvo. Com efeito, ele declara que “mais do que qualquer outra coisa, é o desprezo pela consciência que explica tanta aridez e esterilidade na psicologia, na filosofia da mente e na ciência cognitiva” (Idem, ibidem: 326). O que Searle quer de fato é que a consciência seja reconhecida como a marca característica do mental e que, portanto, seu estudo seja considerado prioritário pelos pesquisadores da mente. Na segunda crítica de Searle à ciência cognitiva, essa sua motivação está ainda mais patente.

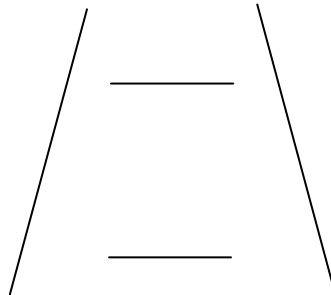
A explicação cognitivista padrão de como a mente executa uma tarefa é bem simples: supõe-se que a mente implementa um certo programa para realizar a tal tarefa e oferece-se uma descrição desse programa. Algumas vezes o programa é simulado em computador, algumas vezes não, mas em qualquer caso sempre se busca evidência empírica de que o programa foi descrito corretamente. Note-se que esse tipo de explicação pressupõe que a mente não é consciente das operações que ela executa para realizar a tal tarefa. Se fosse, não era preciso procurar um modelo das operações mentais. O próprio indivíduo poderia descrever o que se passa na sua cabeça. A segunda crítica searlina concentra-se exatamente sobre essa pressuposição cognitivista. Searle não concorda que existam operações mentais inacessíveis à consciência, e, no sétimo capítulo de *A Redescoberta da Mente*, ele apresenta um sofisticado argumento para mostrar que essa pressuposição é inaceitável.

De acordo com a análise que faço do argumento, ele está todo baseado na premissa de que tudo o que há de mental são os eventos conscientes e os eventos que em princípio podem vir a ser conscientes. Dessa forma, se fazemos operações mentais, elas são conscientes em princípio; se fazemos operações que não podem ser trazidas à consciência, elas não são mentais. Trata-se, portanto, de um argumento que depende de uma certa definição de “mental”.

A princípio, pode parecer que o argumento é uma consequência do princípio de conexão, que Searle enuncia anteriormente e que estabelece que “todos os estados intencionais inconscientes são em princípio acessíveis à consciência” (Idem, ibidem: 224). De fato, ao introduzir o argumento, Searle afirma: “Começo por uma discussão do princípio da conexão e suas implicações” (Idem, ibidem: 327), e, na sequência, ele cita o princípio várias vezes, muitas delas com a clara intenção de tentar justificar sua rejeição da tese de que há

operações mentais inacessíveis à consciência. Searle evidencia essa intenção, por exemplo, quando comenta a ilusão de Ponzo.

A ilusão é produzida por um desenho assim:



Tenho que confessar que não tenho ilusão nenhuma, mas o efeito esperado é que a linha horizontal superior pareça maior que a inferior. A explicação cognitivista para isso, conforme Searle relata, é que:

[...] o agente está inconscientemente seguindo duas regras e fazendo duas inferências inconscientes. A primeira regra é que linhas convergentes de baixo para cima no campo visual sugerem distância maior na direção da convergência, e a segunda é que objetos que ocupam porções iguais da imagem da retina variam em tamanho percebido, dependendo da distância percebida a partir do observador (lei de Emmert). Segundo esta explicação, o agente inconscientemente infere que a linha paralela superior está mais longe por causa de sua posição em relação às linhas convergentes, e, em segundo lugar, infere que a linha de cima é maior porque está mais longe (Idem, *ibidem*: 332).

Depois de relatar a explicação do cognitivista, Searle arremata: “Não há meio pelo qual esse tipo de explanação possa se tornar compatível com o princípio de conexão” (Idem, *ibidem*), o que deixa óbvia sua intenção de fazer parecer que seu ponto de vista é justificado com base no princípio de conexão. Mas a realidade é que não é. O princípio é sobre estados mentais intencionais, ao passo que a tese que Searle está contestando é sobre operações ou processos mentais. O próprio Searle reconhece a diferença entre estados mentais e processos mentais. Estados mentais são condições em que a mente está, são coisas como crenças, desejos, medos, dores etc. Operações ou processos mentais são atividades que a mente realiza. Um exemplo que Searle dá de processo mental é “cantarolar uma melodia silenciosamente para nós mesmos” (Idem, *ibidem*: 342). Dada essa diferença, temos que admitir que o fato do princípio de conexão proibir a

existência de estados mentais intencionais inacessíveis à consciência não implica que ele proíba também a existência de operações ou processos mentais inacessíveis à consciência; processos como as inferências postuladas na explicação cognitivista da ilusão de Ponzo, por exemplo. Dessa forma, embora Searle recorra várias vezes ao princípio de conexão para justificar sua crítica às explicações cognitivistas, o princípio não é suficiente para lhe dar essa justificação. É por isso mesmo que ele tem necessidade de fazer outras manobras para levar seu argumento adiante. A principal delas consiste em, senão definir o adjetivo “mental”, pelo menos insinuar como devemos interpretá-lo.

Searle identifica o domínio do mental com o domínio da consciência. Para ele, o cérebro comporta eventos que são puramente do nível do *hardware*, esses eventos produzem fenômenos conscientes e esses últimos são as únicas coisas que podem ser qualificadas propriamente de “mentais”. Ele afirma, por exemplo: “Sabemos que há experiências visuais conscientes, e sabemos que são causadas por processos cerebrais, mas onde se supõe que esteja o nível mental adicional neste caso?” (Idem, *ibidem*: 332). Aqui Searle expõe sua imagem da mente. Nessa imagem simplesmente não há lugar para processos mentais inconscientes; os únicos eventos mentais são precisamente as experiências conscientes. Segundo Searle, quando falamos que o cérebro realiza operações como inferir, interpretar e planejar, estamos simplesmente antropomorfizando o cérebro. Ele reconhece que o cérebro realiza certas operações, mas suas operações não são essas supracitadas, são antes operações neurofisiológicas não intencionais. Dessa forma, a essência do argumento searliano é dizer que processos mentais inconscientes não existem porque os únicos processos que o cérebro realiza são processos físicos brutos, processos sem conteúdo mental. É verdade que esses processos causam eventos mentais, mas esses eventos são sempre conscientes.

Note que se o argumento de Searle está correto, não só a ideia cognitivista de que há processos mentais inconscientes é refutada, mas na realidade toda a concepção cognitivista de *mente* é colocada em xeque. Com efeito, a ideia é consequência necessária da concepção. Se a mente é um programa, temos que admitir que a maior parte das linhas desse programa nos é desconhecida. A tarefa do cientista cognitivo é justamente descobrir essas linhas. Se não há linhas desconhecidas, então não resta nada para o cientista cognitivo fazer.

Como um adepto da concepção cognitivista de *mente* (como eu) poderia responder ao argumento de Searle? Bem, eu diria que, quando o cientista cognitivo supõe a existência de processos mentais inconscientes, ele está fazendo algo semelhante ao que um físico faz quando supõe a existência de partículas não observáveis. Uma partícula subatômica é postulada porque, com ela, a explicação de certos fenômenos se torna possível, e embora a partícula seja em si mesma inobservável, seus efeitos devem em princípio ser observáveis. Com os processos mentais inconscientes postulados pelo cientista cognitivo, ocorre a mesma coisa. Eles têm uma utilidade explanatória evidente e, apesar de eles não poderem ser diretamente observados, pode ser encontrada evidência empírica indireta de sua existência. Por exemplo, podem-se fazer previsões sobre o que aconteceria em certa situação se um indivíduo executasse uma determinada operação mental. Dada a situação, se a previsão é confirmada, isso conta como uma evidência de que a tal operação foi executada.

Alguém poderia alegar que a utilidade explanatória de processos mentais inconscientes não é tão evidente assim. Por que precisamos supor processos mentais abstratos para explicar um dado comportamento? Por que não podemos falar diretamente dos processos neurofisiológicos que causam aquele comportamento? Duas coisas podem ser ditas aqui. Em primeiro lugar, precisamos de explicações que se limitam a descrever o nível dos processos abstratos porque elas depois podem ser usadas na pesquisa empírica sobre o cérebro. Com efeito, se eu descrevo as operações mentais que uma pessoa faz em determinada situação, eu posso depois me guiar por essa descrição para tentar descobrir quais eventos neurofisiológicos são responsáveis por aquelas operações. É a descrição que dá sentido à minha investigação do cérebro. Em segundo lugar, precisamos de explicações dos processos mentais abstratos porque assim podemos encontrar formas de manipular esses processos. Em outras palavras, podemos criar métodos de controle desses processos. Se entendemos as operações que alguém faz para memorizar uma lista de números em dadas condições, por exemplo, poderemos talvez desenvolver um método de memorização mais eficiente. Essas são algumas das vantagens do método de pesquisa cognitivista. Essas vantagens Searle não cita na sua crítica.

Na minha opinião, o problema com a visão searlina do mental é que ela ainda está muito atrelada à metáfora do suporte. Para ele, as operações que

explicam o nosso comportamento não autômato ou se dão no cérebro ou na consciência. O cérebro é o *espaço* dos eventos neurofisiológicos inconscientes e a consciência é o *espaço* dos eventos mentais, e fora desses dois *espaços* não há nada. Por isso não pode haver nenhuma operação mental que seja inconsciente nem nenhuma operação inconsciente que seja mental. Essa visão fica clara na última passagem citada. A visão cognitivista, por outro lado, pode se abster de achar um “lugar” para as operações mentais inconscientes. Uma vez que tenhamos incorporado a metáfora do computador, perceberemos que elas não precisam estar em lugar nenhum. O que aconteceu de fato com a revolução cognitivista foi que os limites do mental se alteraram. Agora, para algo ser considerado mental não precisa ser algo *do ponto de vista do sujeito consciente*. As operações mentais postuladas pelo cognitivista são operações abstratas objetivas que só contingentemente calham de ser executadas por um cérebro humano, mas poderiam a princípio ser executadas em qualquer computador (no sentido amplo de “computador”). O cognitivista não se abala com o argumento searliano porque ele não compra as pressuposições metafísicas de Searle, ele não tem de dizer “onde se supõe que esteja o nível mental adicional”.

Ora, quando não há um compartilhamento mínimo de pressupostos não há realmente possibilidade de argumentação. Há apenas um paradigma contra o outro. Como podemos decidir qual o melhor? Bem, nesses casos, é certo que nossa escolha pode se basear nas razões mais arbitrárias, em todo caso, acho que, sempre que possível, deveríamos observar qual paradigma tem mais utilidade prática. Nesse ponto, não há dúvida quanto ao valor do modelo cognitivista. O progresso da ciência cognitiva é inegável. O próprio Searle concorda com isso, embora pontue que “é possível fazer bom trabalho com base em hipóteses falsas” (Idem, *ibidem*: 281). Esse, porém, não é o caso aqui, pelo menos na minha opinião. Aliás, acredito que o caso aqui não seja nem de verdade ou falsidade, já que quem escolhe um paradigma em geral escolhe também um critério de verdade.

Entre os resultados práticos produzidos pelo modelo computacional da mente, está o grande avanço no campo da linguística. Desde que Chomsky e outros pioneiros do internalismo linguístico começaram a adotar a estratégia de fazer referências aos nossos processos mentais para explicar o nosso comportamento linguístico, pudemos entender muito melhor o funcionamento da

linguagem. É precisamente isso que vou mostrar no próximo capítulo. E é precisamente essa estratégia que vou adotar para demonstrar minha tese sobre nomes próprios nos capítulos subsequentes.

Feitas essas considerações, é forçoso concluir que a crítica de Searle não atinge o cognitivismo nem em seus fundamentos nem em sua metodologia. A metáfora do computador continua incólume. Como, além disso, ela oferece muitas vantagens práticas tanto a mim e quanto aos cognitivistas, creio que posso considerar que estamos justificados em adotá-la.