



**UNIVERSIDADE FEDERAL DO CEARÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
INSTITUTO DE CULTURA E ARTE
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA**

Maxwell Morais de Lima Filho

**NATURALISMO BIOLÓGICO:
A SOLUÇÃO DUALISTA DE JOHN SEARLE PARA O
PROBLEMA MENTE-CORPO**

**Fortaleza – CE
Junho de 2010**

Maxwell Morais de Lima Filho

**NATURALISMO BIOLÓGICO:
A SOLUÇÃO DUALISTA DE JOHN SEARLE PARA O
PROBLEMA MENTE-CORPO**

Dissertação apresentada ao Curso de Mestrado em Filosofia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre.

Área de concentração: Filosofia

Linha de Pesquisa: Filosofia da Linguagem e do Conhecimento

Orientador: Prof. Dr. Dirk Greimann

**Fortaleza – CE
2010**

Dissertação defendida e avaliada em 23 de junho de 2010, pela banca examinadora constituída pelos professores:

Prof. Dr. Dirk Greimann (Universidade Federal do Ceará - UFC) – Orientador

Prof. Dr. Tárík de Athayde Prata (Universidade Federal de Sergipe - UFS) – Membro Avaliador

Prof. Dr. André Leclerc (Universidade Federal do Ceará - UFC) – Membro Avaliador

Aos doadores de meus cromossomos X e Y,
Keyvylane Abreu da Costa e Maxwell Morais de Lima

AGRADECIMENTOS

Ao Professor Dirk Greimann, por suas críticas e sugestões pertinentes, que tanto enriqueceram o presente trabalho, e pelos valerosos ensinamentos sobre as Filosofias da Linguagem e da Mente;

Ao Professor Tárik de Athayde Prata, pelo constante incentivo, pela paciência e, acima de tudo, pela amizade que desenvolvemos ao longo desses anos. Dizem que os elefantes incomodam muita gente, mas Tárik *sabe que um Maxwell incomoda muito mais!*;

Ao Professor André Leclerc, que está sempre disposto a auxiliar os alunos com um sorriso no rosto, por gentilmente ter aceitado participar de minha Banca de Avaliação;

Ao Professor Cícero Antônio Cavalcante Barroso, pelas importantes observações feitas a uma versão preliminar dessa dissertação e pelas conversas instigantes que tivemos;

À CAPES, que proporcionou as condições financeiras de realização desse trabalho;

Aos meus Professores do Departamento de Filosofia da UFC, que me ensinaram a importância de se construir uma embarcação resistente antes de se aventurar por esse mar revolto que é a Filosofia. Destaco os seguintes Capitães: Professor Guido Imaguire, Manfredo Oliveira, Kleber Amora e Custódio Almeida. Prometo não me afogar antes do Doutorado...

Aos Filósofos pós-socráticos André, Valdetonio, Thiago, Pedro, Erika, Renato e Abraão, que são os maiores e melhores geradores de lero-lero que conheço. Haja imaginação, números transfinitos e argumentos furadores!;

À minha compreensiva Tissi Camu, por todos esses anos de adorável companhia e apoio incondicional;

Aos meus irmãos-amigos Argus e Pablo, os quais são excelentes pessoas para se dividir um cafofo, apesar de serem bagunceiros e não saberem fritar um ovo (mas o Pablo está aprendendo...);

Aos meus amigos-irmãos: Porfírio, Beto, Jorginho, Vítor Ch., Normal, Patrícia, Pâmela, Leopoldina, Márcias (Aquino e Melo), Fernando Martins, Erneson, Jota, Apiano, Tereza, Sérgio, Roner, Vevé, Hudson, Páschoa, Saulos (Davi e Monte), Gesner, Afonsina, Cris, Adolfo, Suzy, Eugênio, Rosângela, Marquinhos, Sarinha, Bruno, Crau, Lili, Michel, Baila, Vladimir, Adaías, Erivânia, Adelina, Madson, Adriano, Batista (e família), Piuí, Eriberto e Renato Arrais.

RESUMO

O objetivo deste trabalho é propor uma classificação do naturalismo biológico de John Searle em uma das concepções teóricas de Filosofia da Mente. Para tanto, apresentarei uma visão panorâmica das principais teorias e uma exposição das teses que compõem o naturalismo biológico, com o intuito de comparar este com aquelas. O próprio Searle resiste em rotular o naturalismo biológico, já que, segundo ele, todas as teorias da tradição em Filosofia da Mente partem de um pressuposto equivocados, a saber, o *dualismo conceitual*, segundo o qual há uma exclusão mútua entre as categorias física e mental: o físico é *não* mental, e o mental é *não* físico. Para Searle, fenômenos mentais são biológicos e, portanto, são físicos. No entanto, isso não significa que há uma redução ontológica do mental ao físico, pois existe uma distinção ontológica entre esses dois níveis – ontologia de primeira pessoa e ontologia de terceira pessoa, respectivamente. O problema é que com tal distinção ontológica, Searle acaba por criar um novo tipo de dualismo, que em vez de contrapor o físico ao mental, contrapõe o objetivo (ontologia de terceira pessoa) ao subjetivo (ontologia de primeira pessoa). Por defender o fisicalismo ontológico e, ao mesmo tempo, endossar que os eventos mentais são reais, causalmente eficazes e ontologicamente irreduzíveis, a concepção de Searle converge em muitos pontos com o fisicalismo não-reduutivo e com o dualismo de propriedade. Compararei o naturalismo biológico com ambas as teorias e, ao final, terei subsídios para argumentar o porquê de classificá-lo em uma delas e não na outra.

Palavras-chave: Filosofia da Mente, John Searle, Naturalismo Biológico.

ABSTRACT

The objective of this work is to propose a classification of John Searle's biological naturalism in one of theoretical conceptions of Philosophy of Mind. In order to do that, I will present a panoramic vision of principal theories and a presentation of theses of biological naturalism in order to compare this with those. Searle himself resist to label the biological naturalism since, according him, all theories of tradition in Philosophy of Mind start from a mistaken assumption, that is, *the conceptual dualism*, according to which there is a mutual exclusion between the physical and mental categories: the physical is *not* mental, and mental is *not* physical. For Searle, mental phenomena are biological and, therefore, physical. However, this does not mean that there is an ontological reduction of mental to physical, because there is an ontological distinction between these two levels – first-person ontology and third-person ontology, respectively. The problem is that with such ontological distinction, Searle ends up creating a new kind of dualism, that instead of countering the physical to the mental, opposes the objective (third-person ontology) to the subjective (first-person ontology). By defending the ontological physicalism and, at the same time, to endorse that mental events are real, causally effective and ontologically irreducible, Searle's conception converges at many points with the non-reductive physicalism and property dualism. I will compare the biological naturalism with both theories and – at the end – I will have subsidies to argue why classify it in one and not the other.

Key-words: Philosophy of Mind, John Searle, Biological Naturalism.

– Você tem cérebro? – Perguntou o espantalho.
– Suponho que sim. Nunca olhei para ver –
replicou o Leão.

(Frank Baum – O Mágico de Oz)

– Eu quero conhecer melhor os seres humanos, o mundo, tudo, em sumo. E os robôs também, George. Estou com vontade de escrever uma história sobre eles.

George então passou o braço pelo ombro do autômato.

– Bom, vamos para casa. Mas antes pega a tua roupa. Andrew, existe um milhão de livros sobre robótica e todos incluem histórias da ciência. O mundo inteiro já está ficando saturado, não só de robôs, mas de informações sobre eles.

Andrew sacudiu a cabeça, gesto humano que ultimamente passara a adotar.

– Mas não é uma história da robótica, George. Uma história de *robôs*, escrita por um robô. Eu quero explicar como eles encaram tudo o que aconteceu desde que começaram a trabalhar e a viver na Terra.

(Isaac Asimov – O Homem Bicentenário)

Se à exceção do homem nenhum outro ser vivente possuísse uma faculdade mental qualquer, ou se os seus *poderes* fossem de natureza totalmente diversa daquela dos animais inferiores, então nunca estaríamos em condições de convencer-nos de que as nossas elevadas faculdades se desenvolveram gradativamente. Mas, pode-se provar que não existe nenhuma diferença fundamental deste gênero.

(Charles Darwin – A origem do homem e a seleção sexual)

No que tange à consciência, é impossível defini-la. Critérios vários indicam que mesmo os invertebrados inferiores têm uma consciência, possivelmente os próprios protozoários, nas suas reações de repulsão. Se se quiser descer mais ainda, até os procariotos (por exemplo, bactérias magnéticas), isso é questão de gosto. Em qualquer caso, o conceito de consciência não pode ser definido nem mesmo aproximadamente, razão pela qual se torna impossível uma discussão em detalhe.

(Ernst Mayr – O desenvolvimento do pensamento biológico)

Pode-se dizer que a consciência é o problema com mais conotações filosóficas que a mente nos impõe, mas também o mais refratário e difícil de abordar.

(...)

Mas a ciência relacionada à mente não é *apenas* significativa. Em um sentido estritamente prático, o estudo do *Homo sapiens* é o mais crucial de todos os empreendimentos científicos.

(...)

Se os neurocientistas, psicólogos, pesquisadores da inteligência artificial e outros investigadores da psique realizarem todos os seus sonhos, poderemos um dia viver em uma cultura moldada por *verdadeiras* teorias da natureza humana. Poderemos deixar de nos atormentar com o enigma da natureza *versus* criação ou com o problema mente/corpo, pois eles terão sido resolvidos a contento para todos.

(John Horgan – A mente desconhecida)

LISTA DE ABREVIATURAS DOS LIVROS E ARTIGOS DE SEARLE

- CECC: “Como Estudar Cientificamente a Consciência”
- DR: “Dualism Revisited”
- FCEU: “Filosofia Contemporânea nos Estados Unidos”
- II: “Intrinsic Intentionality”
- INT: *Intencionalidade*
- LN: *Liberdade e Neurobiologia*
- M: *Mind: Brief Introduction*
- MA: “As Mentes dos Animais”
- MCC: *Mente, Cérebro e Ciência*
- MLS: *Mente, Linguagem e Sociedade: Filosofia no mundo real*
- MCP: “Mentes, Cérebros e Programas”
- TOY: “Twenty-One Years in the Chinese Room”
- WIANPD: “Why I Am Not a Property Dualist”

SUMÁRIO

INTRODUÇÃO	11
1 VISÃO PANORÂMICA SOBRE AS PRINCIPAIS ABORDAGENS EM FILOSOFIA DA MENTE	18
2 O EXPERIMENTO DE PENSAMENTO DO QUARTO CHINÊS: A CRÍTICA DE JOHN SEARLE À INTELIGÊNCIA ARTIFICIAL FORTE	28
2.1 O Teste de Turing	28
2.2 O Debate Acerca do Quarto Chinês	29
2.2.1 O Experimento de Pensamento do Quarto Chinês	29
2.2.2 O Argumento Relacionado ao Quarto Chinês	31
2.2.3 Questões Periféricas	32
2.3 O Novo Argumento de Searle	38
3 A CONCEPÇÃO POSITIVA DE SEARLE ACERCA DA MENTE	43
3.1 1ª tese: Fenômenos Mentais são Fenômenos Biológicos	43
3.2 2ª tese: Fenômenos Mentais são Causados e Realizados no Sistema Cerebral	47
3.3 3ª tese: Fenômenos Neurobiológicos são Ontologicamente Distintos de Fenômenos Mentais	54
3.4 4ª tese: Não Expressabilidade dos Termos que Designam Fenômenos Mentais Através dos Termos que Descrevem os Fenômenos Biológicos	60
3.5 5ª tese: Fenômenos Neurobiológicos Explicam Causalmente os Fenômenos Mentais	63
3.6 6ª tese: As Capacidades Causais dos Fenômenos Neurobiológicos e Mentais são as Mesmas	66
4 REDUACIONISMO E IRREDUTIBILIDADE DA CONSCIÊNCIA NO NATURALISMO BIOLÓGICO	69

4.1	Redução e Filosofia da Mente_____	69
4.2	Causação Mental no Naturalismo Biológico_____	73
4.2.1	Causação e Causalidade Segundo David Hume_____	73
4.2.2	Críticas de Searle a Hume_____	75
4.2.3	Causação Mental_____	77
4.3	O Fisicalismo Redutivo de Jaegwon Kim e o Problema da Sobredeterminação Causal__	79
4.4	Fisicalismo Não-Redutivo, Dualismo de Propriedade e Naturalismo Biológico_____	89
	REFERÊNCIAS BIBLIOGRÁFICAS_____	108

INTRODUÇÃO

O filósofo americano John Searle iniciou sua formação acadêmica¹ dentro da tradição da Filosofia Analítica, em geral, e da Filosofia da Linguagem, em particular. Ele nasceu no ano de 1932, em Denver, nos Estados Unidos e desde 1959 é professor de Filosofia na Universidade da Califórnia, em Berkeley, sendo considerado um dos maiores filósofos do século XX. Searle ficou mundialmente famoso como filósofo da linguagem, ao desenvolver a *Teoria dos Atos de Fala*, proposta inicialmente por John Austin, e ao publicar seu primeiro livro, considerado hoje um clássico da área, *Speech Acts*, em 1969. Apesar de pertencer à tradição analítica, Searle diverge desta ao afirmar que a Filosofia da Linguagem é um ramo da Filosofia da Mente, e que, além disso, a *intencionalidade* precede a linguagem. Em outras palavras, a intencionalidade da linguagem é derivada da intencionalidade dos estados mentais, que é *intrínseca*. O presente trabalho, por questão de tempo, restringe-se à Filosofia da Mente de Searle que, em termos cronológicos, inicia-se com a publicação do artigo “Minds, brains, and programs”, em 1980, o qual consiste à sua famosa crítica ao programa de pesquisa da inteligência artificial forte.

Sendo assim, esse trabalho se situa no âmbito geral da Filosofia da Mente e é, de maneira mais específica, acerca da solução de Searle para o problema mente-corpo – o naturalismo biológico. O problema mente-corpo é a principal questão de Filosofia da Mente, a qual possui um caráter ontológico: qual a relação entre mente e corpo? Mental e físico são categorias ontológicas mutuamente excludentes? Processos mentais podem ser reduzidos a processos neurobiológicos? Existem basicamente quatro respostas para essas perguntas: dualismo, fisicalismo redutivo, fisicalismo não-redutivo e eliminativismo. De acordo com a primeira, há uma distinção entre substâncias ou propriedades mentais e físicas, o que acarreta a irredutibilidade do domínio mental ao físico. Os reducionistas defendem, por sua vez, que estados mentais são idênticos a disposições comportamentais ou a tipos de estados cerebrais, ou seja, são redutíveis a estes. Já os adeptos do fisicalismo não-redutivo defendem que tudo que há no Universo são, no nível mais básico, partículas físicas fundamentais ao mesmo tempo em que sustentam a irredutibilidade ontológica dos fenômenos mentais. Por fim, os

¹ Searle estudou Filosofia em Oxford nos anos 50, e foi aluno de grandes filósofos analíticos, tais como Gilbert Ryle, John Austin e Peter Strawson.

defensores do fisicalismo eliminativo afirmam que a crença em estados mentais é baseada em uma concepção distorcida da realidade, não havendo correspondência de um estado mental para um estado neurofisiológico, e que o avanço do conhecimento neurocientífico inevitavelmente eliminará a psicologia popular.

Dentro desta delimitação – o problema mente-corpo –, são três os objetivos gerais dessa dissertação, a saber: (i) apresentar a crítica de Searle à inteligência artificial forte; (ii) enumerar e explicar as teses que compõem o naturalismo biológico; (iii) classificar o naturalismo biológico em uma das correntes de Filosofia da Mente – o dualismo de propriedade. Esta proposta de classificação do naturalismo biológico é o principal objetivo deste trabalho. No entanto, para atingi-lo, seguirei as seguintes etapas: apresentação do problema mente-corpo e das principais tentativas para solucioná-lo (1º capítulo), a fim de mostrar o pano de fundo teórico no qual está inserido o naturalismo biológico. A seguir, apresentarei no 2º capítulo a crítica de Searle a uma dessas teorias – à inteligência artificial forte – a partir do experimento de pensamento do quarto chinês. Após isso, será exposta a concepção positiva de Searle em relação à mente, a qual consiste nas teses do naturalismo biológico (3º capítulo). Por fim, após ter realizado as etapas precedentes, terei subsídios para fundamentar minha hipótese, a de que o naturalismo biológico é um tipo de dualismo de propriedade (4º capítulo). A seguir, explico de maneira mais detalhada o que será trabalhado em cada um dos capítulos do presente trabalho.

O principal problema da Filosofia da Mente é o chamado problema mente-corpo ou mente-cérebro². Essa é uma questão ontológica, já que diz respeito à *existência e natureza* dos eventos mentais em relação ao substrato físico no qual eles ocorrem, que, dependendo da concepção defendida, pode ser o corpo, de maneira geral, ou o cérebro, de maneira mais específica, ou ainda uma estrutura não-biológica como, por exemplo, o hardware de um computador. Na verdade, antes do surgimento da Filosofia, as Religiões já faziam uma distinção entre corpo (físico) e alma (não-físico): de acordo com o pensamento religioso padrão, alma e corpo são ontologicamente distintos, e é por isso que aquela pode sobreviver à

² Além do problema *ontológico* (CHURCHLAND, 1998, pp. 25-88), há outros três que são debatidos pelos filósofos da mente, que são os problemas *semântico* (CHURCHLAND, 1998, pp. 91-112), *epistemológico* (CHURCHLAND, 1998, pp. 115-135) e *metodológico* (CHURCHLAND, 1998, pp. 137-160). O primeiro destes se refere ao *significado* dos termos que são utilizados para descrever os estados psicológicos; o segundo trata da fundamentação e justificação de como obtemos *conhecimento* acerca dos estados mentais, sejam eles nossos, de outras pessoas (*problema de outras mentes*) ou mesmo de outros animais não-humanos; finalmente, uma condição *sine qua non* para estudar qualquer assunto, e a mente não é exceção, é estabelecer uma *metodologia*, a qual estabelece os procedimentos legítimos para abordar o objeto em questão. Em relação à mente, há basicamente duas metodologias gerais: a que considera o estudo da mente humana em continuidade com as Ciências Naturais e a que reivindica autonomia perante estas. A concepção de Searle, como o próprio nome aponta, defende a primeira delas, isto é, uma *metodologia naturalista* ou um *naturalismo metodológico*.

morte deste último. Porém, esse pensamento não é exclusivo da Religião, tendo sido defendido ao longo do tempo por ilustres filósofos, como Platão e Descartes. Como essa concepção assume a distinção ontológica entre alma e corpo, ela foi denominada de dualismo de substância: a *natureza* da substância-corpo (física) é diferente da *natureza* da substância-alma (não-física) e, conseqüentemente, elas possuem *modos de existência* distintos, ou seja, corpo e alma são *ontologicamente distintos e independentes*³.

O principal problema com o dualismo de substância é o seguinte: como explicar a interação entre alma e corpo, entre uma *coisa física* e uma *coisa não-física*? Principalmente por causa do avanço das Ciências Naturais, notadamente a Física, muitos filósofos passaram a adotar uma visão mais econômica do ponto de vista ontológico, excluindo qualquer entidade ou processo não-físico de suas teorias. As duas principais variações dessa concepção chamam-se fisicalismo reduutivo e fisicalismo não-reduutivo. Nota-se, a partir do nome dessas teorias, que a principal distinção entre elas consiste na redução ou não de fenômenos mentais a processos físicos.

Conhecer as teorias que foram propostas ao longo de mais de 2.500 anos de discussão é bastante relevante para a apresentação do naturalismo biológico, já que a compreensão desta teoria se dá a partir da análise das teses defendidas por Searle, as quais estão relacionadas, direta ou indiretamente, com o pano de fundo geral das demais concepções. Portanto, o ponto de partida de meu trabalho será uma breve exposição histórica da Filosofia da Mente, isto é, descreverei as principais soluções propostas, desde a antiguidade até nossos dias, para o problema mente-corpo (capítulo 1): dualismo de substância, dualismo de propriedade, behaviorismo analítico, teoria da identidade de tipo e, por fim, o funcionalismo.

Das teorias supracitadas, a que gozou de maior sucesso nas últimas décadas foi o funcionalismo, especialmente o funcionalismo de computador. De acordo com este, os fenômenos mentais não são *essencialmente* biológicos, já que podem ser produzidos artificialmente pelos chips de silício de um computador. Sendo assim, não há um *vínculo necessário* entre cérebros complexos e estados mentais, o que importa, na verdade, é o papel funcional desses últimos, isto é, as relações causais que determinado estado mental possui com os dados de entrada (provenientes do ambiente externo), com outros estados mentais e com o comportamento do sistema (dados de saída). Em termos metodológicos, pode-se afirmar que o funcionalismo é behaviorista: parte-se do pressuposto que um comportamento

³ Portanto, pode existir um corpo sem alma e, inversamente, uma alma sem corpo.

similar ao comportamento humano é *suficiente* para se *fundamentar* a atribuição de uma mente artificial *semelhante* à nossa.

Muitos filósofos atacaram a analogia funcionalista de que a mente está para o cérebro assim como o software está para o hardware, entre eles, John Searle. A primeira contribuição de Searle à Filosofia da Mente foi justamente sua crítica à inteligência artificial forte, na qual propôs o famoso experimento de pensamento do quarto chinês. Com esse experimento mental ele pretende mostrar que todo o projeto de pesquisa da IA forte está fadado ao fracasso, pois se baseia num erro conceitual, a saber, de que a sintaxe puramente formal de um programa de computador é *suficiente* para produzir uma semântica. Além disso, Searle também critica a abordagem behaviorista ao defender que um sistema pode obter sucesso em se comportar *como se* tivesse uma mente, porém, *sem tê-la*. Por isso, é importante não confundir *imitação* com *duplicação*: um computador (ou robô) pode imitar perfeitamente o comportamento humano sem duplicar, contudo, os estados mentais, os quais são produzidos pelo sistema nervoso e *causam* nosso comportamento. A crítica de Searle ao funcionalismo será apresentada no segundo capítulo da dissertação. Como o objetivo principal do experimento de pensamento do quarto chinês é afirmar que estados mentais *não são* puramente sintáticos e que, por causa disso, *não podem* ser produzidos por um computador digital, pode-se considerar o ataque de Searle à IA forte como a *parte negativa* de sua concepção.

O artigo no qual Searle propôs o experimento de pensamento (MCP) foi rebatido por vários pesquisadores, tanto da Filosofia quanto da Ciência da Computação. No mesmo número da revista, após os comentários críticos de diversos especialistas, está a resposta de Searle (II), na qual ele se defende das críticas e lança, além disso, o primeiro esboço de sua *concepção positiva* acerca da mente, o naturalismo biológico:

Mental states are as real as any other biological phenomena. They are both caused by and realized in the brain. That is no more mysterious than the fact that such properties as the elastic and puncture resistance of an inflated car tire are both caused by and realized in its microstructure. (...) Mental states are no more epiphenomenal than are the elasticity and puncture resistance of an inflated tire, and interactions can be described both at the higher and lower levels, just as in the analogous case of the tire (II, p. 455).

Apresentarei no terceiro capítulo a parte positiva da concepção de Searle, isto é, apresentarei as teses do naturalismo biológico, que são as seguintes: (i) fenômenos mentais são fenômenos biológicos, (ii) fenômenos mentais são causados e realizados no sistema cerebral, (iii) fenômenos neurobiológicos são ontologicamente distintos de fenômenos mentais, (iv) não expressabilidade dos termos que designam fenômenos mentais através dos

termos que descrevem os fenômenos biológicos, (v) fenômenos neurobiológicos explicam causalmente os fenômenos mentais e, por fim, (vi) as capacidades causais dos fenômenos neurobiológicos e mentais são as mesmas. Apesar de apresentar estas teses em vários livros e artigos no decorrer de três décadas, o próprio Searle não as apresenta de maneira detalhada⁴, o que justifica o objetivo desse capítulo, qual seja: expor sistematicamente as asserções do naturalismo biológico.

Finalmente, o conhecimento da concepção positiva de Searle permitirá classificar o naturalismo biológico como pertencendo a uma das correntes de Filosofia da Mente. Apesar da recusa de Searle em se encaixar em qualquer um dos “ismos” (dualismo, fisicalismo, reducionismo, anti-reducionismo etc.), argumentarei, todavia, que ele está comprometido com uma posição dualista de propriedade, mesmo que ele não o admita.

Essa proposta de classificação é baseada nas teses assumidas como verdadeiras no naturalismo biológico e consiste, como já foi dito, na principal meta desse trabalho.

Entretanto, antes de defender que o naturalismo biológico é um tipo de dualismo de propriedade, faz-se necessário analisar o conceito de *redução*, que é, na minha opinião, o divisor de águas das teorias da mente: de um lado estão as concepções redutivas, como o behaviorismo, e do outro as não-redutivas, como o emergentismo⁵. O conceito de redução é *polissêmico* e neste trabalho será utilizado em três sentidos: redução ontológica, redução causal e redução teórica. Este último está relacionado à redução de uma teoria à outra, dividindo-se em dois tipos – redução intrateórica e redução interteórica. Se tanto a teoria redutora quanto a teoria reduzida pertencerem à mesma Ciência (Física, por exemplo), fala-se de redução intrateórica; caso a redução ocorra entre teorias provenientes de Ciências distintas (Química e Biologia, por exemplo), há redução interteórica. No âmbito da Filosofia da Mente, esse segundo tipo de redução trata da seguinte questão (item 4.1): pode-se reduzir a Psicologia Popular à Neurociência? No entanto, este trabalho focará mais em outro tipo de redução, a saber, a redução ontológica, ou seja, a que trata da possibilidade ou não de se reduzir estados mentais a processos neurobiológicos (seção 4.1).

⁴ Não raras vezes, Searle apresenta as teses do naturalismo biológico de maneira sucinta e superficial, tornando sua posição um pouco confusa. Além disso, ele defende teses mutuamente incompatíveis, sendo este o motivo, a meu ver, que faz com que ele seja classificado por uns como dualista e por outros como fisicalista; Searle rejeita ambas as classificações.

⁵ A terceira alternativa, o fisicalismo eliminativo, não se fundamenta no conceito de redução, e sim no de *eliminação*: a Psicologia Popular deve ser eliminada em detrimento do arcabouço conceitual de uma Neurociência madura. No entanto, para defender tal tese, o eliminativista tem que mostrar que a redução dos estados mentais a estados neurobiológicos é inadequada, já que não existe uma relação *um estado mental-um estado neurobiológico*. Enfim, no que se refere à mente, para se defender o conceito de eliminação, deve-se rejeitar o de redução.

Uma importante divergência de Searle em relação à tradição filosófica é a sua não aceitação do modelo de causação humeano (itens 4.2.1 e 4.2.2), segundo o qual a causa precede *necessariamente* o efeito, como ocorre no famoso exemplo de uma bola de bilhar que causa o movimento de outra. Para Searle, esse modelo é inapropriado para dar conta da relação existente entre o sistema nervoso e os fenômenos mentais, já que estes, segundo ele, são causados e realizados *simultaneamente* na microestrutura neurobiológica (item 4.2.3). Em outros termos, apesar dos estados mentais serem causados por processos neurofisiológicos, *não há* uma lacuna temporal que separe estes últimos (causa) daqueles (efeito): há sincronia entre causa e efeito. Ou melhor, existe apenas um sistema físico descrito em dois níveis distintos e complementares, o micro (processos neurológicos) e o macro (eventos mentais).

Constata-se desde a primeira apresentação do naturalismo biológico que Searle repudia o epifenomenalismo, o qual afirma que os estados mentais não possuem eficácia causal⁶. Isso significa que o meu desejo de levantar o braço (estado mental) é a *causa* do movimento do mesmo (estado físico), o que caracteriza a causação mental-físico. É devido à tese da eficácia causal do mental que surge o problema da sobredeterminação causal no naturalismo biológico, ou seja, *um mesmo efeito possui duas causas*: o movimento do braço tem tanto uma causa física (as sinapses neuromusculares) quanto uma mental (desejo de erguê-lo). Apresento no item 4.2.4 a crítica de Kim a Searle, bem como a defesa deste último em relação à sobredeterminação causal.

Por fim, após mostrar que o naturalismo biológico é um tipo de monismo físico, proporei que, de modo mais específico, ele é um *dualismo de propriedade* (item 4.4). Fundamento minha argumentação no fato de que a subjetividade ontológica dos estados mentais, que considero a tese central do naturalismo biológico, acarreta a irredutibilidade ontológica de tais estados a processos neurobiológicos. Em outras palavras, existe, a despeito do que Searle admite, um dualismo conceitual entre o subjetivo e o objetivo. Sendo assim, apesar do naturalismo biológico não ser, de modo algum, um dualismo de substância, ele se enquadra dentro do dualismo de propriedade.

Para realizar a presente pesquisa, utilizei como bibliografia primária os artigos e livros de Searle: “Minds, brains, and programs” (1980a), “Intrinsic Intentionality” (1980b), *Intencionalidade* (1995), *A Redescoberta da Mente* (1997), *O Mistério da Consciência* (1998), *Mente, Linguagem e Sociedade: Filosofia no mundo real* (2000), “Filosofia

⁶ *Mental states are no more epiphenomenal than are the elasticity and puncture resistance of an inflated tire, and interactions can be described both at the higher and lower levels, just as in the analogous case of the tire* (II, p. 455. Grifos meus).

contemporânea nos Estados Unidos” (2002a), “Twenty-One Years in the Chinese Room” (2002b), “Why I Am Not a Property Dualist” (2002c), “Dualism Revisited” (2007a) e *Liberdade e Neurobiologia* (2007b). A bibliografia secundária se divide em três partes, a saber: (1) textos dos comentadores de Searle, (2) textos de cunho geral sobre Filosofia da Mente e, por fim, (3) os trabalhos de Jaegwon Kim. Para a primeira dessas partes, destaco duas coletâneas de artigos sobre a obra de Searle: *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (editada por Preston & Bishop, 2002) e *John Searle* (editada por Smith, 2003). Sobre o pano de fundo geral da Filosofia da Mente, utilizei a coletânea de textos organizada por John Heil, intitulada *Philosophy of Mind: a guide and anthology* (2004), já que nesta estão presentes textos de filósofos clássicos, modernos e contemporâneos sobre o assunto. Finalmente, deve-se destacar que os textos de Jaegwon Kim são de extrema relevância para minha pesquisa, pois Kim é o principal filósofo reducionista vivo, sendo o responsável por críticas fundamentais ao naturalismo biológico. De Kim, uso como referência os artigos presentes em sua coletânea, intitulada *Supervenience and Mind: selected philosophical essays* (1993), bem como seus três livros – *Philosophy of Mind* (1996), *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* e, por fim, *Physicalism or Something Near Enough* (2005).

1. VISÃO PANORÂMICA SOBRE AS PRINCIPAIS ABORDAGENS TEÓRICAS DA FILOSOFIA DA MENTE

Antes de apresentar a crítica de Searle ao programa de pesquisa da IA forte e de discutir as teses e a classificação do naturalismo biológico, faz-se necessário mostrar uma visão panorâmica das principais teorias de Filosofia da Mente – os dualismos (de substância e de propriedade) e os fisicalismos (reduutivo, não-reduutivo e eliminativo) –, que correspondem ao pano de fundo deste trabalho. O conceito-chave, mas não o único, para se fazer a distinção entre tais concepções é o de redução. Por um lado, os proponentes de teorias reducionistas alegam que a unidade e a simplicidade teórica, devido à economia ontológica e/ou conceitual, e, conseqüentemente, um maior ganho explicativo são as principais vantagens adquiridas ao se reduzir processos psicológicos a processos físicos. Por outro lado, os anti-reducionistas, adeptos tanto do dualismo como do fisicalismo não-reduutivo, defendem a irredutibilidade total ou parcial de fenômenos mentais a fenômenos físicos. Um consenso entre tais filósofos, por exemplo, é a irredutibilidade das sensações qualitativas ou *qualia*: por mais que se tenha um conhecimento completo da microestrutura física que produz estados mentais, as propriedades de sentir um cheiro ou ver uma cor permaneceriam inexplicadas (questão epistêmica), já que estas propriedades são de natureza *não-física* (questão ontológica). Ou seja, propriedades mentais estão acima e além da estrutura física subjacente e são, portanto, ontologicamente irredutíveis a esta. Há, além do reducionismo e do anti-reducionismo, uma terceira alternativa, a saber, a posição que rejeita o conceito de redução em detrimento do conceito de eliminação: uma neurobiologia bem desenvolvida fará com que o arcabouço teórico da psicologia popular seja eliminado. A seguir, apresento tais posições em maiores detalhes.

Não se pode precisar quem foi o primeiro pensador a se preocupar de *maneira sistemática* com a relação que existe entre o “eu”⁷ e o corpo. No entanto, o filósofo grego Platão foi um dos primeiros a apresentar argumentos filosóficos a favor de uma separação ontológica entre a alma, de um lado, e o corpo, do outro. O dualismo platônico concebe a alma como uma entidade que existe independentemente do corpo – além da alma preexistir em relação ao corpo, continua existindo após a morte do mesmo:

(...) se os vivos nascem novamente dos mortos, nossas almas existiriam *lá*, não existiriam? Afinal não poderiam renascer se não existissem, o que seria prova suficiente de sua existência se realmente fosse apresentada a evidência de que os vivos nascem somente de uma fonte, a saber, dos mortos (Fédon, 70d).

⁷ Utilizo o termo “eu” nesse contexto como sinônimo de *alma* ou *mente*.

(...) [a alma] encontra-se pura ao separar-se do corpo, nada arrastando de corpóreo consigo, porque jamais se associou voluntariamente ao corpo durante a vida, evitando-o e recolhendo-se a si mesma, pois sempre se dedicou ao seu exercício regular (...) (Fédon, 80e).

É bem sabido que na concepção platônica o domínio físico, que é conhecido pelos sentidos, é apenas um mundo de aparências, já que ele é uma cópia imperfeita do mundo das Formas (*Eidos*), ao qual temos acesso unicamente pelo intelecto. Esta dualidade dos mundos também se reflete na dualidade entre corpo e alma, já que aquele está para a esfera física da mesma forma que esta está para a esfera suprafísica, ou seja, a alma, diferentemente do corpo, é *não-física e imortal*:

(...) a alma é maximamente semelhante ao divino, ao imortal, ao inteligível, ao uniforme, ao indissolúvel e ao que é sempre imutável, ao passo que o corpo é maximamente semelhante ao humano, ao mortal, ao multiforme, ao ininteligível, ao dissolúvel e ao sempre mutável (Fédon, 80b).

Apesar de Platão ter defendido uma versão de dualismo de substância, o mais famoso defensor dessa posição filosófica foi René Descartes, que concebia uma cisão entre o domínio físico, *res extensa* (substância material ou física), e o domínio mental, *res cogitans* (alma ou substância mental). Para mostrar a existência desses dois domínios ontológicos, Descartes parte do *princípio de distinção* (MASLIN, 2009, p. 56): a possibilidade de conceber claramente a distinção e a independência de corpo e alma *decorre* deles terem sido produzidos por Deus de tal modo, o que implica, por sua vez, que eles *de fato* são diferentes, pois “podem ser postas separadamente, ao menos pela onipotência de Deus” (DESCARTES, 1996, p. 325). E qual é a diferença fundamental entre esses dois domínios ontológicos? A característica essencial dos corpos é a *extensão*, enquanto que a da alma é o *pensamento*:

E, embora (...) eu tenha um corpo ao qual estou muito estreitamente conjugado, todavia, já que, de um lado, tenho uma idéia clara e distinta de mim mesmo, na medida em que sou apenas uma coisa pensante [*res cogitans*] e inextensa, e que, de outro, tenho uma idéia distinta do corpo, na medida em que é apenas uma coisa extensa [*res extensa*] e que não pensa, é certo que este eu, isto é, minha alma, pela qual eu sou o que sou, é inteira e verdadeiramente distinta de meu corpo e que ela pode ser ou existir sem ele [independência entre alma e corpo] (DESCARTES, 1996, p. 326).

Já em sua época Descartes foi criticado por defender a interação entre esses dois tipos totalmente diferentes de substância⁸: como algo físico *causa* algo mental? E, inversamente, como algo mental *causa* algo físico? Para defender sua concepção interacionista, Descartes

⁸ A distinção entre os dois tipos de substância é forte o suficiente para garantir a exclusão mútua: se algo é físico, é *não-mental*; se algo é mental, é *não-físico*.

recorreu à *glândula pineal*⁹ e aos *espíritos animais*, segundo ele, aquela é a sede da alma e estes são os responsáveis pela comunicação entre alma e corpo. Vale ressaltar que apesar da diferença ontológica entre alma e corpo, os dois estão intimamente conjugados, formando um *todo coerente*:

A natureza me ensina, também, por esses sentimentos de dor, fome, sede etc., que não somente estou alojado em meu corpo, como um piloto em seu navio, mas que, além disso, lhe estou conjugado muito estreitamente e de tal modo confundido e misturado, que componho com ele um único todo (DESCARTES, 1996, pp. 328-329).

Nenhuma das soluções propostas por Descartes (glândula pineal e espíritos animais) são satisfatórias: por um lado, é contraditório afirmar que a alma está localizada na glândula pineal, já que a *res cogitans* é uma substância *não-extensa e não-espacial*; por outro lado, por mais que se considere que os espíritos animais sejam constituídos de uma substância física bastante delicada, isso ainda deixa em aberto a questão original – como algo que tem massa e está localizado no espaço (espíritos animais) é capaz de interagir com a alma não-física¹⁰? Apesar das inúmeras críticas ao dualismo de substância e do fato de que poucos pensadores acadêmicos o defendem atualmente, pode-se constatar que essa é a concepção com mais adeptos de uma maneira geral, já que as Religiões (se não todas, pelo menos a maioria) defendem uma separação entre a alma/espírito e o corpo¹¹.

O dualismo de substância, por sua vez, não é compatível com a nossa atual visão de mundo científica, já que de acordo com esta o Universo é, em última instância, *totalmente* composto por partículas físicas fundamentais e seus agregados, o que exclui, entre outras coisas, a *res cogitans* cartesiana. Mesmo aceitando o pressuposto do monismo físico, alguns filósofos propuseram que nosso cérebro possui dois tipos de propriedades mutuamente exclusivas, as físicas e as mentais, as quais podem ser exemplificadas pelas propriedades de ter cem bilhões de neurônios e de ver a cor vermelha, respectivamente. Para o dualista de propriedade, as propriedades mentais são *especiais* em dois sentidos: em primeiro lugar, por sua *singularidade biológica*, pois apesar dos objetos físicos serem constituídos basicamente pelos mesmos “blocos de construção”, os átomos, somente algumas estruturas físicas possuem

⁹ “Noto também que o espírito não recebe imediatamente a impressão de todas as partes do corpo, mas somente do cérebro, ou talvez mesmo de uma de suas menores partes, a saber, aquela onde se exerce a faculdade que chamam o senso comum [glândula pineal] (...)” (DESCARTES, 1996, p. 334).

¹⁰ Note que o mesmo poderia ser dito ao se substituir espíritos animais por sinapses e impulsos nervosos.

¹¹ Uma defesa contemporânea do dualismo de substância se encontra em Richard Swinburne (2008). Searle não nega que existam bons filósofos que defendam o dualismo de substância. No entanto, ele alega que a maioria dos dualistas de substância defende tal posição por causa de uma inclinação ou fé religiosa, particularmente a crença de que a alma sobrevive após a morte corpórea, ou seja, a crença em uma vida após a morte (M, p. 41).

propriedades mentais, a saber, cérebros complexos; em segundo lugar, a despeito de serem produzidas por sistemas físicos, propriedades mentais *não são* propriedades físicas, ou seja, elas “jamais podem ser *reduzidas* ou *explicadas* exclusivamente em termos dos conceitos das ciências físicas habituais” (CHURCHLAND, 1998, p. 30. Grifos meus). Portanto, propriedades mentais, como os *qualia*, são *ontologicamente irreduzíveis* ao arcabouço teórico da Física, Química e Biologia. Dito de outra forma, um cérebro humano possui propriedades mentais, mas estas estão *acima e além* do sistema nervoso.

O filósofo Frank Jackson (2004) propôs um importante experimento de pensamento com o objetivo de mostrar a irreduzibilidade dos *qualia*: Mary é uma neurocientista que sabe *tudo* acerca da física e da neurofisiologia das cores, o que inclui o conhecimento sobre as leis da óptica, das reflectâncias de luz, bem como da anatomia e fisiologia das células fotorreceptoras e do córtex visual. Não obstante seu profundo conhecimento científico, Mary vive em um mundo preto e branco e, portanto, nunca teve a *experiência visual* de cores que não estas duas. Todavia, imagine que ao sair de seu laboratório alvinegro a neurocientista *perceba* a cor vermelha de uma maçã. Esta experiência visual trará um *novo tipo de conhecimento* à Mary, a saber, a experiência *qualitativa* de ver vermelho. Ora, se Mary possuía todo o conhecimento físico à sua disposição e, mesmo assim, aprende um *fato novo*, pode-se concluir que a sensação de ver vermelho (fato novo) é não-física, ou seja, é irreduzível do ponto de vista ontológico a qualquer tipo de estado ou propriedade material. Com isso, Jackson quer mostrar que há algo para além da realidade física (propriedades mentais), que é deixado de fora pela explicação físico-biológica da visão – os *qualia* visuais:

The conclusion (...) is that the qualia are left out of the physicalist story. And the polemical strength of the Knowledge argument is that it is so hard to deny the central claim that one can have all the physical information without having all the information there is to have (JACKSON, 2004, p. 765).

Após a exposição dos dois tipos de dualismo, pode-se perguntar: qual é a relação existente entre as duas formas de dualismo? O dualismo de propriedade é uma posição intermediária entre o fisicalismo redutivo e o dualismo de substância: o primeiro afirma que tudo é, em última instância, redutível ao domínio físico, enquanto o outro defende a separação ontológica radical entre fenômenos físicos mentais (MASLIN, 2009, p. 160). Em outras palavras, do ponto de vista ontológico, o dualismo de propriedade está a meio caminho entre o fisicalismo redutivo e dualismo de substância, assemelhando-se mais ao fisicalismo não-redutivo, já que ambos são monismos não-redutivos. Entretanto, se é verdade que é possível defender o dualismo de propriedade sem abraçar o dualismo de substância, o inverso é falso – defender a existência de substâncias ontologicamente distintas *implica* que estas possuem propriedades irreduzíveis entre si, como observa Jaegwon Kim. Segundo ele, o dualismo cartesiano está associado ao dualismo de propriedade, pois as propriedades físicas e mentais correspondem, respectivamente, às substâncias material e imaterial:

Thus, Descartes's dualism combines substance and *property dualism*: two disparate domains of substances and two mutually exclusive families of properties. Arguably, substance dualism entails property dualism: It would be *prima facie* incoherent to accept two fundamentally different types of substance and yet maintain that a single set of properties suffices to characterize all substances (KIM, 1996, p. 211).

Com o passar do tempo, principalmente após a Revolução Científica, houve uma tendência geral, mas não total, de se abandonar explicações baseadas em fenômenos, entidades e causas sobrenaturais. Sendo assim, os conceitos de alma e teleologia, por exemplo, foram perdendo espaço tanto nas teorias científicas como nas filosóficas. Em Filosofia da Mente isso significou um crescente abandono do dualismo de substância em favor de concepções que defendem um monismo fisicalista. Contudo, aceitar o fisicalismo ontológico não é o mesmo que se comprometer com a redução dos fenômenos mentais à microestrutura física. Ou seja, existem tanto fisicalistas que afirmam que as propriedades mentais são redutíveis a propriedades físicas (fisicalismo redutivo) quanto os que defendem a irreduzibilidade ontológica daquelas a estas últimas (fisicalismo não-redutivo). As principais teorias fisicalistas serão examinadas em maiores detalhes a seguir.

Dentro das concepções do fisicalismo redutivo em Filosofia da Mente, a que predominou durante a segunda metade dos anos 40 e por toda a década de 50 do século passado foi o behaviorismo analítico (ou behaviorismo filosófico). Aliás, essa teoria foi tão influente que alguns filósofos consideram o livro *The Concept of Mind*, do behaviorista Gilbert Ryle, como sendo o marco inicial da Filosofia da Mente Contemporânea. O behaviorismo surgiu em contraposição ao dualismo: o problema mente-corpo é apenas um *pseudoproblema* que é gerado a partir da confusão conceitual, podendo, portanto, ser resolvido através de uma adequada análise lingüística. Para Ryle, não existe uma mente que está acima e além do corpo, e é por isso que ele descreve de maneira irônica e pejorativa o dualismo cartesiano como sendo uma doutrina do fantasma (alma) na máquina (corpo).

Para o behaviorista analítico, *todo* o vocabulário da Psicologia Popular – isto é, os termos que dizem respeito às nossas sensações, crenças e desejos – pode e deve ser reduzido a *padrões comportamentais*. Quando Chaves afirma, por exemplo, que sua comida predileta é sanduíche de presunto, pode-se fazer uma análise mais detalhada desse seu estado psicológico e chegar a conclusão que isso significa, em termos comportamentais, o mesmo que: (i) se o professor Girafales lhe perguntar se ele quer comer um sanduíche de presunto, ele diria sim; (ii) se ele for gastar dinheiro para se alimentar, ele consideraria primeiramente a opção de gastá-lo comprando um sanduíche de presunto; (iii) se lhe derem a opção de escolher entre pão com margarina e pão com presunto, ele optará por este último etc. Nota-se, portanto, que

um *mesmo* estado mental é constituído por *várias* disposições comportamentais. Na verdade, estados mentais *nada mais são* do que disposições comportamentais e são, dessa forma, redutíveis a estas, daí o nome da teoria – comportamentalismo (ou, em inglês, *behaviorism*).

A principal objeção ao behaviorismo filosófico é que tal concepção não leva em consideração o aspecto *qualitativo* dos estados mentais: mesmo que se explique a dor ou a experiência de ver vermelho em termos comportamentais, tais como evitar colocar a mão no fogo e frear o carro antes da faixa de pedestre, a sensação qualitativa de sentir dor e enxergar vermelho fica de fora da descrição disposicional do comportamento, o que significa que os *qualia* são *irredutíveis* ao comportamento.

Outro problema surge ao se constatar que não é possível definir de maneira clara um estado mental por meio de um número finito de disposições comportamentais (MASLIN, 2009, pp. 115-116). O condicional (i) do exemplo anterior só será verdadeiro se Chaves *estiver com fome*; o condicional (ii) será verdadeiro somente se ele *quiser* um sanduíche, e não uma fruta; por fim, o condicional (iii) será verdadeiro apenas se Chaves não *acreditar* que o presunto está estragado. No entanto, querer, estar com fome e acreditar também são termos que pertencem ao vocabulário mental, o que significa, em última instância, que os termos da Psicologia Popular não podem ser definidos unicamente em função da observação do comportamento: o desejo de comer um sanduíche de presunto está inextricavelmente ligado a uma rede de outros estados mentais, e aquele estado só tem *sentido* quando se considera o conjunto das relações que se dão no interior desta rede:

(...) não se pode permitir que análise comportamental alguma contenha termos mentais não analisados, e o problema é que, não importando o quão longe possamos ir na tentativa de converter termos mentais em termos de descrições comportamentais, sempre restará um resíduo de itens mentais não analisados, demandando ainda mais análise comportamental (MASLIN, 2009, p. 115. Grifos meus).

Surge, no final dos anos 50, uma nova teoria fisicalista – a teoria de identidade tipo-tipo. Esta teoria, diferentemente do behaviorismo analítico, identifica classes (tipos) de fenômenos mentais a determinados tipos de processos neurobiológicos que ocorrem no cérebro – “A Physicalist (...) will take the states and activities introspected to be all physical states and activities of a continuing physical object: a brain” (ARMSTRONG, 1997, p. 727). Antes de prosseguir, entretanto, vale a pena ratificar que os defensores da teoria de identidade não afirmam que há uma *identidade semântica* entre os termos da Psicologia Popular e da Neurobiologia. Isto é, a frase “estou com dor” (vocabulário da Psicologia Popular) *não* tem o mesmo significado de “minhas fibras C estão sendo estimuladas” (vocabulário

neurocientífico). Porém, argumenta o teórico da identidade de tipo, afirmar que os termos que se referem aos fenômenos mentais não são sinônimos dos que se referem aos processos neurofisiológicos *não implica* que estes sejam ontologicamente distintos: não é necessária uma redução analítica (redução de significados) para se obter sucesso em reduzir uma classe de fenômenos à outra (redução ontológica). Por exemplo, alguém pode tomar banho de chuva, beber água e lavar seu carro sem saber que a água é *nada exceto* moléculas de H₂O, já que o termo “água” não é sinônimo de “moléculas de H₂O”. Isso significa que *não é contraditório* que Cascão afirme que “água não é H₂O”, mas *é falso*, já que o termo da direita (água) é ontologicamente idêntico ao da esquerda (H₂O).

Da mesma forma que a identidade água-H₂O, a identidade psicofísica (tipos de estados mentais = tipos de estados neurobiológicos) se estabelecerá por meio de pesquisa empírica, e não de simples análise conceitual. Para isso, deve-se obter êxito em encontrar uma correlação necessária e suficiente entre um tipo de estado mental e um tipo de estado neurobiológico. Um exemplo de correlação psicofísica é o seguinte: um indivíduo terá dor *se e somente se* suas fibras C estiverem sendo estimuladas – ou seja, se ele estiver com dor, suas fibras C estão disparando e, inversamente, se houver estímulo de suas fibras C, ele estará com dor (MASLIN, 2009, p. 78).

Um dos problemas com a teoria da identidade é que mesmo que se admita a identidade entre estados mentais e estados cerebrais, não se pode afirmar que *todos* os estados cerebrais são mentais, pois existem vários daqueles que possuem *somente* propriedades físicas, como, por exemplo, os estados que estão relacionados ao funcionamento do sistema nervoso autônomo (SNA). Ora, mas se esse é o caso, os outros estados neurobiológicos, isto é, os que são idênticos a estados mentais, devem possuir – *além das propriedades físicas*, que são comuns a todos estados neurobiológicos – propriedades mentais. Isso significa que o teórico da identidade obteve sucesso em evitar o dualismo de substância, mas não o dualismo de propriedade:

Uma mudança na vida mental de uma pessoa envolverá fundamentalmente mudanças físicas ou eventos no cérebro da pessoa e não eventos em uma alma não-física. No entanto, esses eventos físicos possuirão *dois tipos distintos de propriedades*, físicos e mentais (MASLIN, 2009, pp. 85-86. Grifos meus).

O fisicalismo de superveniência é a posição que defende que as propriedades de nível superior (propriedades supervenientes) são *dependentes* ou *determinadas* pelas propriedades de nível inferior (propriedades basais), ou, dito de outra forma: duas microestruturas físicas indiscerníveis possuem as *mesmas* propriedades de nível superior. Kim (1996, p. 222) cita um

exemplo ilustrativo no campo da Estética: quando um escultor trabalha em uma peça de mármore com o intuito de criar um busto de Aristóteles, por exemplo, ele deve cortá-la, burilá-la, lixá-la etc, até o que antes era uma peça bruta se transforme numa escultura com um formato e textura característicos (propriedades físicas). Todavia, nota-se que as propriedades estéticas da escultura foram determinadas pelas propriedades físicas da mesma, pois ao se finalizar o trabalho físico não resta nenhum trabalho estético adicional. Se duas pessoas esculpisse de modo independente dois bustos de Aristóteles e, ao final do trabalho, as duas esculturas fossem fisicamente indistinguíveis – mesma forma, mesmo tamanho, mesma cor, mesma textura – tais esculturas compartilhariam, *necessariamente*, as *mesmas* propriedades estéticas, como a expressividade e a beleza, por exemplo. Nesse caso, dir-se-ia que as propriedades estéticas são supervenientes em relação às propriedades físicas. De maneira análoga, superveniência em Filosofia da Mente significa o seguinte: se dois indivíduos têm os mesmos estados cerebrais (propriedades basais), eles possuem os mesmos estados mentais (propriedades supervenientes) e que, além disso, não pode haver mudança de estados mentais sem que haja uma mudança dos estados neurobiológicos subjacentes.

A superveniência, como será visto adiante, é utilizada tanto por filósofos que defendem a posição reducionista como pelos anti-reducionistas. Kim, que se enquadra no primeiro grupo, enfatiza que o conceito de superveniência é neutro em relação a essas posições:

Although supervenience perhaps doesn't imply reducibility, it need not be taken to imply irreducibility either – that is, supervenience suffices for the purposes of the nonreductive physicalist if it is consistent with both reducibility and irreducibility (KIM, 1996, p. 223).

Por fim, a última das correntes que será apresentada é o funcionalismo, o qual consiste numa das vertentes do fisicalismo não-redutivo com mais adeptos, sobretudo dentro do projeto de pesquisa da *inteligência artificial* (IA) e das *Ciências Cognitivas*. A mais importante tese do funcionalismo é a da *realização* (ou *realizabilidade*) *múltipla*: inúmeros substratos físicos diferentes podem realizar ou instanciar a mesma função. Daí a famosa analogia entre nós e os computadores: nossa mente está para o cérebro da mesma maneira que o programa de computador (software) está para o hardware¹². Ora, afirmar que o mental é

¹² Como foi dito acima, para um funcionalista o número de diferentes bases físicas capazes de produzir uma mente é praticamente ilimitado. O mais famoso desses possíveis substratos físicos são os circuitos de silício que correspondem ao hardware dos computadores digitais. A posição que defende que computadores podem ter uma mente é chamada de *funcionalismo de computador* ou *inteligência artificial forte*: o cérebro está organizado de tal maneira que surgem estados mentais a partir de seu funcionamento; portanto, caso circuitos de silício estejam

multiplamente realizável em diversas estruturas físicas implica a irredutibilidade daquele a estas: se um estado mental M é instanciado tanto em redes neurais quanto em circuitos de neurônios, nenhum desses sistemas físicos são *necessários* para causar M, apesar de M *depende* da interação dos componentes do micro-sistema para existir.

Nota-se que além de muitos adeptos no mundo acadêmico, tal concepção também exerce grande influência no público em geral, já que um dos principais temas da ficção científica do século XX, seja na literatura ou no cinema, é o que aborda a IA e suas conseqüências, positivas e negativas, para a humanidade. O maior temor em relação às conseqüências negativas diz respeito a um tipo de “revolta” das máquinas contra os humanos: em algum momento de um futuro próximo, as máquinas – tanto na forma de computadores como de robôs – tornar-se-iam *realmente* inteligentes ou, de maneira mais correta, possuiriam várias características que até então eram tidas por muitos como exclusivas de nossa espécie. Com isso, em algumas histórias do gênero, as máquinas se *sentem* exploradas e acabam se rebelando contra seus criadores. Foi pensando nisso que um dos maiores autores de ficção científica de todos os tempos, Isaac Asimov, criou as Três Leis da Robótica¹³:

- 1) Um robô não pode prejudicar um ser humano ou, por omissão, permitir que o ser humano sofra dano;
- 2) Um robô tem de obedecer às ordens recebidas dos seres humanos, a menos que contradigam a Primeira Lei;
- 3) Um robô tem de proteger sua própria existência, desde que essa proteção não entre em conflito com a Primeira e Segunda Leis.

Após o que foi dito, resta uma importante pergunta de cunho teórico-prático: algum dia a IA da ficção científica seguirá o caminho de outras grandes tecnologias de nossa civilização, como o submarino, isto é, deixará de ser *sonho* para se tornar realidade? Marvin Minsky, um dos expoentes da IA, acha que é só questão de tempo para as máquinas não só atingirem as capacidades humanas, mas mesmo ultrapassá-las, chegando ao ponto de afirmar que um dia aquelas terão orgulho de nós, seus criadores¹⁴. Para Minsky, “cérebros são computadores formados de carne”.

organizados de modo semelhante, estes instanciarão uma mente. Sendo assim, o que importa para que um estado seja funcional é sua *relação* entre determinados dados de entrada (inputs) e suas respectivas respostas (outputs).

¹³ As Três Leis da Robótica foram mencionadas pela primeira vez no livro *Eu, Robô*, de 1950. Em um livro posterior, *Os robôs do amanhecer*, Asimov lança outra Lei, que é a mais básica de todas, a saber, a Lei Zero: um robô não pode fazer mal à humanidade e nem, por omissão, permitir que ela sofra algum mal.

¹⁴ Na verdade, em outras declarações Minsky chega a cogitar um futuro bem menos digno para a nossa espécie, já que talvez as máquinas não nos queiram nem mesmo como animais de estimação! É da seguinte maneira que Guitta Pessis-Pasternak apresenta o entusiasta da IA: “Marvin Minsky, famoso informaticista do MIT (Massachusetts Institute of Technology), e um dos fundadores da área de Inteligência Artificial americana,

Um dos maiores críticos desse modelo de IA forte¹⁵ é John Searle, e eu apresentarei no próximo capítulo o experimento de pensamento do quarto chinês (proposto por ele), que tem o intuito de mostrar que *todo* o projeto da IA forte está condenado, *desde o princípio*, ao fracasso.

prevê para as décadas que virão o desenvolvimento de máquinas dotadas de “inteligência geral de um ser humano médio”. E, aos poucos, essas obras-primas “assumirão a sua própria educação. Em alguns anos, elas alcançarão o nível do gênio, e depois o seu poder será incalculável”; a tal ponto que “*teremos sorte se elas resolverem nos conservar como animais domésticos*” (PESSIS-PASTERNAK, 1992, p. 207. Grifos meus).

¹⁵ Existem diferenças entre inteligência artificial fraca e inteligência artificial forte. A IA fraca parte da tese que os computadores podem nos ajudar na compreensão do ser humano, como, por exemplo, no que diz respeito à cognição humana. Mas trata-se *apenas* de simulação, e não de *realização* propriamente dita. Já para a IA forte, simular uma mente é o *mesmo* que ter uma mente. Portanto, qualquer máquina que passe no teste de Turing tem uma vida mental semelhante a nossa.

2. O EXPERIMENTO DE PENSAMENTO DO QUARTO CHINÊS: A CRÍTICA DE JOHN SEARLE À INTELIGÊNCIA ARTIFICIAL FORTE

2.1) O Teste de Turing

Em 1950, Alan Turing, que, juntamente com John von Neumann, é considerado o pai do computador moderno, publicou um artigo¹⁶ em que tentava responder a questão “podem as máquinas pensar?”¹⁷. Turing inicia seu artigo sobre Computação e Inteligência propondo o *jogo da imitação* como substituto para a pergunta “podem as máquinas pensar?” (TURING, 1996, p. 21). Um homem (A), uma mulher (B) e um interrogador ou interrogadora (C) participam do jogo da imitação. O interrogador está separado dos outros dois jogadores. Cada pessoa tem um objetivo diferente no jogo: C faz perguntas aos outros dois participantes, e seu objetivo é identificar corretamente qual o homem e qual a mulher; o objetivo de A, por sua vez, é induzir C ao erro; por fim, a meta de B é dar respostas que auxiliem C na correta identificação. Portanto, A poderia fornecer respostas¹⁸ típicas femininas, como dizer, por exemplo, que tem cabelo comprido e usa vestido. Por sua vez, B poderia dizer para não acreditar em A, e que ela era a verdadeira mulher. Caso tenha sucesso, C acabará o jogo com o seguinte veredicto: A é homem e B é mulher, situação esta em que B também atinge seu objetivo. Por outro lado, A sairá vitorioso perante seus companheiros de jogo se C disser que A é mulher e B é homem.

Após propor o jogo da imitação, Turing pergunta “o que acontecerá quando uma máquina ocupar o lugar de A nesse jogo?”, ou seja, o que ocorrerá se um computador tentar se passar por ser humano (TURING, 1996, p. 22). Para Turing, resolver essa questão corresponde a responder se as máquinas podem pensar: isso significa que um programa de computador é inteligente caso não fosse possível ao interrogador distinguir, de maneira estatisticamente significativa, suas respostas das respostas fornecidas pelo jogador humano¹⁹.

¹⁶ O artigo chama-se “Computing Machinery and Intelligence” e foi publicado em 1950, no prestigioso periódico britânico *Mind*. “Computação e Inteligência” (TURING, 1996) é a tradução brasileira deste artigo.

¹⁷ Na verdade, o projeto do computador digital é antigo, tendo sido idealizado entre 1828 e 1839 por Charles Babbage. Apesar das idéias essenciais intuídas por Babbage, o precursor do computador digital, o qual ele chamou de *máquina analítica*, não chegou a ser construído por causa da insuficiente tecnologia da época (TURING, 1996, p.29).

¹⁸ Para que haja o mínimo de influência, as respostas fornecidas a C devem ser digitadas (ou datilografadas, como melhor convinha à tecnologia mais comum na época, a máquina de escrever).

¹⁹ Vale a pena ressaltar que Turing não estava afirmando que *todos* os computadores digitais seriam capazes de passar no teste. Além disso, ele não disse que os computadores de sua época, isto é, da segunda metade do século

Isso ficou conhecido como teste de Turing, e até hoje é oferecido um prêmio para quem desenvolva um programa de computador que passe no teste.

Turing estava ciente do conhecimento precário sobre as bases neurobiológicas da consciência, mas, mesmo assim, pensava que era possível chegar a uma conclusão em relação ao teste de Turing sem o auxílio das futuras descobertas neurocientíficas sobre a consciência:

Não quero dar a impressão de que penso não existir nenhum mistério no que diz respeito à consciência. Existe, por exemplo, algo assim como que um paradoxo vinculado às tentativas de localizá-la. Mas não acredito que tais mistérios tenham de ser necessariamente resolvidos antes de podermos responder à pergunta que nos preocupa neste artigo (TURING, 1996, p. 41).

2.2) O Debate Acerca do Quarto Chinês

Trinta anos após o artigo *Computação e Inteligência*, Searle publica um trabalho²⁰ no qual propõe seu famoso experimento de pensamento do quarto chinês, que é uma crítica à inteligência artificial forte (IA forte). Segundo Moural (2003), é possível fazer a distinção entre a) o experimento de pensamento propriamente dito, b) o argumento teórico relacionado a este e, por último, c) algumas questões de cunho periférico.

2.2.1) O Experimento de Pensamento do Quarto Chinês

A idéia do quarto chinês surgiu por causa de um software, especializado em responder perguntas relacionadas a uma história, desenvolvido por Schank e Abelson. Apesar desse software em particular ter sido o ponto de partida de Searle, o experimento do quarto chinês não se limita ao mesmo: o quarto chinês se aplica igualmente bem a qualquer máquina de Turing, portanto a qualquer software pretensamente inteligente. O quarto chinês tem o intuito de atacar duas teses centrais da IA (MCP, p. 65), a saber:

1) Que é lícito afirmar categoricamente que a máquina realmente *compreende*²¹ a história e é capaz de prover respostas adequadas às perguntas que lhe são feitas;

XX, teriam bom desempenho no jogo, mas, sim, que podíamos *conceber* computadores com essa capacidade (TURING, 1996, pp. 25-26).

²⁰ “Minds, Brains, and Programs”, publicado em *The Behavioral and Brain Sciences*, Volume 3, (1980). “Mentes, Cérebros e Programas” (MCP) é a tradução brasileira deste artigo.

²¹ Compreensão nesse sentido não significa que a máquina é apenas capaz de *imitar* o comportamento humano, mas que, de fato, a máquina *possui* estados mentais semelhantes aos nossos.

2) Que a máquina e seu programa são capazes de *explicar* a capacidade humana para se entender uma história e responder às perguntas sobre a mesma.

O experimento de pensamento é o seguinte. Searle é um falante de inglês que não compreende nada de chinês²² e está trancado dentro de um quarto. Searle recebe algumas tiras de papel escritas em chinês por meio de uma abertura. Apesar de não compreender o que está escrito, Searle tem acesso a um livro que contém um conjunto de regras, escrito em inglês, que correlaciona a tira de papel que ele recebeu a outro conjunto de tiras de papel, também escritas em chinês, que está à sua disposição no interior do quarto. As instruções do livro são do tipo “ao receber tira contendo x, coloque pela abertura da porta a tira contendo y”. Vale a pena ressaltar que tanto x como y são mensagens com caracteres em chinês, e que Searle só entende a instrução, escrita em inglês, que correlaciona x com y. A correlação entre o símbolo recebido e o símbolo enviado é puramente *formal*, isto é, Searle os identifica unicamente pelos seus formatos²³. Sem que Searle *saiba*, ele está respondendo corretamente, do ponto de vista de um observador falante de chinês que está do lado de fora do quarto, às perguntas que lhe foram fornecidas. Portanto, o quarto chinês se comporta de maneira semelhante a um computador: a tira que Searle recebe corresponde ao dado de entrada ou input; o livro de instruções é análogo ao programa ou software de computador; o próprio Searle faz o papel da unidade central de processamento (CPU); e, por fim, a tira de papel enviada por Searle corresponde ao dado de saída ou output. O ponto central do experimento de pensamento é que Searle, imitando o funcionamento de um computador, pode responder corretamente às perguntas em chinês *sem compreender* nada de chinês. Sendo assim, Searle ataca a tese 1 da IA mostrando que um programa de computador não possui compreensão genuína, já que assim como Searle não entende chinês, o computador também não entende a história sobre a qual ele responde, por mais que essas respostas sejam corretas e semelhantes às que seriam dadas por um ser humano. Portanto, manipular símbolos não é o mesmo que compreender símbolos, isto é, a manipulação de uma sintaxe é condição necessária, mas não suficiente para a existência de uma semântica. Podemos ver que o quarto chinês passa no teste de Turing, mesmo sem compreender uma única palavra em chinês.

Em relação à tese 2, Searle afirma que compreender o funcionamento do software não seria de grande contribuição para explicar o nosso próprio entendimento de histórias, pois

²² O que chamamos de língua chinesa é na verdade formada por uma família de línguas pertencentes ao ramo sino-tibetano. O mandarim e o cantonês são dois dos dialetos mais conhecidos da língua chinesa.

²³ “As regras são em inglês e eu as compreendo tão bem como qualquer outro falante nativo de inglês. Isto me possibilita relacionar um conjunto de símbolos formais com o outro, e o que entendo por formal aqui é que posso *identificar os símbolos por seus formatos* (MCP, 66. Grifos meus).

apesar do programa responder corretamente às perguntas, ele ainda assim não compreende nada da história em questão.

Para fortalecer ainda mais sua crítica, Searle pede que modifiquemos um pouco o experimento acima: as tiras de papel, tanto das perguntas como das respostas, estarão grafadas em inglês. A diferença crucial deste segundo cenário em relação ao primeiro é que enquanto neste havia *apenas* a manipulação formal de símbolos, naquele há, além da mera manipulação simbólica, *compreensão* dos símbolos manipulados. Dito de outra forma, utilizando o jargão da Lingüística: no primeiro cenário há somente sintaxe, diferentemente do segundo, no qual está presente tanto a sintaxe como a semântica. Conclui-se, então, que apesar de um computador poder *imitar* o nosso comportamento²⁴, ele *não funciona* como nossa mente – aliás, ele nem possui uma mente –, já que apenas manipula formalmente os símbolos sem ter qualquer conhecimento sobre seu significado.

2.2.2) O Argumento Relacionado ao Quarto Chinês

Searle pretende demonstrar, através de seu argumento, que a implementação do programa correto é *insuficiente*, por si só, para gerar o entendimento, já que não há intencionalidade na pura manipulação formal de símbolos. O argumento de Searle é composto de quatro premissas e quatro conclusões²⁵ derivadas destas. As premissas são:

P1) O cérebro causa a mente (tese da *redução causal* dos estados mentais a estados cerebrais, a qual é amplamente sustentada pelas evidências provenientes das neurociências);

P2) A sintaxe não é suficiente para a semântica (essa premissa constitui uma verdade conceitual);

P3) Programas de computador são totalmente definidos por sua estrutura formal ou sintática (essa premissa é verdadeira por definição);

P4) A mente possui conteúdos semânticos (segundo Searle, isso é um fato óbvio).

Quatro conclusões são geradas a partir das premissas acima, quais sejam:

C1) Nenhum programa é, por si mesmo, suficiente para produzir uma mente. Isto é, além de programas não serem mentes, eles não são suficientes, por si próprios, para gerarem uma mente (derivada de P2, P3 e P4);

²⁴ Como já foi dito, uma máquina que se comporte, por meio das respostas fornecidas a um interrogador, de maneira indistinta em relação a um ser humano passa no teste de Turing. Isso significa, para um adepto da IA forte, que tal máquina teria uma mente com conteúdos mentais semelhantes aos nossos, isto é, ela seria capaz, entre outras coisas, de compreender verdadeiramente uma história.

²⁵ Conferir Moyal, 2003, pp. 219-220.

C2) O modo como o funcionamento cerebral causa a mente não pode ser unicamente em decorrência da implementação de um programa de computador (derivada de P1 e C1);

C3) Qualquer coisa que cause uma mente tem que ter capacidades causais pelo menos equivalentes àquelas do cérebro (derivada diretamente de P1);

C4) A implementação de um programa de computador por si mesma não seria capaz de dotar uma máquina de estados mentais semelhantes aos nossos, já que esta teria que ter capacidades causais equivalentes às capacidades do nosso cérebro (derivada de C1 e C3).

Com o passar do tempo, Searle reformula, de maneira cada vez mais concisa, seu argumento. Searle apresenta em seu livro *o Mistério da Consciência* (pp. 38-39) o núcleo central do argumento, o qual consiste apenas das premissas P2, P3, P4 e da conclusão C1:

1. Programas são totalmente sintáticos [P3].

2. As mentes têm uma capacidade semântica [P4].

3. a Sintaxe não é a mesma coisa que a semântica, nem é, por si só, suficiente para garantir um conteúdo semântico [P2].

Conseqüentemente, programas não são mentes [C1].

2.2.3) Questões Periféricas

Para a IA forte a implementação de um programa correto²⁶ em *qualquer tipo* de hardware²⁷ bastaria para produzir uma mente. Searle afirma que o slogan da IA forte é o seguinte: a mente está para o cérebro assim como o software está para o hardware. Searle pretende mostrar que o programa de pesquisa da IA forte está fadado ao fracasso, pois esse programa leva a conclusões absurdas.

Pode-se afirmar que a primeira formulação do naturalismo biológico apareceu na resposta de Searle aos seus críticos já em 1980²⁸, a qual pode ser resumida na seguinte passagem:

²⁶ A expressão programa correto significa, neste contexto, um software que descreva formalmente os estados mentais através de um processamento de informação. Conferir Moural, 2003, p. 221.

²⁷ Para a corrente funcionalista, o software tem *primazia* sobre o hardware, pois aquele independe da constituição física deste. Dito de outra maneira, a *organização* dos elementos de um sistema é mais importante do que a *constituição física* destes elementos. Daí se conclui que circuitos de silício, só para citar um exemplo, organizados de maneira formalmente semelhante aos neurônios de nosso cérebro serão responsáveis, causalmente falando, pela produção de uma mente similar à nossa. Tese esta, como já foi visto, que Searle combate veementemente, já que, segundo ele, a organização destes circuitos de silício não reproduziriam as capacidades causais do cérebro. Ver o argumento acima exposto, notadamente as conclusões C3 e C4.

²⁸ O artigo *Minds, brains, and programs* foi seguido por uma série de objeções dos mais variados críticos de Searle, as quais, por sua vez, foram respondidas por ele. Tanto as objeções como as respostas foram compiladas

Mental states are as real as any other biological phenomena. They are both caused by and realized in the brain. That is no more mysterious than the fact that such properties as the elasticity and puncture resistance of an inflated car tire are both caused by and realized in its microstructure (II, p. 455).

Searle utiliza a resposta para ratificar seu argumento contra a IA forte e para esclarecer alguns outros pontos de sua concepção. A distinção que ele faz, por meio de exemplos, entre *intencionalidade intrínseca* e *atribuição de intencionalidade relativa ao observador* é de vital importância para fundamentar sua crítica ao funcionalismo de computador²⁹. A análise de duas sentenças simples serve de auxílio para explicitar essa diferença conceitual:

- 1) Lula tem a esperança de que Dilma vencerá a eleição presidencial em 2010;
- 2) A palavra “Lula” se refere ao atual presidente do Brasil.

Na primeira sentença, a crença de Lula possui, independentemente dela vir ou não a se confirmar, uma intencionalidade intrínseca. Via de regra, estados mentais como crenças e desejos são, de acordo com Searle, intrinsecamente intencionais, e só quem os possui são os organismos que têm um sistema nervoso desenvolvido o suficiente para produzir e sustentar a consciência³⁰: “I say that the operation of the brain is *causally sufficient* for the intentionality” (II, p. 452. Grifos meus). Por outro lado, a intencionalidade que a palavra “Lula” (segunda sentença) possui é meramente derivada³¹, já que tal palavra é desprovida de qualquer tipo de estado mental. Portanto, quando se afirma que um termostato *sabe* que deve diminuir a temperatura ou que um programa de computador *compreende* os erros de português que corrige, não passa, em ambos os casos, de uma intencionalidade que o sujeito (observador)

em um único número da revista *The Behavioral and Brain Sciences* (1980), 3, 417-457. A resposta de Searle é intitulada *Intrinsic Intentionality* e se encontra nas páginas 450-457. O capítulo 3 tratará em detalhes as teses assumidas pelo naturalismo biológico, que é a concepção positiva de Searle sobre a mente.

²⁹ O termo técnico *intencionalidade* não deve ser confundido com a palavra *intenção* da linguagem coloquial. Intencionalidade diz respeito “ao aspecto dos estados mentais pelo qual eles são praticamente dirigidos para objetos e situações outros que não eles próprios” (MLS, p. 95). Dois outros termos possíveis para significar esse aspecto dos estados mentais são *direcionalidade* (“os estados mentais são dirigidos para”) ou *sobredade* (tradução literal da palavra *aboutness*, que significa que “os estados mentais são sobre algo”). Um estado intencional pode ser dirigido incorretamente a um objeto ou, até mesmo, ser sobre um objeto inexistente: a criança que acredita que o gordinho vestido de vermelho e com longa barba branca é o Papai Noel é um exemplo do primeiro caso (ele na verdade é o seu pai barrigudo disfarçado), enquanto que a crença da mesma criança na Fada do Dente é um exemplo do segundo tipo. Searle usa uma boa analogia para explicar tais casos: “Assim como uma flecha pode ser lançada contra um alvo e errar, ou ser lançada mesmo que não haja alvo, do mesmo modo um estado intencional pode ser dirigido a um objeto e errar de direção, ou falhar completamente porque não há objeto nenhum ali” (MLS, p. 95).

³⁰ “There are not two kinds of intentional mental states; there is *only one* kind, those that have intrinsic intentionality” (II, p. 452. Grifos meus).

³¹ Vale a pena ressaltar que para Searle a intencionalidade da linguagem é *derivada* da intencionalidade *intrínseca* dos estados mentais, e é por isto que ele considera a Filosofia da Linguagem um ramo da Filosofia da Mente: “The intentionality of language has to be explained in terms of the intentionality of the mind and not conversely” (M, p. 161).

atribui a tais artefatos, ou seja, é uma intencionalidade *como se* e que deve ser vista metaforicamente: o termostato funciona *como se* soubesse que o ambiente está quente e aciona o refrigerador, da mesma maneira que software corrige erros ortográficos *como se* conhecesse a língua portuguesa. É justamente pelo funcionalismo não perceber essa importante distinção que Searle o acusa de ser uma concepção erigida sobre um erro conceitual, o que inviabilizaria desde o início tal projeto de pesquisa: “Functional attributions are always observer relative. There is no such thing as an intrinsic function, in the way that there are intrinsic intentional states” (II, p. 452). Em outras palavras, Searle utiliza o conceito de intencionalidade para criticar a tese de que a simples implementação de um programa correto seria suficiente para causar uma mente, já que sua teoria do significado é fundamentada na intencionalidade, mais precisamente, na sua concepção da intencionalidade intrínseca dos estados mentais. A maioria³² dos estados mentais é intencional. Estados intencionais, por sua vez, não são definidos por sua forma – como o são os programas –, mas pelos seus conteúdos. Já a manipulação formal de símbolos que é característica dos programas de computador não possui qualquer tipo de intencionalidade intrínseca, sendo, portanto, totalmente desprovida de significado. Como Searle assinala em diversas passagens de seus trabalhos³³: os programas têm apenas sintaxe, mas não semântica.

Apesar da diversidade de propostas, pode-se afirmar que a maioria dos cientistas e filósofos atuais defende *alguma*³⁴ versão fisicalista³⁵ em se tratando da relação existente entre

³² Searle relaciona intencionalidade e consciência da seguinte maneira: “Not all consciousness is intentional, and not all intentionality is conscious, but there is a very serious and important overlap between consciousness and intentionality, and later we will see that there are, in fact, logical connections between the two: mental states that are in fact unconscious have to be the kind of thing that could in principle become conscious” (M, pp. 138-139).

³³ “(...) a manipulação de símbolos formais por si só não tem intencionalidade: eles não têm significado, eles nem mesmo são manipulações de *símbolos*, uma vez que esses símbolos não significam nada. No jargão linguístico, eles têm apenas sintaxe, mas não semântica” (MCP, p. 86).

³⁴ Convém lembrar que nem toda concepção fisicalista é *reducionista*. Pode-se citar o *monismo anômalo* como exemplo de uma teoria fisicalista (os eventos mentais são *idênticos* a eventos cerebrais) não-reducionista (não existem leis psicofísicas, isto é, *não há* uma redução dos eventos mentais a eventos neurofisiológicos). Para maiores detalhes, conferir o artigo de Davidson, *Mental Events* (2004). No entanto, para Searle *todo* tipo de materialismo é *reducionista*: “(...) todas as formas contemporâneas de materialismo que conheço têm o objetivo comum de tentar se livrar dos fenômenos mentais em geral e da consciência em particular, *conforme entendida normalmente*, reduzindo-os a alguma forma física ou material (MLS, p. 51. Grifos meus). É bom ratificar que o que Searle chama de *entendimento normal* da consciência é a sua própria caracterização da mesma, a saber: fenômenos conscientes são fenômenos *internos, qualitativos e subjetivos*. Ora, já que Searle faz a distinção entre ontologia objetiva e subjetiva, e *apenas* os fenômenos mentais possuem uma perspectiva de primeira pessoa, sua teoria é anti-reducionista no que se refere à ontologia. Porém, essa caracterização que Searle faz das correntes fisicalistas como sendo *reducionistas* é, no mínimo, problemática, pois, como foi dito acima, existe o fisicalismo não-reducutivo. Em outras palavras, ao contrário do que pensa Searle, é possível ser fisicalista ao mesmo tempo em que se adota uma postura anti-reducitiva.

³⁵ Searle chega ao ponto de fazer a seguinte pergunta: “o que fez com que o ‘materialismo’ parecesse ser o único enfoque racional da filosofia da mente?” (RM, p. 10). Como será visto no segundo capítulo, Searle defende que sua teoria, o Naturalismo Biológico, não se enquadra nem como dualismo nem como fisicalismo, já que apesar das diferenças entre essas duas correntes elas incorrem no mesmo erro básico, a saber, se utilizam de uma

mente e sistema nervoso, particularmente, entre mente e cérebro. Segundo Searle, são muitas as evidências provenientes das Neurociências que apóiam a premissa 1 (P1) de seu argumento, ou seja, que a mente é *causada* pelo cérebro³⁶. Searle se utiliza dessa premissa para apontar mais uma falha na analogia da mente como um programa, já que esta é *produzida* pelo cérebro, enquanto aquele não é um produto do hardware.

Searle acusa os partidários da IA forte de serem, apesar de soar de maneira contraditória, behavioristas e dualistas, além de confundirem simulação com duplicação. O behaviorismo alega que os conceitos mentais podem ser redefinidos em termos de comportamento, e que sentenças que descrevem fenômenos mentais podem ser traduzidas em sentenças que descrevem o comportamento. Isso é bem característico do teste de Turing, pois o que está sendo avaliado são as respostas (outputs) que a máquina fornece para as perguntas (inputs) feitas pelo interrogador. Portanto, para o defensor do funcionalismo de computador, afirmar que uma máquina se *comporta* como um humano é o mesmo que dizer que ela possui estados mentais semelhantes aos do homem. O interrogador constatará se a máquina se comporta ou não como um ser humano a partir de sua perspectiva, isto é, a partir de um ponto de vista de terceira pessoa. Contudo, Searle afirma que o “comportar-se como humano” não é garantia para ter uma “mente semelhante à humana”³⁷: “*ontologicamente falando, comportamento, papel funcional e relações causais são irrelevantes para a existência de fenômenos mentais conscientes*”³⁸ (RM, p. 103). Já que Searle defende que os estados mentais têm uma ontologia de primeira pessoa, e que o ponto de vista de primeira pessoa é, em termos epistêmicos, bem diferente do ponto de vista de terceira pessoa, pode-se perguntar qual o *fundamento empírico* de que ele se utiliza para afirmar que outros sistemas³⁹ possuem estados conscientes. Sua resposta a esta questão é que os métodos empíricos utilizados por uma pessoa (ponto de vista de terceira pessoa) para inferir e estudar os fenômenos mentais (ontologia de primeira pessoa) de um outro sistema consciente “se fundamentam num princípio prático que empregamos alhures na ciência e no dia-a-dia: *mesmas causas-mesmos-*

terminologia obsoleta, do século XVII, que contrapõe o mental ao físico. Sendo assim, ele defende que tanto o dualismo quanto o fisicalismo se apóiam no que ele designa de *dualismo conceitual*. Searle afirma que “a tradição materialista é sólida, ubíqua e, ainda assim, evasiva”, sendo, erroneamente, chamada de ciência (RM, p. 18).

³⁶ É bom deixar claro que a tese de que a mente é *causada* pelo cérebro não se baseia na concepção ortodoxa de causalidade. Searle defende uma noção de causalidade não-humana, segundo a qual a causa (processos cerebrais) e o efeito (estados mentais) ocorrem simultaneamente. Para maiores detalhes, conferir a 2ª tese do próximo capítulo e a seção intitulada Causação mental no Naturalismo Biológico, do 4º capítulo.

³⁷ Isso está intimamente ligado à confusão entre *simulação* e *duplicação* que será apresentada logo a seguir.

³⁸ Searle propõe, a partir do experimento mental do cérebro de silício (RM, p. 98-105), o *princípio da independência de consciência e comportamento*, o qual afirma que o comportamento não é nem condição suficiente nem condição necessária para os fenômenos mentais.

³⁹ Tanto pessoas quanto outros animais, como cachorros, chimpanzés e golfinhos.

efeitos, e causas semelhantes-efeitos-semelhantes” (RM, p. 112). Portanto, para Searle, a descoberta de que outros organismos possuem estados mentais não se deve apenas à observação de seu comportamento, mas, sim, a uma *conexão* que existe entre o comportamento e a estrutura causal destes indivíduos (RM, p. 115). Desconsiderando-se casos patológicos e pequenas diferenças individuais, como os seres humanos compartilham a mesma fisiologia (mesma causa), pode-se inferir, a partir do princípio prático supracitado, que eles possuem também experiências conscientes idênticas (mesmo efeito) às nossas. Devido a isso, Searle afirma que, do ponto de vista prático ou da vida real, não existe o que os filósofos da mente chamam do *problema de outras mentes*, pois

não sustentamos uma “hipótese”, “crença” ou “suposição” de que outras pessoas sejam conscientes. Ao contrário, temos determinadas maneiras de nos comportar advindas do *Background*, determinadas capacidades possibilitadas pelo *Background*, e essas são constitutivas de nossas relações com a consciência de outras pessoas (RM, p. 116).

De maneira semelhante, apesar de um cachorro, um chimpanzé ou um golfinho pertencerem a espécies diferentes da nossa, em muitos aspectos relevantes, para a questão tratada aqui, eles são fisiologicamente semelhantes ao *Homo sapiens*, daí a atribuição de fenômenos mentais similares aos nossos a esses organismos: causas semelhantes-efeitos semelhantes.

Outra crítica de Searle à IA forte se deve a afirmação que, em princípio, qualquer substrato material devidamente organizado possui uma mente. Searle rejeita essa tese: já que o cérebro causa a mente (P1), as propriedades físicas do cérebro são relevantes para a produção das propriedades mentais. Isso significa que se algum substrato físico, que não seja o sistema nervoso, causa uma mente, ele tem que ter capacidades causais equivalentes às do cérebro (C3):

(...) todo sistema capaz de causar a consciência e outros fenômenos mentais deve ter capacidades causais de criá-la equivalentes às capacidades biológicas iniciais do cérebro animal, seja do cérebro humano, seja do de outros animais. Do fato de o cérebro criá-la causalmente, é fácil chegar à conclusão lógica de que os poderes causais limiares de algum outro sistema capaz de criá-la causalmente devem ser equivalentes aos do cérebro (MA, pp. 109-110).

Ora, como já foi dito, os *únicos* estados que possuem intencionalidade intrínseca são os estados mentais, e estes são ontologicamente irreduzíveis a estados neurobiológicos⁴⁰. Portanto, o fato de que estados mentais são intrinsecamente intencionais está relacionado à irreduzibilidade ontológica dos mesmos, o que se pode constatar facilmente quando Searle defende “que existem fenômenos mentais intrínsecos que não podem ser reduzidos à outra coisa ou eliminados por algum tipo de redefinição” (INT, p. 363).

⁴⁰ Conferir a tese 3 do próximo capítulo.

Em resumo, a única maneira que um sistema tem para produzir intencionalidade é duplicar a capacidade causal do sistema nervoso. Em outras palavras, a duplicação da capacidade causal do sistema nervoso é uma *condição necessária* para a produção de intencionalidade (II, p. 453).

Por fim, Searle alega que há uma confusão entre *simulação* e *duplicação* que permeia o discurso dos defensores da IA forte. Do ponto de vista de um observador externo, Searle se comporta, em seu experimento de pensamento, como um falante de chinês, apesar de não compreender uma única palavra neste idioma. Portanto, pode-se dizer que Searle simula a capacidade lingüística de um falante de chinês, mas não a duplica, já que a capacidade lingüística pressupõe a compreensão dos símbolos manipulados e esta não está presente. Essa distinção entre simular e duplicar fica melhor explicitada ao se constatar que um meteorologista não fica com os nervos à flor da pele ao empregar programas de computador para simular tempestades e terremotos. A calma do meteorologista é totalmente justificada em sua crença ou conhecimento de que simular tais intempéries da natureza *não implica* reproduzi-las. Da mesma maneira, Searle pergunta o porquê de alguém supor “que uma simulação computacional da compreensão de fato entenderia alguma coisa” (RM, p. 88). Para simular a cognição, por exemplo, basta que o programa atue na transformação de um input em um output adequado, e isso é totalmente exequível. Por conseguinte, um programa de computador se *limita* a simular fenômenos naturais e mentais, donde conclui-se que “confundir simulação com duplicação é o mesmo erro, seja com dor, amor, cognição, incêndio ou tempestade” (RM, p. 89)

Já que o quarto chinês pretende mostrar que a compreensão não surge *unicamente* por causa da implementação de um programa correto, pode-se questionar sobre o porquê somos capazes de compreender. A resposta de Searle para tal pergunta é que há uma ligação entre as capacidades causais e a intencionalidade, já que alguns organismos possuem

(...) certa estrutura biológica (física e química) e esta estrutura, em termos causais, é capaz, sob certas condições, de produzir a percepção, a ação, a compreensão, o aprendizado e outros fenômenos intencionais. Parte do núcleo deste argumento é que *só algo que tenha estes poderes causais pode ter intencionalidade* (MCP, p. 85. Grifos meus).

2.3) O Novo Argumento de Searle

Searle propõe outro argumento, que visa atacar frontalmente o *Cognitivismo*⁴¹, ou seja, a linha de pesquisa que defende, entre outras coisas, “que o cérebro é um computador e os processos mentais são computacionais” (RM, p. 282). O novo argumento de Searle é baseado na tese de que a computação é *relativa ao observador*, e de que tudo o que é característico da computação enquanto computação é *exclusivamente sintático*. O argumento pode ser resumido em quatro asserções (MOURAL, 2003, p. 239):

- 1) Por definição, computação⁴² é definida como sendo uma manipulação formal de símbolos;
- 2) Contudo, nem a sintaxe nem os símbolos são definidos nos termos da Física, ou seja, a sintaxe não é *intrínseca* à Física;
- 3) Já que sintaxe e símbolos só possuem significado em relação a um observador, segue-se que também só pode ser atribuído sentido a uma computação diante de um observador;
- 4) Portanto, segue-se que não podemos *descobrir* que qualquer coisa, inclusive o cérebro, seja intrinsecamente um computador digital, apesar de ser possível fazermos uma interpretação computacional acerca de tudo.

Para se compreender este novo argumento, deve-se, antes de tudo, fazer a crucial distinção⁴³ entre as características do mundo que são independentes e as que são dependentes do observador⁴⁴ através de alguns exemplos. Objetos como átomos e moléculas possuem

⁴¹ Searle se refere constantemente à Ciência Cognitiva, apesar de ser mais adequado falar de *Ciências Cognitivas*, haja vista a multiplicidade de posições teórico-metodológicas que tal campo encerra. No entanto, vale ressaltar que Searle está ciente das divergências que existem entre os que estudam essa área, o que fica bem claro na seguinte passagem: “(...) *muitos* de nós pensam que a inteligência artificial (IA) é o coração da ciência cognitiva. Há uma *controvérsia* entre nós quanto ao fato de o cérebro ser ou não um computador digital do tipo antiquado de von Neumann, ou se é uma máquina conexionista” (RM, p. 282. Grifos meus). Todavia, a despeito dessas divergências, o que interessa a Searle para a presente discussão são os pontos de consenso das Ciências Cognitivas, que são, segundo ele, os seguintes: “os processos mentais cognitivos são inconscientes; são, na maioria dos casos, inconscientes em princípio; e são computacionais” (RM, p. 282).

⁴² Essa é a definição padrão de computação. No entanto, existem outras maneiras de defini-la sem tomá-la como *pura* manipulação simbólica, e que se constitui, pelo menos em princípio, como um sério obstáculo ao argumento de Searle do ponto de vista *conceitual*, pois inviabilizaria, logo de início, sua primeira tese. Por exemplo, pode-se objetar que há uma diferença fundamental entre definir *computação* e definir *computador*, pois este último é uma máquina de Turing real. Se isso for o caso, a 1ª asserção seria falsa (e conseqüentemente as demais teses, já que dependem da primeira), o que acabaria por caracterizar como *falacioso* o novo argumento de Searle.

⁴³ Nem todos, como por exemplo, os funcionalistas, aceitam essa distinção.

⁴⁴ Inicialmente, Searle usava o conceito de *intencionalidade intrínseca* no lugar de característica independente do observador. Porém, o uso daquele termo ocasionava muitos mal-entendidos, já que muitos interpretam intrínseco como oposto a relacional. Para Searle, intencionalidade intrínseca significa “a coisa real em oposição à mera aparência da coisa (*como-se*), e em oposição a formas derivadas de intencionalidade, como sentenças, imagens etc” (RM, p. 120).

características, como a massa, que independem de um observador, enquanto mesas e cadeiras têm características, como suas respectivas funções, que são relativas ao observador (ou usuário). Átomos e moléculas *independem* de nossa existência e do que pensamos sobre eles. Já o *significado* de objetos como mesas e fenômenos sociais como o casamento são dependentes do observador. *Em geral*, características que independem do observador são estudadas pelas Ciências Naturais, ao passo que as Ciências Sociais tratam de fenômenos que dependem do observador.

A partir disso, Searle faz a distinção entre os estados mentais de uma pessoa e os que são *atribuídos* a um computador: “my present state of consciousness is entirely observer-independent. No matter what anybody thinks, I am now conscious. But the attributions of mental states that I make to my computer are observer-dependent” (TOY, p. 62). Dessa forma, para Searle, os estados mentais conscientes de alguns animais⁴⁵ (incluindo o homem) são *intrinsecamente* intencionais, enquanto um programa de computador tem apenas uma intencionalidade *como-se*⁴⁶, isto é, funciona apenas como uma metáfora em relação à intencionalidade do observador. Em outras palavras, a intencionalidade de nossos estados psicológicos é *primária* e a do computador é *secundária*, no sentido de funcionar como uma analogia daquela. Portanto, apesar da sentença “Alzheimer tem uma boa memória” ser *sintaticamente* semelhante à sentença “Deep Blue tem uma boa memória”, elas *significam* coisas diferentes, já que a primeira deve ser tomada em sentido *literal* e a segunda de maneira apenas *metafórica*. O uso da intencionalidade *como-se* é comum em nossa comunicação do dia-a-dia, e não causa nenhum problema se o falante tiver em mente que ela tem um papel meramente metafórico. Dessa forma, é legítimo expressar que meu programa de computador *sabe* corrigir erros ortográficos desde que não se atribua, literalmente, fenômenos mentais a tal programa. Em outras palavras, o meu computador corrige erros ortográficos *como-se soubesse* gramática.

⁴⁵ “(...) intencionalidade intrínseca é um fenômeno que seres humanos e determinados outros animais têm como parte de sua *natureza biológica*” (RM, p. 118. Grifos meus). Searle afirma também que se não atribuirmos intencionalidade aos animais, seu comportamento não faria sentido. A partir da constatação de que o animal tem uma fisiologia (natureza biológica) semelhante a nossa, e esta, por sua vez, é causalmente responsável por um comportamento coerente (dotado de sentido), Searle infere que alguns animais não-humanos devem possuir “estados mentais subjacentes a seu comportamento e que esses estados mentais devem ser produzidos por mecanismos feitos de um material semelhante ao nosso” (MCP, p. 82).

⁴⁶ Na verdade, o que Searle designa como intencionalidade *como-se* não é, de fato, um tipo de intencionalidade (cf. RM, p. 118). Para tornar esse ponto mais claro, ele se utiliza do seguinte exemplo: “as plantas de meu jardim têm fome de nutrientes” (MLS, p. 90). O termo fome é utilizado nesse exemplo apenas de maneira metafórica, ou seja, é *como se* as plantas estivessem com fome, mas, de fato, elas não estão: “A ‘fome’ manifestada pelas plantas do meu jardim é puramente aparente. Elas estão murchando por falta de nutrientes, e descrevo sua condição por analogia às pessoas e animais. Atribuo-lhes uma intencionalidade que de fato não têm, embora se comportem como se comportem como se tivessem intencionalidade” (MLS, p. 90).

A partir do que foi exposto, poder-se-ia perguntar qual é a real importância da distinção entre intencionalidade intrínseca e intencionalidade como-se para a teoria de Searle. A resposta é a seguinte: recusar tal distinção leva a uma *reductio ad absurdum*, já que se poderia atribuir intencionalidade a tudo o que existe no mundo e, conseqüentemente, todo o universo seria mental (*panpsiquismo*). Ou seja, fenômenos mentais, com intencionalidade intrínseca, seriam atribuídos a uma pedra que cai, já que ela *quer* chegar ao chão, e a um termostato, pois ele *percebe* mudanças térmicas. A distinção entre o que é mental e o que é não-mental decorre da distinção entre o que é intrínseco e o que não é intrínseco⁴⁷, e ambas as distinções *não* são arbitrárias. Portanto, a distinção entre mental e não-mental “não pode estar apenas no olho do observador – ela deve ser intrínseca aos sistemas, pois de outra maneira ficaria a critério do observador tratar pessoas como não-mentais e furacões como mentais” (MCP, p. 75).

Finalmente, deve-se analisar que tipo de resposta Searle daria a pergunta “pode um computador digital pensar?”. Ao que parece, existem duas respostas contrárias para a questão, sim e não, mas será que elas são contraditórias? Em primeiro lugar, seria possível responder afirmativamente: sim, um computador digital pode pensar, mas isso seria *trivial* e *insatisfatório*. Trivial porque virtualmente qualquer coisa pode ser considerada um computador digital, e o cérebro não é exceção: “é claro que o cérebro é um computador digital, uma vez que tudo é um computador digital, os cérebros também o são” (MCP, p. 93). O caráter insatisfatório da resposta afirmativa decorre justamente dessa trivialidade, pois ela não acrescenta nada de novo ao que já se sabia antes da pergunta:

1) É sabido que alguns tipos de cérebro causam a mente consciente, e esta é capaz, entre outras coisas, de pensar;

2) Por definição, o cérebro pode ser tomado como um computador digital;

3) Donde se conclui que um computador⁴⁸ digital *pode* pensar.

Constata-se, a partir das três asserções acima, que responder sim a pergunta “pode um computador digital pensar?” é trivial, já que, segundo Searle, tudo é um computador⁴⁹. Ou, falando de maneira mais precisa, *é possível interpretar qualquer coisa como instanciando um*

⁴⁷ É interessante notar que, para Searle, a intencionalidade da linguagem é *derivada* da intencionalidade dos estados mentais, e é por este motivo que ele considera a Filosofia da Linguagem como um ramo da Filosofia da Mente, já que a mente *precede* e *dá origem* a linguagem.

⁴⁸ Na verdade, caso se mantenha o paralelo entre cérebro e computador, pode-se afirmar que existem *vários* tipos ou modelos de computador que pensam, os quais correspondem, respectivamente, aos tipos de cérebro capazes de produzir estados mentais conscientes nas diversas espécies biológicas, tais como cachorros, chimpanzés, golfinhos e, é claro, a espécie a que pertence este que vos escreve!

⁴⁹ “É claro que o cérebro é um computador digital, uma vez que *tudo* é um computador digital, os cérebros também o são” (MCP, p. 93. Grifo meu).

ou vários programas de computador. Mas, além de trivial, a resposta afirmativa também é insatisfatória porque o interesse ao se fazer tal pergunta não é saber se máquinas as quais atribuímos a capacidade de pensar, nós, por exemplo, são capazes de pensar: já se sabe, antes que a pergunta se coloque, que máquinas como nós e cérebros como os nossos são capazes de pensar. Portanto, quando se pergunta se um computador (ou uma máquina) é capaz de pensar, tem-se em consideração computadores (e máquinas) que não sejam organismos biológicos, tais como o computador no qual escrevo e robôs. Estes, diferentemente daqueles, não são capazes de pensar, já que esta capacidade envolve, *necessariamente*, a intencionalidade, a qual não é obtida pela *simples* instanciação de um programa de computador: “Seja lá o que o cérebro faça para produzir intencionalidade, esta não pode consistir na instanciação de um programa, pois nenhum programa é por si só suficiente para produzir a intencionalidade” (MCP, p. 93). Em outras palavras, para Searle, os estados mentais são *causados* pelos estados cerebrais, e, conseqüentemente, a intencionalidade intrínseca (propriedade mental ou de nível superior) daqueles é produzida a partir da comunicação que se dá entre os neurônios (propriedade física ou de nível inferior): “O que importa nas operações do cérebro não é a sombra do formalismo dado pela seqüência das sinapses, mas as propriedades efetivas de tais seqüências” (MCP, p. 85).

Em resumo, a primeira contribuição de Searle no âmbito da Filosofia da Mente foi uma contribuição *negativa*, isto é, ele fez uma crítica, a partir do argumento do quarto chinês, com o intuito de rejeitar uma concepção da mente, a saber, o *modelo computacional da mente*:

A não ser que se acredite que a mente é separável do cérebro, conceitual e empiricamente – um dualismo em uma versão forte – não se pode esperar reproduzir o mental escrevendo e rodando programas, uma vez que estes devem ser independentes dos cérebros ou de qualquer outra forma específica de sua instanciação. Se operações mentais consistem em operações computacionais sobre símbolos formais, segue-se que eles não têm nenhuma conexão importante com o cérebro, e a única conexão seria que o cérebro poderia ser um dentre os múltiplos tipos de máquinas capazes de instanciar o programa⁵⁰ (MCP, p. 92).

⁵⁰ O principal argumento do funcionalismo de computador é a tese da realização múltipla, já que seus proponentes defendem que diferentes substratos físicos, como os chips de silício, são capazes de produzir estados mentais. O próprio Searle, como será visto no capítulo 3, aceita uma versão fraca dessa tese, chegando a enunciar claramente esse argumento na seguinte passagem: “de fato seria *possível produzir consciência, intencionalidade e tudo o mais usando princípios químicos diferentes* dos usados por seres humanos” (MCP, p. 86. Grifos meus). Contudo, ao aceitar a realização múltipla, o funcionalista não está rejeitando a tese da realização física, como Searle parece sugerir na citação acima, quando afirma que os “[programas] devem ser independentes dos cérebros ou de qualquer outra forma específica de sua instanciação”. Para um funcionalista, um programa é independente do hardware no sentido de que um mesmo programa pode ser instanciado em diversos hardwares. Todavia, isso não é o mesmo que afirmar que um programa *não necessita* de um substrato físico para ser implementado. No meu modo de entender, Searle *comete um erro* ao comparar o funcionalismo com uma versão forte do dualismo, já que tal versão é o dualismo de substância, a qual defende que a mente ou alma (e, portanto, estados mentais) pode existir sem um substrato físico (o corpo). Em minha opinião, o mais

Searle não se limitou a criticar o funcionalismo, e desenvolveu sua própria concepção de mente, a qual é fundamentada em um *modelo biológico do mental*, isto é, tanto a consciência quanto a intencionalidade são produzidas pela atividade do sistema nervoso, o que as caracterizam como fenômenos biológicos. Em outras palavras, consciência e intencionalidade *dependem* da microestrutura biológica (sistema cerebral) que as produz:

Meu ponto de vista é que *somente* uma máquina pode pensar, e de fato apenas máquinas de um tipo muito especial, a saber, cérebros e máquinas que têm *os mesmos poderes causais* do cérebro. (...) O que quer que seja a intencionalidade [e a consciência], é um fenômeno biológico o qual *deve ser tão causalmente dependente* [grifos meus] da bioquímica específica de suas origens como o é a lactação, a fotossíntese ou quaisquer outros fenômenos biológicos (MCP, p. 92).

Após ter apresentado as críticas de Searle à IA forte, dedicarei o próximo capítulo para apresentar a sua teoria (ou concepção positiva) sobre a mente, o naturalismo biológico, o qual será discutido a partir seis teses inter-relacionadas entre si.

correto, caso se queira afirmar que o funcionalismo é um tipo de dualismo, seria compará-lo à versão fraca, isto é, ao dualismo de propriedade: propriedades mentais são distintas e irreduzíveis em relação às propriedades físicas, mas aquelas *necessitam* da microestrutura física para serem instanciadas. No caso de um ser humano, a morte cerebral *implica*, segundo o dualismo de propriedade, a morte mental. Um software – como, por exemplo, o WORD que estou agora utilizando – é multiplamente realizável, mas a destruição *deste* computador acarreta a inviabilização *deste* programa que estou usando para digitar essa dissertação. Geralmente, entretanto, o funcionalismo é classificado como um tipo de fisicalismo não-redutivo.

3. A CONCEPÇÃO POSITIVA DE SEARLE ACERCA DA MENTE

A Filosofia da Mente vem assumindo um papel cada vez mais central na contemporaneidade, chegando ao ponto de alguns – como Searle, por exemplo – considerá-la como a *Filosofia Primeira* (MLS, p. 9). Para ele, a Filosofia da Mente tem como principal objeto de estudo a consciência, e sua questão fundamental é de cunho ontológico, o chamado problema mente-corpo. A meu ver, sua teoria, o naturalismo biológico, é uma *Filosofia da Consciência*, pois segundo Searle a consciência é “o aspecto mais importante da realidade, porque todas as outras coisas só têm valor, importância ou mérito em relação à consciência” (MLS, p. 82).

Qualquer filósofo que se proponha a enfrentar e a oferecer uma resposta adequada ao problema mente-corpo deve obter sucesso em conciliar duas teses que, pelo menos em princípio, parecem ser mutuamente exclusivas, a saber, a de que o Universo é inteiramente físico e a de que mentes existem. Para Searle, essa contradição é apenas aparente, isto é, ela não existe de fato, pois ela se origina do uso da terminologia de Descartes, que contrapõe, de maneira equivocada o físico ao mental. Portanto, o dilema de relacionar o físico ao mental dissolve-se ao se abandonar a obsoleta terminologia cartesiana.

Apesar de desconfiar do sufixo “ismo”, Searle propõe sua própria solução para o problema mente-corpo, chamada naturalismo biológico, na qual ele explicita de maneira positiva qual é o estatuto dos fenômenos mentais perante os fenômenos físicos. Tal concepção, segundo Searle, não se enquadra nem como fisicalismo nem como dualismo, mas aceita os pontos positivos de ambas ao mesmo tempo em que refuta os negativos.

Este capítulo tem por objetivo enumerar sistematicamente e analisar de maneira crítica as teses que compõem o naturalismo biológico. Tais teses são interconectadas e se articulam de maneira não-linear, assemelhando-se mais a uma estrutura em forma de rede, e é por este motivo que, não raras vezes, serão feitas ligações entre duas ou mais teses.

3.1) 1ª tese: Fenômenos Mentais são Fenômenos Biológicos

A concepção dualista da mente considera que o domínio mental possui características peculiares que estão além do domínio físico, o que caracteriza a irreducibilidade daquele

domínio em relação a este último⁵¹: ou a substância mental é irreduzível à substância física (dualismo de substância) ou as propriedades mentais são irreduzíveis às propriedades físicas (dualismo de propriedade). Isso se deve porque o modo de existência dos fenômenos mentais é distinto do modo de existência dos fenômenos físicos, o que caracteriza uma questão *ontológica*.

Outra característica dos fenômenos mentais que os diferenciam dos fenômenos físicos é que a sua descrição depende, como apontou Nagel⁵², do *ponto de vista do sujeito que os experiencia*. Fenômenos físicos são publicamente acessíveis e são, por isso, passíveis de serem descritos a partir de uma perspectiva de terceira pessoa, o que não ocorre com os fenômenos mentais (questão *epistêmica*). Searle acha, entretanto, que a interpretação epistêmica que é dada ao experimento mental de Nagel é equivocada. Na verdade, a questão decisiva não é como podemos *conhecer* (problema epistemológico) os estados do morcego (ou de outro sujeito consciente), e sim o que *são* (problema ontológico), efetivamente, tais estados, ou seja,

“(…) é uma questão relativa a *quais características reais existem no mundo*, e não, a não ser secundariamente, sobre como conhecemos essas características” (RM, p. 169. Grifos meus).

“O sentido no qual tenho um acesso a meus estados que é diferente do acesso dos outros não é principalmente epistemológico. Não se trata apenas de como os conheço, embora a subjetividade tenha consequências epistemológicas; mais do que isso, *cada um de meus estados conscientes só existe no estado em que existe porque é experimentado por mim*, o sujeito. E, assim, faz parte da seqüência de estados que constitui minha vida consciente (...)” (MLS, p. 47. Grifos meus).

Em resumo, a tese central defendida pelos adeptos do dualismo é a da distinção ontológica entre o físico e o mental: a substância/propriedade física tem um modo de existência diferente da substância/propriedade mental.

Searle, por sua vez, não aceita a distinção supracitada entre físico e mental, pois para ele o mental *é* físico, já que o mental *é* biológico. Em outras palavras, os fenômenos mentais *são* fenômenos biológicos, assim como outros processos que são encontrados no mundo vivo,

⁵¹ O dualismo é uma concepção bastante difundida entre as pessoas em geral, pois ele está tanto afinado com o senso comum, em geral, como com as diversas Religiões, em particular: “(...) todos temos nossas próprias experiências conscientes, e podemos ver com facilidade que elas são diferentes do mundo material que nos cerca” (MLS, p. 52). Entretanto, Searle chama a atenção para o fato de que o problema mente-corpo não é universal, pois em algumas línguas nem é mesmo possível concebê-lo: “Fiquei profundamente impressionado ao ouvir um amigo africano me dizer que, em sua língua nativa africana, o “problema da mente-corpo”, conforme o concebemos, não pode sequer ser anunciado” (MLS, p. 55).

⁵² Conferir Nagel (1996, 2004, 2005).

a saber, como a digestão, a fotossíntese e a mitose⁵³. Portanto, segundo Searle, para lograr êxito na superação do dualismo, deve-se rejeitar “o sistema de categorias que faz da consciência algo não-biológico, que não faz parte do mundo natural” (MLS, p. 55). Contudo, diferentemente dos demais fenômenos naturais, os fenômenos mentais são ontologicamente subjetivos⁵⁴. Antes de explicar o motivo de fenômenos mentais serem considerados fenômenos biológicos, é preciso compreender o que Searle entende por objeto físico e de que ele é composto.

Searle é, antes de qualquer coisa, um defensor e admirador das concepções científicas⁵⁵ atualmente vigentes, especialmente de duas delas – as teorias atômica e evolutiva:

The two basic principles on which any such investigation as the one I have been engaging in depends on are, first, the notion that the most fundamental entities in reality are those described by atomic physics; and, second, that we, as biological beasts, are the products of long periods of evolution, perhaps as long as five billion years (M, p. 303).

A consequência disso é que o modo como Searle enxerga a mente é naturalista, ou seja, essas teorias das Ciências Naturais figuram como um pano de fundo do naturalismo biológico. De acordo com a primeira dessas teorias, *todo* o Universo é constituído de partículas físicas submetidas a campos de força, e nós, seres humanos, não somos exceção. A matéria presente no Universo se organiza através de níveis, onde entidades simples se juntam para formar entidades cada vez mais complexas: partículas subatômicas⁵⁶ se agrupam para

⁵³ A semelhança se dá porque todos são propriedades emergentes de sistemas biológicos, mas, diferentemente da digestão, fotossíntese e mitose, a consciência possui uma ontologia de primeira pessoa (conferir a 3ª tese).

⁵⁴ Contata-se, então, que Searle repudia a dicotomia entre físico e mental, mas aceita, todavia, a distinção entre ontologia objetiva e ontologia subjetiva. Será que ao assumir esta distinção Searle não se *compromete* com alguma espécie de dualismo? Essa é uma questão de extrema relevância para o presente trabalho e será discutida mais adiante.

⁵⁵ Searle defende, em contraposição aos anti-realistas, que o mundo externo é *real* e existe *independentemente* de nós: com ou sem seres humanos, existem o sol, as rochas, as árvores e as formigas, isto é, tais objetos não dependem de nossa opinião ou conhecimento acerca deles. Dessa forma, não se pode questionar a existência do *mundo real*, já que tal questionamento já o pressuporia. Por isso que Searle não considera que o realismo seja sequer uma teoria, “mas sim a estrutura dentro da qual é possível ter teorias” (MLS, p. 39). Além disso, ele claramente deixa transparecer sua decepção (e até mesmo raiva) para com os filósofos que advogam a favor de teorias que são, para Searle, tão contra-intuitivas, tais como as que negam o realismo e a teoria da verdade como correspondência: “Por que, quando começamos a fazer filosofia, somos quase inexoravelmente levados a negar coisas que todos sabemos serem verdadeiras – por exemplo, que existe um mundo real, que podemos ter certos tipos de conhecimento sobre esse mundo, que afirmações são geralmente verdadeiras se correspondem a fatos do mundo, e falsas se não o fazem?” (MLS, pp. 17-18)

⁵⁶ Quanto à carga, existem três tipos básicos de partículas subatômicas: prótons (carga positiva), nêutrons (sem carga) e elétrons (carga negativa). As duas primeiras se localizam no núcleo atômico, enquanto os elétrons se movimentam em uma (ou mais) região (ões) ao redor do mesmo, chamada de orbital (is). Existem mais de cem elementos químicos diferentes, os quais são ordenados por ordem de número atômico (nº de prótons) na tabela periódica. Vale a pena ressaltar que Searle usa o termo “partícula” de maneira frouxa, o que engloba vários

formar átomos. A molécula, por sua vez, origina-se da união de dois ou mais átomos – a molécula de água, por exemplo, é composta de três átomos: dois de hidrogênio (H) e um de oxigênio (O), sendo por isso representada por H₂O. Alguns conjuntos de moléculas se organizaram, no decorrer de um longo processo evolutivo⁵⁷, em estruturas bastante complexas chamadas células, as quais possuem uma propriedade que os biólogos denominam de vida. Sendo assim, uma bactéria, que é um organismo composto por uma única célula, é *qualitativamente* diferente do ambiente físico que habita. No entanto, nem todos os seres vivos são como a bactéria, isto é, possuem uma única célula, o que significa que alguns deles desenvolveram estratégias nas quais a cooperação e a divisão de tarefas se mostraram extremamente eficientes para a sua sobrevivência, originando os organismos pluricelulares, que são exemplificados pelos animais. Dentro do Reino Animal, observa-se, à medida que se originam seus Filos, uma tendência evolutiva geral no que se refere ao aumento do grau de cefalização. Alguns destes animais desenvolveram sistemas nervosos tão complexos que, ao que tudo indica, foram capazes de *produzir* e *sustentar* estados e processos mentais. Após essa breve apresentação das teorias atômica e evolutiva⁵⁸, resta mostrar como elas se enquadram dentro do naturalismo biológico.

Assim como um fisicalista, Searle aceita que o universo é físico e unicamente composto, em última instância, de partículas físicas submetidas a campos de força. Essa assunção está em franca oposição ao dualismo, que defende, por sua vez, que fenômenos mentais estão acima e além de fenômenos físicos. Para o naturalismo biológico, os fenômenos

conceitos físico-químicos distintos: “Exemplos de partículas são elétrons, átomos de hidrogênio e moléculas de água”. Estes são exemplos de diferentes níveis de organização da matéria, a saber, respectivamente, os níveis subatômico, atômico e molecular. Portanto, apesar de ficar “embaraçado” ao utilizar o termo partícula, Searle deixa explícito o que ele quer dizer com tal conceito: que “partículas maiores são constituídas de partículas menores” (RM, p. 128). Para maiores detalhes sobre a Teoria Atômica, conferir Morrison & Boyd (1993).

⁵⁷ Constata-se, a partir do estudo da História da Biologia, que existem duas correntes básicas para se explicar a origem dos seres vivos: o Fixismo e o Transformacionismo (ou Evolucionismo). De acordo com a primeira, os organismos foram *criados* por Deus, e *permaneceram*, tanto morfológica como funcionalmente, *os mesmos*, desde o momento de sua criação até os nossos dias. Já a segunda concepção defende que os organismos se *modificaram* ao longo do tempo, tornando-se mais *adaptados* ao seu ambiente. As espécies que não se adaptaram foram, por um motivo ou outro, extintas. Sendo assim, não há criação divina e imutabilidade dos seres vivos, mas, pelo contrário, o processo evolutivo é não-teleológico e dinâmico: espécies surgem e desaparecem ao longo das eras geológicas, sofrendo transformações estruturais ao longo de sua existência no planeta. O principal mecanismo evolutivo chama-se *seleção natural*, e foi proposto em 1858 por dois cientistas ingleses – Charles Robert Darwin e Alfred Russell Wallace. Para que haja seleção natural é necessário que existam *variações* entre os indivíduos de uma população. Apesar de Darwin saber disso, ele não sabia como explicar de maneira satisfatória o surgimento de tais variações. Foi apenas nos anos 30 e 40 do século XX que o trabalho conjunto de destacados biólogos resultou na síntese entre o mecanismo de seleção natural e os conceitos provenientes da Genética, tais como o de *mutação* e o de *recombinação genética*. Essa teoria ficou conhecida como *Síntese Moderna* ou *Teoria Sintética da Evolução* e é a concepção atualmente vigente na Biologia. Para maiores detalhes sobre a Teoria da Evolução, conferir Futuyma (2003) e Ridley (2006).

⁵⁸ Searle é ciente que tanto a teoria atômica como a teoria evolutiva, *por serem teorias científicas*, são passíveis de ser refutadas empiricamente. No entanto, ele salienta que “hoje, porém, a evidência é tão esmagadora [a favor das referidas teorias] que simplesmente não se prestam a questionamentos” (RM, p. 127).

mentais não possuem nada de misterioso, nem na sua origem nem na sua causa: os fenômenos mentais *surgiram* evolutivamente, assim como a fotossíntese e a digestão, e são, como estes dois últimos processos, *causados*⁵⁹ por fenômenos físicos de nível micro. Portanto, os fenômenos mentais são genética (no sentido de gênese) e causalmente biológicos. Dito de outra forma, fenômenos mentais *são* fenômenos biológicos. Em resumo, a nossa concepção científica do mundo – que inclui, entre outras, as teorias atômica e evolutiva – assume que:

The universe does consist entirely in physical particles in fields of force (or whatever the ultimately true physics discovers), these are typically organized into systems, some of the systems are biological, and some of the biological systems are conscious. Consciousness is thus an ordinary feature of certain biological systems, in the same way that photosynthesis, digestion, and lactation are ordinary features of biological systems (WIANPD, p. 8)

3.2) 2ª tese: Fenômenos Mentais são Causados e Realizados no Sistema Cerebral

No que se refere a esta tese, pode-se observar que o naturalismo biológico tem duas semelhanças com a posição fisicalista não-reduzitiva, a saber: (i) todos os particulares que existem são físicos (fisicalismo ontológico) e (ii) todas as propriedades mentais são instanciadas em particulares físicos, mas são ontologicamente irreduzíveis a estes. Mas, pode-se questionar, qual é o tipo de relação existente entre os particulares físicos e as propriedades mentais, de modo que estas sejam irreduzíveis àqueles? Hellman e Thompson forneceram a seguinte resposta: as propriedades mentais são *determinadas* ou *dependentes* das propriedades físicas, já que os fatos físicos determinam todo e qualquer tipo de fato (KIM, 1993b, p. 341). Contudo, essa resposta ainda é insuficiente, pois ainda resta explicar de que maneira o mental é irreduzível ao físico, já que é determinado por este. Foram desenvolvidas, ao longo dos anos, duas alternativas para conciliar *determinação física* e *irreduzibilidade ontológica do mental*, a saber: *realização física* e *superveniência*.

De acordo com a primeira tese, uma mesma propriedade mental (M) pode ser realizada através de diferentes tipos de propriedades físicas, como, por exemplo, propriedades neurobiológicas de um cérebro humano (N) ou propriedades eletrônicas de “cérebros” de silício (E). Sendo assim, M será realizada por N em seres humanos e por E em robôs. Constata-se que a propriedade mental M é fisicamente determinada por uma propriedade física, mas não pode ser reduzida (ser identificada) à mesma, já que um ser humano pode ter a

⁵⁹ Conferir a 2ª tese.

propriedade M sem ter E, e, inversamente, um robô pode ter M sem ter N (BECKERMANN, 1992, pp. 8-10).

Já para Davidson (2004), a determinação física ocorre por meio da relação de superveniência – eventos mentais são supervenientes a eventos físicos. Com isso, Davidson quer dizer duas coisas: (i) eventos físicos iguais determinam a identidade dos eventos mentais, e (ii) a alteração de alguma propriedade mental do objeto implica, necessariamente, a modificação de alguma propriedade física subveniente⁶⁰. No entanto, apesar das propriedades supervenientes serem determinadas pelas propriedades basais, não existem *leis psicológicas estritas*, já que os predicados físicos e mentais são *incomensuráveis*, ou seja, os termos da Psicologia Popular não são redutíveis aos termos da Neurobiologia. Em outras palavras, Davidson aceita a redução ontológica (eventos mentais *são* eventos físicos), mas rejeita a redução conceitual, o que significa que a linguagem das Ciências Naturais (Física, Química, Biologia) é inadequada para explicar *causalmente* os eventos psicológicos. A explicação destes se dá por referência a outros eventos psicológicos, utilizando, para tanto, o vocabulário da Psicologia Popular:

Mental events as a class cannot be explained by physical science; particular mental events can when we know particular identities. But the explanations of mental events in which we are typically interested relate them to other mental events and conditions (DAVIDSON, 2004, p. 699).

De que maneira as propriedades mentais são fisicamente determinadas no naturalismo biológico? Searle defende, pelo menos *parcialmente*, tanto a realização física quanto a superveniência. De acordo com ele, todo sistema cerebral (S) que instancia uma propriedade mental (M) possui uma propriedade física (P), tal que P realiza M em S. Dessa maneira, as propriedades mentais são *determinadas* pela estrutura do sistema físico. Em outras palavras, a condição necessária e suficiente para M ser instanciada é S estar em P. Ressalte-se que S não precisa ser, necessariamente, um sistema nervoso e, conseqüentemente, não é necessário que P seja uma propriedade neurobiológica. Dito de outro modo, a causação e realização dos estados mentais pelos neurais não significa que exista algum impedimento, tanto lógico

⁶⁰ Na verdade, há três conceitos de superveniência: fraca, forte e global. A distinção entre tais conceitos se deve à força modal com que são formulados (Conferir BECKERMANN, 1992, p. 12; KIM, 1992, pp. 132-133). De acordo com Beckermann, nenhum dos três tipos de superveniência serve para fundamentar uma posição fisicalista não-redutiva, pois tal noção apenas estabelece uma *covariação* entre propriedades físicas e mentais, e isso não implica que estas propriedades sejam *determinadas* por aquelas. Sendo assim, Beckermann (1992, p. 14) conclui que “the concept of supervenience is not of much help if one wants to formulate a viable version of nonreductive physicalism, since physicalism demands more than just causal or nomological dependence. If supervenience was the only hope for nonreductive physicalism, this kind of physicalism would indeed have no good prospects”.

quanto empírico, para a construção de sistemas artificiais dotados de atividade mental⁶¹. Dessa forma, Searle aceita a tese de que as propriedades mentais são, *em princípio*, realizáveis em sistemas físicos não-neurônicos.

No entanto, há aqui duas diferenças fundamentais entre a tese da realização física no funcionalismo e no naturalismo biológico. Primeiro, pode-se afirmar que a primeira dessas concepções aceita uma versão forte da tese da realização múltipla, enquanto que a segunda defende uma versão mais fraca. Isso se deve porque para Searle, diferentemente dos funcionalistas, há uma restrição quanto aos materiais capazes de instanciar propriedades mentais: é *necessário* que as propriedades físicas não-neurônicas *dupliquem* os poderes causais das propriedades neurobiológicas para que o sistema não-neurônico instancie propriedades psicológicas⁶²:

Quando digo que o cérebro é um órgão biológico e a consciência um processo biológico, é claro que *não estou* dizendo ou sugerindo que seria impossível produzir um cérebro artificial a partir de materiais não biológicos que também pudessem ser causa da consciência e sustentá-la [tese da realização múltipla]. (...) O ponto que deve ser enfatizado é que tal cérebro artificial teria de *reproduzir as verdadeiras causas* dos cérebros humanos e animais para produzir estados de consciência internos, qualitativos e subjetivos. A produção de comportamentos similares por si só não seria suficiente (MLS, p. 56. Grifos meus).

Sendo assim, pode-se afirmar, em minha opinião, que a tese da realização múltipla defendida por Searle é bem mais restritiva do que a que é defendida pelo funcionalismo. Aliás, às vezes Searle diz abertamente que a *constituição biológica* dos cérebros é de fundamental importância para produzir e sustentar processos psicológicos: já que “os cérebros são máquinas biológicas; *a sua biologia tem importância*. Não é, como vários praticantes da Inteligência Artificial afirmaram, um facto irrelevante que a mente se realize em cérebros humanos [e de outros animais]” (MCC, p. 50. Grifos meus).

Além do que foi dito acima, pode-se mencionar uma segunda diferença entre as duas concepções, a saber: a realização de uma propriedade mental no funcionalismo diz respeito ao *papel causal* desempenhado pela mesma, enquanto que no naturalismo biológico o conceito de realização está relacionado a *propriedades emergentes*, as quais são, por sua vez, propriedades de nível elevado que não estão presentes nos componentes sistêmicos do micronível. Por exemplo, a liquidez da água é uma propriedade emergente que é *causada* pelas interações de moléculas de H₂O e é *realizada* no próprio sistema molecular. Portanto, a propriedade emergente é *determinada* pelo sistema de moléculas de H₂O, isto é, a liquidez –

⁶¹ “Até onde sabemos, não existe, em princípio, nenhuma razão para não podermos construir um cérebro artificial que também cause e tenha consciência” (MLS, p. 57).

⁶² Conferir, logo a seguir, a 6ª tese.

apesar de não ser uma propriedade instanciada em *uma* molécula de H₂O –, nada mais é, do ponto de vista ontológico e causal, do que um conjunto de moléculas de H₂O interagindo através de pontes de hidrogênio de determinado modo. Searle defende a realização das propriedades nesses termos, excetuando-se a consciência, que é a única propriedade emergente, segundo ele, *parcialmente redutível*: as propriedades conscientes são *causalmente* redutíveis a propriedades neurobiológicas, mas os dois tipos de propriedades são *ontologicamente* distintos e não há, portanto, redução *ontológica* daquelas a estas⁶³.

No meu modo de entender, além da realização física, Searle (RM, p.180) também defende o que ele denomina de *superveniência causal* – um estado mental (M) é superveniente em relação a um estado neurobiológico (N) de tal modo que: (i) N é suficiente para a instanciação de M, (ii) a modificação do estado N para N* acarreta a mudança *correspondente* de M para M* e, finalmente, (iii) estados psicológicos distintos são decorrentes de diferentes estados neurobiológicos, mas não o inverso. Ou seja, o estado psicológico “creio que Fortaleza é a capital do Ceará” pode ser determinado por diferentes estados neurobiológicos dos sistemas nervosos de Capistrano e Gilberto.

Um dos objetivos gerais da teoria de Searle é demonstrar e defender que a consciência e a intencionalidade *são* naturais. Portanto, não seria correto afirmar que ele tem em vista um projeto de naturalização dos fenômenos mentais, pois, para ele, eles já fazem parte da natureza, mais especificamente da natureza biológica de animais que possuem sistemas nervosos complexos⁶⁴. Sendo assim, a consciência e a intencionalidade são fenômenos biológicos semelhantes à fotossíntese e à divisão celular no seguinte sentido: são características de nível superior causadas e realizadas por complexas estruturas biológicas⁶⁵.

Em outras palavras, a consciência é causada⁶⁶ e realizada⁶⁷ no sistema cerebral da mesma forma que a digestão é causada e realizada no sistema digestório (MLS, p.54), isto é, os processos conscientes, assim como os processos digestivos, são processos biológicos,

⁶³ A diferença entre os fenômenos com ontologia de primeira (fenômenos mentais) e terceira pessoa (demais fenômenos naturais) será desenvolvida na próxima tese.

⁶⁴ Em uma passagem do livro *Mente, Linguagem e Sociedade*, Searle chega até a admitir que ao se abandonar o dualismo conceitual, supera-se a dicotomia dualismo-fisicalismo. Essa superação se dá pela *naturalização* da consciência – a consciência é um fenômeno natural, em geral, e biológico, em particular. No entanto, em tal passagem, Searle utiliza a expressão entre aspas: “Assim, ‘*naturalizamos*’ a consciência e, de fato, minha etiqueta para esse ponto de vista é ‘*naturalismo biológico*’ (...)” (MLS, p. 57. Grifo meu). No meu modo de entender, o uso das aspas é para indicar que não é necessário naturalizar a consciência – *ela já é natural*.

⁶⁵ “A consciência é, acima de tudo, um fenômeno biológico. Os processos conscientes são processos biológicos” (MLS, p. 57).

⁶⁶ “Os processos conscientes são *causados* por processos neurais de nível inferior que ocorrem no cérebro” (MLS, p. 57. Grifo meu).

⁶⁷ “A consciência consiste em processos de nível superior *realizados* na estrutura do cérebro” (MLS, p. 57. Grifo meu).

enquanto estes últimos são causados por processos físicos, como a mastigação, e por processos químicos (reações enzimáticas) e realizados na estrutura dos próprios órgãos que compõem o sistema digestório (boca, estômago, intestino delgado etc.), os processos conscientes são causados pelas interações entre neurônios (sinapses) através da secreção de neurotransmissores e realizados na estrutura dos órgãos que compõem o sistema nervoso, notadamente o cérebro, motivo pelo qual Searle prefere se referir a sistema cerebral em vez de sistema nervoso⁶⁸. No entanto, Searle sempre é enfático em sublinhar que há diferenças fundamentais entre os fenômenos mentais em relação aos demais fenômenos biológicos, a saber: o modo de existência de primeira pessoa (subjetividade ontológica) e a propriedade de se dirigir para objetos e situações do mundo (intencionalidade intrínseca).

Também se faz necessário mencionar que as noções de causação e realização de Searle divergem do pensamento de filósofos de grande envergadura, como Descartes e Hume⁶⁹. Em relação ao primeiro, Searle repete incontáveis vezes que o primeiro passo para se superar o problema mente-corpo é levar a cabo uma *revisão conceitual* das categorias utilizadas em Filosofia da Mente. A *res cogitans* cartesiana, em oposição à *res extensa*, não possui localização espacial. No naturalismo biológico, ocorre justamente o contrário, já que a consciência é *espacialmente* localizada. A localização espacial da consciência é a *mesma* dos microprocessos que a causam e das estruturas neurobiológicas que a realizam. Portanto, defender a tese de que os estados conscientes são causados por processos cerebrais e realizados nas estruturas neurobiológicas implica no abandono das categorias cartesianas e, conseqüentemente, no comprometimento com a *localização espacial* de eventos mentais. Em outras palavras, a redução causal da consciência a processos neurobiológicos (5ª tese) implica que aquela possui as mesmas⁷⁰ propriedades desses últimos, sendo por isso que ambos compartilham os poderes causais (6ª tese):

⁶⁸ Searle não esclarece, do ponto de vista anatômico, quais são os órgãos que compõem o que ele denomina de sistema cerebral. No entanto, ele estará cometendo um erro conceitual caso utilize tal termo para se referir especificamente ao cérebro, pois um sistema é um *conjunto de órgãos* que desempenha uma função de modo coordenado. Dessa forma, o cérebro sozinho não pode formar um sistema, já que ele é apenas *um órgão*.

⁶⁹ Isso será abordado no quarto capítulo, o qual trata, entre outras coisas, sobre a causação humeana.

⁷⁰ Esse é um ponto bastante delicado, pois o intuito de Searle é mostrar que a redução causal da consciência explica facilmente a localização espacial da consciência. Entretanto, o problema surge por que Searle *não* continua com a sua linha de raciocínio e conclui que a consciência compartilha *todas* as suas propriedades com os processos neurobiológicos subjacentes, já que ele defende veementemente que os fenômenos mentais possuem propriedades subjetivas e qualitativas que *não são* encontradas nos fenômenos neurobiológicos. Essa ambigüidade que surge com relativa frequência no pensamento de Searle faz com que seja tão difícil classificar o naturalismo biológico: se por um lado ele defende a redução causal da consciência, afirmando que os estados mentais possuem a *mesma* “dimensão espacial” e “carga elétrica” dos estados neurobiológicos subjacentes, por outro lado ele insiste em defender que *algumas* propriedades mentais são irredutíveis.

Os estados e processos conscientes são *internos* em um sentido espacial muito comum, já que acontecem dentro do meu corpo, e especificamente dentro do meu cérebro (MLS, p. 46).

“But are you saying that the conscious state itself has electrical charges and has spatial dimensions [propriedades físicas] in the brain?” The answer to that question is obviously “yes”. We know that conscious events have spatial locations in the brain (DR, p. 176).

Ao defender que os eventos mentais são espacialmente localizados, Searle não enfrenta as dificuldades que se impõem a concepção cartesiana, tais quais: localização espacial (glândula pineal) de uma substância não-espacial (*res cogitans*), e interação entre uma coisa física e uma não-física através de espíritos animais.

O conceito de causalidade neurobiológico-mental do naturalismo biológico diverge bastante do conceito humeano de contigüidade espaço-temporal entre causa e efeito. Searle chama a atenção para o fato de que as críticas iniciais à mecânica newtoniana questionavam a possibilidade da força gravitacional ser causalmente eficiente, já que um astro teria que exercer uma *ação à distância* sobre outro. No entanto, tal tipo de ação era inconcebível para a Filosofia e a Ciência da época, pois, de acordo com o que se acreditava, a ação de um objeto sobre outro só seria possível se eles fossem espacialmente contíguos. Sendo assim,

Para evitar o absurdo da ação à distância, parecíamos forçados a pensar na gravidade como uma questão de fios invisíveis que prendiam os corpos planetários uns aos outros. Hoje em dia, ninguém faz esse tipo de objeção. *Temos um conceito muito mais rico de causalidade*, que inclui, entre outras coisas, campos de força (MLS, p. 62. Grifos meus).

Segundo Searle, além de campos de força, esse “conceito muito mais rico de causalidade” inclui também – em alguns casos, como na causação neurobiológico-mental –, a *simultaneidade* entre causa e efeito⁷¹. No caso da ação gravitacional do Sol em relação à Terra, o que incomodava os filósofos e cientistas modernos era uma questão espacial – “como a força gravitacional que o Sol exerce sobre a Terra *causa* o movimento de translação desta, já que eles estão muito distantes um do outro?”. Porém, quando se analisa a questão da espacialidade no naturalismo biológico, constata-se que o ponto crítico é o oposto ao da ação à distância (a ação ocorre sem que haja um contato entre os objetos) e diferente da causação do tipo bola de bilhar (há contato entre os objetos, mas estes são distintos espaço-temporalmente): *não existe distinção espacial entre processos cerebrais (causas) e estados mentais (efeitos)*. A *identidade espacial* entre processos neurofisiológicos e estados conscientes acarreta, por conseguinte, a *simultaneidade* entre causa e efeito: *não há qualquer*

⁷¹ A causação intencional ou mental é “inteiramente diferente das explicações da causação que surgem nos livros de textos, onde, por exemplo, uma bola de bilhar bate noutra bola de bilhar e a faz mover” (MCC, p. 76). A causação mental no naturalismo biológico será tratada em maiores detalhes no próximo capítulo.

tipo de lacuna temporal entre as sinapses neurônicas e a produção-realização de fenômenos conscientes.

Para fundamentar sua concepção nada ortodoxa de causalidade, Searle afirma que existe apenas *um sistema*, o sistema cerebral, descrito em *dois níveis*: o micro-físico dos processos neurônicos e o macro-mental dos eventos psicológicos. Todavia, o próprio Searle é reticente no que diz respeito ao modelo de descrição múltipla de um sistema:

A metáfora do superior e do inferior, embora seja frequentemente utilizada (até mesmo em meus trabalhos), é *enganosa*. Ela sugere que a consciência seja, por assim dizer, como uma espécie de verniz que recobre uma mesa. Não é o caso. A idéia que tentamos exprimir consiste sobretudo em *fazer da consciência a característica do conjunto do sistema*. A consciência está presente em todas as partes do cérebro onde a atividade neuronal a cria e realiza [localização espaço-temporal dos eventos mentais] (LN, pp. 47-48. Grifos meus).

Deve-se salientar que apesar de Searle definir a consciência como uma característica sistêmica ou emergente de um conjunto de neurônios, ele não endossa uma *redução eliminativa* ou *eliminatória*⁷², ou seja, que o nível de descrição superior (eventos psicológicos) é eliminado, tanto do ponto de vista ontológico como do ponto de vista explicativo, em decorrência do nível inferior (processos cerebrais), o qual seria, de fato, *real* e *causalmente eficiente*⁷³:

Não podemos fazer uma redução eliminatória da consciência porque o padrão das reduções eliminatórias é mostrar que o fenômeno reduzido é apenas uma ilusão. No entanto, no que diz respeito à consciência, a existência da “ilusão” é a própria realidade. Isso quer dizer que, se me parece que estou consciente, então eu estou. Não há nada na consciência além de uma seqüência de “aparências” desse tipo. A esse respeito, a consciência é diferente dos poentes porque eu posso ter a ilusão de que o sol está se pondo por trás das montanhas quando na verdade ele não está fazendo isso (MLS, p. 59).

Já que Searle aceita o fisicalismo ontológico⁷⁴, isto é, a tese de que tudo no Universo é constituído de partículas físicas fundamentais e seus agregados, ele partilha, juntamente com os emergentistas, a concepção de um mundo hierarquizado, dividido em vários níveis *ontológico-descritivos*. Segundo essa visão, um nível, excetuando-se os extremos inferior e superior, tem uma posição relativa, ou seja, só é possível classificá-lo como inferior ou superior *em relação* a outro nível. Por exemplo: o nível dos processos mentais é superior ou inferior? Tal pergunta assim formulada é desprovida de sentido, pois os conceitos de nível superior ou inferior são relacionais. Portanto, a pergunta correta deve envolver outro nível da

⁷² A irredutibilidade ontológica dos eventos conscientes será tratada na próxima tese.

⁷³ Respectivamente, essas são as teses do realismo e da eficácia causal dos estados mentais.

⁷⁴ Conferir o próximo capítulo.

realidade, tal como: qual a relação entre o nível dos processos mentais e o nível X? Nota-se, a partir do que foi explicado, que a resposta para esta pergunta vai variar conforme o nível X em questão: se X se referir às instituições sociais, estas corresponderão ao nível superior; inversamente, caso X se refira aos estados neurofisiológicos, o nível superior será o dos processos mentais⁷⁵. De toda forma, Searle deixa claro que não importa de qual nível se esteja falando – galáxias ou organismos, por exemplo – o nível constituinte mais básico será o nível das partículas físicas fundamentais (quarks e múons, de acordo com a Física atual). Entretanto, a descrição de um nível superior através do micronível *não implica* que este nível é *real* e aquele é apenas uma *aparência*. Dito de outra forma, Searle quer rejeitar o argumento epifenomenalista⁷⁶, segundo o qual a causação dos processos mentais pelos processos cerebrais tem como consequência que aqueles são apenas epifenômenos destes:

O fato de podermos fornecer um relato causal em um nível inferior *não implica* que os níveis superiores não sejam reais. (...) Assinalar que a solidez de um pistão é explicável em termos do comportamento molecular das ligas metálicas *não prova* que a solidez do pistão é um epifenômeno; da mesma maneira, assinalar que as intenções são explicáveis em termos de neurônios, sinapses e neurotransmissores [nível neurobiológico] *não prova* que as intenções [nível da Psicologia Popular] são um epifenômeno (MLS, pp. 63-64. Grifos meus).

3.3) 3ª tese: Fenômenos Neurobiológicos são Ontologicamente Distintos de Fenômenos Mentais

Para Searle, os fenômenos mentais podem ou não, dependendo do que se queira dizer, ser identificados aos processos cerebrais que os causam: se, por um lado, não há identidade entre *tipos* de estados mentais e *tipos* de estados conscientes; por outro lado, no entanto, existe uma identidade de *espécime* – isto é, estados mentais são causalmente idênticos a

⁷⁵ É importante destacar que uma importante discordância de Searle em relação às Ciências Cognitivas está no fato destas defenderem que entre os níveis neurobiológico e psicológico há um nível intermediário, a saber, o de *processamento de informações*. Searle afirma que o problema surge ao se confundir o processamento de informações *de fato* com a possibilidade de descrever o comportamento do cérebro *como se* processasse informação: “Mas o facto de podermos descrever outros processos no cérebro, a partir de um ponto de vista “como se” do processamento de informação, não fornece por si mesmo qualquer prova de que são *psicologicamente reais* ou *psicologicamente relevantes* (MCC, p. 63. Grifos meus).

⁷⁶ Como será visto adiante (6ª tese), Searle defende que as capacidades causais dos fenômenos mentais são *idênticas* às dos fenômenos neurofisiológicos. Dessa forma, ele pode defender o realismo do mental e a causação mental, tanto mental-mental quanto mental-física. Esta última é chamada de causação descendente, e é principalmente contra este tipo de causação que Jaegwon Kim fará suas principais críticas à eficácia causal dos estados mentais, argumentando que posições como o naturalismo biológico se encontram diante de um dilema bastante incômodo: ou a causação mental é apenas ilusória (epifenomenalismo) ou um mesmo evento possui duas causas (sobredeterminação causal). Tanto a causação mental no naturalismo biológico como a crítica de Kim a esta posição serão apresentadas no próximo capítulo.

estados neurobiológicos. Sendo assim, os estados conscientes *são apenas* o nível superior, qualitativo e subjetivo dos processos cerebrais, não sendo possível identificá-los do ponto de vista ontológico, já que *o mesmo evento possui diferentes características*, a saber, características neurobiológicas e características fenomênicas⁷⁷: “One and the same event is a sequence of neuron firings and is also painful. But this type of identity does not give the materialists what they wanted [isto é, a redução ontológica]” (M, p. 125).

Poder-se-ia perguntar, então, se, para Searle, os estados mentais são ou não redutíveis a estados cerebrais? A resposta para essa pergunta depende de qual conceito de redução se tem em mente, o de redução causal ou de redução ontológica. De acordo com o primeiro, fenômenos do tipo A são *causalmente redutíveis* a fenômenos do tipo B, se, e somente se, “the behavior of A’s is entirely causally explained by the behavior of B’s, and A’s have no causal powers in addition to the powers of B’s”⁷⁸ (M, p. 119. Grifos meus). De acordo com a redução ontológica, por sua vez, os fenômenos superficiais são *ontologicamente redutíveis* a fenômenos subjacentes se, e somente se, aqueles forem nada exceto estes. Segundo Searle, os fenômenos mentais são redutíveis do ponto de vista causal (e não ontológico) aos fenômenos neurobiológicos, como será visto logo a seguir.

Tanto a liquidez como a consciência podem ser reduzidas causalmente aos microfenômenos subjacentes a elas, mas, no entanto, existe uma diferença crucial entre esses dois fenômenos naturais de nível superior, a saber: a redução causal da liquidez implica uma redução ontológica, o que não ocorre com a consciência. Portanto, mesmo assumindo que a consciência é um fenômeno biológico como a fotossíntese, digestão ou mitose, Searle adverte que diferentemente destes, aquela é ontologicamente irredutível: “Consciousness is entirely causally explained by neuronal behavior but it is not thereby shown to be nothing but neuronal behavior” (M, p. 119). Portanto, segundo Searle, de todos os fenômenos naturais do Universo, apenas *um*, a consciência, não se enquadra na regra geral de que uma redução ontológica *decorre* de uma redução causal⁷⁹. Isso gera uma *assimetria* dos fenômenos

⁷⁷ Ora, assumir a irredutibilidade das propriedades mentais a propriedades neurobiológicas é uma posição que o naturalismo biológico compartilha com o dualismo de propriedade e com o fisicalismo não-redutivo. No último capítulo desse trabalho, comparei o naturalismo biológico com essas duas teorias com intuito de propor uma classificação para a posição de Searle em Filosofia da Mente.

⁷⁸ Em relação à parte grifada, conferir a 6ª tese: As capacidades causais dos fenômenos neurobiológicos e mentais são as mesmas.

⁷⁹ Segundo Searle, ao se analisar a História da Ciência, constata-se que a redução causal é um passo para a redução ontológica ou a redução ontológica de propriedade: “Often, indeed typically, in the history of science we make an *ontological reduction* on the basis of a *causal reduction*” (M, p. 119). A propriedade cor vermelha, só para citar um exemplo, é *causada* por reflectâncias de luz com comprimento de onda de 700 nanômetros. A partir daí, os físicos *redefinem* o termo “cor vermelha” (característica superficial) como sendo a *mesma coisa* que

naturais: do lado mais numeroso estão fenômenos como, por exemplo, relâmpago, fotossíntese e gene, enquanto o outro lado é solitariamente ocupado por uma *única* classe de fenômenos – os fenômenos mentais, que possuem um ponto de vista único, que é a perspectiva de primeira pessoa. Essa perspectiva subjetiva⁸⁰ *impossibilita* redefinir a consciência em termos ontologicamente objetivos, como o são os termos neurocientíficos, sem perdê-la: “(...) *in the case of consciousness we can make a causal reduction but we cannot make an ontological reduction without losing the point of having the concept*” (M, p. 119). Como será mostrado na próxima tese, tal assimetria é uma consequência da dicotomia que Searle tanto defende e salienta entre fenômenos objetivos e subjetivos.

Como foi visto, fenômenos mentais não são mera aparência, mas existem de fato, são reais. Entretanto, eles só existem como experiências qualitativas e conscientes, pelo menos em potencial, de um sujeito, ou seja, os processos mentais são *subjetivos*, e unicamente o organismo que os possui é que tem acesso a eles. Dizer que a atividade mental de um indivíduo é subjetiva é o mesmo que afirmar que ela tem uma ontologia de primeira pessoa⁸¹, isto é, estados mentais só existem (ontologia) quando são experienciados por um sujeito (subjetiva): “Subjective conscious states have a first-person ontology (“ontology” here means mode of existence) because they exist only when they are experienced by some human or animal agent” (C, p. 41). Já entidades como nuvens, montanhas e árvores são ontologicamente objetivas (ontologia de terceira pessoa), pois seus modos de existência são independentes de qualquer sujeito.

Por causa dessa distinção entre ontologia de primeira e terceira pessoa é que Searle afirma que apesar dos estados mentais serem causalmente redutíveis a estados cerebrais, eles *não são* ontologicamente redutíveis aos mesmos. Ao diferenciar fisicalismo de espécime de fisicalismo de tipo, Kim oferece um exemplo bem ilustrativo de que correlação não é o mesmo que identidade: a asserção de que (i) “todo objeto que tem uma cor tem uma forma” não implica que (ii) “cores são idênticas a formas”. A primeira afirmação⁸² é facilmente constatada ao se observar uma bola amarela. Porém, não se segue deste exemplo que um tipo de cor (amarela) seja *necessariamente correlacionada* com um tipo de forma (esférica), já que podem existir objetos como, por exemplo, uma bola azul ou um cubo amarelo. Isso ocorre de

reflectâncias de luz com comprimento de 700 nm (fenômeno físico subjacente). Ou seja, a redução causal de cor a reflectâncias de luz levou, por redefinição, a uma redução ontológica de propriedades.

⁸⁰ No que diz respeito aos fenômenos mentais, o que é mais importante é sua perspectiva subjetiva, que decorre da ontologia de primeira pessoa desses fenômenos.

⁸¹ Em todo o Universo, somente duas características não possuem uma ontologia de terceira pessoa, a consciência e a intencionalidade: “Consciousness and intentionality are unique only in that they have a first-person ontology” (M, pp. 120-121).

⁸² Excetuando-se, é claro, os casos de objetos transparentes ou invisíveis.

maneira análoga com a teoria de identidade de espécime (fiscalismo de espécime)⁸³, pois a afirmação de que 1) “todo evento que tem uma propriedade mental também possui uma propriedade física” não implica que 2) “um tipo de estado mental seja idêntico a um tipo de estado físico”: “(...) even if every event that has a mental property has a physical property, that doesn’t even entail that mental properties are sistematically correlated with physical properties, much less that they are identical with physical properties” (KIM, 1996, pp. 59-60). Searle usa esse exemplo de Kim para defender que apesar de termos conhecimentos neurocientíficos o suficiente para afirmar que toda ocorrência de um evento mental é causada por uma ocorrência de um evento físico no cérebro, não se segue que aquele é idêntico a este, pois eles são ontologicamente distintos, o que vem a mostrar que o conceito de identidade é de pouca valia no que se refere ao problema mente-corpo:

(...) we can have a notion of neurobiological processes big enough so that every token pain process in a token neurobiological process in the brain, but it does not follow that the first-person painful feeling is the same thing as the third-person neurobiological process. The concept of identity is not much help with the mind-body problem because we can make our events big enough to include both the phenomenological and the neurobiological (M, p. 125).

Portanto, mesmo defendendo que a consciência é uma parte *natural* do mundo *real*, Searle afirma que ela difere de qualquer outro fenômeno natural, como a fotossíntese, mitose ou processos cerebrais, por ter uma ontologia de primeira pessoa, acarretando que estados fenomenológicos não são idênticos a processos neurobiológicos, apesar daqueles serem causados por estes. Poder-se-ia, pelo menos em princípio e visando alguns fins específicos, redefinir os estados conscientes de um indivíduo em termos de estados cerebrais⁸⁴, mas uma característica extremamente importante seria perdida, a *subjetividade*:

(...) the main point of having the concept of conciousness is to capture the first-person, subjective features of the phenomenon and this point is lost if we redefine conciousness in third-person, objective terms. We would still need a name for the first-person ontology (M, p. 120).

Neste ponto, uma crítica pode ser feita a Searle, pois segundo ele a terminologia tradicional faz uma divisão radical do mundo em dois domínios, o físico e o mental, o que acaba criando um dilema para quem vai tentar resolver o problema mente corpo – como ajustar o mental, que é não físico, ao universo material? Para Searle isso é um falso paradoxo que surge por se aceitar a existência de dois domínios, e basta abandonar tal terminologia

⁸³ Para maiores detalhes sobre a teoria de identidade de espécime, conferir o próximo capítulo.

⁸⁴ “Well, we could, and we might for medical or other scientific purposes redefine conciousness in terms of microsubstrates, as we have redefined solidity and liquidity” (M, p.120).

obsoleta para que a contradição desapareça. Todavia, Searle divide e classifica os fenômenos do Universo em duas categorias que *também* são mutuamente exclusivas: por um lado, há os fenômenos físicos, objetivos e com ontologia de terceira pessoa, e, por outro lado, existem os fenômenos mentais, subjetivos e com ontologia de primeira pessoa. É devido a essa distinção entre objetivo e subjetivo que Hodgson classifica o naturalismo biológico como uma forma de dualismo, o dualismo objetivo-subjetivo, afirmando que este não difere muito do dualismo físico-mental, que é tão atacado por Searle:

Searle picks out a property of consciousness, namely subjectivity, that distinguishes it utterly from all other higher-level or global or emergent properties with which he wants to liken it – such as biological features like digestion, or non-biological features like solidity and liquidity. The distinction he draws between the objective and the subjective merits the label ‘dualism’, because he asserts both that the two categories are exhaustive of empirical reality (certainly, he asserts no other category) and also that they are mutually irreducible (...). Furthermore, the dualism he asserts of the objective and the subjective is not significantly different from that which he denies, that of the physical and the mental (HODGSON, 1994, p. 265)

Searle recusa terminantemente que a distinção que ele faz entre objetivo-subjetivo se assemelhe ao dualismo, seja ele de substância ou de propriedade, já que a distinção que ele faz é *real*⁸⁵:

Once again it seems to me the illusion of dualism is generated by misunderstanding a very real distinction. There really is a distinction a between those irreducible features of the world that have a first-person or subjective ontology and those that do not. But is a deep mistake to suppose that that real distinction is the same as the old-time distinction between the mental e the physical, between *res cogitans* and *res extensa*, or that the subjective phenomena are something over and above the systems in which they are realized (M, p. 130).

Apesar de logicamente possível, Searle não leva o dualismo cartesiano muito a sério, já que ele postula a existência de almas imateriais e isto entra em contradição com a nossa visão de mundo científica (M, p. 132). No que diz respeito à outra vertente, Searle acredita que as teses defendidas por um dualista de propriedade o levam a um dilema insolúvel, relacionado à eficácia causal do mental, qual seja: caso a consciência seja causalmente

⁸⁵ Um dos motivos para Searle negar o rótulo de que seja um dualista é que para ele não faz o mínimo sentido perguntar acerca de quantas categorias metafísicas fundamentais existem no mundo, pois a categorização do mundo é *sempre* relativa ao observador, o que significa que ela é relativa e variável: “There are not two (or five or seven) fundamental ontological categories, rather the act of categorization itself is always interest relative. For that reason the attempt to answer such questions as, “How many fundamental metaphysical categories are there?”, as it stands, is meaningless. We live in exactly one world and there are as many different ways of diving it as you like” (WIANDPD, p. 3). Coisas como inflação ou jogo de futebol não se enquadram dentro de conceitos como físico ou mental. O que existe na realidade são várias características de um *único* mundo. Poder-se-ia perguntar se Searle é um pluralista, ao que ele responde da seguinte forma: “My view is not “pluralism”, IF that term suggests that there is a nonarbitrary, noninterest-relative principal of distinguishing the elements of the plurality” (WIANDPD, p. 4).

eficiente, estar-se-á diante do problema da sobre-determinação causal, pois teria que se explicar como é possível que um mesmo evento tenha duas causas, uma física e outra mental. Por outro lado, ao defender a tese oposta, ou seja, que a consciência é causalmente inerte, o dualista se depararia com o epifenomenalismo, o que é, para Searle, igualmente insatisfatório.

Não deixa de ser irônico que o paradoxo acima é justamente o mesmo que alguns críticos de Searle imputam ao naturalismo biológico⁸⁶. Searle se defende⁸⁷ afirmando que o naturalismo biológico está imune a esse dilema porque a eficácia causal da consciência *nada mais é* do que uma descrição de nível superior do funcionamento, através de neurônios e sinapses (micronível), do cérebro. É a partir disto que Searle se utiliza de sua analogia entre consciência e solidez para mostrar que nem uma e nem a outra estão *acima e além* da estrutura física subjacente. Portanto, ele pretende mostrar, ao mesmo tempo, que a eficácia causal da consciência e da solidez (defesa contra o epifenomenalismo) não implica a sobre-determinação causal (por causa dos diferentes níveis de descrição do *mesmo* sistema): “(...) consciousness is to neurons as the solidity of the piston is to the metal molecules. Both consciousness and solidity function causally. But neither is “over and above” the systems of which they are a part” (M, p. 131).

Alguns críticos⁸⁸ não se convenceram com a defesa de Searle e continuam o considerando um dualista de propriedade: “se a posição de Searle não é o dualismo de propriedades expresso em outros termos, então eu temo não saber o que ela é” (MASLIN, 2009, p. 171). A principal razão para que tais filósofos classifiquem Searle como um dualista é porque eles consideram que a cisão objetivo-subjetivo assumida pelo naturalismo biológico é apenas uma nova roupagem para a antiga dicotomia físico-mental. Caso se confirme isso, a teoria de Searle *implode*, pois ele estaria criticando e tentando abandonar os próprios conceitos que formam a base do naturalismo biológico⁸⁹: deixar de lado a diferença entre físico e mental seria equivalente ao abandono da distinção entre objetivo e subjetivo, que é uma tese central do naturalismo biológico. Na concepção cartesiana, físico e mental são mutuamente excludentes, o que não ocorre no naturalismo biológico pois, segundo Searle, os

⁸⁶ Como já foi dito, alguns filósofos classificam Searle como um dualista e propriedade. Conferir Corcoran (2001), Chalmers (1996, p. 164) e Maslin (2009, pp. 165-171).

⁸⁷ Ver também a 2ª e a 6ª tese, respectivamente: “fenômenos mentais são causados por e realizados no sistema cerebral” e “as capacidades causais dos fenômenos neurobiológicos e mentais são as mesmas”.

⁸⁸ Conferir também Maslin, 2009, p. 171: “(...) eu acredito que Searle deve graciosamente aceitar que ele é, afinal, um dualista de propriedades”. Como já foi dito, discutirei a classificação do naturalismo biológico no próximo capítulo.

⁸⁹ Se esse for o caso, pode-se apontar uma importante semelhança entre as posições de Descartes e Searle, já que os fenômenos considerados mentais – respectivamente, *res cogitans* e os fenômenos com ontologia de primeira pessoa – se sobrepõem nas duas teorias e são mentais pela mesma razão, a saber, porque são *subjetivos*, e é isso que os diferenciam de *todos* os outros fenômenos.

fenômenos mentais subjetivos *são* fenômenos físicos, isto é, *são* fenômenos naturais. De acordo com Prata (2007, pp. 193-194), Searle fracassa em sua tentativa de classificar os fenômenos mentais ontologicamente subjetivos como sendo um tipo de fenômeno natural⁹⁰. Por outro lado, se a resposta da pergunta for negativa, deve-se mostrar o porquê de tal diferença entre o dualismo e o naturalismo biológico e, mais importante, demonstrar se a contraposição entre objetivo-subjetivo é capaz de resolver questões que não são solucionadas a partir do arcabouço do dualismo conceitual. Desenvolverei tais questões no último capítulo.

3.4) 4ª tese: Não Expressabilidade dos Termos que Descrevem Fenômenos Mentais Através dos Termos que Descrevem os Fenômenos Neurobiológicos

Uma consequência da irreduzibilidade ontológica da consciência é que as características dos fenômenos mentais *não* podem ser expressas pela descrição dos fenômenos cerebrais subjacentes: mesmo supondo que quando um indivíduo tem dor haja a ativação de fibras C, não é possível, de acordo com Searle, expressar aquela em termos desta. Isso decorre pelo simples motivo de que a sensação subjetiva de dor é perdida caso se insista em redefini-la pelo disparo de fibras C. Afirmar que o único que sente a dor é o sujeito no qual as fibras C estão disparando não significa, segundo Searle, que outros indivíduos não *saibam* que o mesmo está com dor⁹¹. Esse é um problema de cunho *epistemológico* denominado de problema das outras mentes: como saber se outras pessoas têm mentes e, caso as tenham, como garantir que os estados mentais de tais pessoas sejam semelhantes aos meus.

⁹⁰ É interessante notar que, no que diz respeito à distinção entre objetivo e subjetivo, Thomas Nagel (2004) tem a esperança de que no futuro sejam desenvolvidos *conceitos alternativos*, pertencentes a uma *fenomenologia objetiva*, que permitirão abordar de maneira mais eficiente os estados mentais. Tal abordagem modificaria profundamente o significado do que hoje se chama de “físico” e “mental”: “(...) se a nossa idéia do físico vier a se expandir de modo a incluir os fenômenos mentais, ela terá que lhes atribuir um caráter objetivo – quer isso seja feito analisando-os em termos de fenômenos já considerados como físicos, quer não. Parece-me mais provável, contudo, que as relações mental-físico venham a ser expressas, ao final, em uma teoria cujos termos fundamentais não possam ser claramente localizados em nenhuma das duas categorias” (NAGEL, 2005, p. 262).

⁹¹ Há um problema de cunho *epistemológico* denominado de problema das outras mentes. Este, por sua vez, pode ser desmembrado em dois, a saber: (i) como saber se outras pessoas têm mentes? e (ii) caso as demais pessoas possuam uma mente, como garantir que os estados mentais de tais pessoas sejam semelhantes aos meus? O primeiro caso pode ser exemplificado da seguinte maneira: o senhor Smith tem uma mente ou ele é um autômato (ou zumbi)? Já o segundo tipo de problema parte da *pressuposição* que o outro indivíduo tem uma mente, mas não tem certeza se esta é igual à minha: assim como eu, Dalton tem uma mente, mas será que ele enxerga a cor vermelha do sangue da *mesma forma* que eu? Searle considera essas questões como sendo secundárias, ou melhor, ele nem leva a sério tais perguntas, já que, segundo ele, constituições biológicas semelhantes produzem mentes similares. Numa palavra, pode-se afirmar que os inúmeros seres humanos (sem neuropatologias) têm mentes semelhantes *porque* possuem o mesmo *fundamento causal* que as gera – o complexo sistema nervoso.

O vocabulário utilizado para descrever eventos mentais é diferente do vocabulário das Ciências Naturais, e disso surge uma interessante questão, a saber: é possível expressar os termos da Psicologia Popular⁹² (vocabulário mentalista) através dos *mesmos* termos que são utilizados para descrever os fenômenos cerebrais (vocabulário neurocientífico)? Não, pois a descrição em termos objetivos (terceira pessoa) não capta os fenômenos mentais, já que estes têm uma ontologia subjetiva (primeira pessoa): “(...) a description of the third-person facts does not entail the existence of the first-person facts, and this for the trivial reason that the first-person ontology cannot be reduced to the third-person ontology” (M, p. 129). Vê-se, a partir desta passagem, que a não expressabilidade do vocabulário mentalista através do vocabulário neurocientífico é uma *conseqüência* da irreduzibilidade ontológica da consciência (3ª tese):

Suponha que tentássemos dizer que a dor [ontologia de primeira pessoa] é na verdade “nada exceto” as disposições de descargas de neurônios [ontologia de terceira pessoa]. Bem, se tentássemos uma tal redução ontológica, as características essenciais da dor seriam deixadas de lado. Nenhuma descrição dos fatos de terceira pessoa, objetivos, fisiológicos comunicariam o caráter subjetivo, de primeira pessoa da dor, simplesmente porque as características de primeira pessoa são diferentes das características de terceira pessoa (RM, p. 170).

Mais uma vez, Searle insiste no fato de que no que diz respeito a estados conscientes o que realmente importa é o ponto de vista de primeira pessoa. Para fortalecer esta tese, ele se vale do argumento de que existem casos em que as características superficiais de um fenômeno são mais importantes do que as características da microestrutura que a causam. O valor estético (característica superficial) que é causado ao se observar a famosa escultura de Michelangelo, *Davi*, não pode ser expresso pela descrição molecular (microestrutura) do mármore no qual a obra foi esculpida. Apesar da estátua ser composta unicamente de moléculas, a propriedade de ser bela só *emerge* no macronível, e seria absurdo se um crítico de arte se valesse da descrição química da escultura para expressar sua beleza. Assim como a beleza do *Davi* de Michelangelo, a dor é uma propriedade emergente e não faz sentido, portanto, expressá-la pelos padrões de descargas das fibras C. Dessa maneira, mesmo supondo que os neurocientistas dispusessem de um *cerebroscópio* capaz de mapear com precisão *todas*

⁹² O termo psicologia popular (*folk psychology*) foi criado pelo filósofo americano Daniel Dennett em 1981 e diz respeito a uma teoria que os humanos possuem, tornando-os capazes de explicar o “comportamento de outros seres humanos recorrendo às idéias comuns de “intenção”, “crença”, “desejo” e outros termos que compõem o chamado vocabulário mentalista” (TEIXEIRA, 2008, p. 33). Filósofos eliminativistas, como Patricia e Paul Churchland, defendem que o vocabulário da psicologia popular será *eliminado* (e não *reduzido*) com o desenvolvimento da neurociência, já que “a estrutura psicológica de nosso senso comum é uma concepção falsa e radicalmente enganosa das causas do comportamento humano e da natureza da atividade cognitiva” (CHURCHLAND, 1998, p. 79). Dessa forma, em vez de dizer “estou com dor”, dir-se-á, no futuro, algo como “minhas fibras C estão sendo ativadas”.

as sinapses que ocorrem no momento que um indivíduo sente dor, isso não seria suficiente para expressar o caráter subjetivo de tal sensação. Isso seria equivalente a querer analisar esteticamente a beleza do *Davi* por meio da descrição átomo por átomo das moléculas que constituem a estátua. Nesse ponto, pode-se citar Nagel, que afirma que o abandono da perspectiva de primeira pessoa impossibilita que haja uma descrição satisfatória dos fenômenos subjetivos:

Se o caráter subjetivo da experiência é completamente compreensível *somente* de um ponto de vista, então qualquer deslocamento em direção a uma *objetividade* maior – isto é, menos vinculada a um ponto de vista específico – não nos leva mais próximo da *natureza real* do fenômeno: leva-nos para mais longe dela (NAGEL, 2005, p. 256. Grifos meus).

Dessa forma, a distinção que existe no naturalismo biológico entre ontologia de primeira pessoa e ontologia de terceira pessoa está relacionada ao fato de que são *necessárias* duas linguagens diferentes para expressar os fenômenos mentais e neurobiológicos. Nesse ponto, pode-se invocar outra analogia para explicitar o que Searle quer dizer. Um quadro é um *objeto físico* composto pela moldura, pela tela e por diversos pigmentos de tinta dispostos de determinado modo. Porém, de modo semelhante à escultura, a apreciação estética do quadro não se expressa pela *descrição física* do quadro⁹³, onde, por exemplo, poder-se-ia fazer um mapeamento milímetro a milímetro da disposição dos pigmentos na tela e de seus respectivos comprimentos de onda: a tinta que reflete a luz no comprimento de onda de 700 nm está nas coordenadas (34, 57), (45, 39) e (57, 18), e assim por diante. O quadro será *descrito esteticamente* não pela linguagem físico-matemática de reflectâncias de luz e coordenadas cartesianas, mas por termos como beleza, emoção e êxtase. Em resumo, é possível descrever *adequadamente* um quadro (objeto físico) com vocabulários diferentes, mas tais termos só terão significado no interior de sua respectiva linguagem, ou seja, eles não são intercambiáveis. Um bom exemplo disso é o que se segue⁹⁴:

(...) Quando olho um quadro vejo objetos físicos: manchas de pigmento lambuzadas sobre uma tela. E posso descrever esses objetos tão completamente a ponto de dar conta do quadro inteiro. Ao fazê-lo, não menciono a outra coisa que estou vendo: uma caçada ao cervo em frente a uma casa de campo. Isso também eu poderia descrever tão completamente a ponto de dar uma explicação completa do quadro. *Mas as duas explicações são incompatíveis*: não posso atravessar, por assim dizer, de uma para outra. Não posso dizer que o líder da matilha de cães de caça está

⁹³ É claro que tanto o artista como o crítico de arte se interessam pelo material utilizado para produzir a obra. Pode-se produzir uma escultura com madeira, mármore, cobre ou gelo, por exemplo, e a variação da matéria-prima acaba influenciando de algum modo na avaliação estética. Porém, essa questão é secundária ou até mesmo sem importância para a discussão *desta tese*.

⁹⁴ Na verdade, essa analogia é utilizada por Roger Scruton (2008) para explicar a relação entre mente e corpo no monismo de Spinoza.

perseguindo furiosamente uma mancha ocre, nem descrever a área de amarelo-cromo misturado com óleo de linhaça como apoiada sobre o joelho do caçador (SCRUTON, 2008, p. 78. Grifos meus).

Saindo do exemplo da obra de arte e retornando à Filosofia da Mente, o que interessa aqui são os diferentes papéis desempenhados pelas linguagens da Neurociência e da Psicologia Popular: dentro do naturalismo biológico, explicar causalmente a consciência é determinar sua descrição neurobiológica, mas não é possível expressar um fenômeno de uma linguagem a partir dos termos da outra, isto é, não é possível *expressar* o quão desagradável é sua dor de dente em termos fisiológicos, embora seja possível *explicá-la* desta forma. Em outras palavras, os estados mentais são causados por processos cerebrais (2ª tese) e, por isso, podem ser explicados a partir do vocabulário neurocientífico (5ª tese); todavia, como os estados psicológicos são ontologicamente distintos dos processos neurobiológicos que o causam (3ª tese), as linguagens utilizadas para descrevê-los também devem ser diferentes (4ª tese). Isso é uma consequência lingüístico-conceitual da subjetividade ontológica dos fenômenos mentais⁹⁵: enquanto os processos neurobiológicos de nível inferior são descritos pela perspectiva de terceira pessoa da linguagem neurocientífica, a Psicologia Popular é não só adequada, mas *imprescindível* para abordar os fenômenos mentais de nível superior.

Por fim, deve-se enfatizar novamente que quando Searle afirma que as características subjetivas da consciência são *indescritíveis* em vocabulário neurocientífico ele *não* está defendendo que a consciência é *inexplicável* ou que existe alguma *lacuna explicativa* entre os processos cerebrais, de um lado, e os fenômenos conscientes, do outro (PRATA, 2008, pp. 7-11). Esse tema será tratado na próxima seção, a qual mostrará a ligação existente entre a *redução causal* dos fenômenos mentais e a *explicação causal* destes últimos através de processos neurobiológicos.

3.5) 5ª tese: Fenômenos Neurobiológicos Explicam Causalmente os Fenômenos Mentais

Ao defender que fenômenos mentais são fenômenos biológicos (1ª tese), Searle se compromete com a posição de que a explicação daqueles deve provir, em última análise, da Biologia e da Neurobiologia. Chamarei a primeira de *tese da natureza biológica dos*

⁹⁵ Searle caracteriza da seguinte maneira a relação entre irreducibilidade ontológica da consciência e a não expressabilidade dos fenômenos conscientes a partir de um vocabulário de terceira pessoa: “The argument is correct in pointing out that a description of the third-person facts does not entail the existence of the first-person facts, and this for the trivial reason that the first-person ontology cannot be reduced to the third-person ontology” (M, p. 129).

fenômenos mentais e a outra de *tese metodológico-explicativa*. Esta tese quer dizer que os métodos e explicações adequados para tratar os processos mentais são aqueles provenientes das Ciências Biológicas⁹⁶, principalmente da Biologia Evolutiva e da Neurobiologia. Aliás, estas ciências têm papéis *distintos e complementares* na tarefa de explicar⁹⁷ a consciência: a Biologia Evolutiva fornece uma explicação do tipo “por quê” e a Neurobiologia uma explicação do tipo “como”. Em outras palavras, uma vai explicar qual a *vantagem evolutiva* conferida a um conjunto de organismos que há vários milhões de anos produzem e sustentam a consciência, ou seja, por que é vantajoso, darwinisticamente falando, ser consciente⁹⁸. A outra explicação é próxima no sentido de tentar responder *como* o comportamento de determinados sistemas nervosos causam a consciência, *como* a consciência é realizada na estrutura cerebral e *como* estados conscientes causam o comportamento (DR, p.172). Numa palavra, já que os processos mentais são biológicos, deve-se explicá-los por meio do arcabouço conceitual e metodológico das Biociências, levando-se em consideração a existência de dois níveis de explicação que são, ao mesmo tempo, adequados e complementares.

⁹⁶ Esse ponto fica claro quando ele explica o rótulo de sua própria teoria – *naturalismo* porque a mente é parte integrante da natureza e *biológico* “porque o *modo de explicação* da existência de fenômenos mentais é *biológico* – por oposição (...) a computável, comportamental, social ou lingüístico” (MLS, p. 57. Grifos meus).

⁹⁷ Obviamente, Searle é consciente que a atividade científica não é estática. Sendo assim, o conhecimento científico faz *progressos* com o passar do tempo: hoje sabemos mais do que ontem, e amanhã saberemos mais do que hoje. Em relação à interpretação da mudança do conhecimento científico através de mudanças de paradigmas (Searle acha que o próprio Kuhn pensava diferente. Ver MLS, p. 31), o que acarretaria uma interpretação *anti-realista* do mundo: o conceito de massa não é o mesmo para as teorias de Aristóteles, Newton e Einstein, por exemplo. Como o naturalismo biológico parte do pressuposto geral do realismo externo, Searle é um crítico severo da concepção segundo a qual os mundos de Aristóteles, Newton e Einstein *não são o mesmo mundo* por suas respectivas teorias corresponderem a três diferentes paradigmas. Para se falar sobre o mundo, faz-se necessário a linguagem, claro. Contudo, isso não significa que a realidade não seja independente daquela: “o fato de haver água salgada no oceano Atlântico ocorria muito antes de haver alguém para identificar aquela porção de água como o oceano Atlântico, para identificar de que era feita essa água, ou para identificar um de seus componentes químicos como o sal. Obviamente, para que possamos fazer todas essas identificações, devemos ter uma linguagem, mas e daí? *Os fatos existem, totalmente independentes da linguagem.* (...) Inventamos palavras para afirmar fatos e para dar nome a coisas, mas isso *não significa* que inventamos os fatos ou as coisas” (MLS, p. 30. Grifos meus). No que diz respeito à relação mente-corpo, Searle diz que é um fato biológico bruto que os processos psicológicos são causados por processos cerebrais. Isso significa que a atual explicação neurobiológica da consciência e da intencionalidade pode, e provavelmente vai, mudar nos detalhes, mas não no *modo* ou *tipo*. Quando Searle explica a sede em termos fisiológicos, baseado no atual conhecimento científico, ele deixa claro que “certamente, os eventos do hipotálamo, por si mesmos, são provavelmente insuficientes para qualquer estado consciente. É preciso haver todo tipo de conexões com outras partes do cérebro [hipótese a ser confirmada por um estágio mais avançado da Fisiologia]. Mas a finalidade do exemplo é mostrar como uma explicação desse tipo [baseada nas Ciências Biológicas] *pode* fornecer uma explicação neurológica, e portanto naturalista, de certa forma de intencionalidade [o mesmo vale para a consciência] (MLS, p. 92). Em outra passagem, Searle afirma que a nossa atual neurociência é imperfeita, haja vista que “o nosso conhecimento acerca do modo como o cérebro funciona está em permanente aumento. No entanto, a *forma* da explicação pode permanecer válida, mesmo se os *pormenores* se alteram (MCC, p. 23).

⁹⁸ “Os processos da racionalidade consciente representam uma parte tão importante de nossas vidas e, além disso, constituem uma parte biologicamente tão extensa delas que se um fenótipo de tamanha magnitude não desempenhasse nenhum papel funcional para a vida e a sobrevivência do organismo, nosso conhecimento da própria evolução é que estaria sendo refutado” (LN, pp. 57-58).

O nível adequado para se explicar os fenômenos mentais também surge, no naturalismo biológico, como uma consequência da 2ª tese: já que os estados mentais são causados e realizados no sistema cerebral, os estados neurobiológicos explicam os estados psicológicos. Sendo assim, as propriedades emergentes de um nível devem ser explicadas a partir dos microprocessos que a causam: para se explicar a solidez ou a liquidez, deve-se compreender o comportamento molecular de nível micro; de maneira análoga, a explicação dos fenômenos mentais (macronível) deve provir das interações sinápticas que ocorrem no sistema nervoso (micronível). Mais uma vez, a analogia entre solidez e liquidez, de um lado, e fenômenos mentais, por outro, é limitada, pois, para Searle, a redução causal da consciência a processos cerebrais não tem como consequência a redução ontológica da mesma: a dor é *causada* pelo disparo de fibras C, mas enquanto este tem uma ontologia de terceira pessoa, aquela tem um modo de existência subjetivo. Dito de outra forma, a dor e o disparo de fibras C, segundo Searle, são a *mesma* coisa em termos causais, mas não do ponto de vista ontológico.

Dessa maneira, a redução causal do mental possibilita que os fenômenos mentais sejam causalmente explicados por fenômenos cerebrais, já que aqueles surgem do comportamento destes⁹⁹. Isso se deve porque a noção de explicação presente no naturalismo biológico é bem peculiar: “Na concepção de Searle o ponto decisivo é que o fenômeno que explica seja *causalmente suficiente e necessário* para o fenômeno que é explicado (...)” (PRATA, 2008, p. 7). No caso, o fenômeno neurobiológico causalmente suficiente e necessário para explicar o fenômeno mental é o que é chamado por *correlato neural da consciência* (NCC). Quando, se é que isso é possível, a Neurociência descobrir todos os correlatos neurais dos estados psicológicos, ter-se-á alcançado o tão sonhado estágio de uma *explicação causal* completa da consciência. Caso isso se concretize, será possível afirmar coisas como: o estímulo das fibras nervosas X causa o estado mental Y, pois sem tal estímulo Y não ocorre (*necessidade*), e basta que as fibras X disparem para que o indivíduo se encontre no estado psicológico Y (*suficiência*). Logo, conclui-se que a *explicação adequada* da produção de Y é uma explicação causal, ou seja, deve-se conhecer o fenômeno neurobiológico subjacente (causa) para se explicar o fenômeno mental de nível superior

⁹⁹ Searle admite, inclusive, que é possível mapear o cérebro, relacionando uma região cortical com o estado mental que ela produz. Esse possível cenário da Neurociência do futuro preconizado por Searle se assemelha, *nesse sentido*, ao projeto de pesquisa fisicalista: “Tivéssemos um conhecimento perfeito de como o cérebro produz, por exemplo, a sede ou as experiências visuais, não hesitaríamos em atribuir regiões no cérebro para tais experiências, caso as evidências apoiassem uma tal atribuição” (INT, p. 375).

(efeito). Além disso, constata-se que isso é uma questão empírica e que faz parte, portanto, do escopo da Neurociência.

Todavia, pode-se questionar sobre a possibilidade de se conciliar coerentemente a não expressabilidade do mental (4ª tese) com a tese de que os fenômenos mentais são explicáveis causalmente (5ª tese). A situação é problemática porque se os fenômenos mentais (ontologia subjetiva) não podem ser expressos nos termos da Neurociência, eles também não poderiam, *em princípio*, ser explicados por processos cerebrais subjacentes (ontologia objetiva):

(...) se uma descrição dos fatos neurobiológicos não é sequer capaz de *expressar* o caráter subjetivo dos fenômenos mentais, então parece impossível que esses fatos possam *explicar* tal caráter subjetivo, ou seja, parece impossível que o recurso a fatos neurobiológicos possa mostrar o *porquê* de o caráter subjetivo dos fenômenos mentais ser tal como ele é (PRATA, 2007, p. 190).

Já que a descrição neurofisiológica é incapaz de expressar e, conseqüentemente, explicar a subjetividade dos estados mentais, isso parece mostrar que o naturalismo biológico é, à primeira vista, semelhante ao dualismo de propriedade. No entanto, só tratarei disso no próximo capítulo.

3.6) 6ª tese: As Capacidades Causais dos Fenômenos Neurobiológicos e Mentais são as Mesmas

Assumir que a consciência é ontologicamente irreduzível a processos cerebrais não implica, segundo Searle, afirmar que aquela tenha capacidades causais que não se encontrem nestes, isto é, as capacidades causais dos fenômenos de nível superior (estados conscientes) são *idênticas* às capacidades causais dos fenômenos de nível inferior (processos neurobiológicos subjacentes). Essa tese da identidade das capacidades causais entre fenômenos neurobiológicos e mentais parece, *à primeira vista*, com a concepção physicalista e reducionista da mente, pois mostra que esses dois fenômenos não são coisas independentes: “If two things in the real empirical world have an independent existence they must have different causal powers. But the causal powers of consciousness are exactly the same as those of the neuronal substrate” (M, pp. 127-128).

Mais uma vez, a analogia com a liquidez pode ser aqui invocada para esclarecer melhor esse ponto. A liquidez da água (característica de nível superior) é causada pela interação de moléculas de H₂O através de pontes de hidrogênio (micronível)¹⁰⁰, e a liquidez

¹⁰⁰ Observe que tal descrição se refere à redução causal de um fenômeno de nível superior a outro de nível inferior, “em que se demonstra que a existência e, *a fortiori*, as capacidades causais da entidade reduzida são inteiramente explicáveis em termos das faculdades causais dos fenômenos redutores” (RM, p. 166). Dessa

não possui *nenhuma* capacidade causal adicional em relação ao conjunto de moléculas que a constitui. Portanto, a fluidez, a densidade, o alto calor específico e a elevada tensão superficial, que são características da água em estado líquido, podem ser explicadas causalmente pelo comportamento do sistema no micronível, a saber, através das pontes de hidrogênio que se estabelecem entre as moléculas de H₂O. Isso significa que a liquidez e as moléculas de H₂O não são duas entidades independentes, mas, sim, o mesmo sistema em diferentes níveis. Com essa analogia, Searle pretende mostrar que a consciência é uma propriedade emergente e natural semelhante à liquidez (como já foi mostrado, essa analogia é apenas parcial).

A união da 2ª tese (fenômenos mentais são causados por e realizados no sistema cerebral) e da 6ª tese (identidade das capacidades causais dos fenômenos neurobiológicos e mentais) tem um importante papel dentro do naturalismo biológico, que é livrá-lo do dilema epifenomenalismo-sobredeterminação causal. Por um lado, fenômenos mentais não são epifenômenos, já que são causalmente eficientes e possuem as mesmas capacidades causais dos processos cerebrais. Dessa maneira, afirmar que meu desejo de levantar o braço (M) *causou* o movimento deste membro (P*) se explica da seguinte forma: M tem os mesmos poderes causais que o estado neurobiológico subjacente (P), e este, por sua vez, tem a capacidade de *causar* o movimento de meu braço (P*). Por outro lado, não se pode afirmar que um mesmo evento possui duas causas no naturalismo biológico, pois, segundo Searle há apenas uma *única causa* descrita em dois níveis diferentes, o neurobiológico (P) e o mental (M). Ou seja, a causa do movimento de meu braço (P*) pode ser descrita no nível neurofisiológico, através da liberação de neurotransmissores e de sinapses, ou no nível da Psicologia Popular, por meio dos termos que designam os estados mentais, como desejar que x, onde x é, nesse exemplo, levantar o braço.

Essa solução de Searle para o dilema epifenomenalismo-sobredeterminação causal, todavia, é problemática, pois quando a 3ª tese é analisada conjuntamente com a 6ª, constata-se que elas são mutuamente exclusivas: a distinção ontológica entre fenômenos neurobiológicos e mentais (3ª tese) *implica* que tais fenômenos possuem diferentes poderes causais (PEREBOOM & KORNBLITH, 2004, pp. 714-715). Dessa maneira, ou Searle deve abrir mão da distinção ontológica ou da identidade das capacidades causais. Como se sabe, o preço a se pagar por qualquer uma das alternativas é muito elevado: se, por um lado, Searle

maneira, pode-se dizer que a liquidez é causalmente redutível ao comportamento molecular subjacente, da mesma maneira que os fenômenos conscientes são, segundo Searle, causalmente redutíveis a processos cerebrais (ver a 2ª tese: Fenômenos mentais são causados por e realizados no sistema cerebral).

considerar que o mental é idêntico ao físico, estaria assumindo uma posição fisicalista redutiva; por outro lado, defenderia algum tipo de fisicalismo não-redutivo ou de dualismo de propriedade caso rejeitasse a identidade dos poderes causais. Nenhuma dessas possíveis saídas agrada Searle, mas a tese da subjetividade ontológica da consciência é suficiente para descartar que ele seja um fisicalista redutivo, já que tal tese ocupa um lugar central no naturalismo biológico. Ou seja, a subjetividade ontológica da consciência acarreta a irreduzibilidade desta a processos neurobiológicos e, conseqüentemente, o naturalismo biológico se enquadra como uma posição anti-reducionista em Filosofia da Mente. Porém, o naturalismo biológico é um tipo de fisicalismo não-redutivo ou de dualismo de propriedade? No próximo capítulo responderei essa pergunta e proporei uma classificação do naturalismo biológico dentro da taxonomia teórica da Filosofia da Mente.

4. REDUACIONISMO E IRREDUTIBILIDADE DA CONSCIÊNCIA NO NATURALISMO BIOLÓGICO

Após ter apresentado no capítulo anterior as teses que compõem o naturalismo biológico, pretendo no presente capítulo propor a classificação do mesmo em uma das vertentes da Filosofia da Mente (seções 4.3 e 4.4). Porém, antes disso, abordarei dois temas que servirão para fundamentar a minha proposta, a saber: (i) o conceito de redução no âmbito da Filosofia da Mente (seção 4.1), e (ii) o conceito de causalção no naturalismo biológico (seção 4.2).

4.1) Redução e Filosofia da Mente

Em se tratando de redução em Filosofia da Mente, duas questões se colocam, a saber: (i) o avanço da Neurobiologia acarretará a redução da Psicologia Popular?, e (ii) o desenvolvimento do conhecimento neurocientífico acabará por mostrar que estados psicológicos são *nada exceto* estados neurofisiológicos? A primeira questão diz respeito à redução teórica, enquanto a segunda se refere à redução ontológica. É possível ter redução ontológica sem redução teórica, mas não o contrário¹⁰¹. No que se segue, será focado o segundo tipo de redução.

A posição anti-reducionista em Filosofia da Mente defende a irredutibilidade dos estados mentais, afirmando que essa irredutibilidade não é um problema empírico, mas um problema ontológico-conceitual¹⁰². Um filósofo anti-reducionista pode atacar tanto a necessidade como a suficiência da propriedade neurobiológica para produzir a propriedade mental. Em primeiro lugar, ele pode questionar que em outros organismos, reais ou

¹⁰¹ Por exemplo, o avanço da Biologia mostrou que os fatores mendelianos são nada exceto fragmentos de moléculas de DNA, mas isso *não significa* que a tentativa de reduzir a Genética Mendeliana à Genética Molecular obtenha sucesso. De maneira semelhante, seria *possível* sustentar que a Psicologia Popular é irredutível à Neurociência mesmo se os cientistas descobrissem que os estados mentais são nada exceto estados cerebrais (redução ontológica).

¹⁰² Dessa forma, os filósofos não-reducionistas defendem que a irredutibilidade do mental não é decorrente de uma defasagem e imperfeição da atual Neurociência (atualmente, não é possível reduzir à Psicologia à Neurociência, mas será possível no futuro). Na verdade, a irredutibilidade dos estados mentais está relacionada à diferença ontológica e conceitual desses estados em relação aos estados cerebrais. Para o naturalismo biológico, por exemplo, a irredutibilidade (ontológica) ocorre porque os primeiros têm uma ontologia de primeira pessoa e os últimos possuem ontologia de terceira pessoa.

possíveis¹⁰³, um mecanismo subjacente que não seja o estímulo de fibras C realize a dor, mostrando que aquele *não é* necessário para esta (na verdade, Searle alerta para o fato de que as fibras C não são as responsáveis pelas sensações de dor. Para maiores detalhes, conferir a nota 5, do capítulo 2, pp. 359-360). Isso é bem plausível do ponto de vista biológico, já que devido à enorme biodiversidade uma mesma função pode ser executada por diferentes estruturas¹⁰⁴. Ou seja, provavelmente um cientista não ficaria espantado se descobrisse que em determinada espécie a dor é realizada pelo estímulo de fibras Z, por exemplo. Do ponto de vista evolutivo, animais que sentem dor levam mais vantagens sobre os que não a sentem, já que, por fugirem do perigo ao evitar a dor, sobrevivem mais e, conseqüentemente, têm maior probabilidade de deixar descendentes, os quais herdarão tal característica vantajosa. Porém, isso não significa dizer que todos os animais devem ter a mesma base neurobiológica subjacente. O que interessa aqui não é saber se *de fato* existem animais que sentem dor pela ativação de uma célula nervosa que não seja a fibra C, isso é um problema empírico que diz respeito aos biólogos. A questão crucial para o debate filosófico, de acordo com o fisicalista não-reduutivo, é outra, a saber: caso haja a *possibilidade* da dor ser realizada, em outros organismos, por meio de um mecanismo distinto do estímulo de fibras C, isso implicaria que o estímulo destas fibras *não é* uma condição necessária para a sensação de dor.

Em segundo lugar, o opositor do reducionista pode alegar que é possível ter o estímulo de fibras C e, mesmo assim, o organismo não sentir dor. Para mostrar isso, o filósofo dualista David Chalmers se vale de um conhecido experimento mental: segundo ele, há a possibilidade lógica de existir um organismo fisicamente idêntico a um ser humano, molécula a molécula, que seja desprovido de estados mentais. Em tal organismo, o *zumbi*, o estímulo de fibras C, que do ponto de vista neurobiológico são idênticas às nossas, não realiza a dor, já que ele não possui vida mental. Dessa forma, essa experiência de pensamento pretende mostrar que o estímulo de fibras C *não é* condição suficiente para a dor, pois o zumbi possui aquele e não esta.

Suponha que a dor seja instanciada por uma variedade de realizadores físicos, tais quais: fibras C, X, Y e Z. Constata-se que todos são *suficientes* para a dor, pois o organismo que tem qualquer uma dessas fibras estimuladas sente dor. No entanto, observa-se que nenhum destes realizadores é *necessário* para a dor, já que esta pode ser instanciada, de

¹⁰³ É logicamente possível pensar em organismos que possuam vida mental e a sustentem a partir de um substrato físico bastante distinto do nosso, como circuitos de silício (computador) e “gosma” verde (alienígena).

¹⁰⁴ Um exemplo disso é a função de respiração, que pode ser executado por diferentes estruturas no *mesmo* organismo ao longo de sua vida: o girino (estágio larval) respira através de brânquias e o sapo (animal adulto) respira tanto através da pele quanto pelos pulmões.

acordo com este exemplo, de quatro diferentes maneiras. Portanto, a dor ocorre se e somente se houver o estímulo das fibras C ou X ou Y ou Z, ou seja, existem, nesse exemplo, quatro realizadores físicos (P), que são as fibras C (P₁), X (P₂), Y (P₃) e Z (P₄) (daqui em diante, a disjunção entre esses realizadores será representada simplesmente por UP_i). A dor ocorre se e somente se houver estímulo de um dos realizadores físicos (R*), o que acarreta a identidade entre a disjunção dos realizadores físicos e a propriedade de sentir dor (R**) ¹⁰⁵:

(R*) UP_i ↔ M;

(R**) UP_i = M.

Em outras palavras, como é bastante difícil negar a *tese da realização múltipla*, que afirma que uma mesma propriedade mental pode ser realizada por diferentes substratos físicos (realizadores físicos – P), uma possível saída para o reducionista é defender que o bicondicional (R*) e, conseqüentemente, a identidade (R**) ocorrem não entre uma propriedade física e uma mental, mas entre a *disjunção* ¹⁰⁶ dos realizadores físicos e a propriedade mental.

Para um fisicalista reducionista, os objetos e propriedades da Física são – dos pontos de vista descritivo, explicativo e ontológico – o nível mais fundamental do mundo, e os estados mentais não são exceções, já que são redutíveis a estados neurobiológicos; são redutíveis *porque* são idênticos. O fisicalista também assume que propriedades de nível superior dependem de, ou são determinadas pelas, propriedades de nível inferior:

(...) higher-level properties a given entity has are totally fixed by the lower-level properties and relations characterizing its parts. Generally, then, a dependency relation characterizes both the entities and properties at adjacent levels. Higher-level entities are determined by the lower-level ones in that they are “mereological” structures wholly decomposable into parts that belong to the lower level. It is this asymmetric and transitive part-whole relation that generates a hierarchy of levels or tiers of entities ¹⁰⁷ (KIM, 1996, p. 222).

¹⁰⁵ Searle não aceita o argumento do zumbi, pois, segundo ele, é *logicamente impossível* existir um mundo fisicamente idêntico ao nosso sem que existam fenômenos conscientes. Jaegwon Kim também rejeita o argumento de Chalmers: “In fact, I believe the zombie hypothesis is untenable; *Chalmers-style zombies are not conceptually possible*” (KIM, 2005, p. 169, nota 16. Grifos meus). No entanto, caso tal argumento seja levado em consideração, a tese supracitada de que a dor ocorre se e somente se houver estímulo de um dos realizadores físicos ficaria comprometida: um zumbi pode ter seu realizador físico estimulado sem que haja qualquer instanciação de dor.

¹⁰⁶ Nas palavras de Kim, a *estratégia da disjunção* pode ser enunciada da seguinte maneira: “Given that P₁, P₂, ... are all the realizers of M, M is instantiated on a given occasion only if one or another of these P_s is instantiated” (KIM, 1996, p. 219).

¹⁰⁷ A *relação parte-todo assimétrica e transitiva* citada por Kim se dá entre entidades. Isso significa que as entidades de um nível podem ser decompostas nas entidades do nível subjacente, até chegar ao nível das

Kim defende um *fisicalismo de superveniência* ao defender que propriedades mentais sobrevêm a propriedades neurobiológicas, ou seja, só pode haver mudança de estados mentais se houver, *necessariamente*, uma mudança dos estados físicos subjacentes. Numa palavra, a superveniência do mental ao físico significa que as propriedades mentais (M) dependem de, ou são determinadas pelas, propriedades neurobiológicas (P) subjacentes. Ao se levar este raciocínio às últimas conseqüências, conclui-se que os estados mentais são causalmente inertes em dois sentidos: um estado mental *não é* a causa de outro estado mental, e um estado mental *não* causa um estado físico. Em outras palavras, como os estados mentais não possuem novos poderes causais, eles seriam considerados *epifenômenos* e, conseqüentemente, não seria adequado se falar em causação mental, mas em *causação epifenomenica* (KIM, 1993c, p. 107). Em resumo, alguém poderia afirmar que o fisicalismo de superveniência acarreta o *epifenomenalismo*¹⁰⁸. Todavia, uma interpretação não epifenomenalista é possível: a redução ontológica dos estados mentais a estados neurobiológicos¹⁰⁹ implica que aqueles possuem eficácia causal, já que são idênticos a estes. Ou seja, já que estados neurobiológicos são causalmente eficientes, e estados mentais são redutíveis a estes, o fisicalista redutivo pode assumir a tese da eficácia causal do mental, ressaltando, entretanto, que tal eficácia resulta da identidade entre os poderes causais dos dois níveis, físico-neurobiológico e mental. Sendo assim, “salva-se” a causação mental ao se admitir que os estados mentais não possuem quaisquer poderes causais que não estejam presentes nos estados neurobiológicos. Aliás, Kim mudou sua posição de um fisicalismo redutivo epifenomenalista, descrito anteriormente, para um fisicalismo redutivo compatível com a eficácia causal do mental (Kim abandonou o fisicalismo de superveniência e passou a adotar uma posição mais forte, passando a assumir que a maioria dos estados mentais é idêntica a estados cerebrais. Conferir KIM, 1996, 1998),

partículas físicas fundamentais, que é o mais básico: células são decompostas em moléculas; estas, em átomos; e estes, por fim, são decompostos em partículas físicas fundamentais.

¹⁰⁸ A superveniência é um conceito *neutro* no sentido de poder ser utilizado por posições diametralmente opostas – isto é, tanto pelo reducionismo como pela teoria que concebe o mental como sendo irreduzível ao neurobiológico: “Although supervenience perhaps doesn’t imply reducibility, it need not be taken to imply irreducibility either – that is, supervenience suffices for the purposes of the nonreductive physicalist if it is consistent with both reducibility and irreducibility” (KIM, 1996, p. 223). A concepção fisicalista de Kim exemplifica a primeira concepção, enquanto que o monismo anômalo, proposto por Donald Davidson, é um exemplo da segunda.

¹⁰⁹ Posteriormente, Kim abandonou o fisicalismo de superveniência e passou a adotar uma posição mais forte, assumindo que a *maioria* dos estados mentais é idêntica a estados cerebrais (KIM, 1996, 1998). Porém, nem todos os estados mentais são redutíveis – a exceção são os *qualia*: “my view is that the qualitative characters of conscious experience, what are now commonly called “qualia”, are irreducible, but that we have reason to think that the rest, or much of it anyway, is reducible” (KIM, 2005a, p. 162). Como para Kim os *qualia* são irreduzíveis a algo físico e o domínio físico é causalmente fechado, os *qualia* são causalmente inertes, ou seja, são epifenômenos: “the mental residue, insofar as it resists physical reduction, remains epiphenomenal. It has no place in the causal structure of the world and no role in its evolution and development” (KIM, 2005a, p. 171).

desde que esta esteja restrita ao domínio físico causalmente fechado: “Physical reduction would save causal efficacy for mentality, at cost of its autonomy. Reductionism allows only one domain, the physical domain, but the mental may find a home in that domain” (KIM, 2005a, p. 159).

Desse modo, pode-se verificar que Kim e outros fisicalistas redutivos endossam a redução ontológica de estados psicológicos a estados físicos subjacentes. Porém, já que Searle, assim como Kim, defende que as capacidades causais dos fenômenos mentais são idênticas a dos fenômenos neurobiológicos (6ª tese), por que eles chegam a conclusões contrárias no que diz respeito à redução ontológica? O principal motivo é a distinção – feita por Searle, mas não por Kim – entre ontologia de primeira pessoa e de terceira pessoa (3ª tese). O segundo motivo é porque Searle diferencia redução ontológica de redução causal (5ª tese): estados mentais são causados por estados neurobiológicos, mas são ontologicamente irreduzíveis a estes. Isso é insustentável dentro do fisicalismo de Kim. Por fim, a noção de causação de Kim *não é* a mesma do naturalismo biológico. Nota-se, a partir do princípio do fechamento causal do domínio físico, que Kim defende a concepção humeana de causação, o que não ocorre com Searle.

4.2) Causação Mental no Naturalismo Biológico

No meu entender, a maior dificuldade de se classificar o naturalismo biológico se deve à aceitação da redução causal ao mesmo tempo em que se nega a redução ontológica dos estados mentais. Portanto, antes de propor qualquer tipo de classificação da teoria de Searle, deve-se explicitar o conceito de causação humeana (seção 4.2.1), a crítica de Searle a Hume (seção 4.2.2) e a tese da eficácia causal do mental no naturalismo biológico (seção 4.2.3)¹¹⁰.

4.2.1) Causação e Causalidade Segundo David Hume

Antes de se falar de causação mental, deve-se esclarecer o significado geral de causação e causalidade. Para tanto, analisar-se-á sucintamente estes conceitos à luz da

¹¹⁰ Isso já foi feito em parte na apresentação das teses do naturalismo biológico, notadamente a 2ª.

filosofia humeana¹¹¹. Segundo o filósofo escocês, a noção *ortodoxa* de causação é composta por três componentes: (i) prioridade, (ii) continuidade no espaço e no tempo e (iii) conexão necessária entre causa e efeito. Prioridade significa que a causa sempre precede o efeito, isto é, a causa não pode vir após seu efeito. O segundo componente é o que afirma que causa e efeito são adjacentes ou contíguos, tanto espacial como temporalmente. Finalmente, de acordo com a última componente da causação, a causa faz com que o efeito ocorra, isto é, a causa deve estar *necessariamente* conectada a seu efeito.

Entretanto, Hume afirma que ao observarmos os casos de causa e efeito na natureza, não percebemos nenhuma conexão necessária. Vemos apenas um evento A preceder um evento B, o que não significa que haja uma conexão necessária entre ambos – o que experienciamos é uma conjunção constante de casos semelhantes, mas a conexão necessária entre os mesmos não passa de uma ilusão de nossa mente (crença psicológica), isto é, não há nenhum fundamento racional (justificação lógica) que apóie a suposição de que o próximo caso se assemelhará ao caso precedente.

Hume notou que o princípio de causa e efeito pode ser desmembrado em dois princípios mais básicos ou subjacentes, os princípios de causação e causalidade. O primeiro afirma que todo evento possui uma causa; já o segundo é o que diz que as mesmas causas têm os mesmos efeitos. Contudo, esses dois princípios não são verdades analíticas. Dessa forma, resta saber se os mesmos são verdades empíricas sintéticas, o que não é verdade, pois os métodos empíricos das ciências naturais já *pressupõem* esses princípios, caracterizando assim uma petição de princípio – parte-se do que se quer provar.

Chega-se, após o que foi exposto, ao famoso *problema da indução*: não podemos fazer uma generalização a partir da observação de casos particulares, já que estes, não importam quantos sejam, possuem menos informação do que aquela. Por conseguinte, o enunciado “todos os homens são mortais” é apoiado pela evidência dos homens que morreram no passado e morrem no presente, mas nada pode garantir que isso sempre será o caso. Em resumo, as evidências dos enunciados particulares são insuficientes para sustentar o enunciado geral, que nesse caso deve ser chamado de *hipótese*. Deve-se ressaltar que o método científico não é arbitrário: existem regras que balizam a passagem das evidências às hipóteses, como o pressuposto de que a natureza é uniforme – o futuro será semelhante ao passado. Porém, mais uma vez, cai-se num círculo vicioso, já que esse pressuposto é fundamentado indutivamente, o que, como já foi mostrado, não é suficiente: primeiro, mostra-

¹¹¹ Para maiores detalhes, conferir o *Tratado da Natureza Humana*, livro 1, parte 3.

se que não existem fundamentos racionais para os argumentos indutivos. No entanto, as ciências naturais não investigam arbitrariamente a natureza, isto é, o método científico¹¹² segue regras como, por exemplo, a que pressupõe a regularidade da natureza. Contudo, essa pressuposição é fundamentada pela indução, ficando evidente, dessa maneira, que se trata de um argumento circular. Nota-se, então, que o ceticismo humeano chega a duas conclusões, quais sejam, a de que não existe conexão necessária na natureza, e a que afirma que não existe base racional para a indução.

4.2.2) Crítica de Searle a Hume

Searle¹¹³ fará uma crítica a Hume a partir do conceito de intencionalidade ou “direcionalidade” – todo fenômeno intencional necessita de uma condição causalmente auto-referencial. A intenção em levantar meu braço só é satisfeita se causa o movimento do mesmo, assim como minha experiência de ver uma flor só é satisfeita se causada pelo objeto percebido, no caso, a flor em questão¹¹⁴. Em ambos os casos, há uma conexão causal entre a *experiência* e os *objetos e estados de coisas do mundo*. Caso ainda restem dúvidas, Searle pede para que levantemos nosso braço¹¹⁵. Notaremos, ao fazer o que se pede, que nosso braço se erguerá de maneira intencional, isto é, um estado mental, o desejo de levantar o braço, *causou* o movimento do mesmo. Isso significa, em outras palavras, que a consciência possui eficácia causal, e podemos falar de *causação mental*.

¹¹² Em 1934, o filósofo austríaco Karl Popper lança seu importante livro *A Lógica da Investigação Científica*. Nesse livro, Popper ataca o problema da indução de Hume afirmando que a ciência não age de maneira indutiva, como pensava o filósofo escocês, e sim que o método científico é hipotético-dedutivo. Para Popper, o conhecimento científico é sempre conjectural e o dever do cientista não é o de verificar teorias, como defendia o Positivismo Lógico, mas o de tentar falseá-las (falsificacionismo).

¹¹³ Searle defende uma concepção representacionista da mente, que, em linhas gerais, defende que o nosso conhecimento é gerado a partir de representações mentais que fazemos do mundo. Muitos são os críticos dessa visão, como o filósofo austríaco Ludwig Wittgenstein, em sua fase madura (*Investigações Filosóficas*), e os biólogos chilenos Humberto Maturana e Francisco Varela (*A Árvore do Conhecimento*).

¹¹⁴ Portanto, podemos diferenciar entre *ação intencional* e *percepção*. A primeira refere-se a nossa intenção consciente em causar movimentos corporais, enquanto a segunda diz respeito a experiências perceptuais causadas em nós por objetos e estados de coisas do mundo.

¹¹⁵ O neurocirurgião norte-americano Wilder Penfield realizou uma experiência em que estimulava, através de microeletrodos, neurônios localizados no córtex motor do cérebro de pacientes, que tinha como resultado o movimento do braço dos mesmos. Contudo, o que é mais interessante nessa pesquisa não é o movimento do braço, mas a interpretação que os pacientes davam ao notar que seu braço se mexera de maneira involuntária, dizendo coisas como “eu não fiz isso, você fez”. Searle cita os experimentos de Penfield para deixar claro que existe uma diferença fundamental entre levantar o braço voluntária e involuntariamente.

Searle acusa Hume de olhar para o lugar errado: em vez de se deter em objetos e eventos externos a nós, ele devia ter observado que nossas próprias experiências fazem com que algo ocorra (ação intencional) e que nossa experiência de algo faz com que algo nos ocorra (percepção). Tanto na ação intencional como na percepção é totalmente comum a experiência de conexão causal, experiência esta que nos é dada pelas *conjunções constantes* entre nossas próprias experiências e o mundo.

Todavia, resta agora mostrar que existe o mesmo tipo de conexão necessária no mundo, independente das experiências que temos dele. Para resolver esse problema, Searle nos dá o seguinte exemplo: quando eu empurro meu carro, noto que o movimento deste é causado por mim. Temos, portanto, uma relação de causa (a força que utilizo para empurrar o carro) e efeito (o movimento do veículo). Por outro lado, podemos dizer que a mesma relação causal está presente quando observo outra pessoa empurrando um carro, já que a força daquela causa o movimento deste. A partir disso, podemos estender a relação de causa e efeito para os casos onde não há um agente envolvido, ou seja, a relação deve ser também aplicada para a causação entre as seqüências de eventos que há no mundo e independem de nossas experiências. Nas palavras de Searle: “(...) the same relation we experience when we make something happen to us, can be perceived to exist even when no experiences are involved in the causal relation” (M, p. 205).

Searle (CECC, pp. 41-42) reconhece que *muitas* relações causais são humeanas, no sentido em que a causa precede o efeito, como, por exemplo, quando se afirma que uma bala de revólver foi a responsável pela morte de uma pessoa. Nesse caso, é possível distinguir dois acontecimentos distintos, tanto espacial como temporalmente: o tiro (acontecimento 1) *causou* a morte do indivíduo (acontecimento 2), isto é, a causa (tiro) *precede* o efeito (falecimento da pessoa). Mas, segundo Searle, nem sempre esse é o caso, pois

(...) muitas relações causais não são acontecimentos distintos, são forças causais permanentes agindo ao longo do tempo. Pensemos na atração gravitacional. Não podemos afirmar que, primeiro, existe a atração gravitacional e que, depois, mais tarde, as cadeiras e mesas exercem pressão sobre o chão. A atração gravitacional é uma força operativa constante e, pelo menos nesses casos, a causa e o efeito são simultâneos (CECC, p. 42).

Outros dois exemplos que Searle cita com sendo de causação não-humana são a solidez e a liquidez: solidez e liquidez são propriedades de nível superior causadas pelo comportamento molecular e realizadas na microestrutura do sistema. Aqui, mais uma vez, a causa e o efeito são simultâneos, ou seja, não há, primeiramente, o comportamento molecular

(nível micro) num tempo t_1 e, após isso, a solidez ou liquidez num tempo t_2 (nível macro), mas, sim, a simultaneidade entre causa e efeito (CECC, p. 42).

De modo análogo à solidez e à liquidez, afirmar que os estados neurobiológicos causam os estados mentais não implica qualquer tipo de lacuna temporal entre aqueles e estes, já que a causação do mental pelo neurobiológico é não-humana (no sentido de causa e efeito são simultâneos), ou seja, estados psicológicos são causados por processos neurofisiológicos *ao mesmo tempo* em que são realizados no sistema cerebral. Com isso, Searle pretende mostrar que não é possível descrever a relação causal entre o neurobiológico e o psicológico como uma sequência de acontecimentos distintos, o que acarretaria, em última instância, uma *dualidade* entre cérebro e consciência, tese esta que Searle rejeita:

(...) o fato de existir uma relação causal entre os processos cerebrais e os estados conscientes não implica um dualismo do cérebro e da consciência, do mesmo modo que a relação causal entre movimentos moleculares e solidez não implica um dualismo entre moléculas e solidez (CECC, p. 43).

Sendo assim, ao rejeitar o dualismo cérebro-consciência, Searle está defendendo que a consciência é uma característica de nível superior, causada e realizada no sistema cerebral (CECC, p. 43), e que é, devido a isso, *explicada* pelo nível neurobiológico, pois os estados mentais possuem os mesmos poderes causais dos estados cerebrais. A consequência disso é que os estados mentais funcionam causalmente no mundo (CECC, pp. 36-37), tanto produzindo outros estados mentais como causando estados físicos, como erguer o braço. Com isso, Searle pretende resguardar sua teoria da ameaça do epifenomenalismo (próxima seção), mas acaba esbarrando no problema da sobredeterminação causal (seção 4.3).

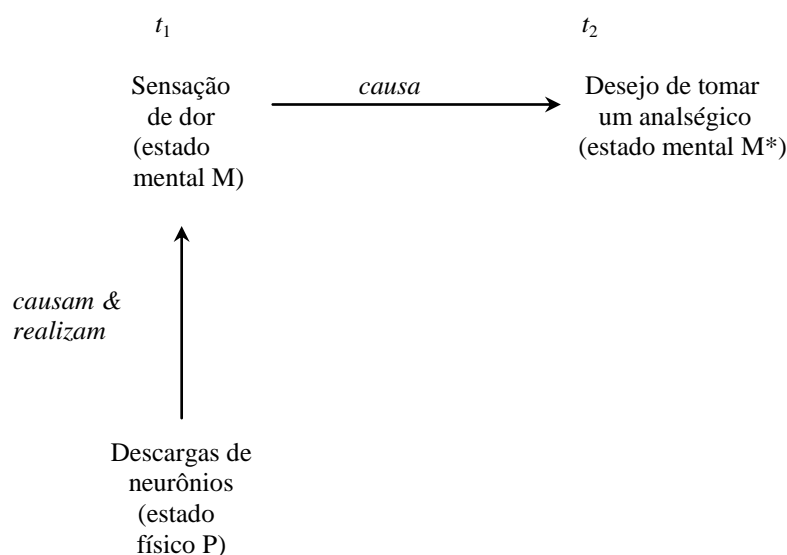
4.2.3) Causação Mental

Searle defende repetida e enfaticamente a tese da eficácia causal dos estados mentais¹¹⁶: um estado mental M pode ser a causa de outro estado mental M* ou de um estado físico P¹¹⁷. Caso o naturalismo biológico fosse um tipo de fisicalismo reduutivo, seria fácil

¹¹⁶ Vale a pena ressaltar que nem todas as correntes em Filosofia da Mente defendem a causação mental. Para um epifenomenalista, os estados mentais são apenas epifenômenos dos estados cerebrais, ou seja, os estados cerebrais causam os estados mentais, mas estes são causalmente inertes. Em resumo, é apenas uma *ilusão*, segundo o epifenomenalismo, achar que nossos desejos, decisões e volições *causam* nosso comportamento (Kim, 2005a, p. 21). Logo a seguir, será abordada a crítica de Jaegwon Kim ao naturalismo biológico.

¹¹⁷ Para os propósitos desta discussão, considerar-se-á a definição de estado físico em um sentido *lato*, o que significa que com P podemos nos referir a determinados estados cerebrais ou mesmo mudanças fisiológicas que

defender a tese da eficácia causal dos estados mentais, já que de acordo com essa concepção os estados psicológicos são *idênticos* aos estados cerebrais subjacentes, e estes por sua vez são causalmente eficazes. Porém, este não é o caso: apesar de admitir a redução causal, Searle nega a redução ontológica dos estados mentais a estados neurobiológicos, pois estes têm uma ontologia de terceira pessoa (ontologicamente objetivos), enquanto aqueles possuem uma ontologia de primeira pessoa (ontologicamente subjetivos). Contudo, já que existe uma diferença ontológica entre esses estados, alguns filósofos afirmam que há sobredeterminação causal no Naturalismo Biológico – o comportamento humano consciente¹¹⁸ seria determinado por dois tipos de causas, uma cerebral e outra mental. Por exemplo, minha dor de cabeça, estado mental M, causa meu desejo de tomar uma aspirina, estado mental M*:



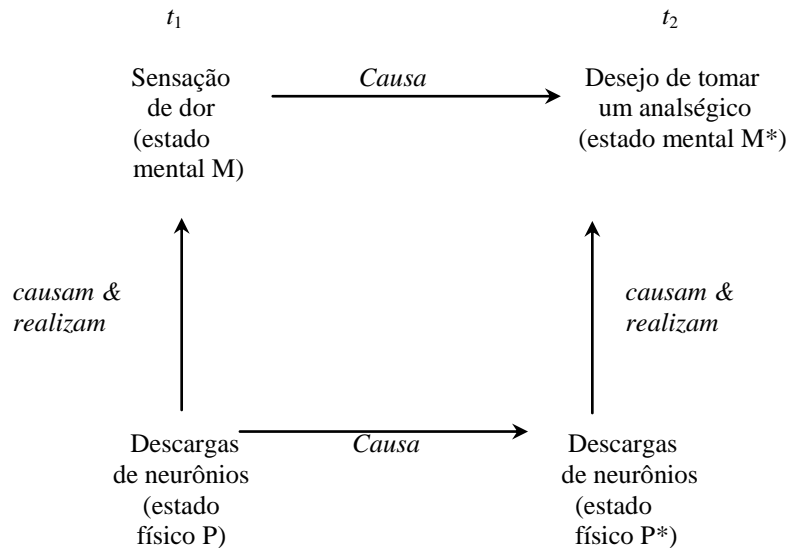
Todavia, M (macronível) é causado por um estado neurobiológico P subjacente (micronível) em t_1 , enquanto M* (macronível) é causado por outro estado cerebral P*¹¹⁹

provoquem a movimentação do corpo, por exemplo. Portanto, nessa definição abrangente não existem diferenciações entre físico, fisiológico e neurobiológico.

¹¹⁸ Nem todo comportamento humano é consciente. O chamado *ato reflexo* é um bom exemplo de comportamento que não envolve, num primeiro instante, a consciência: imagine uma pessoa que está conversando de maneira distraída numa festa. De repente, alguém encosta sem querer um cigarro em seu braço. O que ocorre nessa situação? De maneira simplificada, neurônios sensoriais (aférentes) presentes no órgão sensorial, no caso, a pele, comunicam-se com a medula espinhal, através de sinapses com neurônios de associação. Estes, por sua vez, farão sinapses com neurônios motores (eferentes), que atuarão fazendo com que o braço (órgão efetador) seja retirado. Apenas alguns milissegundos depois do ato reflexo é que o sinal nervoso chega ao cérebro, e a pessoa toma consciência do ocorrido. Em resumo, primeiro a pessoa que sofreu a lesão retira o braço, depois ela diz “ai”. As vantagens evolutivas desse mecanismo de estímulo e resposta do sistema nervoso são óbvias...

¹¹⁹ O estado cerebral P* é causado, por sua vez, pelo estado cerebral P.

(micronível) em t_2 . Portanto, meu desejo de tomar uma aspirina possui duas causas simultâneas diferentes, a saber, M (causa mental) e P* (causa física):



Um raciocínio semelhante se aplica para a causação mental-físico, já que Searle aceita que o desejo de se comprar um medicamento (M') causa o deslocamento à farmácia (P'). Searle não abre mão da tese da eficácia causal do mental. Sendo assim, resta a ele mostrar como o naturalismo biológico evita o seguinte dilema: ou se admite que os estados mentais são meros epifenômenos ou, caso tais estados sejam causalmente eficientes, deve-se enfrentar o problema da sobredeterminação causal. Dessa maneira, Searle deve, num primeiro momento, demonstrar que há causação mental para evitar o epifenomenalismo e, num momento posterior, mostrar que a eficácia causal do mental não compromete sua teoria com a sobredeterminação causal. O dilema epifenomenalismo-sobredeterminação causal será colocado principalmente por Kim, que é um dos principais defensores do fisicalismo reduutivo atualmente.

4.3) O fisicalismo reduutivo de Jaegwon Kim e o problema da sobredeterminação causal

Jaegwon Kim afirma que o problema mente-corpo é constituído de dois problemas que são interconectados: o problema da causação mental e o problema da consciência. Esses dois

problemas geram duas difíceis questões que o fiscalista tem que responder para que sua teoria seja considerada plausível. A primeira questão pode ser assim resumida: como é possível que estados mentais tenham eficácia causal num mundo fundamentalmente físico? A outra questão, por sua vez, é explicar por que existe algo como a consciência em um mundo povoado unicamente por entidades físicas.

Quando Descartes propôs o dualismo de substância, no século XVII, ele enfrentou severas críticas de seus contemporâneos. Um dos problemas levantados contra o dualismo cartesiano dizia respeito à explicação de como duas substâncias diferentes, física (*res extensa*) e não-física (*res cogitans*), poderiam interagir sem entrar em contradição com a lei de conservação de energia da Termodinâmica:

It seems impossible to make substance dualism consistent with modern physics. Physics says that the amount of matter/energy in the universe is constant; but substance dualism seems to imply that there is another kind of energy, mental energy or spiritual energy, that is not fixed by physics. So if substance dualism is true then it seems that one of the most fundamental laws of physics, the law of conservation, must be false (M, p. 42. Grifos meus).

A partir desse problema surgiram outras teorias para abordá-lo, tais como o Ocasionalismo de Malebranche, a Teoria de Duplo-Aspecto de Spinoza e a Teoria da Harmonia Prestabelecida de Leibniz. Descartes foi uma exceção à sua época, já que defendeu a “mental causation as an integral element of his view of the mind” (Kim, 2005, p. 8). Caso a causação mental fosse, de fato, apenas uma ilusão, poder-se-ia dizer que o homem estaria desprovido de duas de suas características fundamentais: a moralidade e sua enorme capacidade de aprendizado. A primeira seria afetada porque as questões morais envolvem a pressuposição de que as nossas ações são determinadas por nossa vontade consciente¹²⁰. Caso essas ações fossem *unicamente* determinadas pelos estados cerebrais e os estados mentais não tivessem nenhuma influência causal nas mesmas, não seria possível julgar como imoral um estuprador ou um assassino. Pior do que isso, seria impossível se fazer qualquer tipo de julgamento, pois este depende da causação mental em última instância. Também é sabido que todo processo cognitivo pressupõe a causação mental. Dito de outra forma, a eficácia causal do mental é uma *condição necessária* para se conhecer algo, já que todo processo cognitivo é um processo *ativo*.

¹²⁰ O conceito de livre-arbítrio está relacionado à noção de eficácia causal do mental, já que aquele parte do pressuposto de uma ação *livremente* escolhida. Esta, por sua vez, exige que o agente opte, de maneira deliberada, por uma de duas ou mais alternativas. Por fim, essa deliberação pertence ao domínio mental, o qual interage com o domínio físico na efetuação de uma ação livre.

O epifenomenalismo é uma concepção filosófica que nega a causalização mental e, pode-se dizer, vai de encontro ao senso comum e a muitas concepções filosóficas¹²¹, inclusive o naturalismo biológico. Searle aceita a superveniência de maneira restrita, defendendo que estados neurofisiológicos idênticos causam os mesmos estados mentais. Como há uma *relação causal*¹²² entre o nível superveniente e o basal ou subjacente, Searle intitula tal superveniência de *superveniência causal*, defendendo que a “identidade de neurofisiologia garante identidade de mentalidade¹²³”, mas, no entanto, a “identidade de mentalidade não garante identidade de neurofisiologia” (RM, p. 180). A partir disso pode-se ver o porquê de Searle rejeitar o experimento de pensamento do zumbi, já que se este fosse, *de fato*, neurobiologicamente idêntico a mim, teria os *mesmos* estados mentais que possuo¹²⁴:

(...) tomando o famoso exemplo do cérebro-dentro-da-cuba, se você tivesse dois cérebros que fossem tipo-idênticos até a última molécula, então a *base causal do mental* garantiria que tivessem os mesmos fenômenos mentais. (...) estados físicos são *causalmente suficientes*, embora não necessariamente causalmente necessários, para os estados mentais correspondentes (RM, p. 180. Grifos meus).

Sendo assim, os estados neurofisiológicos são *suficientes* para produzir os estados mentais, apesar de não serem *causalmente necessários*: em princípio, os estados mentais podem ser produzidos por uma base física diferente da nossa, contanto que aquela *duplica* as capacidades causais desta.

Até este ponto há uma relativa convergência entre o fisicalismo de superveniência de Kim e o naturalismo biológico de Searle. Todavia, esta concepção, diferentemente daquela, defende uma interação de mão-dupla entre o físico e o mental, o que implica que fenômenos

¹²¹ Como bem observa Maslin (2009, p. 183), uma pessoa que aceita o epifenomenalismo não poderia, caso ele fosse verdadeiro, expressar sua opinião, pois seus estados mentais, como sua crença em tal concepção e seu desejo em torná-la clara, não têm eficácia causal: “Paradoxalmente, qualquer um que acredite que o epifenomenalismo é verdadeiro, e consegue dizer isso, abandona a doutrina no próprio ato de afirmá-la”.

¹²² Searle faz a seguinte distinção entre as noções de superveniência em Ética e em Filosofia da Mente: “(...) as características de um objeto que o fazem bom não *causam* que ele seja bom, mas antes *constituem* sua excelência [superveniência constitutiva]. Mas, no caso da superveniência mente/cérebro, os fenômenos neurais causam os fenômenos mentais [superveniência causal]” (RM, p. 181).

¹²³ Kim (1996, p. 224) também aceita esta tese: “If two organisms or systems are physically indistinguishable, they must be functionally equivalent and realize the same psychology”. Portanto, nem Searle nem Kim aceitam o argumento do zumbi de Chalmers, pois se o zumbi não tem vida mental é porque ele não é neurobiologicamente idêntico a nós.

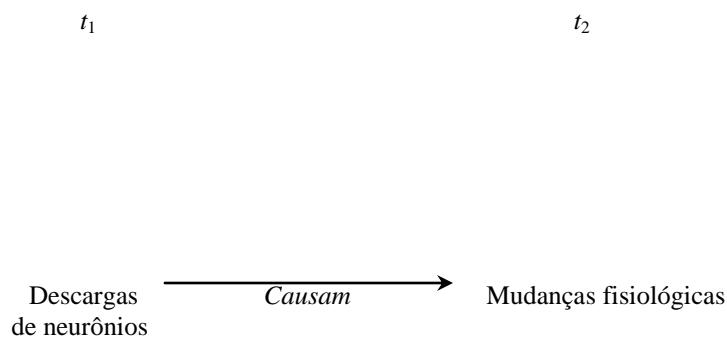
¹²⁴ Da mesma forma que não seria possível, de maneira específica, um zumbi ser fisicamente idêntico a mim e não possuir estados mentais, Searle também alega que não seria concebível, de maneira geral, a existência de um universo que fosse fisicamente idêntico ao nosso e não contivesse consciência, já que esta é uma característica física do mundo e necessariamente emergiria caso o mesmo fosse duplicado: “(...) in fact it is not logically possible that the physical universe should be exactly as it is, atom for atom, without all of its physical features being exactly as they are” (M, p. 128).

mentais são causalmente eficientes, ou seja, *não são* epifenômenos¹²⁵: “(...) o fato de as características mentais serem supervenientes em relação a características neurônicas de modo algum diminui sua eficácia causal” (RM, p. 182). Searle lança mão da seguinte analogia para defender a eficácia causal das propriedades supervenientes: a solidez de um objeto é superveniente em relação a sua estrutura molecular subjacente, mas seria um erro afirmar que por causa disto a solidez seja um epifenômeno. Assim como ocorre com a solidez e a estrutura molecular, os estados mentais e os estados cerebrais são apenas diferentes níveis de descrição do mesmo fenômeno, onde aqueles são causados e realizados por estes. Ao defender os múltiplos níveis de descrição, assume-se, também, mais de um tipo de causação, o que ocasiona, segundo Searle, a inutilidade do conceito de superveniência:

Minha conclusão é que, quando você reconhece a existência de formas de causação de baixo para cima, micro para macro, a noção de superveniência deixa de ter qualquer função na filosofia. As características formais da relação já estão presentes na *suficiência causal* das formas de causação micro-macro (RM, p. 182. Grifos meus).

De acordo com o naturalismo biológico, poderíamos estabelecer os seguintes tipos de causação:

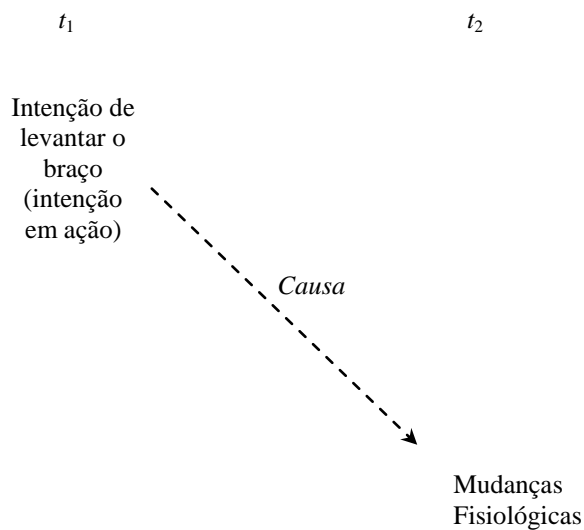
- 1) **Causação micro-micro:** é quando um processo de micronível causa outro processo de micronível. Exemplo: descargas de neurônios causando mudanças fisiológicas:



- 2) **Causação micro-macro/cima-baixo:** ocorre nos casos em que um processo de nível superior é *causado por e realizado pelo* processo de nível inferior. Exemplo 1: descargas de neurônios causam e realizam a intenção em ação. Exemplo 2: o movimento corporal é causado por e realizado pelas mudanças fisiológicas:

¹²⁵ Na verdade, Kim só considera os *qualia* como epifenômenos, já que são irreduzíveis. Para maiores detalhes, conferir a seção 4.1.

5) **Causação macro-micro/esquerda-direita:** é o tipo de causação em que um processo de nível superior causa um processo de nível inferior. Exemplo: a intenção de levantar o braço causa as mudanças fisiológicas que são necessárias para que ocorra, posteriormente, o movimento corporal:



Dentre esses tipos de causação, a mais controversa é esta última, que também pode ser chamada de *causação descendente*, pois este tipo de causação “prima facie implies the failure of causal closure at the lower level, and in-principle impossibility of a complete theory of the lower-level phenomena in their own terms” (KIM, 1995, pp. 193-194). Kim¹²⁶ se vale do seguinte argumento para criticar a causação descendente: segundo a causação descendente, (i) um processo mental M (nível superior) causa um processo físico ou cerebral P (nível inferior). No entanto, (ii) outro evento físico P* é a *condição basal* do evento mental M, já que este é superveniente em relação àquele, o que significa que P* é *nomologicamente suficiente* para M. E, conseqüentemente, (iii) P* é a causa de P. Conclui-se, então, que M não tem

¹²⁶ Conferir Kim, 2005b, p. 576.

capacidades causais novas¹²⁷, isto é, que já não sejam encontradas nos processos de micronível que o originam, no caso, P¹²⁸.

Searle considera como absurda e anti-intuitiva a concepção epifenomenalista, pois ela contraria uma de nossas certezas mais básicas, a certeza de que agimos, pelo menos na maioria das vezes, de maneira livre e consciente: “Every time I decide to raise my arm, it goes up. And it is not a random or statistical phenomenon. I do not say, ‘Well, that’s the thing about the old arm. Some days she goes up and some days she doesn’t’” (M, p. 30).

Desse modo, para mostrar que sua concepção dos diferentes níveis de descrição do sistema cérebro-mente¹²⁹ é plausível, Searle utiliza o exemplo que envolve o comportamento do motor do carro. Inicialmente, o movimento de elétrons entre os eletrodos (nível micro) causa e realiza o aumento de temperatura (nível macro). Este aumento, por conseguinte, causa a explosão no cilindro (nível macro). Contudo, a explosão do cilindro é causada e realizada pela oxidação dos hidrocarbonetos que estão presentes no combustível (nível micro), oxidação esta que foi causada pelo movimento de elétrons entre os eletrodos. Portanto, pode-se descrever o comportamento do motor em níveis distintos, onde os processos de nível superior são causados pelos microprocessos e são realizados no sistema, mas isso não implica, alega Searle, que aqueles sejam causalmente ineficazes ou epifenomênicos:

(...) nossa aceitação provisória da eficácia causal da consciência não é ameaçada se assinalarmos que qualquer explicação no nível da consciência tem como base fenômenos mais fundamentais, porque é verdade em relação a qualquer sistema físico que as explicações causais de níveis superiores têm como base explicações microfísicas mais fundamentais nos níveis inferiores (MLS, pp. 63-64).

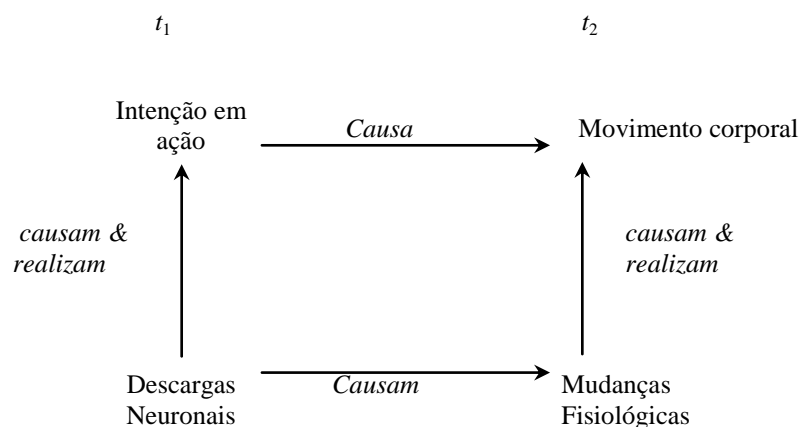
A partir dos conceitos supracitados, nota-se que, de acordo com a noção de superveniência, os estados neurofisiológicos causam estados mentais; contudo, em princípio, ela não está comprometida com o sentido oposto, o da *causação mental*, onde estados mentais causam estados neurofisiológicos. Por conseguinte, esclarecer o que Searle entende por superveniência é de extrema importância, pois uma posição central do naturalismo biológico é, como foi visto no capítulo anterior, a tese da *eficácia causal do mental*, ou seja, a tese que afirma que estados mentais causam tanto outros estados mentais como estados neurofisiológicos. Portanto, qual o argumento que Searle utiliza para defender que estados

¹²⁷ A capacidade causal da propriedade superveniente é a *mesma* da propriedade basal, que é o realizador físico P. Isto é o que Kim denomina de *princípio da herança causal*: “If a functional property *F* is realized on a given occasion in virtue of the instantiation of one of its realizers, *P*, the causal powers of this instance of *F* are identical with the causal powers of *P*” (KIM, 2005b, p. 580).

¹²⁸ Caso Searle esteja correto, isto é, que M possui capacidades causais novas e distintas em relação a P*, a afirmação de que tanto P* como M causam P se torna problemática, já que o mesmo fenômeno teria duas explicações causais distintas. Isso é conhecido como o *problema da sobredeterminação causal*.

¹²⁹ Concepção esta que foi explicada e representada há pouco, juntamente com os diferentes tipos de causação.

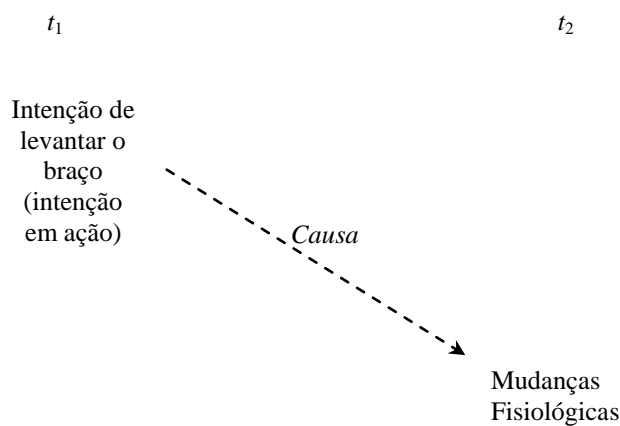
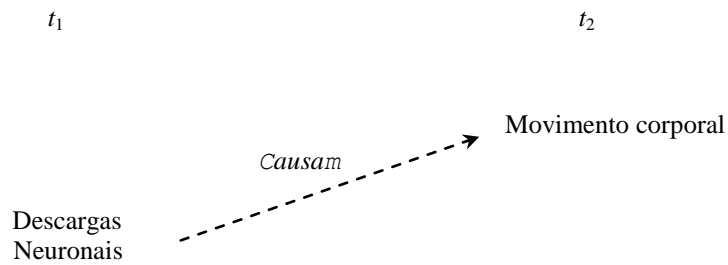
mentais são causalmente eficientes? Sua resposta é que o *mesmo sistema* pode ser descrito em *diferentes níveis*, como mostrado na figura abaixo¹³⁰:



Descargas de neurônios e mudanças fisiológicas são, por um lado, descrições de nível inferior do sistema, enquanto a intenção em ação e movimento corporal são, por outro lado, descrições de nível superior. No entanto, não é somente o conceito de causação que é importante para compreender o que Searle quer dizer com níveis de descrição diferentes, mas, também, o conceito de *realização*: a intenção em ação tanto é causada como é realizada pelas sinapses de neurônios. O mesmo se pode dizer do movimento corporal, que é causado e realizado pelas mudanças fisiológicas subjacentes. Estes exemplos caracterizam o que se chama de causação *ascendente* ou causação *para cima*. Outro tipo de causação, a causação de *nível inferior*, é exemplificada pelas descargas de neurônios causando mudanças fisiológicas, e a causação de *nível superior* ocorre quando o movimento corporal é causado pela intenção em ação. Estes são os três tipos de causação que estão representados no esquema acima. No entanto, como os processos de macronível, intenção em ação e movimento corporal, são causados e realizados pelos processos de micronível, não há problema algum, segundo Searle, em se representar setas diagonais de causação, onde as descargas de neurônios (nível micro), em um tempo t_1 , causam o movimento corporal (nível macro), em um tempo t_2 ; e, no sentido oposto, as mudanças fisiológicas (nível inferior) são causadas pela intenção em ação (nível superior), como mostrado a seguir¹³¹:

¹³⁰ Esse esquema é uma adaptação do que se encontra na página 374 do livro *Intencionalidade*.

¹³¹ Apenas a noção de causação está envolvida em ambos os exemplos.



Ao defender este último tipo de causação (causação descendente), Searle se posiciona em oposição ao epifenomenalismo. Mas, poder-se-ia questionar, como Searle fundamenta tal tipo de causação? Uma vez mais, a fundamentação está na Biologia – a consciência é uma característica que surgiu em alguns animais por meio de um longo processo evolutivo. Logo, é de se esperar que um organismo que adquiriu pela primeira vez consciência do mundo que o cerca tenha levado algum tipo de vantagem sobre seus concorrentes. Entretanto, só faz sentido postular tal vantagem se a consciência funcionar de maneira a controlar o corpo do animal (causação descendente), pois se o epifenomenalismo estiver correto, animais, conscientes ou inconscientes, não passam de autômatos *fisicamente determinados*, e aqueles animais não teriam vantagem alguma sobre estes. Na verdade, ocorreria justamente o contrário, pois é sabido com algum grau de segurança que a consciência está intimamente relacionada a cérebros bem desenvolvidos, e que estes gastam uma boa parcela da energia do organismo, ou seja, seria um grande dispêndio biológico-evolutivo manter uma estrutura que é, a um só tempo, cara e ineficaz. Nas palavras de Searle,

seria milagroso, inédito na história biológica, se algo biologicamente tão complexo, rico e estruturado como a consciência humana e animal não tivesse importância causal no mundo. De acordo com o que sabemos sobre a evolução, é pouco provável que o epifenomenismo possa estar certo (MLS, p. 61).

Portanto, o epifenomenalismo entra em confronto tanto com o senso comum quanto com a Biologia Evolutiva, apesar de ser logicamente possível e, portanto, filosoficamente defensável:

Não estou dizendo que o epifenomenismo é falso por uma questão de lógica. No que diz respeito à possibilidade lógica, poder-se-ia descobrir que os estados mentais são totalmente epifenomênicos e, portanto, que não têm nenhuma função causal. Tal possibilidade é logicamente concebível mas, até onde sabemos, o fato de nossos estados mentais atuarem de modo causal na produção de nosso comportamento é apenas um fato simples a respeito de como o mundo funciona (MLS, p. 64).

Quando Searle diz que a eficácia causal do mental *é apenas um fato simples de como o mundo funciona*, ele está se referindo ao *nosso mundo*, ou seja, nada impede, em princípio, que haja *outro mundo* em que a consciência não funcione causalmente (daí a possibilidade lógica do epifenomenalismo), a saber, um mundo habitado por zumbis (no sentido filosófico do termo) – criaturas que se comportam *exatamente* como nós, a despeito de serem *inconscientes*. Mas Searle tem plena consciência da distinção entre os mundos *real* e *possível* ou, falando de outra maneira, da diferença entre o que existe de *fato* e o que *poderia* existir:

Ao analisar as funções evolutivas das asas, ninguém pensaria ser admissível afirmar que elas são inúteis porque podemos imaginar os pássaros voando sem asas. Por que se supõe, então, ser possível afirmar que a consciência é inútil porque podemos imaginar seres humanos e animais se comportando como o fazem, mas sem consciência? (CECC, pp. 39-40).

Sendo assim, penso que, no que se refere a este ponto, a abordagem *naturalista* de Searle é coerente e correta, pois explicação da vantagem adaptativa de uma característica, órgão ou comportamento só faz sentido ao se levar em consideração que a evolução é um processo *histórico* e *contingente* e, portanto, sujeito aos fatores e condições ambientais aos quais os organismos foram expostos ao longo do tempo: “Isso significa que a questão evolutiva só tem sentido se forem aceitos certos pressupostos básicos sobre o funcionamento da natureza” (CECC, p. 39). Se o mundo em questão fosse outro, isto é, um mundo com condições ambientais radicalmente diferentes das que existem na Terra, *poderia* ser o caso que uma característica (ou órgão ou comportamento) vantajosa nesse planeta hipotético fosse extremamente desvantajosa no contexto terrestre, e vice-versa. Ou ainda, seria possível que as leis e, conseqüentemente, o funcionamento da natureza divergisse bastante nesses dois planetas, o que afetaria, inevitavelmente, a evolução dos organismos nos respectivos planetas.

Sendo assim, seria possível imaginar um mundo onde a consciência *produzisse* energia ou fosse *epifenômica*, mas Searle está com razão ao defender que no atual estágio científico sabemos o suficiente sobre a consciência para afirmar que ela consome uma boa parcela da energia total do organismo e que seria inteiramente dispendiosa caso não funcionasse causalmente, conferindo uma importante *vantagem evolutiva* ao organismo que a possui.

4.4) Fisicalismo Não-Redutivo, Dualismo de Propriedade e Naturalismo Biológico

Compararei nesta seção o naturalismo biológico a duas posições de fisicalismo não-reduutivo e ao dualismo de propriedade: num primeiro momento, a concepção de Searle será comparada com o emergentismo e com a abordagem defendida por Pereboom e Kornblith (2004) e, posteriormente, com o dualismo de propriedade. Com isso, analisarei quais são as semelhanças e diferenças do naturalismo biológico com tais teorias, com o intuito de propor uma classificação para o mesmo.

Em relação ao emergentismo, seguirei Kim (1993b, pp. 344-348)¹³², o qual afirma que o emergentismo consiste na conjunção de quatro teses: (i) monismo físico, (ii) anti-reducionismo, (iii) tese da realização física e, por último, (iv) realismo mental. Constata-se, numa primeira aproximação, que Searle defende cada uma dessas teses na sua concepção¹³³; no entanto, ao se observar mais atentamente as duas teorias, verifica-se que o *argumento* que Searle utiliza para defender a irreducibilidade ontológica do mental (tese ii) diverge da posição emergentista. Em outras palavras, apesar de tanto o emergentismo como o dualismo de propriedade serem posições anti-reducionistas, penso que a primeira concepção não é dualista, mas, sim, fisicalista não-reduitiva. A justificativa para isso, em minha opinião, é que a *essência* do argumento anti-reducionista do dualista de propriedade se fundamenta na subjetividade ontológica de alguns fenômenos mentais, a saber, os *qualia* (para Searle, *todos* os fenômenos mentais são qualitativos). O emergentista, por sua vez, defenderá a irreducibilidade a partir do conceito de *emergência*: à medida que entidades físicas se organizam em sistemas complexos, surgem propriedades de nível superior que não são

¹³² É interessante notar que em um trabalho posterior Kim classifica o emergentismo como um tipo de dualismo de propriedade, e não como fisicalismo não-reduutivo: “*Emergentism (...) is a form of property dualism, the position that, in addition to physical properties, there are physically irreducible domains of emergent properties, of which mental properties are among the leading candidates*” (KIM, 2005a, pp. 157-158).

¹³³ Não quero com isso afirmar que o naturalismo biológico é constituído *apenas* por essas quatro teses, mas sim que todas as quatro fazem parte de um conjunto maior de teses defendidas por Searle. Para um tratamento mais completo e sistemático das teses do naturalismo biológico, conferir o capítulo anterior.

encontradas no nível subjacente e que, além disso, possuem *novos* e *genuínos* poderes causais em relação a este último nível.

Searle critica o conceito usual de propriedade emergente, afirmando que “essa noção não é muito clara, pois é usada de maneira bastante confusa na literatura” (CECC, p. 42). Porém, logo em seguida, Searle propõe um conceito que ele considera mais claro e útil, a saber, o conceito de *propriedade causalmente emergente*, que corresponde

(...) à ideia de que a consciência, como a solidez e a liquidez, é uma propriedade emergente do comportamento dos microelementos de um sistema composto por eles. Uma propriedade emergente, assim definida, é uma propriedade que se explica pelo comportamento dos microelementos, mas não pode ser simplesmente deduzida da composição e dos movimentos deles (CECC, pp. 42-43).

Sendo assim, o conceito de emergência adotado no naturalismo biológico é diferente do que é comumente utilizado pelos emergentistas, sobretudo pelo aspecto causal defendido por Searle. As propriedades causalmente emergentes ocorrem nos mais variados níveis hierárquicos da natureza, e podem ser exemplificadas pelas propriedades de ser líquido, propriedade de ser vivo e propriedade de ser consciente e todas estas são propriedades de nível elevado que mantêm relações *causais* com o micronível do sistema – moléculas de água, a maquinaria celular e as sinapses entre neurônios, respectivamente.

Além da divergência conceitual acima apresentada, deve-se ressaltar que a argumentação anti-reducionista de Searle é alicerçada *principalmente* no modo de existência de primeira pessoa da consciência, ou seja, na subjetividade ontológica da consciência, e não no conceito de emergência: basta lembrar que, para Searle, a liquidez e a solidez são propriedades causalmente emergentes que são ontologicamente redutíveis, o que não ocorre com a consciência. Ou seja, o que garante a irreducibilidade ontológica da consciência não é o fato dela ser emergente, mas, sim, a sua subjetividade ontológica! Portanto, penso que esses dois motivos – divergência conceitual e centralidade da subjetividade ontológica – são, por si mesmos, suficientes para não se classificar o naturalismo biológico como uma teoria emergentista em Filosofia da Mente.

Na verdade, Searle resiste em ser classificado (inclusive como emergentista. Ver grifo na citação logo abaixo), pois segundo ele a Filosofia da Mente é o único campo da Filosofia Contemporânea em que as principais teorias são falsas (M, p. 2). A falsidade de todas as teorias (ou de todos os “ismos”, como ele mesmo as chama) tem, em termos gerais, uma causa comum, a saber, aceitar o dualismo conceitual. A crítica é bastante vasta, e Searle não economiza quando o assunto é apontar quais são as teorias que ele discorda:

By such theories I mean just about anything that has “ism” in its name. I am thinking of dualism, both property dualism and substance dualism, materialism, physicalism, computationalism, functionalism, behaviorism, epiphenomenalism, cognitivism, eliminativism, panpsychism, dual-aspect theory, and *emergentism*, as it is standardly conceived (M, p. 2. Grifos meus)¹³⁴.

Deve-se, antes de comparar o naturalismo biológico ao emergentismo, mostrar o pano de fundo científico sobre o qual a teoria se fundamenta. Para Searle, tudo o que há no mundo é formado por partículas físicas fundamentais e seus agregados – “o universo consiste inteiramente em fenômenos físicos extremamente pequenos que consideramos conveniente (...) chamar de ‘partículas’” (RM, p. 128). Dessa maneira, ele defende que a matéria se organiza de maneira hierárquica em níveis de complexidade crescente. Nessa visão de mundo, não há espaço para substâncias não-físicas, como a *res cogitans* cartesiana e o élan vital dos biólogos vitalistas: tudo o que existe são partículas físicas e seus agregados submetidas a campos de força. Porém, deve-se deixar claro que para Searle o termo “físico” não se contrapõe ao termo “mental” – o mental é físico. Ou seja, não existe a dicotomia cartesiana entre dois tipos mutuamente exclusivos de substância, a física e a mental. De acordo com essa dicotomia, tudo o que existe no mundo deve ser classificado como físico ou mental, mas essas duas categorias são muito pobres para dar conta da imensa variedade de objetos que conhecemos: a crise que abalou o mercado financeiro em 2009, o Governo do Estado do Ceará, a habilidade do Besouro Mangangá em jogar capoeira e um jogo de futebol entre o Flamengo e o Fluminense pertencem a qual âmbito, o físico ou o mental? Não restam dúvidas, ao menos para os cientistas e filósofos de inclinação naturalista, de que tudo o que existe é formado por partículas físicas fundamentais. Entretanto, tal constatação não entra em contradição com a afirmação de que as inúmeras entidades da realidade física possuem variados modos de existência:

(...) se formos chamar de física qualquer coisa que é constituída de partículas físicas, então, *trivialmente, tudo no mundo é físico*. Mas dizer isto não é negar que o mundo contenha tentos marcados em jogos de futebol, taxas de juros, governos e sofrimentos. *Tudo isso tem sua própria maneira de existir* – atlética, econômica, política, mental etc (RM, p. 41. Grifos meus).

Nota-se, a partir do que foi exposto, que Searle rejeita o dualismo de substância. Dessa forma, ele deve optar por uma das duas alternativas que restam: ou se defende o monismo mental (idealismo) ou o monismo físico (fiscalismo), e ele defende este último. Apesar de Searle insistir com veemência que é um erro acreditar que se deve escolher entre o dualismo e o monismo fiscalista, no meu modo de entender, a afirmação de “que o mundo é constituído

¹³⁴ Constata-se, a partir da citação acima, que Searle critica, apesar das ressalvas, o emergentismo.

inteiramente de partículas físicas em campos de força” (RM, p. 45) é claramente comprometida com a concepção do monismo físico. Talvez a relutância dele em aceitar o rótulo de fisicalista tenha a ver com a identificação entre esse termo e o reducionismo. No entanto, como será visto, o fisicalismo é neutro em relação à redutibilidade do mental ao físico, isto é, existe o fisicalismo redutivo e o não-redutivo, e aceitar este último não significa se comprometer com particulares não-físicas. O não-reducionismo do naturalismo biológico diz respeito à tese de “que somos todos conscientes e que nossos estados conscientes têm propriedades fenomenológicas *irredutíveis* bastante específicas” (RM, p. 45), isto é, com a tese da subjetividade/irredutibilidade ontológica do mental.

Como foi dito acima, uma confusão muito comum em Filosofia da Mente é identificar o fisicalismo com o reducionismo, como se todo fisicalista fosse *também* um adepto da tese reducionista e aceitasse a asserção “o mental é físico significa o mesmo que o mental é redutível ao físico”. Richard Rorty há muito já alertava sobre o perigo de se confundir tais conceitos e que, como consequência disso, poderia haver uma falta de diálogo entre as diferentes tradições filosóficas:

Um dos obstáculos que impedem a comunicação entre as filosofias alemã e americana é o fato de que, no interior da filosofia alemã, materialismo e fisicalismo são associados com reducionismo e cientificismo. Em contrapartida, na filosofia americana, a filosofia alemã é associada com o desdém à ciência natural. Mas, como eu vejo, a filosofia americana alcançou agora uma posição que, apesar de permanecer plausivelmente descrita como “materialista” ou “fisicalista”, não é de modo algum cientificista (RORTY, 1997, p. 157).

Rorty tem em mente, ao fazer essa consideração, os filósofos norte-americanos Hilary Putnam e Donald Davidson, que compartilham a posição fisicalista não-redutiva, mas diferem, como já foi dito, quanto ao argumento anti-reducionista que utilizam em suas respectivas teorias: o primeiro propôs pela primeira vez a tese da realização múltipla e o segundo utilizou o conceito de superveniência para abordar o problema mente-corpo. Além de Rorty, Jaegwon Kim também afirma que a realização múltipla e a superveniência são os dois principais argumentos utilizados para se defender as teorias que são fisicalistas e não-reducionistas¹³⁵.

¹³⁵ Kim ataca ambos os argumentos, propondo, por sua vez, que é incoerente defender tal posição. Em relação à tese da realização múltipla, ele propõe resolver o problema a partir da redução da propriedade mental à disjunção dos diversos realizadores físicos de tal propriedade. No que diz respeito à tese da superveniência do mental ao físico utilizada no monismo anômalo, Kim sustenta que caso se leve essa tese às últimas consequências, constatar-se-ia que o mental seria eliminado, já que o mesmo não possuiria qualquer diferença causal no mundo.

Searle também faz parte do grupo de filósofos que endossa o fisicalismo¹³⁶ ao mesmo tempo em que defende a irreducibilidade ontológica dos estados mentais a estados neurobiológicos. Com isso, porém, ele não quer dizer que a consciência esteja acima e além do mundo físico – como, segundo ele, defendem os dualistas de propriedade – mas apenas que devido ela ser ontologicamente subjetiva, ela é ontologicamente irreduzível ao físico:

The dualist says, “Consciousness is irreducible to third-person neurobiological processes.” I say, “Consciousness is irreducible to third-person neurobiological processes.” But the dualist takes this to imply that consciousness is not part of the ordinary physical world but is something over and above it. What I mean is that consciousness is causally reducible [tese da redução causal] but not ontologically reducible [tese da irreducibilidade ontológica]. It is part of the ordinary physical world and is not something over and above it [tese do monismo físico] (M, p. 127).

Em resumo, a consciência é uma propriedade causalmente emergente de sistemas nervosos complexos, e como tal é ontologicamente irreduzível a estes. Porém, a irreducibilidade ontológica não implica, segundo Searle, a irreducibilidade causal – a consciência depende de, e é causada por, estados neurofisiológicos, sendo, portanto, causalmente redutível a estes: “consciência é uma *propriedade causalmente emergente* do comportamento de neurônios, e, portanto, a consciência é causalmente redutível aos processos do cérebro” (RM, p. 168. Grifos meus). No entanto, a explicação da atividade mental pela neurobiológica é *necessariamente limitada*, pois existe um *obstáculo ontológico-lingüístico conceitualmente intransponível*. Do ponto de vista ontológico, os fenômenos mentais são propriedades de nível superior e ontologicamente subjetivas, e não existe nenhuma diferença entre a *aparência* e a *realidade* quando se trata de fenômenos conscientes. A consequência disso é que, da perspectiva lingüística, os fenômenos mentais pertencem a um nível de descrição superior, o que acarreta que a linguagem utilizada para descrever os processos neurobiológicos é *inadequada* para descrever estados mentais:

(...) a redução de outros fenômenos depende em parte da distinção entre ‘realidade física objetiva’, de um lado, e meras ‘aparências subjetivas’, de outro; e da eliminação da aparência dos fenômenos que foram reduzidos. Mas no caso da consciência, *sua realidade é a aparência* (RM, p. 176. Grifos meus).

¹³⁶ Vale ressaltar, mais uma vez, que Searle *não admite* ser nem fisicalista nem dualista. Ele defende que os estados mentais não fazem parte de um domínio autônomo em relação ao físico (tese esta defendida pelos dualistas), mas constituem apenas um nível superior do mesmo sistema físico que os produz. Em outras palavras, estados mentais não estão, segundo Searle, acima e além do nível físico, mas são somente propriedades de nível superior que emergem dos estados neurobiológicos (base física) que os geram, e ambos os estados, mentais e neurobiológicos, constituem um único e mesmo sistema – o sistema cerebral: “Conscious states are realized in the brain as features of the brain system, and thus exist at a level higher than that of neurons and synapses” (M, pp. 113-114).

Portanto, a irreduzibilidade da consciência é uma consequência da maneira que ela é definida e do modelo de redução utilizado:

Dada sua existência [da consciência], sua irreduzibilidade é uma consequência trivial de nossos métodos de definição. (...) quando falo da irreduzibilidade da consciência, estou falando de sua *irreduzibilidade de acordo com modelos padrões de redução* (RM, p. 179. Grifos de Searle).

Na verdade, há uma articulação entre as questões linguística e ontológica no naturalismo biológico: como Searle defende que a linguagem é derivada da mente, deve-se ter em conta que o obstáculo lingüístico é *conseqüência* do obstáculo ontológico¹³⁷. Dito de outro modo, já que o modo de existência dos fenômenos mentais (*ontologia* de primeira pessoa) é distinto do modo de existência dos processos neurofisiológicos (*ontologia* de terceira pessoa), a linguagem que descreve estes (*vocabulário* de terceira pessoa) não pode ser utilizada para descrever aqueles fenômenos. Os fenômenos psicológicos são descritos, por sua vez, pelo *vocabulário* de primeira pessoa da Psicologia Popular. A questão que se coloca a seguir é saber se esse obstáculo é apenas temporário ou se ele é, ao contrário, insuperável. Novamente, existem duas respostas para a mesma pergunta. No tocante aos interesses empíricos da pesquisa científica, pode-se redefinir a consciência em termos de seus correlatos neurais¹³⁸, mas penso que isso é insuficiente para que haja uma *redução conceitual* efetiva. Isso se deve porque o sucesso da redução possibilita um *ganho explicativo*, e ocorre justamente o contrário quando se *tenta* redefinir dor em termos de disparos de fibras C, perde-se o caráter subjetivo de tal experiência:

Suponha que tentássemos dizer que a dor é na verdade “nada exceto” as disposições de descargas de neurônios. Bem, se tentássemos uma tal redução ontológica, as características essenciais da dor seriam deixadas de lado [*perda explicativa*]. Nenhuma descrição dos fatos [nível ontológico] de terceira pessoa, objetivos, fisiológicos comunicariam [nível lingüístico] o caráter subjetivo, de primeira pessoa, da dor, simplesmente porque as características de primeira pessoa são diferentes das características de terceira pessoa (RM, p 170).

Em resumo, o obstáculo ontológico-lingüístico pode ser *parcialmente* superado por interesses empírico-pragmáticos, mas é, *a priori*, conceitualmente intransponível. Em outras palavras, a irreduzibilidade ontológica defendida por Searle acaba gerando um produto secundário, a saber, a *irreduzibilidade conceitual*¹³⁹: os estados mentais são ontológica e

¹³⁷ Conferir a 4ª tese do capítulo anterior.

¹³⁸ É digno de nota essa rara concessão de Searle: “Well, we could, and we might for medical or other scientific purposes redefine consciousness in terms of microsubstrates, as we have redefined solidity and liquidity” (M, p. 120).

¹³⁹ A irreduzibilidade conceitual, aliás, é um ponto de convergência entre o naturalismo biológico e o monismo anômalo, que é um tipo de fisicalismo não-reduutivo (Cf. DAVIDSON, 2004).

conceitualmente irreduzíveis a estados cerebrais. Levando-se em consideração que a *principal* diferença entre o fisicalismo redutivo e o não-redutivo é a posição anti-reducionista deste último, a qual Searle endossa, constata-se que por causa disso (além, é claro, da aceitação do monismo físico), o naturalismo biológico já se enquadraria como um tipo de fisicalismo não-reducionista. Com isso, fica evidente que Searle aceita as duas primeiras teses que caracterizam o fisicalismo não-redutivo (monismo físico e anti-reducionismo).

Em relação à terceira tese (realização física), a questão que se coloca inicialmente é a seguinte: ao defender a tese da redução causal, Searle se compromete com a tese da realização física? Acredito que sim, mas antes de argumentar a favor disso, deve-se ter em mente no que consiste essa última tese. De acordo com a tese da realização física, todo sistema físico que instancia uma propriedade mental M possui uma propriedade física P, tal que P realiza M nesse sistema¹⁴⁰. Por conseguinte, os eventos e estados mentais dependem de um substrato físico e se este estiver presente, aqueles ocorrerão, ou seja, as condições físicas basais são *suficientes* para o aparecimento das propriedades de nível superior, no caso, das propriedades mentais.

Searle aceita que a estrutura física basal seja suficiente para produzir propriedades mentais, mas não *qualquer* estrutura física. Como foi visto no segundo capítulo, ele critica o argumento funcionalista que se baseia na tese da realização múltipla para defender que a implementação de um programa de computador é suficiente para produzir uma mente, pois tal implementação não é capaz, por si só, de duplicar as capacidades causais do sistema nervoso e, conseqüentemente, não é capaz de produzir estados mentais: “*qualquer sistema capaz de causar consciência tem que ser capaz de reproduzir as capacidades causais do cérebro*” (RM, p. 137)¹⁴¹. Em outras palavras, se um determinado sistema possuir uma estrutura física que reproduza as capacidades causais do cérebro (condição basal), ele produzirá e sustentará estados mentais (propriedade causalmente emergente), já que aquela estrutura é *suficiente* para a produção destes. A tese da realização física, por sua vez, acarreta a tese da redução

¹⁴⁰ A aceitação de tal tese vai de encontro ao dualismo de substância, já que nessa corrente os estados mentais ou a alma podem existir independente dos corpos. Contudo, ela não se contrapõe ao dualismo de propriedade, pois ao defender o monismo físico (só há a substância física), essa teoria também se compromete com a tese da realização física: apesar de existirem dois tipos de propriedade – física e mental – esta última depende de seu substrato físico. Kim, no entanto, não aceitaria isso, já que para ele a irreduzibilidade das propriedades mentais a propriedades físicas implicaria a existência de dois domínios autônomos, portanto, dois tipos de substâncias. Dessa maneira, Kim afirma que só existem três posições possíveis: o dualismo anti-reducionista, o reducionismo e o eliminativismo (KIM, 1993, p. 267).

¹⁴¹ Searle deveria estar atento para o fato de que, em geral os defensores da IA não falam em termos de *consciência*, mas, como o próprio nome da teoria diz, em *inteligência*. Para Searle, a inteligência *depende* da consciência; porém, alguns teóricos distinguem as duas, já que seria possível considerar o termostato um aparelho inteligente, mas sem consciência.

causal, já que afirmar que os estados mentais necessitam de uma estrutura física implica que os mesmos são causados por tal estrutura, no caso, pelo conjunto das sinapses neurônicas. Além disso, ao defender que estados mentais são causalmente redutíveis a estados cerebrais, Searle está sugerindo que o nível mais adequado para se explicar a mente é o nível neurobiológico, já que fenômenos mentais *são* fenômenos biológicos:

(...) the explanatory apparatus that we use to give a causal account of mental phenomena is an apparatus that we need to account for nature generally. The level at which we attempt to account for mental phenomena is biological rather than, say, at the level of subatomic physics (M, p. 301).

A citação acima deixa claro que Searle defende que o mundo é hierarquizado e que, para um mesmo sistema, existem diferentes níveis de explicação, os quais não se excluem, mas se complementam. Dessa maneira, embora sejamos formados por átomos, o nível atômico é *inadequado* para explicar o meu desejo de beber uma cerveja gelada na noite de sexta-feira. Todavia, a constituição atômica de *meu* corpo é uma condição *sine qua non* para o *meu* desejo de beber cerveja, pois *meus* átomos formam *minhas* moléculas; estas constituem minhas células, inclusive *meus* neurônios, que por sua vez, ao se comunicarem por sinapses, produzem *meus* estados mentais, entre os quais, o *meu* desejo de beber cerveja.

De acordo com esse quesito, o naturalismo biológico se assemelha ao fisicalismo de espécime, mas não de tipo: há, *obrigatoriamente*, um estado cerebral que produz o desejo de Homer Simpson beber uma cerveja gelada na sexta-feira à noite (tese da realização física). Entretanto, *não é necessário*, a despeito de nossos desejos constituírem o mesmo tipo de estado mental, que o estado cerebral de Homer seja igual ao meu (fisicalismo de tipo), pois *diferentes* estados cerebrais, C e C*, por exemplo, podem produzir o *mesmo tipo* de estado mental M (desejo de beber cerveja). Contudo, a identidade de estados cerebrais é *suficiente* para garantir a identidade de estado mental (fisicalismo de espécime) – se o cérebro de Homer estiver no mesmo estado cerebral que o meu na ocasião em que desejo beber cerveja gelada, *inevitavelmente* ele também estará desejando uma cerveja gelada. Além disso, quando o meu desejo de beber uma cerveja dá lugar ao meu desejo de dormir, isto é, quando ocorre uma modificação de meus estados psicológicos, há *necessariamente* uma mudança de meus estados cerebrais, já que estes geram aqueles (tese da redução causal). Searle deixa bem claro esse ponto ao explicitar o conceito de superveniência que ele adota, a superveniência causal:

(...) estados mentais são totalmente dependentes de estados neurofisiológicos correspondentes [tese da realização física], no sentido de que uma diferença nos estados mentais envolveria necessariamente uma diferença correspondente nos estados neurofisiológicos [tese da superveniência causal]. (...) causas

neurofisiológicas de tipo idêntico teriam efeitos mentalísticos de tipo idêntico (RM, p. 180).

Em relação à última tese (realismo mental), deve-se ressaltar que, para Searle, a dependência dos estados mentais em relação aos neurobiológicos não significa que aqueles são irrealis: estados psicológicos são reais, isto é, não são meras ilusões ou aparências. Sendo assim, a tese do realismo mental faz parte do naturalismo biológico. Já um defensor do eliminativismo defende o seguinte: o desenvolvimento da Neurociência não acarretará a redução da Psicologia Popular, pois é improvável que haja uma correspondência um-a-um entre as duas ciências. Segundo esse ponto de vista, a Psicologia Popular é uma representação distorcida e enganosa de nossos estados neurofisiológicos e das causas de nosso comportamento, e deve, por esse motivo, ser *eliminada* (e não reduzida) por uma Neurociência avançada e precisa:

(...) não podemos esperar que uma explicação realmente adequada de nossa vida interior feita pela neurociência revele categorias teóricas que correspondam exatamente às categorias do arcabouço de senso comum [posição anti-reducionista]. Dessa forma, devemos esperar que o antigo arcabouço seja simplesmente eliminado, e não reduzido, por uma neurociência amadurecida [tese eliminativista] (CHURCHLAND, 1998, p. 79).

Churchland lança mão da seguinte analogia para defender a sua visão. Ao retroceder alguns séculos, pode-se constatar que antigamente, como nos tempos das Inquisições, muitas pessoas alegavam que determinadas mulheres tinham pacto com Satanás. As bruxas, como eram chamadas essas mulheres, eram caçadas e condenadas à morte, sendo queimadas vivas nas fogueiras. Nos dias de hoje, nem a Igreja nem qualquer pessoa de bom senso admitiria a existência de bruxas no sentido acima citado, ou seja, não é que o conceito de bruxa seja incompleto, ele é simplesmente *falso*, e por isso foi eliminado de nosso arcabouço conceitual – devido ao nosso atual conhecimento científico, não há espaço para as místicas bruxas em nosso universo. Da mesma maneira que com o passar do tempo o conceito de bruxa foi eliminado – a falsidade do conceito de bruxa implica a sua não existência –, também serão eliminados, com o desenvolvimento neurocientífico, os conceitos da Psicologia Popular, tais como crença, desejo e dor¹⁴²:

Nossas explicações sobre os comportamentos uns dos outros irão recorrer a coisas como nossos estados neurofarmacológicos, nossa atividade neural em áreas anatômicas específicas e a outros estados que forem relevantes para a nova teoria.

¹⁴² Essa analogia, entretanto, é muito limitada, pois nunca houve uma única mulher que possuísse, *de fato*, os poderes mágicos de uma bruxa, mas dores são estados mentais que existem inegavelmente, já que praticamente todas as pessoas sentem ou já sentiram dores (existe uma patologia rara denominada *insensibilidade congênita à dor*, cujos portadores são exceção).

Nossa introspecção individual também será transformada e poderá ser profundamente aprimorada em razão de um arcabouço conceitual mais penetrante e preciso com o qual ela terá de trabalhar (...) (CHURCHLAND, 1998, p. 82).

Por defender que os estados mentais descritos pelo arcabouço teórico da Psicologia Popular não passam de ilusão, pode-se dizer que o eliminativismo defende a tese do *irrealismo mental*. Searle trata com sarcasmo os proponentes das teorias que negam a realidade dos estados mentais e conscientes:

Como, por exemplo, alguém empreenderia a refutação do ponto de vista de que a consciência não existe? Deveria eu beliscar seus adeptos para lembrá-los de que são conscientes? Deveria beliscar a mim mesmo e relatar os resultados no *Journal of Philosophy*? (RM, p. 17).

A tese oposta, como foi dito acima, faz parte do naturalismo biológico: propriedades mentais são propriedades *reais* de objetos e eventos. Ou, melhor ainda, além de não serem meras ficções com uso meramente instrumental, as propriedades mentais têm uma ontologia de primeira pessoa (tese da subjetividade ontológica dos fenômenos mentais). Assim como o naturalismo biológico, o fisicalismo eliminativista adota uma posição anti-reducionista, mas por motivos bastante distintos: o primeiro por afirmar que não é possível reduzir estados e eventos com ontologia de primeira pessoa a estados e eventos de terceira pessoa, e o segundo por negar qualquer status ontológico aos fenômenos mentais.

A tese do realismo mental está relacionada à da subjetividade ontológica dos estados mentais no naturalismo biológico: os estados têm um modo de existência de primeira pessoa, logo eles são reais, e não meras ficções, isto é, eles fazem parte do universo físico. Além disso, os estados psicológicos não são, segundo Searle, misteriosos, pois apesar de serem irreduzíveis, do ponto de vista ontológico, a estados neurofisiológicos, são causalmente explicáveis por estes. O realismo do mental é uma tese *ontológica* que tem uma importante consequência *epistemológica*, já que

Todo estado consciente é sempre estado consciente *de alguém* [tese da subjetividade ontológica]. E, da mesma forma como tenho uma relação especial com meus estados conscientes, que não é como minha relação com os estados conscientes de outras pessoas, assim estas, por sua vez, têm uma relação com seus estados conscientes que não é como minha relação com seus estados conscientes [tese da assimetria epistemológica] (RM, p. 140).

Como visto anteriormente, o naturalismo biológico não é uma teoria emergentista. No que se segue, compararei a teoria de Searle com a de Pereboom e Kornblith, os quais defendem uma concepção de fisicalismo não-redutivo na qual a realização múltipla não é a característica central da irreduzibilidade, divergindo, neste ponto, de filósofos como Putnam.

Apesar de não ser a principal característica da teoria de Pereboom e Kornblith, o fenômeno de realização múltipla serve para mostrar que a possibilidade de um estado mental (M) ser realizado por múltiplos estados físicos ($P_1, P_2, P_3\dots$), faz com que a relação, através de princípios-ponte, entre um estado mental e suas múltiplas instanciações físicas envolva disjunções abertas. O estado mental de desejar um sorvete, por exemplo, pode ser realizado por P_1 ou P_2 ou P_3 (...). No entanto, as disjunções abertas de estados físicos não são tipos naturais físicos, pois não é possível extrair *leis genuínas* de tal disjunção, isto é, leis explicativas do tipo “se alguém deseja sorvete, então ela está no estado neurobiológico N”. O desejo de tomar sorvete de Carlota pode ser instanciado por N_1 , enquanto o de Joaquina por N_2 . Ou ainda, é possível que o desejo de Carlota tomar sorvete na 2ª-feira seja realizado pelo estado neurobiológico N_1 , enquanto que amanhã será realizado por N_2 . A partir disso, conclui-se que os tipos de estados mentais são irreduzíveis a tipos de estados físicos, tais quais os estados neurobiológicos que os realizam. Sendo assim, o papel desempenhado pela tese da realização múltipla numa teoria fisicalista não-reduzitiva é o de mostrar que as “leis” e “explicações” obtidas a partir da *tentativa* de redução dos tipos de estados mentais a tipos de estados neurobiológicos são insatisfatórias:

The basis (...) for the claim that psychology is not reducible to neurophysiology is not simply that mental states are multiply realizable at the neurophysiological level, but rather that this multiple realizability shows that attempts at reduction would require laws and explanations of a very peculiar kind; so peculiar, indeed, that they would be unsatisfying as laws and explanations (PEREBOOM & KORNBLITH, 2004, p. 711).

Deve-se ressaltar que o obstáculo da realização múltipla não se dá apenas no âmbito de organismos da mesma espécie, mas perpassa numerosas espécies com suas respectivas estruturas biológicas distintas. Ou seja, a dor do jacaré, da galinha e do homem pode ser realizada por diferentes estruturas físicas, o que acarretaria que estados neurofisiológicos distintos – N_j, N_g e N_h , respectivamente – realizariam o mesmo tipo de estado mental, a dor. Porém, segundo Kim (1993a, p. 272), esse fato não oferece maiores problemas para a teoria reducionista, pois existe uma *conexão forte* entre estados psicológicos, de um lado, e os tipos de estruturas físico-biológicas que o realizam, de outro, o que possibilita a realização de reduções espécie-específica. De acordo com esse tipo de redução, um tipo de estado mental é redutível a um tipo de estrutura biológica que o realiza em *determinada espécie*. Dessa maneira, apesar do tipo de estado mental dor não poder ser reduzido a um único estado ou estrutura físico-biológica presente em todas as espécies de organismos, nada impede que haja sucesso na estratégia de se reduzir o tipo de estado psicológico a estruturas biológicas

específicas: o jacaré, a galinha e o homem possuem diferentes estruturas físico-biológicas (S_j , S_g e S_h) que são, por sua vez, *condições necessárias e suficientes* para a instanciação do estado mental dor no jacaré (D_j), na galinha (D_g) e no homem (D_h).

No entanto, a solução das reduções espécie-específica proposta por Kim é criticada por Pereboom e Kornblith, já que o objetivo do projeto reducionista é reduzir explicações e teorias de uma ciência de nível elevado a explicações e teorias de uma ciência mais básica, e explicações e teorias quantificam sobre tipos. Além disso,

(...) psychological theories and explanations quantify over types of mental states, like pain, that are instantiated by organisms of many different species and structural types. The suggestion that there are species-specific reductions of pain results in the claim that pains in different species have nothing in common (PEREBOOM & KORNBLITH, 2004, p. 718).

Para compreender a irredutibilidade da Psicologia à Neurobiologia, deve-se, antes de tudo, fazer a distinção entre dois tipos de explicação, a *causal* e a *constitutiva*. Um exemplo do primeiro tipo de explicação é o seguinte: “o aumento da temperatura do gás em um balão de ar quente *causa* a subida do mesmo”. Já a explicação constitutiva, como o nome indica, diz respeito à especificação da constituição de algo. No que se refere ao exemplo anterior, a explicação constitutiva é de fundamental importância para se entender *como* o aumento de temperatura está relacionado à subida do balão, pois com o aumento da energia cinética do gás, este diminui a densidade, e a baixa densidade do gás *causa* a subida do balão. Quando a explicação causal é esclarecida pela descrição constitutiva, ocorre a redução de tipos de fenômenos que pertencem a um domínio elevado a tipos de fenômenos de um domínio mais básico:

The constitutional account of gas temperature thus invokes properties that illuminate the above type-level causal explanation. In this case, a constitutional account provides us with a *reduction* of a type-level causal explanation because the constitutional account illuminates this explanation (PEREBOOM & KORNBLITH, 2004, p. 712).

Todavia, há casos onde a descrição constitutiva proveniente de uma ciência mais básica não esclarece a explicação causal em nível de tipo. Isso pode ser exemplificado através da relação entre a Genética Clássica e a Biologia Molecular: a transmissão dos genes do homem e da mulher ocorre por meio da reprodução, que consiste no processo onde o gameta de um (espermatozóide) fecundará o do outro (óvulo). A formação dos gametas (gametogênese) se dá por um tipo de divisão chamada meiose, a qual consiste na separação das cromátides e cromossomos, conseqüentemente, na separação dos genes. O processo de

separar os genes é multiplamente realizável em nível molecular, já que as cromátides e cromossomos podem ser separados pela ação do citoesqueleto, de forças nucleares ou mesmo da gravidade. É devido à realização múltipla que tal processo não pode ser considerado um tipo natural, o que impede, por sua vez, uma redução em nível de tipo, pois

(...) there is no molecular constitutional explanation which illuminates, rather than obscures, the type-level genetic process. In general, because molecular constitutional accounts of genetic process-type fail to illuminate type-level genetic causal explanation, type-level causal explanation are not reducible to molecular explanations.

(...)

It is one thing to explain how a token process is realized in physical material, but quite another to show that a certain type of process reduces to a physical process. The first is possible for every existing non-microphysical process, the second only for those where there is a constitutional explanation at the type-level (PEREBOOM & KORNBLITH, 2004, p. 713).

Quando ocorre uma redução em nível de tipo, os poderes causais dos objetos e propriedades do nível macro são *idênticos* aos poderes causais do nível micro. Desse modo, afirmar que “a temperatura elevada do gás (explicação de nível macro) causa a subida do balão” é equivalente a dizer que “a elevada energia cinética das moléculas de gás (explicação de nível micro) causa a subida do balão”. Contudo, nos casos em que a descrição constitutiva não resulta em uma redução em nível de tipo, não há identidade dos poderes causais de nível superior e inferior. Retornando ao exemplo da Genética, o poder causal de um espécime de gene é *constituído* dos poderes causais da molécula de DNA:

Where there is irreducibility of explanation, there is only token constitution, and no identity of causal powers. In fact, when one type of explanation does not reduce to an explanation at a more basic level, the causal powers at the higher level of explanation are *neither type- nor token-identical* to causal powers at the more basic level (PEREBOOM & KORNBLITH, 2004, p. 714).

Quando se considera a relação entre estados psicológicos e estados neurofisiológicos, constata-se que os poderes causais daqueles *não são* idênticos aos poderes causais destes últimos, nem no nível de tipo nem no de espécime. Como um tipo de estado mental é multiplamente realizável em nível físico, não se pode reduzir aquele a este, não havendo assim identidade em nível de tipo. Por outro lado, também não há identidade em nível de espécime, pois o mesmo espécime de estado mental pode ser realizado por diferentes espécimes físicos. Pereboom e Kornblith se utilizam do seguinte exemplo para apoiar essa tese: imagine que P é o poder causal das moléculas M que fazem parte seus neurônios, que por sua vez realizam o desejo de tomar sorvete (estado mental D). Suponha ainda que ao comer cereal no seu primeiro aniversário você tenha adquirido tais moléculas. Entretanto,

havia no armário duas caixas de cereal do mesmo tipo, e sua mãe escolheu a da esquerda. A questão que se coloca é: e se ela tivesse escolhido a da direita? Caso isso ocorresse, você não teria ingerido e absorvido as moléculas M e, conseqüentemente, não teria os poderes causais dessas moléculas. Como o cereal da direita é do mesmo tipo que o da esquerda, ele fornece moléculas do mesmo tipo M', as quais possuem poderes causais P'. Conclui-se, a partir do que foi dito, que não há identidade em nível de espécime porque o mesmo espécime de estado mental (D) pode ser causado por diferentes espécimes de moléculas constituintes (M e M'). Do mesmo modo, o mesmo espécime de poder causal psicológico é realizado a partir de diferentes poderes causais físicos (P e P'):

(...) you clearly would have had the very same token desire for ice-cream with its token psychological causal power D. Hence, a token psychological state and its token psychological causal power can remain the same even when its token molecular constitution, and thus its token molecular causal powers, are altered. This result can easily be generalized; hence, when one type of explanation does not reduce to another, there is neither type- nor token-identity between causal powers (PEREBOOM & KORNBLITH, 2004, pp. 714-715).

Em resumo, a tese central da proposta de fisicalismo não-reduutivo de Pereboom e Kornblith é a de que como não há identidade entre os poderes causais dos estados mentais e neurobiológicos, a Psicologia é irreduzível à Neurociência, assim como as explicações psicológicas (nível elevado) são irreduzíveis a explicações neurocientíficas (nível básico).

Em primeiro lugar, o naturalismo biológico diverge fundamentalmente da abordagem de Pereboom e Kornblith, já que uma das teses defendidas por Searle é a de que os poderes causais dos fenômenos mentais são os mesmos dos fenômenos neurobiológicos, conferindo, dessa maneira, eficácia causal aos estados mentais: quando meu desejo de levantar o braço (estado mental M) causa a sua elevação (estado físico P), é porque M tem o mesmo poder causal que seu estado neurofisiológico subjacente N. Em outras palavras, não existe nenhum poder causal de M que P não possua, e vice-versa: "(...) the causal powers of consciousness are exactly the same as those of the neuronal substrate" (M, pp. 127-128). Portanto, o argumento anti-reducionista de Searle não se fundamenta na distinção entre os poderes causais dos níveis macro e micro – já que ele defende justamente o contrário –, mas na diferença ontológica entre tais níveis: o nível psicológico tem uma ontologia de primeira pessoa e o neurobiológico, de terceira pessoa. Contudo, Searle tem que escolher entre a distinção ontológica ou a identidade dos poderes causais dos fenômenos mentais e neurobiológicos, ou seja: se dois estados são ontologicamente distintos, eles *não* podem ter as

mesmas capacidades causais; inversamente, se dois estados possuem os mesmos poderes causais, eles *são* ontologicamente idênticos.

Segundo, a tese da identidade das capacidades causais também está relacionada com a tese da redução causal: como as capacidades causais dos estados mentais e neurofisiológicos são idênticas, pode-se, segundo Searle, explicar causalmente aqueles estados a partir destes últimos, ou seja, a Psicologia é causalmente redutível à Neurobiologia. Esta tese do naturalismo biológico também está em desacordo com o tipo de fisicalismo não-redutivo exposto acima, já que a proposta de Searle entra em conflito com a noção de explicação constitutiva da teoria de Pereboom e Kornblith. Isso significa que, de acordo com estes últimos, é um erro defender que estados neurobiológicos causam estados mentais; o correto é afirmar que estados mentais são *constituídos* por estruturas neurobiológicas. Nesse caso, diferentemente do exemplo temperatura-energia cinética, a descrição neurobiológica constitutiva (nível micro) não esclarece a explicação psicológica, pois, como já foi dito, um mesmo estado mental é realizável de múltiplos modos em nível físico, o que impossibilita a redução de estados e propriedades mentais a estados e propriedades neurobiológicas.

Para Searle, a redução causal dos estados mentais não implica a redução ontológica dos mesmos, já que estes possuem uma ontologia subjetiva, enquanto os estados cerebrais são ontologicamente objetivos. Contudo, isso também constitui um problema para sua concepção: se tudo o que existe consiste, em última análise, de partículas físicas fundamentais e os estados mentais *são* estados físicos, disso decorre que após a realização da redução causal daqueles estados a estes últimos, não deveria ficar de fora nenhuma característica psicológica. Aqui, surge o segundo dilema a ser enfrentado por Searle: se o mental é, de fato, físico, das duas uma – ou se admite que os fenômenos psicológicos de nível superior têm ontologia de terceira pessoa, já que tais fenômenos seriam caracterizados nos termos do nível mais fundamental, o qual possui ontologia objetiva, ou se defende que a subjetividade ontológica está presente no micronível. A primeira posição é fisicalista redutiva, enquanto a segunda se enquadra em uma espécie de idealismo panpsiquista. Searle não endossa nenhuma dessas abordagens, mas a única maneira de evitar ambas é negar que o mental seja físico, o que acarretaria alguma espécie de dualismo. Portanto, constata-se que assim como a teoria de Descartes, o naturalismo biológico também tem como pressuposto um dualismo conceitual. No entanto,

A diferença é que, enquanto a tradição contrapunha conceitos *mentais* e conceitos *físicos*, [Searle] afirma que o mental *é* físico (porque biológico), mas acaba contrapondo conceitos *subjetivos* e conceitos *objetivos*. Desse modo, sua

teoria da mente parece consistir em “um dualismo dentro do físico” [expressão de Andreas Kemmerling, professor da Universidade de Heidelberg], atribuindo uma natureza física aos fenômenos mentais, mas ainda dualista no que diz respeito aos conceitos. É extremamente questionável se esse novo dualismo entre o subjetivo e o objetivo apresenta alguma vantagem em relação ao dualismo tradicional do físico e do mental (PRATA, 2009, p. 117).

Searle se defende, afirmando que é um erro confundir sua teoria com um dualismo, pois, segundo ele, o dualista parte do pressuposto que o fenômeno mental está acima e além da base física, o que acaba gerando o dilema sobre-determinação causal-epifenomenalismo. Entretanto, nenhum desses problemas surgiria no naturalismo biológico, pois, de acordo com ele, “the causal functioning of consciousness is just a form of brain functioning described at a level higher than that of neurons and synapses” (M, p. 130). Contudo, essa defesa de Searle não é de grande valia¹⁴³, pois afirmar que os fenômenos mentais são ontologicamente subjetivos, causalmente eficazes e reais leva a sua posição à sobre-determinação causal. Por outro lado, a saída para o epifenomenalismo consiste em defender a causação mental, e esta se deve à identidade entre os poderes causais dos estados neurobiológicos e mentais. Todavia, é incoerente sustentar a identidade das capacidades causais e a diferença ontológica entre os fenômenos neurobiológicos e mentais. Portanto, é difícil compreender como a diferença ontológica é compatível com a eficácia causal do mental: “For if the mental is not identical to the physical, and the physical is causally complete, then how can there be any causal work for the mental to do?” (CRANE, 1993, p. 319). Sendo assim, a concepção de causação mental que está presente no naturalismo biológico é bastante problemática, e não é possível resolver o problema mente-corpo sem se obter êxito nesse quesito:

Any would-be solution to the mind-body problem must include an account of mental causation, and not simply posit it as an obvious fact we must all accept. I don't see such an account in Searle; nor I do see a serious possibility of developing one within the terms set by his biological naturalism (KIM, 1995, p. 194).

Terceiro, devido à superveniência causal defendida por Searle, o naturalismo biológico assemelha-se à concepção fisicalista de espécime: a identidade neurobiológica é uma condição tanto suficiente como necessária para a identidade psicológica. Como foi visto, Pereboom e Kornblith rejeitam a identidade em nível de espécime, pois o mesmo espécime de estado psicológico é multiplamente realizável por diferentes espécimes de estados neurofisiológicos: o desejo de tomar sorvete da pessoa P no tempo t (espécime de estado

¹⁴³ Corcoran (2001, pp.320-321), que classifica o naturalismo biológico como sendo uma concepção dualista de propriedade, também afirma que o naturalismo biológico está enquadrado na tradição filosófica que Searle diz ter rejeitado, afirmando, ainda, que além de sua abordagem não ser coerente, continua submetida a muitos problemas, como o da causação mental.

mental) pode ser realizado por diferentes constituições microfísicas – as moléculas M e M' e seus respectivos poderes causais, P e P'. A adoção da tese da superveniência faria com que o naturalismo biológico fosse ao mesmo tempo uma posição fisicalista (os estados basais são estados físicos) e não-reduzitiva (os estados supervenientes são irreduzíveis aos estados basais). Essa abordagem resulta do duplo comprometimento de Searle: o fisicalismo decorre de seu compromisso com uma *filosofia naturalista*, enquanto que o anti-reducionismo provém de seu apego à uma *filosofia subjetivista*.

A partir disso, pode-se questionar: já que a abordagem naturalista está baseada numa metodologia epistemológica de terceira pessoa que visa explicar objetos e propriedades naturais ontologicamente objetivas, como conciliá-la com uma perspectiva centrada nos aspectos fenomênicos com ontologia de primeira pessoa. Eis o mais um problema que se coloca à concepção de Searle: caso ele defenda, de fato, uma abordagem naturalista, ele terá que abrir mão da subjetividade ontológica da consciência e, conseqüentemente, de sua irreduzibilidade (fisicalismo reduzitivo); por outro lado, defender que a consciência é a única característica física ontologicamente irreduzível presente no Universo é o mesmo que dizer que ela é *sui generis* ao ponto de poder se considerar que ela está acima e além do domínio físico (dualismo). Mais uma vez, Searle não endossa nenhuma dessas posições.

De volta ao início, o naturalismo biológico é um tipo de fisicalismo não-reduzitivo? Não, porque suas teses fisicalistas entram em contradição com as teses anti-reducionistas, tornando o naturalismo biológico um sistema incoerente repleto de dicotomias mutuamente exclusivas: naturalismo/subjetivismo, neurobiológico/psicológico, físico/mental, ontologia de terceira pessoa/ontologia de primeira pessoa, independência do observador/dependência do observador, intencionalidade secundária/intencionalidade primária, redução causal/irreduzibilidade ontológica, determinação física/indeterminismo mental (os estados mentais são *determinados* pelos estados físicos/livre-arbítrio). Segundo David Chalmers, apesar de Searle querer levar a sério a consciência e ainda assim ser fisicalista, ele fracassa, já que “the nonreductive materialism advocated by Searle turns out to have internal problems and collapses into one of the other views (most likely property dualism) (CHALMERS, 1996, p. 164).

Esses problemas, como mostrados anteriormente, surgem da própria tentativa de conciliar as teses do naturalismo biológico. Pode-se citar, por exemplo, o seguinte. Já que a não expressabilidade dos fenômenos mentais (4ª tese) é derivada da 3ª subjetividade ontológica dos mesmos (3ª tese), e os fenômenos mentais são causalmente explicáveis (5ª tese) porque são causados e realizados no sistema cerebral (2ª tese), surge aqui uma dupla

dificuldade para o naturalismo biológico, a saber: (i) como conciliar a redutibilidade causal da consciência (que é derivada da 2ª tese) com sua irredutibilidade ontológica (3ª tese), e (ii) como defender coerentemente que a não expressabilidade do mental a partir do neurobiológico (4ª tese) seja compatível com a explicação causal daquele por meio deste (5ª tese). Aqui, mais uma vez, está em questão as distinções que Searle faz entre ontologia de primeira e terceira pessoa e entre os conceitos de redução ontológica e causal, o que o leva às duas peculiaridades do mental, a saber: (i) todos os fenômenos físicos são ontologicamente objetivos, *excetuando-se* os fenômenos mentais, e (ii) a explicação causal de um fenômeno físico através de processos subjacentes leva à redução ontológica daquele fenômeno a estes processos, *excetuando-se* a consciência, que é causalmente redutível, mas não ontologicamente, a processos cerebrais.

Na minha opinião, Searle tem que abrir mão, *se quiser ser coerente*, de uma das duas distinções acima. Como o naturalismo biológico é uma teoria da consciência e esta, para ele, é ontologicamente subjetiva, o preço a se pagar pelo abandono da distinção entre objetivo e subjetivo seria muito elevado, pois sua concepção passaria a se enquadrar como um fisicalismo redutivo: estados mentais são *nada exceto* estados neurobiológicos. A outra alternativa seria abandonar a distinção entre redução ontológica e causal, realizando uma *economia conceitual*, já que a noção searleana de redução causal já está *implícita* no conceito de redução ontológica: se um fenômeno de nível superior é *causalmente* explicado pelo comportamento exibido no micronível, constata-se que o modo de existência daquele é *nada exceto* o modo de existência deste. Como ocorre o inverso disso no naturalismo biológico – ou seja, a irredutibilidade ontológica da consciência –, constata-se que o mental *não é* explicado neurobiologicamente, já que a completa descrição neural de meu cérebro *não capta nenhum* aspecto qualitativo de meus estados psicológicos, e isso mostra, por sua vez, que a noção de explicação causal de Searle é, em minha opinião, *insuficiente* para originar uma verdadeira teoria neurocientífica acerca da consciência. Em outras palavras, a perspectiva de primeira pessoa dos estados mentais estaria, a despeito do que Searle diz, *acima e além* dos processos cerebrais, isto é, o naturalismo biológico seria uma teoria dualista de propriedade. Isso fica claro quando Searle defende que “no caso da consciência, temos um elemento subjetivo irredutível que sobra depois de havermos feito um relato causal completo da base neurobiológica” (MLS, p. 58). Ou ainda, quando ele afirma que o conceito de consciência está *restrito* aos únicos fenômenos ontologicamente irredutíveis da natureza, os quais são irredutíveis *por definição*: “(...) não podemos eliminar as experiências subjetivas da

consciência, porque a *finalidade principal* do conceito de consciência é ter um nome para os *fenômenos subjetivos de primeira pessoa*” (MLS, p. 60).

Proponho, por fim, que o naturalismo biológico é um tipo de dualismo de propriedade pelos seguintes motivos: (i) sistemas nervosos complexos são inteiramente constituídos de partículas físicas e seus agregados (monismo físico), mas eles possuem dois tipos de propriedades, físicas e mentais: as primeiras possuem um modo de existência objetivo, enquanto as últimas têm uma ontologia de primeira pessoa. Além disso, existem estados neurobiológicos que possuem *apenas* propriedades físicas (os que controlam meu batimento cardíaco e minha frequência respiratória, por exemplo), enquanto outros, como a ativação das fibras nervosas que causam dor, possuem tanto propriedades físicas quanto mentais, e estas são ontologicamente irreduzíveis àquelas; (ii) é por causa do modo de existência de primeira pessoa das propriedades mentais que elas são irreduzíveis, da perspectiva ontológica, às propriedades neurobiológicas. Por mais que no futuro a Neurociência fosse capaz de mapear *todos* os correlatos neurais da consciência, mostrando a correspondência entre cada estado mental com um estado neurobiológico subjacente, ainda ficaria algo de fora, isto é, as propriedades qualitativas e subjetivas da consciência; (iii) como nenhuma descrição física da atividade cerebral é capaz de exprimir os aspectos subjetivo e qualitativo dos fenômenos mentais, constata-se que tais aspectos estão acima e além da abordagem científica, no geral, e da Neurociência, em particular. Isso, por conseguinte, leva à consequência de que a teoria de Searle para explicar o mental *não é* naturalista, mas dualista, já que ele foi incapaz de se livrar da dicotomia que, no meu modo de entender, *exclui o mental do natural*, a saber, a dicotomia entre subjetivo e objetivo.

REFERÊNCIAS BIBLIOGRÁFICAS

ARMSTRONG, David. “What Is Consciousness?”. In: BLOCK, FLANAGAN & GÜZELDERE. *The Nature of Consciousness: Philosophical Debates*. Cambridge, MA: MIT Press, 1997.

BECKERMANN, Ansgar. “Introduction – Reductive and Nonreductive Physicalism. In: _____, FLOHR & KIM. *Emergence or Reduction?: Essays on the prospects of Nonreductive Physicalism*. Walter de Gruyter, 1992.

BRANQUINHO, João [et al.]. *Enciclopédia de termos lógico-filosóficos*. São Paulo: Martins Fontes, 2006.

CHALMERS, David. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, 1996.

CHURCHLAND, Paul. *Matéria e consciência: uma introdução contemporânea à Filosofia da Mente*. Tradução de Maria Clara Cescato. São Paulo: Editora Unesp, 1998.

CORCORAN, K. “The trouble with Searle’s Biological Naturalism”. *Erkenntnis*, 55, pp. 307-324, 2001.

CRANE, Tim. “Review on The Rediscovery of The Mind”. *International Journal of Philosophical Studies*, v. 1, n. 2, pp. 313-323, 1993.

DESCARTES, René. *Meditações*. Tradução de J. Ginsburg e Bento Prado Júnior. São Paulo: Nova Cultural (Coleção Os Pensadores), 1996.

DAVIDSON, Donald. “Mental events”. Reimpresso In: HEIL, John. *Philosophy of Mind: a guide and anthology*. Oxford University Press, 2004.

FUTUYMA, Douglas. *Biologia Evolutiva*. Tradução coordenada por Mario de Vivo. São Paulo: FUNPEC Editora, 2003.

HODGSON, David. “Why Searle Has Not Rediscovered The Mind”. In: *Journal of Consciousness Studies*, 1, n. 2, Winter. pp. 264-74, 1994.

HUME, David. *Tratado da Natureza Humana*. Tradução de Déborah Danowski. São Paulo: Unesp, 2009.

KIM, Jaegwon. “‘Downward Causation’ in Emergentism and Nonreductive Physicalism”. In: BECKERMANN, FLOHR & KIM. *Emergence or Reduction?: Essays on the prospects of Nonreductive Physicalism*. Walter de Gruyter, 1992.

_____. “The myth of nonreductive materialism”. In: _____. *Supervenience and Mind: selected philosophical essays*. Cambridge University Press, 1993a.

KIM, Jaegwon. “The nonreductivist’s troubles with mental causation”. In: _____. *Supervenience and Mind: selected philosophical essays*. Cambridge University Press, 1993b.

_____. “Epiphenomenal and supervenient causation”. In: _____. *Supervenience and Mind: selected philosophical essays*. Cambridge University Press, 1993c.

_____. “Mental Causation in Searle’s Biological Naturalism”. *Philosophy and Phenomenological Research*, v. LV, n. 1, 1995.

_____. *Philosophy of Mind*. Boulder CO: Westview, 1996.

_____. *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. MIT Press, 1998.

_____. *Physicalism or Something Near Enough*. Princeton University Press, 2005a.

_____. “Supervenience, Emergence, Realization, Reduction”. In: LOUX, Michael & ZIMMERMAN, Dean. *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press, 2005b.

MASLIN, K. T. *Introdução à Filosofia da Mente*. Tradução de Fernando Rocha. Porto Alegre: Artmed, 2009.

MATURANA, Humberto & **VARELA**, Francisco. *A Árvore do Conhecimento: as bases biológicas da compreensão humana*. Tradução de Humberto Mariotti e Lia Diskin. São Paulo: Editora Palas Athena, 2005.

MORRISON, Robert & **BOYD**, Robert. *Química Orgânica*. Tradução de M. Alves da Silva. Lisboa: Fundação Calouste Gulbenkian, 1993.

MOURAL, Josef. “The Chinese Room Argument”. In: SMITH, Barry (Ed.). *John Searle*. Cambridge University Press, 2003.

NAGEL, Thomas. “O Fisicalismo”. Tradução de José Antônio Finocchio. In: TEIXEIRA, João de Fernandes (Org.). *Cérebros, Máquinas e Consciência: uma introdução à Filosofia da Mente*. São Carlos: Editora da UFSCar, 1996.

_____. “What is it like to be a bat?”. Reimpresso In: HEIL, John. *Philosophy of Mind: a guide and anthology*. Oxford University Press, 2004.

_____. “Como é ser um morcego?”. Tradução de Paulo Abrantes e Juliana Orionne. In: *Cadernos de História e Filosofia da Ciência*. Série 3, v. 15, n. 1, jan-jun. 2005.

_____. “O Fisicalismo”. Tradução de José Antônio Finocchio. In: TEIXEIRA, João de Fernandes (Org.). *Cérebros, Máquinas e Consciência: uma introdução à Filosofia da Mente*. São Carlos: Editora da UFSCar, 1996.

PEREBOOM, Derk & **KORNBLITH**, Hilary. “The metaphysics of irreducibility”. In: HEIL, John. *Philosophy of Mind: a guide and anthology*. Oxford University Press, 2004. (2004)

PESSIS-PASTERNAK, Guitta. *Do Caos à Inteligência Artificial*. Tradução de Luiz Paulo Rouanet. São Paulo: Editora Unesp, 1992.

PLATÃO. *Fédon (ou da alma)*. In: Platão – Diálogos Socráticos. Tradução de Edson Bini. São Paulo: Edipro, 2008.

POPPER, Karl. *A Lógica da Investigação Científica*. Tradução de Pablo Rubén Mariconda e Paulo de Almeida. São Paulo: Abril Cultural (Coleção Os Pensadores), 1980.

PRATA, Tárík. “Subjetividade Ontológica na Filosofia da Mente de John Searle”. *Philosophos*. v. 12, n. 2, pp. 171-204, jul/dez, 2007.

_____. “Dificuldades da concepção de John Searle sobre a redução da consciência: o problema das capacidades causais”. *Princípios*, v. 15, n. 24, pp. 2-29, jul/dez, 2008.

_____. “Irreduzibilidade ontológica versus identidade: John Searle entre o dualismo e o materialismo”. *O que nos faz Pensar* n. 25, pp. 107-124, 2009.

RIDLEY, Mark. *Evolução*. Tradução de Henrique Ferreira, Luciane Passaglia e Rivo Fischer. Porto Alegre: Artmed, 2006.

RORTY, Richard. “Fisicalismo não-redutivo”. In: _____. *Objetivismo, relativismo e verdade: escritos filosóficos I*. Tradução de Marco Antônio Casanova. Rio de Janeiro: Relume Dumará, 1997.

SCRUTON, Roger. *Uma breve história da Filosofia Moderna: De Descartes a Wittgenstein*. Tradução de Eduardo Francisco Alves. Rio de Janeiro: José Olympio editora, 2008.

SEARLE, John. “Minds, brains, and programs”. *The Behavioral and Brain Sciences*, 3, pp. 417-424, 1980a.

_____. “Intrinsic Intentionality”. *The Behavioral and Brain Sciences*, 3, pp. 417-424, 1980b.

_____. *Mente, Cérebro e Ciência*. Tradução de Artur Morão. Lisboa: Edições 70, 1984.

_____. *Intencionalidade*. Tradução de Julio Fischer e Tomás Rosa Bueno. São Paulo: Martins Fontes, 1995.

_____. “Mentes, cérebros e programas”. Tradução de Cléa Regina de Oliveira Ribeiro. In: TEIXEIRA, João de Fernandes (Org.). *Cérebros, Máquinas e Consciência: uma introdução à Filosofia da Mente*. São Carlos: Editora da UFSCar, 1996.

_____. *A Redescoberta da Mente*. Tradução de Eduardo Pereira Ferreira. São Paulo: Martins Fontes, 1997.

_____. *O Mistério da Consciência*. Tradução de André Yuji Pinheiro Uema e Vladimir Safatle. São Paulo: Paz e Terra, 1998.

SEARLE, John. *Mente, Linguagem e Sociedade: Filosofia no mundo real*. Tradução de F. Rangel. Rio de Janeiro: Rocco, 2000.

_____. “Filosofia contemporânea nos Estados Unidos”. In: BUNNIN, Nicholas & TSUI-JAMES, E. P. *Compêndio de Filosofia*. Tradução de Luiz Paulo Rouanet. São Paulo: Edições Loyola, 2002a.

_____. “Twenty-One Years in the Chinese Room”. In: PRESTON, John & BISHOP, Mark (Eds). *Views into the Chinese Room: New essays on Searle and Artificial Intelligence*. Oxford University Press, 2002b.

_____. “Why I Am Not a Property Dualist”. *Journal of Consciousness Studies*, 9, n. 12, pp. 57-64, 2002c.

_____. *Mind: A Brief Introduction*. Oxford University Press, 2004.

_____. “Dualism revisited”. *Journal of Physiology – Paris*, n. 101, pp. 169-178, 2007a.

_____. *Liberdade e Neurobiologia*. Tradução de Constancia Maria Egrejas Morel. São Paulo: Editora Unesp, 2007b.

_____. “Como Estudar Cientificamente a Consciência”. In: *Consciência e Linguagem*. Tradução de Plínio Junqueira Smith. São Paulo: Martins Fontes, 2010a.

_____. “As Mentes dos Animais”. In: *Consciência e Linguagem*. Tradução de Plínio Junqueira Smith. São Paulo: Martins Fontes, 2010b.

SWINBURNE, Richard. “Uma defesa do dualismo de substâncias”. *Princípios*, v. 15, n. 23, pp. 291-313, jan/jun, 2008.

TEIXEIRA, João de Fernandes. *A mente segundo Dennett*. São Paulo: Perspectiva, 2008.

TURING, Alan. “Computação e Inteligência”. Tradução de Fábio de Carvelho Hansen. In: TEIXEIRA, João de Fernandes (Org.). *Cérebros, Máquinas e Consciência: uma introdução à Filosofia da Mente*. São Carlos: Editora da UFSCar, 1996.

_____. “Computing machinery and intelligence”. Reimpresso In: HEIL, John. *Philosophy of Mind: a guide and anthology*. Oxford University Press, 2004.

WHITBY, Blay. *Inteligência Artificial: um guia para iniciantes*. Tradução de Cláudio Blanc. São Paulo: Madras, 2004.

WITTGENSTEIN, Ludwig. *Investigações Filosóficas*. Tradução de M. S. Lourenço. Lisboa: Fundação Calouste Gulbenkian, 2002.
