



UNIVERSIDADE DE BRASÍLIA UNB
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA

ERIC AUGUSTO CAMARGOS

**A CONSCIÊNCIA REPLICADA:
UMA REFLEXÃO ACERCA DA CONSCIÊNCIA NA FILOSOFIA
DE JOHN SEARLE**

BRASÍLIA DF
2023

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

CC173c Camargos, Eric
A CONSCIÊNCIA REPLICADA: UMA REFLEXÃO
ACERCA DA CONSCIÊNCIA NA FILOSOFIA DE JOHN
SEARLE / Eric Camargos;
orientador André Leclerc. -- Brasília, 2023.

43 p.

Monografia (Graduação - Filosofia) -- Universidade de Brasília,
2023.

1. Consciência. 2. Linguagem. 3. Inteligência Artificial.
4. Ciência. I. Leclerc, André, orient. II. Título.

ERIC AUGUSTO CAMARGOS

A CONSCIÊNCIA REPLICADA:
UMA REFLEXÃO ACERCA DA CONSCIÊNCIA NA FILOSOFIA
DE JOHN SEARLE

Monografia apresentada ao departamento
de filosofia da Universidade de Brasília
como requisito parcial para a obtenção de
título de bacharel/licenciado em Filosofia .

Orientador: Prof. André Leclerc

Comissão avaliadora:

Prof. Dr. André Leclerc (orientador)

Prof. Dr. Alexandre Costa Leite (avaliador)

Brasília, _____ de _____ de 2023

AGRADECIMENTOS

Enquanto dava os toques finais neste trabalho (e dar esse retoques para um cérebro inquieto como o meu foi tarefa mais árdua que escrevê-lo) eu fiquei me perguntando se devia escrever alguma nota de agradecimento e me perguntando como faria isso. A conclusão que eu cheguei pareceu um pouco óbvia (negá-la pra mim era o fruto da minha total relutância) é que seria injusto agradecer a pessoas e situações específicas pela confecção deste trabalho. Ele é um parte ínfima de um todo absolutamente intranquilo e igualmente enriquecedor.

Começo agradecendo principalmente a meus professores, pois sem eles, sem a paciência e a aptidão para orientarem essa jornada, tudo isso seria uma completa perda de tempo, esvaziado de conteúdo e significado. Seria também uma grande tortura, de ser jogado às cegas em meio a tantos pensamentos, textos difíceis, contradições, e inflexões anacrônicas por parte de uma tradição permeada por um senso comum um tanto inconveniente.

Agradeço a meus familiares, minha mãe principalmente, por, desde quando eu era criança, encher a minha casa com livros e revistas e me fazer tomar gosto por esses conteúdos. Certamente sem essa disposição da parte dela eu não teria nenhuma disposição hoje para ultrapassar esse senso comum. Minha tia merece agradecimentos pela força que deu desde que eu saí do interior e vim pra essa cidade conturbada e desde então tem me dado certo suporte, enquanto eu redigia com certa pressa esse trabalho.

Por fim agradeço às dezenas de pessoas que entraram, permaneceram e saíram do meu círculo de convivência. Alguns acompanharam de perto esse empreendimento, testemunharam toda a trajetória, outros nem tanto, mas carregam certa contribuição indireta. Agradeço às pessoas que me incentivaram nesse curso e, dentre essas, agradeço principalmente àquela pessoa, em especial, que reviveu o ânimo que foi abalado pelas peripécias de uma época sombria entre os anos de 2020 e 2022. Quando ela ler, vai entender do que eu estou falando.

Resumo: Esse trabalho monográfico visa explicar, à luz da filosofia de John Searle, as noções que permeiam o conceito de consciência. Para Searle, a consciência é um produto fisiológico, como digestão e respiração, sendo assim suscetível de ser inserida na concepção científica do mundo. Para tal, o filósofo estadunidense pretende demonstrar, por meio de reflexões linguísticas que não há a possibilidade de criar consciência a partir de réplicas, pois as inteligências artificiais operam apenas em nível sintático da linguagem, manipulando símbolos.

Palavras-chave: consciência, cérebro, subjetividade, funcionalismo

Abstract: This monographic work aims to explain, in the light of John Searle's philosophy, the notions that permeate the concept of consciousness. For Searle, consciously it is a physiological product, like digestion and breathing, and is also susceptible to inclusion in the world scientific conception. Therefore, the American philosopher intends to demonstrate, through linguistic reflections, that it is not possible to raise consciousness based on replicas, because artificial intelligence operates only at the syntactic level of language, manipulating symbols.

Key words: consciousness, brain, subjectivity, functionalism

SUMÁRIO

INTRODUÇÃO.....	8
CAPÍTULO 1 O PROBLEMA MENTE- CÉREBRO	11
1.1 Consciência.....	13
1.2 Intencionalidade.....	14
1.3 Subjetividade.....	16
1.4 Causação Mental.....	17
CAPÍTULO 2 CONSIDERAÇÕES ACERCA DA CONSCIÊNCIA.....	21
2.1 A unidade da Consciência	21
2.2 O Estudo Científico da Consciência.....	23
2.3 Irredutibilidade da Consciência	25
2.4 A consciência Replicada.....	27
CAPÍTULO 3 SEARLE CONTRA O FUNCIONALISMO	29
3.1 Inteligência Artificial.....	29
3.2 Considerações Sobre a Linguagem.....	33
CONSIDERAÇÕES FINAIS	38
Podem as Máquinas Pensar?	38
REFERÊNCIAS BIBLIOGRÁFICAS	41

ÍNDICE DE FIGURAS

Figura 1. Ilustração da Causação Mental em Searle	18
Figura 2. Modelo conexionista esquematizado	32

INTRODUÇÃO

“Ela acordou, se sentindo
duas pessoas diferentes,
dois estranhos, dois mudos
que compartilhavam
a mesma pele,
mas que não colaboravam,
nem conversavam entre si.”

Alan Moore

Podem as máquinas pensar? Essa pergunta, que dá título ao segundo capítulo do livro *Mente, Cérebro e Ciência* (Searle; 1984), também dá início a uma problemática sobre a qual John Searle (1932-) dedicou muita atenção, ao longo, não só da obra em questão, mas por toda extensão de sua compilação de ensaios, artigos, e livros sobre a mente e as relações desta com restante do universo. A concepção de mente, mesmo para Searle, cujo conceito se estrutura como um dos pilares centrais da sua filosofia se mostra algo de difícil definição e por essa mesma razão contrapontos a campos como Inteligência Artificial Forte (IA forte), cognitivismo e o velho problema do dualismo cartesiano se fazem necessários.

Sob a luz das explanações do próprio Searle, os problemas e desafios de se pensar a IA forte, se concebem aqui, como o objetivo dessa discussão. Digo objetivo porque as questões paralelas levantadas denotam a mesma importância para a discussão de modo que, falar da crítica que Searle faz às noções de IA forte, embora seja o objetivo, não é o pilar central, tampouco o mais importante. E, para tal, foi imprescindível que, nos capítulos que se seguem, eu me ativesse aos principais pontos abordados por Searle em sua discussão no campo da filosofia da mente.

Desse modo, no primeiro capítulo, eu me disponho a tratar do dualismo mente – cérebro com ponto de partida. No meu artigo denominado *A Inviabilidade do Dualismo Cartesiano na Filosofia da Mente*, eu me fiz valer de uma breve abordagem entre os pontos convergentes entre John Searle e o filósofo britânico Gilbert Ryle (1900-1976) para respaldar a ideia da problemática do dualismo cartesiano como um ponto de partida para se discutir a mente, mais precisamente como um contraponto para alinhar a discussão com a concepção científica do mundo. Nas minhas palavras:

O dualismo de substâncias é o ponto de partida para que Searle aborde a consciência, ainda que o filósofo estadunidense alegue que, ao longo da história da filosofia, deram uma atenção exagerada ao problema. Segundo ele, o dualismo poderia ser facilmente resolvido quando consideramos a neurofisiologia e pensamos o problema da consciência como um fenômeno de mesma ordem fisiológica que a digestão, respiração, circulação etc. (Eric Camargos, 2023, p. 7).

E é a partir da afirmação de que o dualismo cartesiano recebeu atenção exagerada por parte de muitos estudiosos, que Searle, ainda em *Mente, Cérebro e Ciência*, estipula uma pequena lista das quatro características acerca dos estados mentais que tornam difícil sua inserção na concepção científica do mundo. São as quatro características: Consciência, Intencionalidade, Subjetividade, Causação Mental e neste primeiro capítulo pretendo discuti-las de forma detalhada, de modo a enfatizar os pontos que amparam a resposta de Searle à noção de IA forte.

No segundo capítulo, eu volto à noção de consciência, que para Searle é a característica mais importante dos estados mentais (cito essa mesma afirmação enquanto apresento essa característica no primeiro capítulo). Aqui, a discussão adentra na perspectiva searleana acerca da mente, a qual ele vai denominar naturalismo biológico e suas implicações..

Searle questiona por que a consciência, ou qualquer evento mental, não pode ser vista com um problema fisiológico, de ordem causal, assim como a circulação e digestão. Para o filósofo estadunidense, a consciência opera de modo a ser caracterizada por interações de ordem causal por elementos do cérebro, bem como a solidez e a liquidez fazem parte de sistemas de moléculas. (Searle; 2006, p. 164).

Nesse capítulo, portanto, eu me disponho a me aprofundar no naturalismo biológico, no qual temos a consciência humana como instâncias dotadas de características peculiares, dificilmente capazes de serem reproduzidas por qualquer outro meio.

Por fim, no terceiro e último capítulo, apresento as noções que permeiam a expressão Inteligência artificial, ressaltando o capítulo 4 do livro *Inteligência Artificial*, onde a autora Margaret A. Boden (1936-) aborda as *Redes Neurais Artificiais* (justamente o título do capítulo em questão), e o capítulo 6, que pergunta: *Mas será que*

isso é inteligência de verdade?

De acordo com as questões levantadas pela autora, essas dimensões colocam a noção de inteligência artificial sob a égide de uma discussão filosófica, estabelecendo uma abertura para o embate entre o naturalismo biológico de Searle e o funcionalismo de Hilary Putnam (1926-2016) e Daniel Dennett (1942-). Para os funcionalistas, a capacidade que uma máquina tem de reproduzir tarefas ditas como humanas, grosso modo, a torna suscetível a ser caracterizada com um ser consciente, o que pode ser sistematizado na seguinte frase: *a mente está para o cérebro como os programas (softwares) estão para a máquina, isto é, computador ou processador (hardware)* (Leclerc, 2018, p.81).

Esse capítulo retoma a crítica que Searle faz a IA forte e tenta responder, dentro do escopo da linguagem, se máquinas são capazes de pensar, ao mesmo tempo em que empreende suas críticas ao funcionalismo e ao teste de Turing. Aqui, as sutis divergências entre filosofia da mente e filosofia da linguagem tendem a se desfazer. No seu ensaio *O que é a linguagem?* Searle evoca a noção de intencionalidade pré linguística, uma característica que a consciência possui de ser acerca de algo e, desse modo, ser capaz de fazer semântica.

Searle tenta demonstrar, portanto, por meio do experimento do quarto chinês que há um equívoco acerca da IA forte no tocante a se estabelecer consciência. Nesse ponto, a principal divergência entre seres conscientes e o pensamento da máquina é que máquinas operam estritamente em parâmetros sintáticos, por mais que em algumas atividades haja uma funcionalidade mais precisa.¹

¹ Em 1996, o famoso enxadrista Garry Kasparov foi derrotado pelo computador da IBM Deep Blue. O ocorrido ficou marcado como um dos primeiros vislumbres de que a inteligência dos computadores superaria a inteligência humana, pelo menos em nível de sintaxe.

CAPÍTULO 1

O PROBLEMA MENTE-CÉREBRO

O dualismo cartesiano sempre funcionou como um ponto de partida nos estudos filosóficos acerca da mente. Não só isso, grande parte dos filósofos da linguagem no século XX, tomaram essa dicotomia promulgada por Descartes (1596-1650) como uma forma de contraponto, uma vez que a dimensão metafísica do cogito (penso logo existo) por ser uma posição metafísica, apresentava, em seu cerne, problemas de cunho ontológico e causal.

Gilbert Ryle foi um dos primeiros pensadores dessa vertente a destituir do dualismo cartesiano a dimensão científica que até então vinha vigorando, sobretudo nas questões epistemológicas. A filosofia da linguagem no século XX sugeria uma proposta que era, basicamente, reescrever uma grande parte dos problemas filosóficos apontando seus erros de categoria..

Nesse sentido, Ryle aponta as inconsistências ontológicas do dualismo cartesiano como oriundas de um mau uso da linguagem; Descartes abusava de forma equivocada das categorias para estabelecer uma interação da mente imaterial com o cérebro material. Em outras palavras, Descartes é acusado de colocar a mente em um mesmo tipo lógico que o corpo.

Foi a essa incompatibilidade ontológica e linguística que Ryle, portanto, veio batizar pelo chavão “Fantasma na Máquina”. Como poderia, dentro de uma perspectiva científica, se aceitar que uma mente imaterial fosse responsável por conceder ao corpo material e mecânico desejos, razão, emoções?

Ryle, na mesma problematização, faz uma reflexão sobre como o mau uso da linguagem reverberava pelo contexto da época. Segundo ele, Descartes se fez valer de uma transcrição das metafísicas teológicas tradicionais para uma ciência que emergia na época, sob a sintaxe de Galileu. Nem a modernidade, nem o renascimento eram fortes o suficiente na época para desvencilhar os problemas filosóficos por completo da influência (diretas e indiretas) da igreja católica.

O dualismo cartesiano se concebe, portanto, como uma posição metafísica e essa

dimensão metafísica teve seu impacto, sobretudo na validação do sujeito como ser pensante e operante ativo na asserção do conhecimento. No entanto, a partir do século XX, os filósofos da linguagem e da ciência não têm a preocupação de manter essa ligação com a igreja. Na realidade, há uma disposição para se destituir de qualquer dimensão metafísica os problemas filosóficos, sobretudo no tocante a seus equívocos linguísticos, os quais nada mais se mostraram que uma dificuldade para alinhar os estudos da mente e consciência à concepção científica do mundo.

Em Searle, essa preocupação se mostra bem recorrente, tanto que ele aborda o dualismo cartesiano como um ponto de partida nos seus principais livros. A problematização se faz valer sob a alegação de que o vocabulário dos estudos sobre a mente ainda sofre influência da tradição cartesiana. Ou seria o dualismo cartesiano fruto de uma concepção arraigada de mundo, onde se assume que o universo seja inteiramente composto por partículas físicas inconscientes, totalmente redutíveis à física e ainda assim podemos assumir que exista o fenômeno da consciência?

Searle vê que a ciência encontra uma certa dificuldade para se apropriar do assunto quando elabora o seguinte questionamento:

Por que é que temos ainda na Filosofia e na Psicologia, após todos estes séculos, um problema da Mente-Corpo de modo que não temos, digamos, um problema da digestão- estômago? Por que é a mente mais misteriosa do que outros fenômenos biológicos? (Searle, 1984, p.19)

E para responder a essa questão, Searle relembra que a abordagem, ainda que latente, do dualismo cartesiano denuncia uma espécie de anacronismo linguístico, pois o problema ainda é abordado no século XX, sendo estruturado em um vocabulário do século XVII. E como já foi dito antes, o problema da mente- corpo, sob esse vocabulário, carece de respaldo de ordem causal.

Desse modo, Searle expõe brevemente algumas respostas excludentes de uma dimensão ou de outra: por um, lado temos o materialismo negando a existência de atos mentais, bem como o funcionalismo categorizando a mente como uma mera estrutura causal (falarei mais detalhadamente sobre isso no capítulo 3) e, por outro, chegamos a noções radicais como o idealismo de George Berkeley (1685-1753), cuja a existência da realidade depende, imprescindivelmente, da percepção.

Searle, portanto, estipula uma lista das quatro características dos eventos mentais que tornaram difícil a sua inserção na concepção científica do mundo em resposta às dificuldades ocasionadas pelo dualismo mente- corpo. São elas:

1. 1 Consciência

Pilar central a filosofia da mente de Searle, a consciência é uma noção que vai ser explorada em várias dimensões ao longo de sua obra. Entretanto, ainda que toda explanação seja absolutamente profunda e ele nos apresenta uma série de experimentos de pensamento, a consciência se mostra um problema de difícil definição; não obstante se mostra na visão do autor suscetível a uma série de equívocos.

Ainda assim, Searle se dispõe a tentar algumas definições, se referindo à consciência como “estados e processos de sensibilidade e ciência, internos, qualitativos e subjetivos.” (Searle, 2010, p. 55). Dentro dessa definição, é apontada duas características da consciência que é a subjetividade, característica a ser tratada ainda nesse capítulo e a dimensão qualitativa chamada *qualia*.

Pelo conceito de *qualia*, entendemos que a consciência é naturalmente qualitativa, mais precisamente os atos mentais específicos são carregados de qualidades subjetivas intrínsecas a sua representação. Para Searle, isso vale tanto para percepções de coisas como cor, cheiro, sabor, quanto para pensamentos abstratos.

Para Searle, no entanto, a consciência não é uma substância, por isso a dificuldade de se tentar uma definição. A consciência é resultado de um processo biológico, tal como digestão e circulação, embora contenha particularidades que podem ser objetivamente observadas como os demais fenômenos; a subjetividade é uma delas.

O problema da consciência pode ser explicado por interações causais a nível de micro ou macrocosmo, da mesma forma que isolar a consciência reverbera em isolar a digestão do seu sistema que o produz. A consciência, portanto, se expressa como um produto do funcionamento de sistemas emergentes e causais dentro do cérebro.

A importância da consciência se dá no sentido de que é ela quem torna a vida humana possível. Sem a consciência, fenômenos como linguagem, pensamentos,

emoções não seriam possíveis. Todas as atividades ou manifestações mentais só se tornam possíveis a partir de uma relação com a consciência. Em outras palavras, a vida mental está totalmente dependente dessa característica, e por essa razão a falta de atenção por parte de estudiosos das áreas de psicologia e da neurobiologia dificultou em muito os estudos sobre a mente humana.

Como o foco deste trabalho, conforme aponte na discussão, objetiva explorar a crítica que Searle faz à noção de IA forte, é válido destacar como principal equívoco, dentro dessa discussão, a noção funcionalista acerca da mente como um programa de computador. A tese funcionalista e as máquinas Turing destituem da consciência qualquer caráter biológico, a submetendo a uma mera causalidade; isto é para que a mente seja reproduzida e qualquer sistema seja consciente, apenas necessita de inputs e outputs corretos.

1.2 Intencionalidade

A segunda característica dos estados mentais é oriunda do psicologismo de Brentano. Por intencionalidade tomamos a característica que os atos mentais têm de se direcionarem a algo, ou se referem ou são sobre algo, necessariamente. E por mais que o radical da palavra seja oriundo da palavra intenção, intencionalidade não se limita (estritamente) a esse conceito. Se refere também a crenças, desejos, sentimentos, emoções, salvo algumas emoções quando não são direcionadas, como nervosismo, ansiedade.

Há uma distinção pouco clara entre consciência e intencionalidade. Talvez isso se dê pelo fato de que a consciência carece de uma definição embora a tomemos como um estado de quem sente e percebe as qualidades das coisas ao seu redor enquanto está acordado, ao passo que intencionalidade se define por uma capacidade de representar (Leclerc, 2018, p. 32).

No entanto, Searle defende que se pode encontrar uma certa identidade entre consciência e intencionalidade, exemplificando que a consciência é sempre consciência acerca de algo. Ainda assim, alguns eventos mentais apesar de intencionais carecem de consciência, como por exemplo determinadas crenças que operam em estado latente,

sem que se pense nelas, ao passo que determinados eventos mentais conscientes não são intencionais, como, por exemplo, uma dor, que não se direciona a algo além de si mesma.

No livro *Consciência e Linguagem* (Martins Fontes, 2010), Searle dá atenção a um tipo especial de intencionalidade: a intencionalidade intrínseca, a que trata especialmente de estados e eventos mentais que ocorrem e existem apenas no cérebro e mente dos agentes. Nesse ponto, tal como ocorre com a consciência, temos a intencionalidade intrínseca como processos neurofisiológicos que ocorrem no cérebro. E esse tipo de intencionalidade, por estar submetida a uma visão naturalista, denota uma característica de nível superior em um sistema de neurônios operando em nível inferior.

Tomemos como exemplo quando indivíduo *x* afirma ver que está chovendo, o ato de *x* ver já concebe uma condição de satisfação para a intencionalidade do estado mental. Ao afirmar que está vendo algo, o agente está veiculando a intencionalidade por meio da visão, uma vez que o lóbulo occipital, no cérebro, é responsável por determinados processos que vão coadunar em uma intencionalidade intrínseca, assim como outras áreas como o hipocampo responsável pela linguagem. Nesse caso temos a fórmula *x vê p implica p*.

Como foi dito antes, as abordagens naturalistas da mente e, mais especificamente, nesse caso, sempre encontraram certa dificuldade por causa de uma tradição respaldada por um vocabulário arraigado do século XVII, com o problema mente- cérebro. Essas dificuldades reverberam inclusive na relutância por parte dos filósofos da linguagem do Século XX em abordar a linguagem como um fenômeno biológico.

Nesse ponto, Searle alega que a linguagem deve ser vista como uma espécie de extensão natural das capacidades biológicas não linguísticas. Pra esse tópico nos interessa, mais precisamente, a relação da linguagem com a intencionalidade. Ele afirma que a intencionalidade precede a linguagem, de tal modo que bebês e animais também possuem intencionalidade, ainda que sejam incapazes de produzir atos de fala.

Nesse contexto, os atos de fala se concebem como uma representação e essa representação por sua vez, está atrelada a conteúdo intencionais, o que diverge da noção

de representação no campo das inteligências artificiais, no qual opera apenas em formalização estrutural. A intencionalidade, portanto, é a capacidade mental de se dirigir a algo, existente ou não, enquanto esse algo seja dotado de conteúdo semântico, suscetível ou não a formulação sintática. Os atos de fala são, portanto, uma expressa realização de um estado intencional, correspondente.

1.3 Subjetividade

Enquanto a consciência se concebe como a característica principal dos fenômenos mentais, a subjetividade se coloca com a principal característica da consciência. Além disso, nessa discussão, é a característica de mais difícil compreensão, e a característica que mais tornou difícil a adequação dos estudos da mente à concepção científica do mundo.

Por subjetividade, entendemos uma condição ontológica de que meus estados mentais são exclusivamente meus. Eu sinto, minhas dores, minhas alegrias, minha raiva, minhas angústias. Apenas eu tenho a vivência e a percepção desses estados, ao passo que eu não tenho acesso aos mesmos estados enquanto manifestos em outras pessoas. Eu posso ter uma ideia do que a outra pessoa sente quando ela me descreve uma sensação, ou esboça na face um trejeito de dor ou desconforto, ou quando derrama lágrimas de alegria. Mas somente a pessoa está totalmente consciente da vivência desse estado.

A dificuldade reside na distinção entre subjetividade epistêmica e subjetividade ontológica. Enquanto a primeira consiste basicamente em um processo de individuação composto por crenças, opiniões, posicionamentos perante estigmas sociais, a subjetividade ontológica dos estados mentais por ter essa característica absolutamente privativa, foge qualquer análise de cunho objetivo. E por ciência designar, por definição, verdades sistemáticas empíricas e objetivamente observáveis, essa asserção do termo subjetividade ontológica denota certa dificuldade.

Entrementes, Searle reforça que tal subjetividade existe e, tal como a consciência e a intencionalidade, ela é resultado de processos ocorridos no cérebro, uma vez que cada indivíduo, sendo epistemologicamente dotado de subjetividade, é por sua

vez dotado de um cérebro particular. Essa afirmação endossa o caráter subjetivo da realidade, cuja objetividade é perpassada por processos de percepção e linguagem, onde a relação mente para o mundo se faz tão presente quanto a relação do mundo para a mente. Por exemplo, nossa capacidade de nomear as coisas mostra a outra face dessa moeda, pois o mundo só é mundo enquanto percebido, compreendido e catalogado de acordo com as nossas necessidades.

A subjetividade ontológica tem sido ignorada não só por decorrência da tradição dualista, mas também por parte do behaviorismo, cuja defesa se baseia na ideia de que certos fenômenos mentais com crenças e desejos podem ser condicionados meramente a processos comportamentais e por experiência direta, e também por parte do funcionalismo, cuja visão de causa/efeito, input/output, reduz a mente a meros processos computacionais, desprovidos de qualquer qualidade. Em todos esses casos, se vê uma disposição em tratar a consciência fora da consciência, sem levar em conta a subjetividade e reduzindo assim a um objeto a ser estudado em terceira pessoa.

1.4 Causação Mental

A quarta e última característica dos fenômenos mentais nessa discussão se dá pelo seguinte questionamento: *Como é que, por exemplo, algo imponderável e etéreo como o pensamento pode causar uma ação?* (Searle; 1984, p. 34). A causação mental consiste basicamente em estabelecer essa ligação entre o que é, à primeira vista, etéreo e imaterial com fenômenos físicos. Como o pensamento de que eu preciso me levantar da cadeira, me leva a levantar da cadeira?

Em primeiro lugar, Searle procura demover o equívoco acerca do mental, respondendo que não há essa dimensão etérea circundando o pensamento ou qualquer outra atividade mental. Aqui, Searle evoca novamente a argumentação de que os eventos mentais são resultado de processos cerebrais. Os pensamentos, portanto, decorrem de uma atividade cerebral. Quando penso em levantar a mão, meu cérebro emite disparos neuronais como forma de mensagem, impulsos elétricos que me impelem a levantar a mão.

Searle alega, sob essa perspectiva, que há dois níveis a serem explorados:

a) nível inferior em termos fisiológicos, onde ocorrem as interações e as atividades neuronais, responsáveis pelo estado mental, tal como as moléculas de água que interagem para criar um estado líquido ou sólido;

b) nível superior em termos mentais, no qual o estado mental, a consciência, o pensamento, podem causar determinadas ocorrências, tal qual o estado sólido da água é uma propriedade pela qual se realiza determinadas funções.

Desse modo, as interações causais de nível inferior, são responsáveis por causar propriedades mentais de nível superior.

Para reforçar a ideia da causação presente nos fenômenos mentais, Searle estipula que há uma causação intencional nesses mesmos fenômenos. Ele alega que, em essência, o próprio estado intencional tem função causal na produção de suas condições de satisfação, ou, alternativamente, suas condições de satisfação têm função causal na sua produção. Desse modo temos a própria representação enquanto representação que reproduz o que representa, ao passo que o objeto representado causa a produção da representação. (Searle, 2010; p. 134)

Tomemos como exemplo o desejo de tomar água, cujo conteúdo para a satisfação do desejo de tomar água seja “que eu tome água”. Então o estado mental “que eu tome água” sugere que eu tome água para que o desejo seja satisfeito. A causa se concebe, portanto, como uma representação daquilo que ela causa.

Figura 1

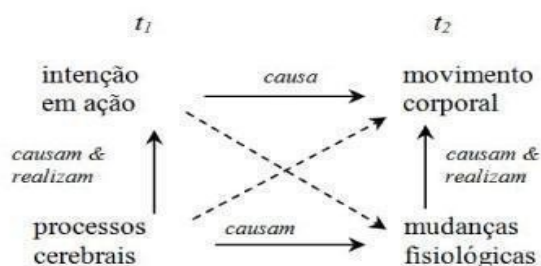


Figura 1. Ilustração da Causação mental em Searle. (Searle, 2002, p.374)

Essa pequena lista nos aponta alguns pontos a serem considerados, no que diz respeito à ontologia dos estados mentais e suas divergências com as noções

funcionalistas acerca da mente, mais especificamente à defesa da Inteligência artificial forte, conforme veremos com mais detalhes no capítulo 3.

O primeiro ponto é o caráter biológico dos fenômenos mentais. A filosofia da mente de Searle tem como propósito naturalizar o conceito de mente, a fim de inseri-lo na concepção científica do mundo e tratá-lo como fenômeno biológico. Para Searle, está evidente que a consciência existe e pode sim ser estudada de modo objetivo; para tanto precisa ser encarada com uma instância central para a discussão e ter seu caráter ontológico considerado.

Dada as vagas definições que o conceito de consciência carrega (pois se trata de algo de difícil definição, uma vez que não se trata de uma substância e nem mesmo a neurociência conseguiu definir), podemos entender, a partir do naturalismo biológico de Searle, que se trata de um evento mental dotado de qualidades (*qualia*). A mente se concebe, nessa discussão, portanto como um fenômeno biológico, um processo que resulta a nível superior a partir de interações causais a nível inferior fisiológico, tal como, digestão, circulação, solidificação.

O segundo ponto é que os estudos sobre a mente sofrem pela tradição arraigada do dualismo mente-cérebro. A discussão se torna difícil a partir dessa perspectiva e, enquanto alguns permanecem no equívoco, outros como os behavioristas lógicos em vez de tratarem o problema, tendem a anular a existência da mente relegando aos fenômenos mentais a mera característica de algo resultado de comportamentos externos; tal como adeptos do materialismo eliminativo, que excluem qualquer concepção de mente; e por fim com os funcionalistas, relegando à mente mero produto de relações de causa e efeito.

Searle entende como necessário a abandono do dualismo cartesiano, paralelo a abordagem da mente como um produto do cérebro, abordagem essa que difere do abandono da mente junto com o dualismo, tal qual sugerem os behavioristas, materialistas e funcionalistas. Só assim o problema do dualismo estaria realmente superado e os estudos da mente avançariam nos campos da neurobiologia e da psicologia.

O terceiro ponto suscitado a partir dessa lista é a relação entre linguagem e

intencionalidade. Essa relação se faz necessária enquanto objetivamos estabelecer uma distinção entre pensamento humano e pensamento de máquina. Essa distinção por sua vez se expressa pela noção de representação.

Tal como a consciência, Searle alega que há uma dimensão naturalista acerca da linguagem; ela também é um fenômeno natural, mais precisamente como uma extensão natural de capacidades biológicas não linguísticas (Searle, 2012, p. 17). Com isso ele quer dizer que os atos de fala por serem representações linguísticas carregadas de significado intencional. Essas representações, por terem conteúdos semânticos divergem das representações no campo das inteligências artificiais: as representações intencionais se concebem por seu conteúdo, ao passo que representações no campo da inteligência artificial operam apenas por estrutura formal.

CAPÍTULO 2

CONSIDERAÇÕES ACERCA DA CONSCIÊNCIA

2.1 A Unidade da Consciência

No capítulo anterior, aponte algumas características da consciência: ela é, de acordo com a filosofia de Searle, um fenômeno biológico, tal qual digestão, circulação; ela é subjetiva e privativa, de modo que somente eu posso ter acesso a meus estados e, da mesma forma, não posso ter acesso aos estados dos outros; ela é qualitativa de modo que cada experiência consciente é dotada de determinado conteúdo e produz determinada impressão.

Foi dito, também, que a subjetividade da consciência incorre em uma dificuldade para que a ciência, que se trata de empreendimentos, em sua grande maioria, empíricos, objetivos e, conseqüentemente, observáveis, aborde e se aproprie do problema da consciência. E em resposta, Searle argumenta que subjetividade, que é notadamente ontológica e não epistêmica, ser o que difere a consciência dos demais fenômenos biológicos, justamente por ser ontológica confere à consciência o status de ser estudada cientificamente. O naturalismo biológico de Searle propõe uma abertura de caminho para que essa ideia se torne uma realidade considerável.

Uma outra característica da consciência da qual não falei ainda é a unidade. Para Searle, todas as experiências fazem parte de um campo consciente unificado. (Searle, 2010, p.59). Enquanto estamos presenciando uma paisagem, temos a consciência de que estamos em pé ou sentados, temos ciência dos barulhos ao redor, dos pensamentos que surgem. Desse modo:

“Experimento todas essas coisas como parte de um único campo consciente unificado. Essa unidade de todo estado de subjetividade qualitativa tem importantes conseqüências para um estudo científico da consciência. (...). Se tentarmos imaginar que meu estado de consciência está dividido em dezessete partes, o que se imagina não é um único sujeito, mas dezessete diferentes centros de consciência. Em suma, um estado consciente é, por definição, unificado e a unidade decorrerá da subjetividade e da qualidade, porque não é possível ter subjetividade e qualidade, exceto nessa forma particular de unidade. (Searle,

2010, p.60)

Nesse contexto, Searle aponta duas áreas de pesquisa para demonstrar a importância de se entender essa concepção de unidade acerca da mente:

A primeira se trata do experimento dos “dois cérebros”, realizado em 1998 pelo psicólogo estadunidense Mike Gazzaniga (1939-). O experimento em questão foi feito em pacientes que haviam sofrido uma comissurotomia, que consiste basicamente em uma cisão entre os dois hemisférios do cérebro, geralmente feita em pessoas acometidas por epilepsia. O que Gazzaniga observou foi que houve uma imperfeita comunicação entre esses dois hemisférios e que cada um demonstrou ser um centro de consciência distinto do outro. Era como se houvesse de fato dois cérebros existentes e operantes dentro do crânio.

O segundo trata do problema da integração. Nesse caso, os eventos mentais são causados por uma compilação de estímulos que são distintos por análise fisiológica, mas convergem em uma espécie de unificação. Searle cita uma hipótese em que os correlatos neurais da consciência², são produzidos por disparos neuronais que se alinham em uma sincronia de 40 hertz de frequência. Essas ações ocorreriam, dentro dessa perspectiva, em conexões entre o tálamo e algumas camadas específicas do córtex. (Searle, 2010; p.61).

Com base na unidade da consciência, Searle propõe a teoria do campo unificado, em resposta à teoria dos blocos, predominante na neurociência. Segundo a teoria dos blocos, o campo da consciência é construído pela interação de micro- consciências. Sendo assim, cada estado consciente é causado pela predominância de NCCs específicos (correlatos da visão são responsáveis pela criação de consciência visual).

A teoria do campo unificado opera de modo inverso. Ela pressupõe uma consciência prévia e unificada onde há um deslocamento da atividade subjetiva e da atenção como modificação do bloco. Ou seja, para que o indivíduo tenha uma experiência visual, ou auditiva, ou experimente um cheiro ele precisa ser consciente.

2 Correlatos neurais dos estados conscientes, ou os NCCs (em inglês neurobiological correlates of conscious states) são atividades neuronais responsáveis por experiências perceptivas específicas de consciência. Por exemplo, atividades neuronais no nervo óptico são correlatos responsáveis pela experiência consciente da visão.

Para Searle, se entendermos que a consciência é um produto em macro- nível das atividades cerebrais, e se for tomada como pressuposto unificado, a inserção do seu estudo na concepção científica do mundo se torna possível. Correlatos neurais isolados, embora sejam pertinentes para a expressão de estados conscientes, não se concebem como detentores de consciência específica e isolada, a menos que resultem do experimento dos dois cérebros.

2.2 O Estudo Científico da Consciência

Searle dedicou o quarto capítulo do livro *A Redescoberta da Mente* (Martins Fontes, 2006) para empreender uma tentativa de situar a consciência na concepção científica do mundo. Nesse mesmo capítulo, ele volta a ressaltar a importância da consciência como pilar central dos fenômenos mentais, bem como ressalta a já comentada dificuldade que o conceito de consciência sofre para receber uma definição.

Tomemos a consciência, então, a partir dos seguintes exemplos para ilustrá-la:

Quando acordo de um sono sem sonhos, entro num estado de consciência, um estado que permanece enquanto estiver acordado. Quando vou dormir, ou sou colocado sob uma anestesia geral, o morro, meus estados conscientes cessam. Se durante o sono tenho sonhos, torno-me consciente, embora formas oníricas de um nível muito baixo de intensidade e vividez do que a consciência desperta ordinária. (Searle; 2006; p.124)

Os exemplos citados apontam uma existência fenomenal da consciência; ela é um estado cuja existência ontológica é apreendida pela sua manifestação. Ou seja, a consciência é qualitativa e sempre acerca de algo. Ela não é uma substância e não admite uma definição. Ainda assim, ela é um estado cuja existência é atestada por suas manifestações. Além disso, é sabido que ela resulta de processos cerebrais, mais especificamente de interações de ordem causal entre neurônios.

Desse modo, como situar a consciência em uma concepção científica do mundo? Como pode haver, em um universo intrinsecamente formado por partículas físicas fenômenos mentais irreduzíveis? Para responder a essa pergunta, Searle recorre a dois aspectos daquilo que ele vai chamar de visão científica do mundo: uma delas é a teoria

atômica da matéria e a outra a teoria evolutiva da biologia.

A teoria atômica da matéria consiste, basicamente, na ideia de que o universo é composto por partículas e essas partículas interagem entre si para formar sistemas. Essas interações se concebem como subsistemas, pelos quais os eventos em nível macro podem ser explicados em nível micro.

As moléculas de água, por exemplo, interagem causalmente entre si para promoverem um estado líquido; as moléculas constituem o micro ao passo que a água no estado líquido, é o macro. Isso vale para outros fenômenos mais complexos. Segundo essa teoria, tudo no universo são sistemas compostos a partir de partículas que por sua vez são compostos por partículas menores organizadas em subsistemas. E essas escalas de interação não se dão apenas em nível de constituição, mas pela teoria atômica da matéria, o macro pode ser entendido no nível dos microsistemas.

Quanto à teoria evolutiva, que consiste basicamente na ideia de que determinados sistemas precisam evoluir de maneiras distintas, Searle alega há sobreposições de ocorrências³, onde uma nova ocorrência surge e a anterior é deixada pra trás ou anulada. A ocorrência que vigora é a que tem mais chances de sobrevivência e tem uma tendência a reproduzir ocorrências da mesma estirpe, com uma certa aptidão natural para sobrevivência.

Como resultado desses processos evolutivos, determinados organismos se constituem a partir de organizações de subsistemas denominados células. Nas palavras do próprio Searle:

*(...), alguns desses organismos desenvolvem subsistemas de células nervosas, que consideramos como “sistemas nervosos”.
(...), alguns sistemas nervosos extremamente complexos são capazes de causar e sustentar estados e processos conscientes.”*
(Searle; 2006; p.132)

Se assume, nesse sentido, que os estados mentais se concebem como resultados

em macronível, decorrentes de operações causais por micro- sistemas ou células. Essas operações ocorrem dentro do cérebro, este por sua vez, produto de milhões de anos de evolução. Características decorrentes dos processos cerebrais, como a consciência e linguagem, são fenômenos biológicos que expressam os aspectos dessa evolução. Não se trata aqui, porém, de uma evolução superior, uma vez que as demais espécies são submetidas a outros tipos de sobreposições de ocorrências.

2.3 A Irredutibilidade da Consciência

Para entendermos esse tópico, é necessário nos atemos ao caráter ontológico que circunda a consciência e, mais precisamente, a subjetividade. Conforme discutido no tópico anterior, o estudo da consciência, para Searle deve ser submetido a teoria atômica da matéria, juntamente com a teoria evolutiva para ser inserida na concepção científica do mundo. Essas teorias conjuntas reforçam o caráter biológico da consciência que é resultado de interações de subsistemas dentro do cérebro.

Entretanto, quando pensamos nessas teorias, principalmente quando falamos em situar algo dentro de uma concepção científica do mundo, automaticamente somos levados a pensar nessas teorias em caráter epistêmico. Desse modo, se levarmos em conta o eventual caráter epistêmico da consciência, deixamos de lado sua característica importante que é a subjetividade ontológica e esse é o cerne do problema.

O reducionismo é uma ideia de identidade, onde determinadas coisas nada são, exceto outros tipos de coisa. Pode ser expresso tanto por “nada- exceto” ou por *A's* podem ser reduzidos a *B's* sse *A's* nada são exceto *B's*. Um exemplo disso é a teoria da identidade mente- cérebro. Para os materialistas, a mente está no cérebro, ou mais especificamente a mente é o cérebro de modo que os estados mentais são apenas processos dos neurônios. A esse tipo de redução damos o nome de redução ontológica de propriedades. Enquanto que por redução ontológica temos a ideia de que certos tipos de objetos são nada mais que outros tipos de objetos (livros são compilações de moléculas), a especificação de propriedades tende a demonstrar que um objeto é idêntico a suas particularidades (um estado mental nada mais é que um evento neural).

A redução ontológica (e/ou de propriedades) representa uma dificuldade para o caráter qualitativo e subjetivo da mente. Se tomarmos, por exemplo, a dor meramente como um processo neurofisiológico observável em terceira pessoa, as características essenciais que qualificam a dor como algo mental seriam anuladas. Haveria, nesse sentido, na visão de Searle, um comprometimento na comunicação com a subjetividade ontológica. Eu poderia, por exemplo, saber o que é uma dor por ter um entendimento de neurofisiologia (terceira pessoa), sem, contudo, ter experienciado uma dor (primeira pessoa).

A teoria da irreducibilidade encontra certa dificuldade quando confrontada com o materialismo eliminativo, que tem como principais representantes Patrícia e Paul Churchlands. Esse tipo de materialismo tem como ponto de partida a negação do dualismo cartesiano e desse modo, a negação da mente. Para o materialismo, não é o caso que exista uma mente qualitativa. Os estados mentais, como percepções, dores, sentimentos carecem de referência e não encontram extensão em um campo ontológico. (Leclerc, 2018, p. 99) Desse modo a irreducibilidade da consciência denota uma outra espécie de dualismo.

Vemos uma disposição parecida, por parte do instrumentalismo de Daniel Dennett, onde há uma negação da importância de se entender a consciência nos moldes de uma ontologia de primeira pessoa. Para Dennett, a validade das teorias científicas não carecem de verdades lógicas; elas precisam ser funcionais, definidas por suas utilidades. Nesse sentido, os estados mentais são concebidos apenas por uma espécie de conveniência, uma forma utilitária, uma espécie de faz de conta funcional.

Para ilustrar isso, Dennett recorre ao exemplo do ato de jogar xadrez com uma máquina. Do ponto de vista prático dos usos da inteligência, não há diferença entre uma máquina e um ser humano. Quando jogamos xadrez com uma máquina pressupomos determinados eventos mentais, de cunho estratégico para que o jogo ocorra. O fato de o jogo ocorrer, entretanto, não assegura que haja uma validade lógica em termos de verdade e falsidade acerca dos eventos mentais.

Em resposta à tese materialista que a mente irreducível é um aspecto de dualismos, Searle aponta para a trivialidade da irreducibilidade. Segundo essa característica, a irreducibilidade não carregaria o pretense peso científico ou metafísico

que a mente cartesiana carregava. Ou seja, ela a irredutibilidade da consciência não incorre em nenhum prejuízo científico se admita sua existência de modo ontológico. Para Searle a irredutibilidade seria, por questões de subjetividade e qualidade (qualia), concernentes aos modelos de reduções já conhecidas no campo da ciência, como por exemplo nos fenômenos biológicos da respiração, a solidez da água

2.4 A Consciência Replicada

Até aqui foram ditas e repetidas as propriedades ontológicas e naturalistas da consciência. Adversamente a noção materialista de que mentes são cérebros, temos, pelo naturalismo biológico “cérebros causam mentes” expressão que denota esse caráter natural, biológico e causal da consciência.

Não obstante, há um esforço por parte dos estudos no campo da inteligência artificial para que a consciência seja replicada. A metáfora proposicional por parte dos funcionalistas “a mente está para o cérebro como os programas está para os computadores” expressa um certo respaldo para esse esforço. Falarei sobre isso no próximo capítulo. Nesse ponto, nos interessa apontar como há uma dificuldade por qualquer campo da ciência em remodelar a consciência.

O que nos interessa neste tópico são dois pontos. O primeiro é o caráter biológico da consciência e seu fundamento da causalidade neurobiológica. Como já disse incansavelmente nos tópicos acima, a consciência é causada no cérebro por interações causais neuronais de nível micro para macro. Essa característica por si só já torna extremamente difícil sua replicação, uma vez que se equiparamos a inteligência artificial à consciência humana, estamos assumindo que basta que qualquer sistema, dotado dos inputs e outputs corretos é o suficiente para se conceber uma mente. Nesse ponto, de certa forma estamos negando a existência da consciência.

No entanto, já assumimos que exista uma existência qualitativa acerca da consciência. Então, desse modo reduzir os processos de consciência a meramente causalidade funcional, sob a alegação de que, se algo parece pensar, então esse algo pensa, denuncia que a própria inteligência artificial peca ao estabelecer os fundamentos

dessa causalidade. Ela carece, portanto, do fundamento neuronal presente na consciência.

Análogo a isso, temos o segundo ponto que estabelece uma divergência entre a consciência humana e inteligência artificial. Se trata da intencionalidade e consequentemente das capacidades linguísticas da consciência. Mais uma vez, como mais detalhadamente nos próximos capítulos, a inteligência artificial falha em estabelecer fundamentos, neste caso, no campo da linguagem. Por intencionalidade se assume que os eventos genuinamente mentais são dotados de significado e sendo assim, fazem semântica, ao passo que as máquinas operam de modo restrito a manipulação de símbolos sintáticos.

Os fundamentos neuronais e as capacidades linguísticas intencionais são o que tornam, portanto, a consciência um evento neurobiológico, que se restringe aos humanos e, até certos pontos, a alguns animais. Afinal as capacidades meramente sintáticas não dão conta de aspectos privativos como subjetividade e intencionalidade. Esses aspectos e suas divergências com relação às inteligências artificiais é que dão norte para esse trabalho.

CAPÍTULO 3

SEARLE CONTRA O FUNCIONALISMO

3.1 Inteligência Artificial

Em seu artigo intitulado *Mentes, Cérebros e Programas* (1981), Searle inicia sua problematização expondo dois tipos de Inteligência artificial : a IA fraca, cuja função é tecnológica e instrumental e a IA forte caracterizada como um dispositivo, cujas pretensões são de replicar a mente, de modo que haja uma identificação da mente com programas de computador. Se, por definição, entendemos inteligência artificial como dispositivos que almejam auxiliar no desempenho de funções que dependem de inteligência humana, muitas vezes com nível de precisão maior, a IA fraca, na concepção de Searle, não apresenta pontos altamente refutáveis, pelo menos não para o trabalho em questão. Ao passo que as noções de IA forte, é o principal alvo da crítica do artigo em questão.

Antes de nos aprofundarmos na crítica, porém, é importante situar as conceituações e históricos que permeiam a inteligência artificial. Estima-se que a IA tenha seu começo por volta do ano 1840 nas mãos de Ada Lovelace (1815-1852) e Charles Babbage (1791-1871). O empreendimento em questão se limitava a questões lógicas, simbólicas e matemáticas. Ao contrário do que os adeptos do legado de Turing defendem posteriormente, o objetivo da máquina de Lovelace era puramente tecnológico.

Alan Turing (1912-1954) por sua vez, além da atenção aos objetivos tecnológicos que exigiriam inteligência, foi um dos expoentes da inteligência artificial que Searle chamaria de forte, ou seja, da inteligência artificial que visava remodelar a mente. Ele foi também responsável por abordagens da IA como as redes neurais (RNA) e abordagem evolutiva.

As redes neurais variam com relação aos princípios da GOFAI⁴ que é basicamente a sigla para se referir à inteligência artificial clássica. Enquanto a GOFAI visa uma instrumentalidade, as redes neurais se expressam como uma espécie de réplica

4 Good Old-Fashioned AI, a boa e velha IA)(Boden, 2020, p.19).

do modelo de interação neural no cérebro humano. Trata-se, nas palavras de Margaret A. Boden, de uma composição de muitas unidades conectadas capazes de processar apenas uma coisa (A. Boden, 2020, p. 111).

Nesse ponto, identificamos uma divergência entre as RNAs e o naturalismo de Searle, sobretudo, quando levamos em conta a unidade da consciência. As RNAs são compostas por unidades que operam em paralelo, como se fossem réplicas de correlatos neurais, mas sem o aparato central que seria a consciência.

As RNAs apresentam uma inclinação funcional no que diz respeito a seus objetivos. Um deles é o aprendizado. Entrementes, deixando de lado as explanações de cunho tecnológico, vemos um paralelo entre as RNAs e a metáfora proposicional. Para os funcionalistas a mente é uma espécie de programa de computador, e os estados mentais dependem de relações causais de entrada (input), processamento interno e saída ou resultado (output). Desse modo, nas palavras de André Leclerc:

“Como o materialista e o behaviorista, o funcionalista estima que o cérebro tem uma importância decisiva na geração de comportamento inteligente; e como o behaviorista, ele também fala de entrada e saída, input, output, ou estímulos e respostas. Para o funcionalista, um conceito mentalista é algo que corresponde a um intermediário causal entre a excitação dos órgãos sensoriais (entrada) e o comportamento público (saída).” (Leclerc, 2018, p. 83)

O célebre teste de Turing não só ilustra bem a metáfora proposicional como serviu de embasamento para ela. Segundo o teste, há uma relação entre três participantes, um deles é uma máquina. Um dos participantes, enquanto interroga a máquina e o outro participante, precisa identificar quem é quem. Caso ele não consiga fazer a distinção, então a máquina é dotada de inteligência humana.

Esse teste confere certo respaldo à negação da mente tal como Searle concebe. Se para os funcionalistas, as validades científicas são qualificadas por sua utilidade, tal como defende o instrumentalismo de Dennett, então basta que uma máquina consiga replicar as funções humanas para que seja dotada de consciência.

Em seu livro *The Consciousness Explained*, Dennett rejeita a ideia de uma subjetividade ontológica acerca dos fenômenos mentais. Para ele, não é o caso que

existam dados, uma consciência qualitativa. Ele procura demonstrar que determinados estados mentais, como a dor provocada por um beliscão, estão submetidos a experiências sensoriais de input e output e são observáveis em terceira pessoa como qualquer outro fenômeno. Segundo Dennett, os processos que circundam os nossos estados mentais são os mesmos que uma máquina pode sofrer. Essa estrutura causal ocorre pela seguinte maneira:

- a) input de estímulos (pressão do beliscão);
- b) estados discriminativos (processamento neurofisiológico);
- c) disposições reativas (dor/desconforto).

Adversa à noção de mente qualitativa, Dennett estipula o modelo de “Esboços múltiplos”. Essa ideia tem como ponto de partida a crítica que ele faz ao que ele chama de “Teatro Cartesiano”, ideia que por sua vez nos leva a crer que há no cérebro um lugar especial para a consciência. Nesse sentido, os esboços se concebem como uma série de estados agindo dentro do cérebro. Essa teoria se respalda no fato de que, dentro do cérebro, determinados órgãos exprimem uma conexão necessária com outros órgãos. Cada lobo tem seu próprio tálamo. (Searle, 1997, p.121).

Segundo Searle, o objetivo da ideia dos “Esboços múltiplos”, ao criticar o teatro cartesiano”, é negar a existência da subjetividade ontológica acerca dos estados mentais, uma vez que para a neurobiologia, é implausível conceber essa instância acerca dos estados mentais de maneira objetiva. Dentro dessa discussão, é possível estabelecer um paralelo com o materialismo eliminativo, de tal modo que, para negar o cartesianismo, deve-se negar também o conceito de mente. Por ser um funcionalista, pouco importa, para Dennett, essas noções e características acerca da mente, uma vez que o que vale na discussão são as implicações práticas da inteligência, de forma que *qualia* não é um termo interessante.

Desse modo, Dennett se mostra um defensor da IA forte. Sua noção de mente parte de quatro noções que circundam os debates acerca da inteligência artificial. São elas:

- a) máquinas de von Neumann: se trata de computadores seriais, com capacidade de processamento de milhões de passos computacionais por segundo, operando em

paralelo e ao mesmo tempo interagindo uns com os outros. Fisicamente mais parecidos com a arquitetura do cérebro;

b) conexionismo: se trata de um modelo neuronal, uma rede onde ocorrem múltiplas interações, a partir dos dados inseridos na máquina (input), para produzir outputs mais potentes e completos. Essa capacidade de processamento, em outras palavras, assegura um aprendizado de máquinas mais preciso;

c) máquinas virtuais: enquanto computadores físicos são compostos por fios, transistores, metal ou qualquer outro material, as máquinas virtuais se concebem como programas e sistemas de processamento, criados por um programador;

d) meme: termo cunhado por Richard Dawkins, pode ser entendido como uma proximidade significativa ao termo gene na biologia. Os memes por sua vez seriam propagados de cérebro em cérebro por uma espécie de imitação, ocorrida em âmbito cultural e linguístico.

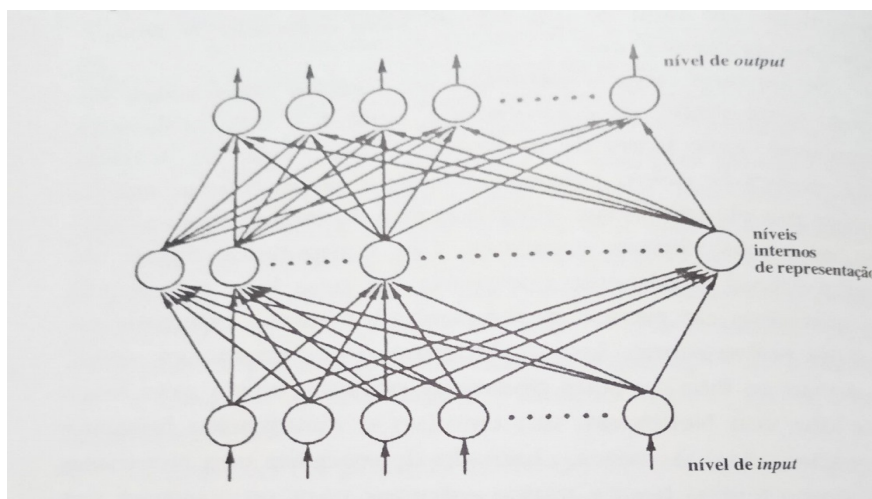


Figura 2. Modelo conexionista esquematizado. Cada unidade se conecta com todas as outras unidades do nível superior. (Searle, 1998, p. 122)

Em resumo, os estados conscientes seriam para Dennett respostas a programações em uma máquina paralela que evolui naturalmente e por meio de memes.

Em outras palavras, mentes são programas remodeláveis, tal como computadores e máquinas virtuais. Dennett vai ainda mais longe na sua negação da consciência ao afirmar que pessoas conscientes não diferem se zumbis e de máquinas pelo simples fato de que, objetivamente falando, não há nenhuma consciência qualitativa em nenhum dos casos, principalmente se um zumbi ou um robô desempenharem uma mesma atividade que uma pessoa. Não é o caso de que seja necessário, na visão de Dennett, uma consciência qualitativa.

3.2 Considerações Sobre a Linguagem

Searle, dedicou parte do seu trabalho como filósofo da mente e, inevitavelmente, da linguagem para se fazer compreender as divergências, que a princípio se mostram sutis (pelo menos para os parâmetros do funcionalismo), entre a inteligência humana e a inteligência artificial, mas que no fim se mostram um tanto complexas e equivocadas. O experimento de pensamento do quarto chinês serve para mostrar a falha na defesa, por parte dos funcionalistas, de que basta que uma máquina module, por meio de uma estrutura de processamento, para que seja dotada de consciência.

Imagine a seguinte situação: um homem muito competente na arte de manipular símbolos é trancado dentro de um quarto, isolado do mundo exterior. Em cada uma das extremidades opostas desse quarto, há uma portinhola e a primeira delas é usada para inserir um texto em chinês. O homem não sabe falar chinês, mas logo em seguida é inserido pela mesma portinhola uma tabela com instruções em inglês para que esse homem possa, a partir do texto inicial em chinês, formular sentenças em chinês. Esse homem então, tem certo êxito nisso e, após um tempo, ele está respondendo às eventuais perguntas feitas em chinês que entram pela primeira portinhola como se entendesse chinês e emitindo essas respostas pela outra portinhola.

O experimento do quarto chinês ilustra um exemplo de estrutura causal, defendida pelos funcionalistas e, dentro dessa mesma estrutura, as contradições que circundam uma noção de consciência por parte da IA forte. Para os funcionalistas, basta que uma máquina module um cérebro e, dispondo dos programas adequados, pode-se dizer que essa máquina possui uma mente.

Submetendo o experimento a parâmetros linguísticos, temos uma asserção um pouco diferente, uma vez que a consciência, subjetiva e qualitativa, tal qual foi explicada ao longo desse trabalho, é uma instância negada por parte dos funcionalistas. O instrumentalismo por parte de Daniel Dennett é um claro exemplo disso. Dessa forma, enquanto os funcionalistas defendem a estrutura causal como suficiente para mostrar que os computadores podem pensar, o naturalismo biológico, como ponto de partida mostra que há mais coisas a se considerar antes de fazer afirmações acerca do assunto.

A primeira coisa que devemos observar por parte do indivíduo no quarto chinês é que ele não entende chinês. Por mais que ele tenha uma aptidão na manipulação de símbolos em chinês, por mais que ele responda em chinês a perguntas feitas em chinês, essas respostas não ocorrem mediante asserção de significado. Sua atividade está amparada pela lista de instruções na sua língua, no caso do experimento, em inglês, no qual ele entende o significado, mas os significados não são diretamente interessantes para as respostas em chinês. Em outras palavras, o homem no interior do quarto chinês, por estar apenas manipulando símbolos vazios de significado, ele está fazendo sintaxe.

Por muito tempo, se acreditou que, na ciência da computação, a sintaxe era o suficiente para compreender a linguagem concernente às distintas máquinas. O SHRDLU de Terry Winograd foi criado como uma forma de sofisticar a PLN⁵. No começo da IA, a linguagem era um tanto quanto tacanha, principalmente quando se levava em conta as estruturações de texto e capacidade gramatical. Se apostou então no aprimoramento das capacidades sintáticas das máquinas, a fim que esses problemas de linguagem fossem resolvidos.

Entretanto, por questões de cunho tecnológico, o SHRDLU e outros programas sintáticos foram abandonados. A sintaxe já se mostrava insuficiente para a exploração de uma linguagem efetiva, ou seja, algo que expressasse determinados significados. O PLN, desse modo, passou a se voltar para dispositivos de cunho pretensamente semântico e contextual, deixando de dar atenção para palavras isoladas, conforme o desempenho da análise sintática, para focar em grupos de palavras e até mesmo período.

Desse modo, a compreensão por parte da máquina foi se apurando à medida que a tecnologia ia avançando e a própria capacidade sintática foi aprimorada.

No entanto, essa “compreensão”, na visão de Searle, se mostra muito mais um dispositivo funcionalista do que uma compreensão de fato. O mesmo com a noção de semântica. Por semântica, entendemos o ramo da linguística que se refere à propriedade das línguas naturais serem sobre algo e acerca de algo. Trata dos significados. E, por definição, no escopo desse trabalho, se concebe como uma propriedade restrita a seres humanos.

Desse modo, já sabemos da dimensão natural da linguagem humana. A linguagem se concebe como uma extensão de capacidades intencionais pré – linguísticas, propriedade presente também em animais⁶. Os atos de fala, por sua vez, são representações um tanto diretas dessas capacidades intencionais. E a semântica por sua vez, por cuidar dos significados, está muito mais atrelada a essas capacidades intencionais e não pode ser amparada por análise sintática.

A compreensão das máquinas, portanto, não tem o mesmo sentido que a compreensão humana. Sua representação, ainda que sua efetividade seja condicionada pela refinação tecnológica, diferente da compreensão humana, é oriunda de uma programação externa. Essa representação continua sendo oriunda de uma mera formalização e símbolos desprovidos de significados subjacentes.

Em resumo, podemos dizer que a compreensão humana é dotada de conteúdo e capaz de fazer semântica. A compreensão da máquina é um nome funcionalista dada a sua capacidade de processamento estipulado por um programador humano, e sua expressão e funcionamento se reduzem à sintaxe.

Searle defende que os significados, por serem intencionais, estão na cabeça. O significado, atrelado a um determinado estado psicológico determina a extensão, e esse é o princípio básico da semântica. Essa ideia se põe contrária a ideia de que o significado, em vez de ser psicológico, é determinado em moldes de indexicalidade, na visão de Hilary Putnam.

⁶ Segundo Searle, em *O que é a linguagem*, animais e recém nascidos são dotados de intencionalidade : contém crenças, desejos, intenções e são capazes de reconhecer quando essas atividades mentais são satisfeitas ou frustradas.

Contra o psicologismo e contra a intencionalidade, Putnam defende que a extensão linguística de cada objeto é definida pelas propriedades superficiais que compõem o objeto, no seu contexto de fala e/ou percepção. Por exemplo, a água, em qualquer mundo possível, é um conceito que se aplica a um líquido transparente, composto por um conglomerado de moléculas, por sua vez compostas por um átomo de oxigênio e dois de hidrogênio.

Nas palavras de Searle:

“A estratégia de Putnam é tentar conceber casos intuitivamente plausíveis em que o mesmo estado psicológico determina extensões diferentes. Se estados psicológicos tipo-idênticos podem determinar extensões diferentes, a determinação da extensão deve transcender os estados psicológicos e, portanto, a visão tradicional é falsa.” (Searle, 2002, p.280).

Em um outro exemplo, um tanto mais sofisticado, Putnam explora o fato de que há divergências entre faias e olmos para demonstrar que os estados psicológicos são insuficientes para determinar a extensão. A percepção dos tipos de árvores diverge por parte de quem tem um conhecimento de árvore de quem não tem o mesmo conhecimento. Desse modo, o que vai determinar a conceituação de uma faia ou um olmo, visto elas possuem características parecidas, é a percepção e conhecimento de suas particularidades. Essa conceituação parte de um mesmo estado psicológico, mas diverge pela extensão.

Searle rebate essa noção alegando haver um sutil paralelismo entre a indexicalidade de Putnam e a intencionalidade que ele critica. A indexicalidade, no tocante a extensão, acaba se colocando como uma forma de intencionalidade, na medida em que o significado é pressuposto a partir do seu contexto de uso. A intencionalidade estabelece certas condições para a satisfação de determinadas extensões. A indexicalidade, nesse sentido, é pressuposta mediante as experiências perceptivas, essas são suas condições de satisfação. Ou seja, o conteúdo da extensão carrega em si uma intencionalidade cuja representação varia de acordo com as propriedades. Um olmo tem conteúdo representacional e uma faia tem outro conteúdo, e ambos são intencionais.

A intencionalidade, portanto, se concebe como uma condição para a asserção semântica e tal como Searle demonstrou, ela é pré-linguística e natural. Desse modo, máquinas não fazem semântica, por carecerem desse aspecto natural da consciência e, conseqüentemente, serem desprovidas de uma linguagem intencional. Suas representações ocorrem ao nível sintático e, desse modo, não correspondem à capacidade humana de ter compreensão e atribuírem significado ao conteúdo da sua programação, pois os significados estão na cabeça dependem de conteúdo intencional.

CONSIDERAÇÕES FINAIS

Podem as Máquinas Pensar?

Por fim, para respondermos a essa questão, temos que levar em conta em que dimensões os conceitos “máquinas” e “pensar” se aplicam. Começando por “pensar”, até aqui, conduzi uma minuciosa discussão acerca da consciência, para demonstrar que ela é qualitativa e é um processo que ocorre no cérebro, por interação causal entre neurônios e conexões sinápticas. Logo, para pensarmos em pensamento, devemos levar em conta essas características.

Há uma série de equívocos acerca do que seria pensar e o principal deles decorre de uma redução funcional da atividade a algo que pode ser representado a nível de sintaxe. Como vimos, a sintaxe não é suficiente para embasar algum tipo de pensamento, sobretudo, quando se leva em conta questões de nível tecnológico. Os PLNs mais arcaicos demonstram que não há muita apuração nem mesmo na composição sintática de certos programas, principalmente naqueles que requerem uma atitude mais sensível e deliberativa, como uma narração. E mesmo as tecnologias mais sofisticadas, ainda não são capazes de expressar em que nível, comparando com a capacidade intencional humana, uma máquina poderia ser capaz de atribuir significados ao conteúdo “pensado”.

Se tomarmos a definição de máquinas, podemos encontrar certos equívocos para responder à pergunta. Há uma dimensão generalista acerca do conceito. Se definirmos máquinas por um instrumento pelo qual se processam e veiculam pensamentos formais, então a resposta à questão pode ser sim, uma vez que, em determinado contexto, os seres humanos processam pensamentos formais. Essa é uma definição conveniente para os adeptos do cognitivismo. Os cognitivistas defendem a ideia de que as atividades cerebrais estão muito bem amparadas pelas ideias funcionalistas de que basta usar de determinada tecnologia para entender a nível de processamento computacional o funcionamento do cérebro.

Para Searle, “Podem as máquinas pensar?” não expressa uma dimensão

adequada de real dúvida. Vimos no tópico acerca da consciência replicada que, para que uma máquina de silício, um computador, possa ter consciência, ela precisa reproduzir as mesmas estruturas causais neurofisiológicas que levam a determinada atividade mental. Nesse sentido, máquinas/computadores podem pensar, desde que haja uma tecnologia que seja eficiente o bastante para remodelar essas estruturas causais e conferir aos estados mentais nas máquinas as mesmas características que os estados mentais em humanos: consciência qualitativa, intencionalidade, subjetividade ontológica, causação mental.

A questão talvez não seja se um computador pode pensar. Se observarmos as disposições funcionalistas e materialistas para eliminar a consciência da jogada, vemos que nem mesmo os defensores do cognitivismo e da IA forte não são capazes de dar uma definição do que seria pensar. Talvez a definição que se preze seja, justamente, a de reduzir a mente como uma máquina computacional e isso não responde à questão. Se computadores são definidos como seres pensantes apenas por carregarem programas adequados, e, a partir desse programa, são capazes de realizarem tarefas que exigem uma dimensão inteligente, então a ausência de consciência, não categoriza computadores e máquinas de Turing como seres pensantes.

Desse modo, a questão correta seria “Pode um computador, em função de ter um programa adequado, ter consciência, capacidade de compreensão?”. E a resposta seria “Não”. O experimento do quarto chinês ilustra como operam as estruturas causais dos programas de computadores. Mostra que a atividade computacional de manipular símbolos não é o suficiente para atribuição de significados. A atividade computacional é meramente sintática; nenhum programa por si só é capaz de fazer semântica, nada além do que a sua própria programação permite.

Para finalizar, a inteligência é um campo vasto, subdividido em categorias como linguística, estratégica, emocional etc. Entretanto, sua asserção, a partir de conceitos naturais e intencionais, são alvo de equívocos. Graças a esses equívocos, vemos a atitude cognitivista de reduzir o fenômenos mentais como programas psicológicos de computador. Segundo essa visão, seríamos todos autônomos meramente probabilísticos tal como defendia Putnam. No entanto, o cognitivismo, respaldado fortemente pelo funcionalismo (e é bem tentador confiar que as funções de

determinados sistemas podem nos fornecer respostas precisas) representa uma digressão no que diz respeito a o estudo científico da consciência que ainda se mostra um problema alvo de complexa discussão, nada suscetível a respostas meramente sintáticas.

REFERÊNCIAS BIBLIOGRÁFICAS

- BODEN, Margaret A., *Inteligência Artificial*, Editora Unesp; 2020
- CAMARGOS, Eric.; *A Inviabilidade do Dualismo Cartesiano na Filosofia da Mente*, Unb; 2023
- COSTA, Cláudio; *Filosofia da Linguagem*; Zahar; 2008
- DESCARTES, René; *Discurso do Método*; Nova Fronteira, 2011
- DESCARTES, René; *Meditações Metafísicas*; Edipro; 2016
- LECLERC, André; *Uma Introdução à Filosofia da Mente*; Appris Editora; 2018
- OLIVEIRA, Victória de; *Uma Introdução à Filosofia da Mente de Gilbert Ryle*; Fi; 2021
- RYLE, Gilbert.; *Categorias*; Os pensadores; Abril Cultural; 1975
- RYLE, Gilbert.; *The concept of mind*; The University of Chicago Press; 1984
- SEARLE, John R.; *A Redescoberta da Mente*, Martins Fontes; 2006
- SEARLE, John R.; *Consciência e Linguagem*; Martins Fontes; 2010
- SEARLE, John R.; *Intencionalidade*, Martins Fontes; 2002
- SEARLE, John R.; *Mente, Cérebro e Ciência*; 70 Biblioteca de Filosofia Contemporânea; 1984
- SEARLE, John R.; *Mente, Cérebros e Programas*, UFScar, 2011
- SEARLE, John R.; *O Mistério da Consciência*, Paz e Terra; 1998
- SEARLE, John R.; *O que é a Linguagem*, Editora Unesp; 2012
- TEIXEIRA, João de Fernandes; *Filosofia do Cérebro*; Paulus; 2018

